



HAL
open science

Rationaliser les pratiques numériques en archéologie : l'exemple des chantiers de fouilles de Bibracte

Quentin Verriez

► To cite this version:

Quentin Verriez. Rationaliser les pratiques numériques en archéologie : l'exemple des chantiers de fouilles de Bibracte. Archéologie et Préhistoire. Université Bourgogne Franche-Comté, 2023. Français. ⟨NNT : 2023UBFCH035⟩. ⟨tel-04546650v2⟩

HAL Id: tel-04546650

<https://theses.hal.science/tel-04546650v2>

Submitted on 8 Apr 2026

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

**THÈSE DE DOCTORAT DE L'ÉTABLISSEMENT UNIVERSITÉ BOURGOGNE
FRANCHE-COMTE PRÉPARÉE À**

l'Université de Bourgogne

Ecole doctorale n°594

Société, Espace, Pratique, Temps

Doctorat d'Archéologie

Par

Quentin Verriez

**Rationaliser les pratiques numériques en archéologie :
l'exemple des chantiers de fouilles de Bibracte**

Thèse présentée et soutenue le 18 décembre 2023 à Besançon

Composition du jury

M. Dominique Garcia	Professeur, Aix-Marseille Université	Rapporteur
Mme Katherine Gruel	Directrice de recherche CNRS émérite, UMR 8546 AOrOc-ENS	Rapporteuse
M. Loup Bernard	Maître de conférences, Université de Strasbourg	Examineur
M. Sébastien Durost	Resp. cellule éditoriale et développement numérique, EPCC Bibracte	Examineur
Mme Réjane Roure	Professeure, Université Paul-Valéry-Montpellier 3	Examinatrice
M. Pierre Nouvel	Professeur, Université de Bourgogne	Directeur de thèse
M. Matthieu Thivet	Ingénieur de recherche, Université de Franche-comté	Co-directeur de thèse

Sauf mention contraire, toutes les illustrations sont de l'auteur

REMERCIEMENTS

Je tiens à remercier toutes les personnes qui ont contribué de près ou de loin à la concrétisation de cette recherche. Je sais que je risque d'oublier certains d'entre vous, mais toutes les discussions, tous les moments de partage ont compté pour l'aboutissement de ce travail.

Ma première pensée va à mes directeurs de thèse, Pierre Nouvel et Matthieu Thivet. Ce travail de recherche leur doit beaucoup. Les opportunités qu'ils m'ont offertes ont fait mûrir ma réflexion scientifique et m'ont amené jusqu'ici. De plus, leurs conseils avisés et leur confiance indéfectible ont été des piliers essentiels qui m'ont permis de repousser mes limites et de mener à bien cette thèse.

Mes remerciements vont ensuite à Vincent Guichard et, par extension, à Bibracte dans son ensemble. Le projet Bibracte Numérique m'a offert un cadre sur mesure pour construire et mener à bien ce sujet, en plus de l'intégrer dans une dynamique de recherche plus large, faite de rencontres toujours riches et stimulantes. Je pense particulièrement à mes collègues Eric Lacombe, Jean-Pierre Girard, Guillaume Reich et surtout Sébastien Durost, pour le temps qu'ils m'ont accordé et la qualité de nos discussions. Je n'oublie pas non plus mes compagnons de route sur PC2, Charlotte Defer, Simon Reymondi, Chiara Martini, Nils Scavone et Oriane Rousselet. Je manque de mots pour exprimer à quel point je leur suis redevable de la confiance qu'ils m'ont accordé. Ils ont fait de ce chantier une expérience remplie de joie, d'entraide et de belles découvertes archéologiques. Qu'ils en soient mille fois remerciés.

Même si j'ai peu fréquenté les locaux, je souhaite également remercier les membres du laboratoire ArTeHiS, qui m'ont toujours réservé un accueil chaleureux, en plus de se montrer disponibles et à l'écoute, à chaque fois que j'en ai eu besoin. Ma gratitude va aussi au laboratoire Chronoenvironnement et la MSHE C.-N. Ledoux, qui m'ont accueilli et m'ont fourni des équipements indispensables durant ces années de recherche.

Pour remonter un peu plus loin dans le temps, je remercie également toutes les personnes qui ont marqué mon parcours professionnel et universitaire. Un merci particulier à Matthieu Poux et Audrey Pranyies, qui m'ont fait pleinement entrer dans le monde de l'archéologie. Merci à Benjamin Clément, à qui je dois beaucoup de mes compétences sur le terrain. Merci à tous mes collègues de chez Antea-Archéologie, dont j'ai également beaucoup appris. Merci

à Philippe Barral et Jean-François Piningre d'avoir supporté mes premières tentatives dans le monde de la recherche.

Certains éléments de cette recherche n'auraient pas non plus pu voir le jour sans Christophe Tufféry et Bertrand Bonaventure. Ils m'ont permis de tester leurs outils, en plus de partager leurs expériences lors de plusieurs discussions déterminantes pour mon travail.

Plusieurs personnes m'ont aussi accordé un temps précieux pour un travail que je n'ai malheureusement pas pu aboutir, par manque de temps. Je remercie donc Delphine Champeaux, Valérie Bel, Samuel Brassaud et Vincent Franchini.

Ce manuscrit n'aurait probablement pas la même allure sans les relectures assidues de plusieurs personnes. Un grand merci à Albane Rossi, Rebecca Perruche, Juliette Hantrais, Julien Soichet, François-Xavier Simon et Eloïse Vial, pour leurs conseils, leur écoute et leur aide précieuse durant cet exercice.

J'ai également une pensée pour toutes les personnes qui ont fait de mon séjour dans le Morvan un souvenir chaud à mon cœur. La famille Goujon, pour la qualité de son accueil et pour les nombreux moments de partage. Taloula Colas, pour sa douceur, son écoute et sa confiance indéfectible dans les moments les plus compliqués. Xavier Degron, parce qu'il n'y a pas deux endroits dans le monde comme les Morins. Mattei Georgiu, qui a toujours pris le temps et à qui l'aboutissement de cette thèse doit beaucoup. Et enfin Anaïs Lachambre, car avec toutes nos discussions, nos séances de motivation, nos fous rires et nos larmes, cette thèse est aussi un peu la sienne.

Je tiens également à remercier mes amis Chiara Piai, Sylvain Gayet, Manu Hamon, Amandine Angeli, Kévin Charrier, Christelle Sanchez, Émilien Estur, Jessy et Edwige Glesson, Kévin Jourda et tous les autres, qui m'ont tant de fois permis de prendre du recul et de traverser les moments de doute.

J'ai une pensée particulière pour ma famille, qui m'a toujours encouragé dans mes choix d'orientation. Je réalise aujourd'hui la chance qu'ils m'ont offert en me permettant de pouvoir me consacrer pleinement à mes études et mon travail. Cette accomplissement est aussi le leur.

Et bien sûr Anna, pour tout. Pour ton soutien, pour ta patience, pour ton aide précieuse et tant d'autres choses.

SOMMAIRE

REMERCIEMENTS.....	4
INTRODUCTION.....	11
PARTIE 1 – Le numérique en archéologie : définition, état de l’art et cadre de l’étude	15
1.1.Les chaînes opératoires numériques en archéologie.....	16
1.1.1.Le numérique : définition.....	16
1.1.2.Le numérique en archéologie	18
1.2.Le programme de recherche du mont Beuvray : Bibracte comme cadre d’expérimentation méthodologique.....	44
1.2.1.De J.-G. Bulliot aux fouilles récentes.....	45
1.2.2.Le programme Bibracte Numérique	65
1.2.3.La nécessité de rationaliser les données numériques archéologiques de terrain	77
PARTIE 2 – Les éléments constitutants du numérique en archéologie : données, science ouverte et Plan de Gestion de Données	81
2.1.Les données numériques en archéologie.....	82
2.1.1.La donnée : définition.....	82
2.1.2.Les types de données	86
2.2.Culture et savoir libre : la démarche <i>Open</i> en archéologie.....	92
2.2.1. <i>Open</i> : le savoir libre	92
2.2.2. <i>Open science</i> : la science ouverte.....	95
2.2.3. <i>Open access</i> : le libre accès, ou accès ouvert, aux publications scientifiques...	96
2.2.4. <i>Open data</i> : les données ouvertes.....	98
2.2.5. <i>Open Source</i> et <i>Free software</i> : deux visions du logiciel libre.....	101
2.2.6.Le libre et la loi : protéger les œuvres de l’esprit et leurs auteurs	102
2.2.7.L’archéologie : une science ouverte ?	105
2.3.Plan de Gestion de Données.....	113
2.3.1.État de l’art.....	113
2.3.2.Création des plateformes de Plan de Gestion de Données	115
2.3.3.Diffusion du Plan de Gestion de Données en Europe.....	116

2.3.4. Diffusion du Plan de Gestion de Données en France	118
2.3.5. Plan de Gestion de Données en archéologie.....	119
PARTIE 3 – Un chantier archéologique numérique à Bibracte.....	123
3.1. Drones : nouveaux vecteurs, nouvelles échelles.....	124
3.1.1. Objectifs.....	124
3.1.2. Choix des équipements.....	124
3.1.3. Acquis des expérimentations	130
3.2. Appareils photo numériques : nouveaux capteurs, nouvelles résolutions	134
3.2.1. Objectifs.....	134
3.2.2. Choix des équipements.....	141
3.2.3. Acquis des expérimentations	145
3.3. Photogrammétrie : nouvelle technique, nouveaux (re)levés.....	151
3.3.1. Objectifs.....	151
3.3.2. Choix et réglages des équipements.....	165
3.3.3. Acquis des expérimentations	183
3.4. Système d'Information Géographique : nouveau logiciel, nouveau relevé de l'US au plan de synthèse.....	187
3.4.1. Objectifs.....	187
3.4.2. Choix des équipements.....	199
3.4.3. Acquis des expérimentations	204
3.5. Outils de saisie : nouvelles interfaces, nouveaux relevés attributaires	211
3.5.1. Objectifs.....	211
3.5.2. Choix des équipements.....	212
3.5.3. Acquis des expérimentations	217
3.6. Bilan sur les expérimentations numériques à Bibracte	223
3.7. Diffusion des données primaires d'un chantier archéologique : l'exemple de la domus PC2	225
3.7.1. Présentation de la donnée	225
3.7.2. Présentation du modèle de Plan de Gestion de Données MASA	226
3.7.3. Le PGD de la fouille de la <i>domus</i> PC2 : retour d'expérience	231
3.8. Le développement des chantiers archéologiques numériques à Bibracte	233

PARTIE 4 – Repenser l'archéologie numérique : analyse et recommandations.....	235
4.1.L'humain e(s)t la machine : critique de l'usage des technologies numériques en archéologie	236
4.1.1.La place de la machine pour l'archéologue	237
4.1.2.Le poids environnemental des technologies numériques	241
4.1.3.La technique comme solution miracle ?.....	245
4.1.4.Le futur de l'archéologie numérique : vers une pratique raisonnée	246
4.2.Recommandations pour des données archéologiques conformes à la science ouverte.....	250
4.2.1.Matériels et logiciels	250
4.2.2.Les pratiques.....	255
4.2.3.La science ouverte en archéologie : un horizon inaccessible ?	276
4.3.L'évolution des Humanités Numériques en archéologie française.....	277
4.3.1.HAL science ouverte	277
4.3.2.ArkeoGIS.....	278
4.3.3.Chronocarto	282
4.3.4.Conservatoire National des données 3D.....	285
4.3.5.Bilan sur les plateformes de dépôt.....	289
CONCLUSION ET PERSPECTIVES.....	293
LISTE DES FIGURES	307
LISTE DES ABRÉVIATIONS.....	313
ANNEXES	319
LISTE DES ANNEXES	384
BIBLIOGRAPHIE.....	385

INTRODUCTION

Depuis une vingtaine d'années, les méthodes appliquées à l'archéologie font l'objet de transformations rapides, dû à l'essor des technologies numériques dans les pratiques quotidiennes des chercheurs. Cette période correspond à l'arrivée sur le marché de la micro-informatique, qui a largement participé à populariser les équipements informatiques, en proposant notamment des ordinateurs personnels à des prix abordables. Ce phénomène a franchi une nouvelle étape il y a une dizaine d'années avec la démocratisation de techniques de représentation numériques telles que la photogrammétrie et les Systèmes d'Informations Géographiques (SIG). Les technologies numériques modifient la collecte, le traitement, la gestion, la conservation et la diffusion des données archéologiques. Grâce aux formats numériques, les volumes de données manipulés et échangés ont considérablement augmenté, en plus de se complexifier à chaque avancée technologique. Cette situation est d'ailleurs loin d'être propre au domaine archéologique, et touche l'ensemble du monde scientifique, voire des sociétés humaines. La transformation des pratiques par les technologies numériques est si marquée que certains n'hésitent pas à la qualifier de révolution (Kitchin 2014).

Face à cette montée en puissance du transit des données numériques s'est développé le mouvement pour la science ouverte. Ce dernier promeut l'idée de l'ouverture et du partage des connaissances, à l'aide de dispositifs numériques, en tant que principe fondamental de la recherche scientifique. Il met ainsi en avant des solutions pérennes pour diffuser les ressources de la recherche. Dans cette optique, la science ouverte représente un enjeu important pour la discipline archéologique, car cette dernière a pour caractéristique de produire une quantité de données significative à conserver sur le long terme. En effet, il s'agit d'une

science qui se fonde avant tout sur la contextualisation de ses objets d'étude, les sites et le mobilier archéologiques. Les archéologues déduisent des hypothèses à partir d'un faisceau d'indices matériels fragmentaires, les vestiges archéologiques, afin de construire leurs interprétations. Or, la qualité de ces hypothèses s'établit sur la mise en réseau des indices matériels, dans leurs contextes proches (sur le site de découverte) et à plus large échelle (régionale, aire culturelle, sphère d'influence entre aires culturelles). La comparaison des données pour caractériser les vestiges est donc un des fondements de l'analyse archéologique. Pourtant, la méthode de la fouille archéologique impose le plus souvent de détruire méthodiquement et irrémédiablement les sites étudiés. Les comparaisons ne sont alors possibles qu'à partir des multiples enregistrements textuels et graphiques produits par les fouilleurs.

L'arrivée de technologies numériques en archéologie permettent de multiplier les méthodes de levé, de description et de représentation des vestiges, et ainsi, de consolider les enregistrements archéologiques. Dans le même temps, le partage de ses enregistrements peut renforcer les processus de comparaison des données archéologiques et ainsi améliorer la qualité des interprétations formulées par les archéologues. Cependant, l'absence de normes de production et de partage peut aussi rapidement mener les archéologues à ne plus trouver ni pouvoir exploiter des données numériques qui pourraient s'avérer utiles pour leurs recherches. La science ouverte défend donc un ensemble de principes qui permettent de créer et de diffuser des données transparentes, souples et structurées, de telle manière à conserver des informations exploitables par les chercheurs, et ce quel que soit le volume de données interrogées.

Dans ce contexte se pose la question de l'intégration des données de fouilles numériques dans une démarche de science ouverte. Autrement dit, comment l'archéologie peut-elle exploiter au mieux les avantages du numérique tout en répondant aux principes de la science ouverte ?

L'Établissement Public de Collaboration Culturel (EPCC) Bibracte, qui porte l'ensemble du projet autour de l'oppidum éponyme, offre un environnement idéal pour aborder cette problématique. Il s'agit en effet d'un acteur singulier de l'archéologie française, qui a la particularité de rassembler une grande partie de la chaîne opératoire archéologique. Il assure à la fois la gestion d'un site archéologique, la capitale du peuple gaulois Éduen au Ier s. av. n. è. sur le mont Beuvray ; d'un centre de recherche européen qui poursuit un programme de recherche international depuis 1984 ; d'un musée labellisé "Musée de France" ; d'une maison d'édition qui gère la collection "Bibracte", qui porte sur les recherches du mont

Beuvray et l'archéologie de l'Europe celtique. Il accueille enfin dans ses locaux un centre de conservation et d'étude archéologique (CCE) du ministère de la Culture. Aussi, Bibracte s'est construit comme un véritable laboratoire méthodologique pour l'archéologie, grâce à la diversité des acteurs du programme de recherche et la pluralité des spécialités exercées. De ce fait, il se place parmi les pionniers de l'informatisation des pratiques archéologiques en France.

Cependant, Bibracte a récemment fait le constat que son infrastructure numérique avait atteint une forme de palier. Bien qu'elle continue d'assurer la gestion et l'archivage des données du programme de recherches, elle ne répond plus pleinement aux besoins numériques émergents des archéologues. De plus, elle ne permet pas de créer un flux de diffusion d'informations archéologiques en accord avec les principes de la science ouverte. Le centre et plusieurs de ses partenaires se sont donc regroupés autour d'un projet ambitieux qui consiste à *“développer un écosystème archéologique numérique, c'est-à-dire un terrain d'expérimentation et d'apprentissage des nouvelles pratiques du numérique dans les différents champs d'activités couverts par le site (archéologie, mais également muséographie, éducation, tourisme) pour faciliter la rencontre, la communication, l'échange et la transmission entre les acteurs, internes, mais également externes, qui fréquentent le site de Bibracte”* (Ayache et al. 2018).

Ce sujet de thèse s'intègre pleinement dans ce projet, sous la forme d'une Cifre (Convention industrielle de formation par la recherche) établie entre Bibracte EPCC, l'UMR 6298 ArTeHiS et l'Université de Bourgogne, en se concentrant sur l'étape à l'origine de la chaîne : la fouille archéologique. L'objectif de notre travail de recherche est de concevoir, d'expérimenter, de valider et de documenter une chaîne opératoire numérique libre sur le chantier de la domus PC2, une vaste maison romaine au cœur de l'oppidum de Bibracte. En d'autres termes, il est question d'éprouver la faisabilité d'une fouille fondée uniquement sur l'utilisation de logiciels libres et de produire des données transparentes, structurées et aux formats ouverts dans le cadre d'une campagne archéologique, depuis la fouille jusqu'à la publication des résultats. Nous avons ainsi codirigé quatre campagnes de fouilles sur la domus PC2, de 2017 à 2020, afin de construire des méthodes d'enregistrement et d'analyse libres et innovantes, adaptées aux conditions réelles d'un chantier archéologique.

Ce travail s'articule en quatre parties. Nous nous sommes demandé en premier lieu si l'utilisation des technologies numériques s'inscrivait réellement dans une forme de rupture méthodologique en archéologie. Nous avons ainsi dressé un état des lieux de ces usages, qui

permet de mettre en évidence la multiplicité des pratiques numériques des années 1960 à aujourd'hui. Il en ressort qu'un des enjeux actuels de l'archéologie numérique réside dans la nécessité de rationaliser les pratiques numériques en appliquant les principes de la science ouverte.

Face à la multiplicité des pratiques, il nous est apparu nécessaire de revenir dans la deuxième partie sur certains éléments fondamentaux. Il est question de définir clairement ce que représentent les données numériques et la science ouverte, et leurs places actuelles en archéologie.

La troisième partie représente le cœur de ce travail de recherche, car après avoir dressé les contextes historique et théorique, elle décrit le cadre pratique de nos quatre années de mise en œuvre d'outils numériques sur le chantier de la domus PC2. Cette partie reprend la même structure pour chaque outil : un rapide état de l'art et une définition des objectifs attendus pour la mise en service de l'outil ; une justification du choix des dispositifs numériques employés ; une critique des avantages et des limites quant à l'utilisation de l'outil sur une fouille archéologique.

Toutes les réflexions et les manipulations menées dans le cadre de cette thèse nous ont permis d'exercer un regard critique sur l'archéologie numérique. C'est pourquoi la quatrième partie propose dans un premier temps de s'éloigner des aspects techniques pour prendre également en compte les dimensions sociales et environnementales de l'utilisation de technologies numériques en contexte archéologique. La partie s'appuie ensuite sur ce constat pour livrer, dans un second temps, un ensemble de recommandations afin de rationaliser les pratiques numériques en archéologie, en anticipant la production de données archéologique en adéquation avec les objectifs de la science ouverte dès la phase de terrain.

PREMIÈRE PARTIE

Le numérique en archéologie : définition, état de l'art et cadre de l'étude

1.1. Les chaînes opératoires numériques en archéologie

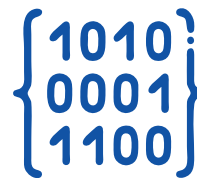
Avant d’aller plus loin dans l’étude des usages technologies numériques en archéologie, il est essentiel de préciser ce que nous entendrons par le terme “numérique” dans ce manuscrit.

1.1.1. Le numérique : définition

Selon l’Académie française¹, l’adjectif numérique, au sens technique du terme, “*se dit, par opposition à analogique, du codage, du stockage, de la transmission d’informations ou de grandeurs physiques sous forme de chiffres ou de signaux à valeur discrète (ou discontinue)*”. En informatique, les chiffres ou signaux se traduisent par le langage binaire, composé de 1 et de 0 (binary digit en anglais), qui se trouve à la base de n’importe quelle opération logique, du simple calcul mathématique à la restitution 3D (fig. 1). Mais avec l’explosion de la micro-électronique et la mise en réseau des infrastructures dans le tournant des années 2000, la dimension du numérique (ou *Digital* en anglais) dépasse maintenant largement le spectre de l’informatique et des ordinateurs. Des concepts aussi larges que les nouvelles technologies (smartphone, réalité virtuelle, intelligence artificielle, etc.) ou les nouveaux modes de communication (Internet, réseaux ou médias sociaux) sont couramment assimilés au numérique (Moatti, 2012, 2016).



Analogique
Papier



Numérique
Binaire

1. Les supports d’informations : analogique ou numérique

1 <https://www.dictionnaire-academie.fr/article/A9N0816> ; consulté le 19/10/2023.

Quand il reprend sa forme d'adjectif, il n'est pas rare de le voir associé à des termes comme "transition", "transformation" ou encore "révolution numérique". Les expressions ainsi composées, même si elles restent très vagues, sous-entendent que nous nous trouvons face à un phénomène croissant et globalisé, à la manière de la "révolution néolithique" et de l'émergence pour les civilisations humaines de l'agriculture et de la sédentarisation. Ne parle-t-on pas d'ailleurs de plus en plus de "culture numérique" ou "culture du numérique" ? Le numérique représente alors un ensemble d'innovations techniques qui, depuis le milieu du XXe siècle, a permis d'accroître de manière exponentielle les interactions machine-humain, machine-machine et humain-humain, au point de bouleverser les codes socio-économiques pour créer une véritable rupture dans les comportements et les pratiques des groupes humains ayant accès à ces technologies (Monnoyer-Smith, 2017).

Il est en tout cas certain que le secteur du numérique est en constante évolution, ce qui explique la difficulté à en donner une définition exhaustive et explicite. Il nous faut toutefois tenter d'en proposer une, pour la bonne compréhension de ce manuscrit. Sous le terme "numérique", nous désignerons plusieurs éléments :

- les informations à base de caractères binaires (les données) ;
- les dispositifs qui les interprètent (les outils matériels ou logiciels) ;
- les procédés qui les exploitent (les méthodes et protocoles) ;
- les canaux par lesquels les informations circulent (les réseaux).

1.1.2. Le numérique en archéologie

Des années 1960 à 1990

Dans les sciences humaines, les archéologues ont toujours été parmi les plus enthousiastes de l'informatique (Cacaly 1989, p. 148). Les avancées significatives dans le domaine ont régulièrement suivies un changement de paradigme scientifique, eux-mêmes influencés par les progrès techniques du matériel et des logiciels (Giligny 2011).

Dès les années 1950, les archéologues ont exploré les possibilités de la mécanographie pour la gestion et la structuration de l'information archéologique, sous l'impulsion de chercheurs comme J.-C. Gardin (Gardin 1976 ; Gardin 1979). Les machines devenaient capables de réaliser des opérations logiques pour traiter, classer ou effectuer des recherches dans les données et la documentation à l'aide de cartes perforées. Le formalisme descriptif nécessaire à la gestion documentaire mécanographique a rendu possible l'émergence des premières banques de données informatisées avec l'apparition de l'ordinateur dans les années 1960 (Giligny, Desachy 2019). Alors appelés calculateurs, les machines pouvaient occuper des salles entières. C'est pourquoi elles étaient essentiellement concentrées dans les universités.

L'utilisation de ces machines en archéologie est liée au développement du courant de pensée de l'archéologie processuelle (ou *New Archaeology*), apparu dans le monde anglo-saxon dans les années 60-70 (Giligny 2011). À l'inverse de l'approche historico-culturelle alors en place, qui consistait à rapprocher dans le temps et dans l'espace les objets archéologiques les plus marquant pour définir des cultures archéologiques, l'archéologie processuelle recherchait à prendre en compte l'intégralité des faits sociaux, technologiques et culturels pour appréhender les civilisations passées. Pour cela, de nombreuses méthodes d'analyses quantitatives (graphiques, mathématiques, statistiques) ont été mis en place, notamment grâce aux nouvelles capacités de stockage et de traitements des ordinateurs.

Dans les années 70-80, les travaux de classification et d'inventaire ont continués, avec la démultiplication des banques de données en archéologie (les futures bases de données). Par exemple, le Ministère de la Culture lance en 1978 SIGAL 1, un inventaire général informatisé des sites archéologiques Français. Cette période connaît aussi un transformation dans les pratiques de terrain. La création de l'archéologie de sauvetage va mener aux décapages de très grandes surfaces, entraînant logiquement le besoin de traiter de grandes quantités de

données dans des temps relativement restreints. L'usage de l'informatique s'est donc renforcé à la fin des années 1970, notamment grâce à l'arrivée sur le marché des micro-ordinateurs, des machines dédiées à un usage personnel, bien plus compactes et abordables que les grands calculateurs accessibles seulement dans les centres de recherche. De fait, la décennie 1980 a opéré un véritable tournant dans le développement de l'informatique, en archéologie comme dans beaucoup d'autres domaines (Giligny, Desachy 2019). Si les pratiques ne sont pas nouvelles (gestion documentaire, banque de données, traitement mathématiques, statistiques, *etc.*), elles prennent une toute autre ampleur. Les archéologues, équipés de leur propre matériel, vont se lancer dans le développement massif de bases de données, "*dans tous les domaines et sur tous les sujets*" (Chaillou 2003, p. 34). La plupart étaient d'ailleurs des autodidactes, ce qui a eu pour conséquence d'éloigner une partie de la discipline des méthodes en science des données. Plusieurs outils d'enregistrement historiques ont été conçus à cette époque, tels que Syslat pour la fouille de *Lattara* ou le système CENTRAR à l'origine de l'actuel système de gestion et d'archivage du centre archéologique européen de Bibracte (Buchsenschutz *et al.* 1986).

La décennie suivante a véritablement marqué l'entrée de l'informatique comme un outil du quotidien pour l'archéologie. L'utilisation des logiciels de bureautique s'est progressivement banalisée, tout comme l'habitude d'emprunter des outils informatiques à d'autres disciplines, telles que les solutions graphiques pour le dessin assisté par ordinateur (DAO). À l'époque, cette informatisation des pratiques a eu pour effet d'améliorer le contenu et la forme des rapports et publications dans la discipline archéologique (Bellan 2010). Un outil en particulier à toutefois marqué la fin des années 1990 en archéologie numérique : le Système d'Information Géographique (SIG). En alliant les possibilités offertes par les bases de données, la cartographie numérique et l'analyse spatiale, les "*SIG, pensait-on, allait permettre de tout faire, de résoudre tous les problèmes*" (Chaillou 2003, p. 35).

Les années 2000

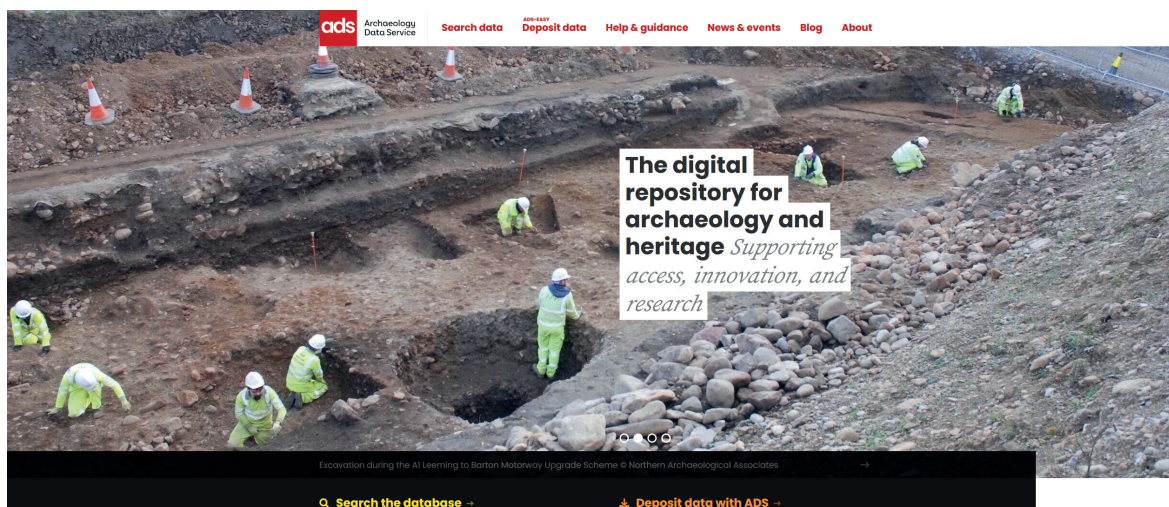
Il est complexe, si ce n'est impossible, de présenter en détail l'évolution du numérique en archéologie tant elle a été rapide après le changement de millénaire. En effet, l'informatisation massive des métiers de l'archéologie dans les années 1990 a été renforcée par l'arrivée du réseau Internet au tournant des années 2000. Les nouvelles opportunités techniques offertes par l'informatique en termes d'automatisation, de classification et de partage ont participé à inscrire les dispositifs numériques (ordinateurs, appareils photo numériques, tablettes, *etc.*) dans l'ensemble de la chaîne opératoire archéologique.

Ainsi, nous proposons de livrer ici une image non pas exhaustive, mais représentative de l'archéologie numérique contemporaine, en nous appuyant sur l'expérience concrète de deux structures emblématiques du début du XXI^e siècle : l'*Archaeology Data Service* (ADS) au Royaume-Uni et l'Institut national de recherches archéologiques préventives (Inrap) en France.

L'exemple de l'Archaeology Data Service

L'ADS est le service d'archives nationales des données numériques archéologiques et patrimoniales du Royaume-Uni (Richards 2017, p. 227) (fig. 2). Il a été fondé le 1^{er} octobre 1996 à l'université de York (Richards 2002, p. 344). Il faisait alors partie des cinq services qui composent l'*Art and Humanities Data Service* (AHDS), un rassemblement des structures d'archivage dans les domaines de l'art et des sciences humaines de plusieurs universités britanniques. L'ADS est donc le plus ancien service d'archives de données archéologiques et patrimoniales numériques au monde (Richards 2021). Dès le départ, son objectif était de “*collecter, décrire, cataloguer, préserver et fournir un support pour la réutilisation des données numériques produites durant les recherches des archéologues britanniques*” (Richards 1997). La volonté d'ADS n'était donc pas de créer un simple entrepôt qui centralise et conserve toutes les données archéologiques du pays, ce qui représentait déjà un réel défi en soi, mais bien de profiter du potentiel des technologies numériques pour optimiser leur réutilisation. Pour cela, plusieurs décisions judicieuses ont été prises, qui ont permis à l'ADS de devenir une référence mondiale dans la gestion des données archéologiques numériques.

Tout d'abord, le service a fait le choix de construire une plateforme transdisciplinaire, capable d'interagir avec d'autres services similaires, sur le territoire comme à l'international (Richards 1997). L'ambition de l'ADS était de devenir la porte d'entrée vers de multiples sources de données, plutôt que de s'imposer comme lieu de stockage unique. Ce type de modèle évite de dupliquer inutilement des données et permet une économie d'échelle significative : chaque service mis en réseau reste garant de l'intégrité des données hébergées, de son infrastructure et de son personnel. Il a ainsi permis d'assurer la viabilité d'ADS, avec des moyens limités – à son lancement, le service n'accueillait que deux personnes, avec 60 000 £ de frais de fonctionnement par an – tout en lui offrant une grande visibilité (Richards 2017, p. 227).



2. Page d'accueil du portail de l'Archaeology Data Service

Ensuite, les outils, les services et les plateformes interrogés par l'ADS avaient besoin d'utiliser un langage commun pour faire communiquer des données provenant de sources différentes. Il devait permettre de localiser les données sur le réseau et de décrire à la fois leurs contenants et leurs contenus. Ce langage commun a été trouvé avec la technique des métadonnées, littéralement des données à propos des données (*cf. infra* 2.1.2. Métadonnées) (Wise, Miller 1997). Elles se structurent selon différents standards et L'AHDS, et par extension l'ADS, a très tôt fait le choix d'utiliser celui des *Dublin Core*. L'initiative était alors relativement récente, mais elle est par la suite devenue la norme mondialement reconnue pour le partage transdisciplinaire de ressources numériques (ISO 2009).

Enfin, face à la durée de vie assez courte des programmes informatiques qui permettaient d'exploiter les fichiers archivés, le service a choisi de mettre en avant les formats ouverts (*cf. infra* 2.2.1. Open : le savoir libre) (Richards 2002, p. 350-351). En effet, il semblait impossible de maintenir chaque version des logiciels ou systèmes d'exploitation couramment utilisés par la communauté archéologique, en vue de la réutilisation future d'un fichier. À l'inverse, la conversion des fichiers dans des formats standardisés, simples et lisibles par tous demandait un investissement certain au moment du dépôt, mais assurait leur viabilité sur le long terme.

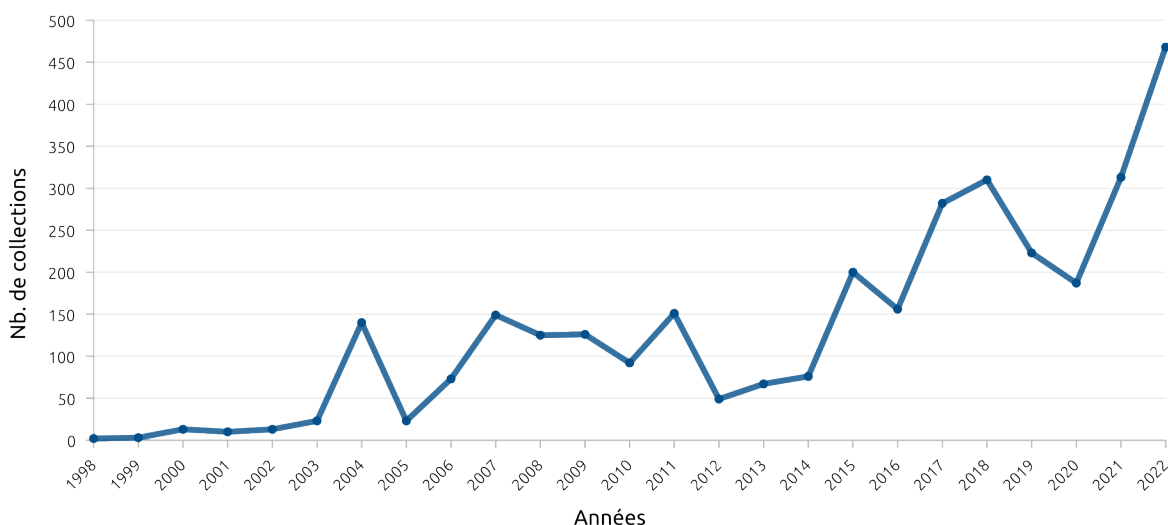
Dans la pratique, l'ADS s'est organisé autour de trois domaines d'activité (Richards 2002, p. 347). D'abord, un service de développement des collections, qui négociait l'accès aux données auprès des institutions tierces et assistait les créateurs de données dans la production et le dépôt d'archives. Ensuite, un service était affecté à la conservation. Son rôle était de

maintenir et de développer les systèmes de versement des données de l'ADS. Il s'occupait également de la validation, de la gestion et de la préparation des données pour leur conservation sur le long terme. Enfin, un dernier service encourageait et aidait à la réutilisation des données numériques par les utilisateurs.

Pour l'archivage, il est nécessaire de bien comprendre l'enjeu de la gestion de la donnée archéologique numérique à l'époque. Pour cela, il suffit de constater le soutien qui a été offert à l'ADS par les principaux organismes de financement de la recherche dès sa création. En effet, ils ont été plusieurs à recommander, voire imposer, l'ADS comme structure dépositaire des données de projets subventionnés dès le milieu des années 1990, alors même que le service était en pleine construction (Richards 1997 ; Richards 2002, p. 347). Afin de s'accorder avec les différents jeux de données envisagés pour le dépôt, l'ADS a fait le choix de travailler par collections. Une collection correspond à un ensemble de données cohérent, tel que des inventaires ou des référentiels, mais aussi la documentation de fouilles ou d'autres productions des programmes de recherche en archéologie et histoire ancienne. L'archivage a commencé concrètement en 1998, avec deux collections, mais très vite les dépôts se sont multipliés, pour atteindre régulièrement plus d'une centaine par an à partir de 2007 (fig. 3). Cela ne signifiait pas pour autant que l'ADS acceptait n'importe quelles données. L'archivage a un coût, tant en personnel qu'en matériel. Bien que la structure fût financée par l'AHDS, du moins dans les premières années, les ressources n'étaient pas pour autant illimitées. L'accent a donc été mis sur la conservation de données numériques à fort potentiel de réutilisation, en portant une attention particulière à la qualité et l'exhaustivité des fichiers archivés (Richards 2002, p. 347). Si la réutilisation d'un document ne présentait que peu d'intérêt, l'ADS conseillait d'ailleurs de le conserver en format papier, quitte à le numériser plus tard en cas de besoin.

Le service de réutilisation des données a été officiellement lancé le 15 septembre 1998, avec la mise en ligne du catalogue *ArchSearch*² (ADS 1999, p. 8-9). À l'époque, le portail interrogeait déjà les métadonnées de plusieurs bases de références, telles que *Royal Commission on the Historical Monuments of England* ou *Sites and Monuments Record du West of Scotland Archaeology Service*. Il était alors associé à un moteur de recherche relativement simple, autorisant la recherche d'un terme ou d'une expression, mais qui permettait déjà d'identifier des ressources archéologiques à travers l'Écosse et l'Angleterre, et ce depuis une simple page web. Le catalogue sera rapidement amélioré, avec l'ajout d'un outil de requête cartographique courant 2001, qui comportait une recherche par coordonnées ou selon un

2 <https://www.archaeologydataservice.ac.uk/archsearch/browser.xhtml> ; consulté le 20/10/2023



3. Évolution du nombre de collections par années

périmètre circulaire. Enfin, une refonte complète de la base de données a été réalisée à la fin de la même année, pour lancer *ArchSearch II* (ADS 2001, p. 5-6 ; ADS 2002, p. 6). Le nouveau système proposait une amélioration sémantique importante, en offrant la possibilité de réaliser des requêtes plus complexes, avec de nouveaux champs de recherche, la combinaison de champs, l'affinement d'une première requête, ou encore la possibilité d'exporter le résultat de la recherche dans un fichier tabulé. En mars 2000, *ArchSearch* a été complété par *ADS Library*³, une compilation de références bibliographiques de livres, monographies, revues, articles, rapports et documents conservés par l'ADS ou ses partenaires (ADS 2001, p. 10). L'ensemble se présentait déjà sous la forme de fiches inventaires, qui renvoyaient vers la ressource accompagnée d'un lien, si cette dernière était disponible en ligne. La bibliothèque contenait également une section pour les publications qui n'étaient plus édités, mais considérés d'intérêt général, disponibles au format PDF.

Néanmoins, inventorier les documents publiés n'était pas suffisant. Déjà à l'époque, la majorité des informations archéologiques primaires se trouvaient dans les rapports de fouilles rarement publiés. L'ADS, avec le soutien de l'*English Heritage*, a donc initié en 2000 le projet OASIS⁴ (*Online Access to the Index of Archaeological Investigations*) afin de créer un

3 <https://archaeologydataservice.ac.uk/library/index.xhtml> ; consulté le 20/10/2023

4 <https://oasis.ac.uk/> ; consulté le 20/10/2023

outil en ligne pour répertorier la littérature archéologique grise (ADS 2000, p. 9) (fig. 4). Là encore, l'objectif de l'ADS était de créer une organisation collaborative, qui mettait en relation les services patrimoniaux du Royaume-Uni, plutôt que d'assumer seuls la gestion des rapports non publiés (Hardman 2009 ; Evans, Gilham 2021, p. 2). Le projet a été officialisé le 1er avril 2004, avec la mise en place, d'abord en Angleterre, d'un formulaire web pour recenser les rapports d'opération et déposer l'archive PDF associée (ADS 2004, p. 12). Une fois le formulaire rempli par un auteur authentifié, le document au format numérique était automatiquement catalogué et rendu disponible par *ArchSearch*, si le dépositaire n'avait pas défini d'embargo.

OASIS
The online system for reporting archaeological investigations and linking research outputs and archives

archaeol20-514926: Data from a Magnetometer Survey at Culham Battery Storage Site (Phase 2), Culham, Oxfordshire, 2023
©Archaeological Surveys Ltd

Update: September 2023

The OASIS Dashboard

OASIS V now live in Scotland

Guidance for linking large-scale projects or phases of work

Funded by

Historic England

HISTORIC ENVIRONMENT SCOTLAND

ÀRAINNEACHD EACHDRAIDHEIL ALBA

SCOTLAND'S ARCHAEOLOGY STRATEGY

4. Page d'accueil du service d'accès en ligne à l'index des fouilles archéologiques du Royaume-Uni (OASIS)

Ce dispositif, a priori basique, a néanmoins fonctionné. À la fin des années 2000, le formulaire OASIS était disponible dans tout le Royaume-Uni et enregistrait plus de 100 dépôts par mois (Hardman 2009). Dans les faits, l'association des dépôts OASIS et du moteur de recherche cartographique sur *ArchSearch* a permis de créer l'équivalent des services de carte archéologique des Services Régionaux de l'Archéologie (SRA) en France, mais avec un accès direct aux informations en ligne. Enfin, toujours pour faciliter la collecte de données

archéologiques et historiques, le service a déployé en 2013 *ADS-easy*, un nouvel outil de dépôt d'archives semi-automatisé (ADS 2013, p. 34 ; Moore *et al.* 2014) (fig. 5). Cette méthode a rendu possible pour le déposant de gérer seul le dépôt, depuis un outil en ligne, sans l'intervention de l'équipe d'ADS. Il permet ainsi de limiter les frais pour les petits et moyens dépôts, avec un système de charges fixes, plus un supplément par fichier (fig. 6). Actuellement, si le dépôt dépasse 1000 fichiers, ou des fichiers de plus de 100 Mo, l'ADS impose de gérer le dépôt directement avec leur équipe, afin d'assurer la bonne gestion de la collection.

ads Archaeology Data Service

ADS Main Website Help | Login

[COSTING CALCULATOR](#)

ads easy

ADS-easy is a faster way to deposit small to medium-sized project archives.

ADS-easy will allow:

- depositions that include **certain data types** (i.e. databases, documents and reports, geophysics, GIS, maps and plans, photographs and images and spreadsheets)
- files found in the [list of accepted formats](#)
- files with a **maximum size of 100MB**
- deposits with a **maximum of 1000 files**
- a **geophysical survey of less than 50Ha**

If you are uncertain about whether you should use ADS-easy please [contact us](#) and we can talk to you about the options. [Find out more about ADS-easy.](#)

Login / Register

[Login to ADS-easy](#)

[Register for ADS-easy](#)

[Forgotten your login password?](#)

Browsers

Please note that this site works best with the Firefox, Chrome and Safari browsers.






Help

For more information about using ADS-easy, help with using ADS-easy and frequently asked questions please [click here.](#)

Scheduled Maintenance

Please note that for this service there is currently a scheduled maintenance period: 2nd Tuesday of the month, 6:30 - 7:00am. The website may be unavailable during this period.

University of York Legal Statements | Website Terms and Conditions | Cookies | Privacy Policy | Accessibility Contact Us

5. Page d'accueil du service de dépôt autonome ADS-easy

Type de données	Extension	Quantité	Prix
Base de données	csv	1 fichier	€2.00
	xls	1 fichier	€6.00
	xlsx	1 fichier	€6.00
	odb	1 fichier	€6.00
	json	1 fichier	€10.00
	rdf	1 fichier	€10.00
	txt	1 fichier	€2.00
Documents et rapports	xml	1 fichier	€10.00
	doc	1 fichier	€4.00
	docx	1 fichier	€2.00
	odt	1 fichier	€6.00
	pdf	1 fichier	€6.00
SIG	txt	1 fichier	€2.00
	shp (et fichiers associés)	1 fichier	€6.00
	gml	1 fichier	€2.00
	jpg et jpw	1 fichier	€4.00
	kml	1 fichier	€6.00
	grd	1 fichier	€10.00
Géophysique	tif et tfw	1 fichier	€2.00
Cartes et plans	N/A	1 hectare	€9.85
	dwg	1 fichier	€4.00
	dxf	1 fichier	€2.00
Photos et images	svg	1 fichier	€4.00
	jpg et tif	151 fichiers (150 gratuits)	€1.50
	png	1 fichier	€1.50
	pdf	1 fichier	€6.00
	gif	1 fichier	€1.50
Tableurs	bmp	1 fichier	€1.50
	xls	1 fichier	€6.00
	xlsx	1 fichier	€6.00
	ods	1 fichier	€6.00
	csv	1 fichier	€2.00
Frais de gestion	txt	1 fichier	€6.00
			€150.00
Total (HT)			€309.85
Total (TTC)			€371.82

Prix valables du 1er avril 2023 au 31 mars 2024
(<https://www.archaeologydataservice.ac.uk/easy/costingCalculator.xhtml>)

6. Exemples de prix pour le dépôt de fichiers grâce au service ADS-easy

En amont de l'archivage et de la réutilisation, l'AHDS et l'ADS ont aussi toujours mis l'accent sur la formation des producteurs de données. Leur philosophie pourrait être résumée ainsi : plus la création d'une archive est anticipée, plus le dépôt est simplifié et le potentiel de réutilisation est multiplié. C'est pour cela que tous les services de l'AHDS se sont donné pour objectif de créer une série de guides de bonnes pratiques, des volumes thématiques destinés aux producteurs de données (Richards 1997). Cinq volumes ont ainsi vu le jour entre 1998 et 2002, traitant de la documentation de fouilles, des données géophysiques, du SIG, du DAO et même de la réalité virtuelle (Mitcham *et al.* 2010). Chacun de ces volumes prodiguait des conseils pratiques pour faciliter l'administration, la préservation et la réutilisation des données, ainsi que la création de métadonnées. Autrement dit, les guides n'imposaient pas de méthode d'enregistrement, ils indiquaient plutôt comment documenter la donnée existante, dans le cas où elle avait vocation être préservée et réexploitée. Le but de l'ADS a toujours été de faciliter les échanges d'information, sans limiter les techniques de production. Il s'agissait donc de définir des formats d'échanges standards, pour les fichiers et les métadonnées, qui puissent s'adapter au maximum de types de données produites par les archéologues durant leurs recherches (Richards 2017, p. 229). Par la suite, pour suivre les développements méthodologiques et technologiques, l'ADS et son homologue américain, *Digital Antiquity*, ont entrepris une importante mise à jour des guides à la fin des années 2000. Sur le fond d'abord, en rectifiant les volumes déjà existants et en proposant du nouveau contenu, mais surtout sur la forme, avec la fin des éditions papier. La nouvelle génération de guides⁵ a été conçue sur la base technologique du *wiki*, une application web qui permet la création et la modification collaborative de pages d'un site web, à la manière de la célèbre encyclopédie libre Wikipédia. Ce nouveau format était plus à même d'évoluer avec la discipline, tout en permettant aux lecteurs de naviguer très simplement entre les différents thèmes abordés.

En parallèle des guides, l'ADS a multiplié, d'année en année, les participations à des ateliers ou des conférences, pour présenter à la communauté scientifique son propre développement, les services proposés, mais également pour former les futurs utilisateurs (ADS 1998 à 2021). Le service a aussi souhaité toucher les étudiants, afin d'intégrer la gestion des données numériques comme une pratique courante pour les nouvelles générations d'archéologues. On peut notamment citer l'exemple du projet PATOIS (*Publications and Archives in Teaching : Online Information Sources*) au début des années 2000, qui a visait la création de tutoriels en ligne pour l'analyse de données primaires archéologiques numériques à des-

5 <https://archaeologydataservice.ac.uk/help-guidance/guides-to-good-practice/> ; consulté le 20/10/2023

mination des étudiants (ADS 2000, p. 10 ; ADS 2003, p. 13). Plus tard, l'ADS a continué de multiplier les supports de formation pour atteindre les différents publics, en utilisant entre autres les MOOC (*Massive Open Online Courses*), créés en collaboration avec les universités de Southampton et Leicester (ADS 2014, p. 14).

La revue *Internet Archaeology*, surnommée la “*sister e-journal*” de l'ADS, a également vu le jour en 1996, (Richards 2017, p. 229). En tant que revue entièrement numérique, son but était de renforcer les liens entre publications et jeux de données. Le premier numéro reflétait déjà ces ambitions, avec par exemple l'article de P. Tyers, qui présentait un atlas numérique interactif des amphores romaines en Grande-Bretagne (Tyers 1996). Pour les premiers numéros, les bases de données liées aux publications étaient hébergées sur le serveur d'*Internet Archaeology*. Il a fallu attendre 2001 pour qu'un article renvoie vers une collection hébergée par l'ADS (Richards 2001). Il s'agit d'un rapport papier entièrement retravaillé, où chaque mention d'une donnée primaire est accessible, par un lien, dans la collection dédiée de l'ADS. Une nouvelle étape a été franchie dans la deuxième moitié des années 2000, avec les projets LEAP, qui ont permis la création de huit exemples de publications numériques natives dans la revue, avec des liens dynamiques vers les jeux de données (<https://intarch.ac.uk/leap/>). La publication dans la revue *Internet Archaeology* n'impose d'ailleurs pas de devoir héberger les données par l'ADS. Ainsi l'article de F. Limp *et al.*, qui explore les possibilités de la 3D à travers l'analyse d'objets précolombiens et de l'Égypte antique, interroge directement tDAR, le service d'archive de *Digital Antiquity* aux États-Unis (Limp *et al.* 2011). Enfin, à partir de 2013, *Internet Archaeology* propose de publier des *Data Papers*⁶, des articles évalués par les pairs décrivant précisément des jeux de données (Richards 2017, p. 229).

Enfin, la position de précurseur de l'ADS en matière de gestion de la donnée numérique archéologique les a amené à être à l'origine de nombreux réseaux nationaux et internationaux. Ils proposent en effet de diffuser leur savoir-faire et de leurs méthodes à travers des initiatives comme HEIRNET (*Heritage Information Ressource*) au Royaume-Uni, ARENA, ARIADNE et ARIADNE+ (*Archaeological Research Infrastructure for Archaeological Data Networking in Europe*) en Europe, ou encore TAG (*Transatlantic Archaeology Gateway*) aux États-Unis (ADS 1999, p. 12 ; ADS 2002, p. 14 ; ADS 2010, p. 20 ; Niccolucci, Richards dir. 2019, p. 7-25). Grâce à leur expertise, le portail ARIADNE⁷ est le projet de référence européen de la diffusion archéologique, soit le plus important au monde. Ces réseaux ont

6 <https://intarch.ac.uk/authors/data-papers.html> ; consulté le 20/10/2023

7 <https://portal.ariadne-infrastructure.eu/> ; consulté le 20/10/2023

permis l'émergence de projets de recherche ambitieux, soutenus par l'Union Européenne, avec des retombées directes pour le service. Parmi les plus innovants, on peut citer le projet *Archaeotools* et ses suites logiques, *STAR (Semantic Technologies for Archaeological Ressources)* et *STELLAR (Semantic Technologies Enhancing Links and Linked Data for Archaeological Ressources)*, qui ont eu lieu de 2007 à 2011 (Jeffrey *et al.* 2009 ; Tudhope *et al.* 2011). L'objectif de ces trois projets était de développer la reconnaissance automatique de vocabulaire dans des corpus non structurés, afin de faciliter l'exploitation, l'extraction et la classification des données gérées par l'ADS. Cette technologie a été une des bases de la troisième génération d'*ArchSearch*, qui est désormais capable de lire le contenu de certains documents et fourni une aide à la saisie pour les déposants, avec par exemple la proposition de vocabulaires contrôlés pour le formulaire OASIS (Richards 2017, p. 229). Plus récemment, le service a fait partie du consortium à l'initiative d'*ArchAIDE*, une application de reconnaissance automatique de fragments de céramiques à partir d'images sur smartphone, liée à une base de données évolutive, pour développer et affiner les connaissances typologiques (Anichini *et al.* 2020).

Durant ses 26 années d'existence, l'ADS a donc construit, étape par étape, une des structures les plus avancées au monde sur la gestion et la diffusion de données numériques archéologiques. Nous n'en avons présenté ici qu'une partie, qui illustre le développement des principales missions du service : l'archivage et l'exploitation de données primaires, la formation des chercheurs, la publication des collections, la création de réseaux d'échange. Dans le même temps, l'ADS a toujours su maintenir une recherche active, afin de conserver une offre innovante, en adéquation avec l'évolution des problématiques archéologiques du début du XXI^e s. Le service comporte aujourd'hui 16 membres, sous la direction de Julian Richards, à la tête du service depuis sa création. Le reste de l'équipe se compose principalement de spécialistes des archives numériques, mais aussi de chargés de développement des collections, d'éditeurs, de gestionnaires de projet et de développeurs. Pour garantir son développement, L'ADS a coordonné ou été partenaire d'une quarantaine de projets de recherches nationaux et internationaux. Il administre quotidiennement plus de 25 To de données archivées, soit environ 3500 millions de fichiers et 3100 collections gérées en interne (ADS 2021, p. 7). Avec les nombreux dépôts partenaires, le moteur de recherche *ArchSearch III*⁸ explore plus de 1,3 million de métadonnées, à travers 4 entrées (basique, navigateur, carte et avancé) (fig. 7). *L'ADS Library*, la bibliothèque en ligne, contient 330883 références bibliographiques, de

1770 à 2022, dont 119262 publications hébergées directement par l'ADS, disponibles au format PDF. Parmi elles, la littérature scientifique grise comporte aujourd'hui plus de 60000 titres. Ce chiffre est en constante augmentation, car le portail OASIS enregistre environ 200 nouveaux rapports par mois. La revue *Internet Archaeology* vient également de publier son 60^e numéro (IA 2022), et l'intégralité des contenus est en libre accès depuis 2014 (ADS 2015, p. 10). La publication est gratuite pour les doctorants, et la revue propose une subvention, l'*Open Donation Fund*, aux auteurs n'ayant pas accès à aux aides à la publication (ADS 2018, p. 9). Enfin, les guides de bonnes pratiques offrent une trentaine d'entrées, selon les différentes étapes et techniques de la chaîne opératoire archéologique. Finalement, l'ADS a concrétisé son ambition de départ, qui consistait à bâtir une infrastructure complète et complémentaire, qui rassemble et partage des ressources précieuses pour la discipline archéologique, sans jamais imposer de pratiques, mais plutôt en fournissant des outils adaptés, accompagnés d'un temps de formation significatif.

ads Archaeology Data Service

ADS Main Website Help

BASIC

ARCHSEARCH is an integrated online catalogue indexing over 1.3 million metadata records, including ADS collections and metadata harvested from UK historic environment inventories.

KEYWORD SEARCH

You can limit your search to just ADS archival collections by using the [ARCHIVES](#) search.

ADVANCED MAP BROWSER

University of York Legal Statements | Website Terms and Conditions | Cookies | Privacy Policy | Accessibility Contact Us

ads ARCHAEOLOGY DATA SERVICE

re3data.org
REGISTRY OF RESEARCH DATA REPOSITORIES
http://doi.org/10.17616/R3MW23
Archaeology Data Service

UNIVERSITY of York

CORE TRUST SEAL

7. Archsearch exploite plus de 1,3 million de ressources

Un dernier aspect intéressant à évoquer à propos du service est son modèle économique. Sur les cinq dernières années, l'ADS a présenté un budget compris entre 621 000 £ et 821 000 £. La grande majorité de cette somme, entre 80 et 90 %, est destinée au salaire du personnel, la gestion du parc informatique étant assurée par plusieurs universités britanniques. Le service doit donc disposer de cette somme chaque année, sans monnayer l'accès aux données hébergées. Pour cela, il présente trois sources de revenus principales. La première, anecdotique, se fait à travers des retours d'expertise rémunérés, sur la gestion d'archives numériques. En second, on trouve les projets de recherche auxquels participe l'ADS. Si l'activité de recherche et développement permet de garantir une haute qualité de service constante, les subventions des projets apportent également plus de 75 % du budget annuel de l'ADS. Sans ces projets, qui durent tout au plus quelques années, le service ne pourrait tout simplement plus exister. Enfin, une dernière partie provient du service de préservation. Comme les données sont pour la plupart en libre accès, c'est au déposant de payer pour leur conservation. Ce modèle économique, s'il peut paraître injuste, n'est en fait pour l'ADS que le cheminement logique des données de la recherche. Les résultats scientifiques étant destinés par nature à être réexploités, c'est aux producteurs d'en anticiper la gestion, la conservation et la diffusion. L'ADS souligne d'ailleurs que le modèle est rentable quand il est pensé dans sa globalité. La conservation des données a certes un coût, mais la collecte aussi. Particulièrement quand la donnée est difficilement accessible, comme c'est bien souvent le cas en archéologie. Il est nécessaire de prendre en compte les mois de recherche nécessaire pour rassembler les corpus en amont ou durant les projets pour en évaluer le coût total. Ainsi, anticiper les frais de gestion des données dès l'origine d'un projet permet de construire des archives de fouilles accessibles et réutilisables, soit une réduction du temps de collecte lors de projets futurs, en plus de l'émergence de problématiques innovantes, à des échelles encore rarement atteintes en archéologie.

En conclusion, l'ADS a toujours été capable de s'adapter aux contraintes structurelles et scientifiques, en cherchant toutes les solutions pour dépasser les barrières techniques, financières ou sociales pour la préservation, le partage et la réutilisation des données archéologiques et historiques.

L'exemple de l'Inrap

Historiquement, la création de l'Inrap s'inscrit dans la continuité de l'Association pour les fouilles archéologiques nationales (Afan). Cette association, créée en 1973 sous le régime de la loi de 1901, gérait les financements des fouilles programmées et de sauvetage, et réalisait les opérations archéologiques prescrites, en relation avec le ministère de la Culture. Pourtant, bien qu'il s'agissait d'une avancée importante pour la reconnaissance des vestiges archéologiques comme bien patrimonial, le cadre législatif et les sources de financement restaient trop faibles et incertains. C'est pourquoi l'État a commandé en 1999 une enquête qui a débouché sur la promulgation de la loi du 17 janvier 2001 relative à l'archéologie préventive (JORF 2001). Elle définissait alors cette dernière comme une mission de service public, entendue comme recherche scientifique, ayant *“pour objet d'assurer, à terre et sous les eaux, dans les délais appropriés, la détection, la conservation ou la sauvegarde par l'étude scientifique des éléments du patrimoine archéologique affectés ou susceptibles d'être affectés par les travaux publics ou privés concourant à l'aménagement. Elle a également pour objet l'interprétation et la diffusion des résultats obtenus”* (art. 1). La loi prévoyait également la création d'une redevance pour financer les diagnostics et les fouilles préventives (art. 9), ainsi que la mise en place d'un établissement public administratif national qui héritait des biens, droits et obligations de l'Afan (art. 4). L'Institut national de recherche en archéologie préventive (Inrap) a ainsi été créé le 1^{er} février 2002, sous la double tutelle du ministère chargé de la Culture et du ministère chargé de la Recherche. Le cadre législatif a ensuite rapidement évolué avec la loi du 1^{er} août 2003, en plaçant l'Inrap sur un marché désormais ouvert à la libre concurrence avec des services d'archéologie territoriaux et des opérateurs privés.

La création de l'institut ne s'est donc pas faite *ex nihilo*. À la différence de la majorité des “jeunes” structures privées de l'époque, l'Inrap accueillait dès le départ plus de 1 000 agents, répartis dans différents centres régionaux et engagés sur les nombreux travaux archéologiques de l'Afan. Ainsi, là où les opérateurs privés ont construit des pratiques numériques en fonction de la progression de leurs activités, l'Inrap a dû rapidement imaginer des solutions pour coordonner une organisation d'ampleur nationale qui possédait un historique scientifique, matériel et structurel (Tufféry 2022, p. 105). Les problématiques auxquelles l'établissement a fait face et les décisions prises en matière de développement numérique sont donc particulièrement éclairantes pour comprendre les enjeux d'une telle démarche en archéologie.

Le premier plan d'établissement de 2004-2005 affichait déjà une volonté de structurer les pratiques numériques, avec l'intention de développer une infrastructure réseau généralisée et de fournir le personnel en équipement informatique, ainsi que les services gestionnaires de l'information (Inrap 2007, p. 34-35, 54-55). Cet élan a été soutenu par la mise en place de deux services au sein de l'institut : la Direction Scientifique et Technique (DST) et la Direction des Systèmes d'Information (DSI). Le premier avait pour objectif, entre autres, d'inventorier les méthodes d'acquisition et d'analyse, et d'identifier les leviers pour construire des pratiques communes ; et le second de construire l'infrastructure technique nécessaire.

Parmi leurs premières actions, les deux directions ont mis en place des séminaires autour des pratiques numériques, afin d'ouvrir le débat sur les attentes et les besoins des agents de l'Inrap (Koehler, Tufféry 2012, p. 230-231). Lors de ces rencontres, les participants ont majoritairement exprimé un désir de cohérence dans la mise en œuvre et l'utilisation des solutions numériques de l'institut. En effet, Ch. Tufféry explique dans sa thèse qu'en tant que chargé de mission à la DST, il a recensé plus de 70 outils numériques, obsolètes ou non, utilisés à l'Inrap pour l'enregistrement de données de terrain (Tufféry 2022, p. 125-131). Parmi les plus anciens fichiers consultés, certains dataient de l'Afan, à la fin des années 1990, mais c'est surtout avec l'arrivée des tablettes et des smartphones dans les années 2010 que ces outils se sont démocratisés sur le terrain. En constatant la forte évolution technologique, la capacité d'emport et la baisse des prix de ce type de dispositif, la DST s'est procuré des tablettes en vue de conduire les premières expérimentations en 2009. Après plusieurs années de tests rigoureux, la direction de l'Inrap a validé le déploiement de 60 tablettes pour les activités de terrain en 2014. Parmi les solutions numériques utilisées, on peut notamment citer les exemples de CADoc⁹, Datadiag, ou encore ArcheoDB développés par des archéologues de l'institut (Bryas, Carpentier 2012 ; Bolo *et al.* 2014 ; Tufféry 2022, p. 131). Les programmes ou applications sont souvent le fruit d'initiatives personnelles, plusieurs agents ayant développé des outils pour répondre à leurs besoins spécifiques d'enregistrement, de gestion ou d'analyse (Tufféry *et al.* 2021, § 25). Ch. Tufféry qualifie d'ailleurs la très grande majorité d'entre eux de "*systèmes spécifiques pas ou faiblement adaptables*" : ils fonctionnent uniquement dans le cadre de certains types d'opérations (diagnostics ou fouilles), certains contextes archéologiques (peu stratifié ou dense) ou en accord avec des méthodes et habitudes de travail localisées (Tufféry 2022, p. 129).

La multiplicité des outils utilisés à l’Inrap implique également l’enregistrement de données hétérogènes, tant dans le contenu que dans les formats. Les formulaires des systèmes d’enregistrement ne possèdent pas tous les mêmes champs descripteurs, et ces derniers sont plus ou moins permissifs (par exemple, case à cocher *vs.* texte libre). Le format, quant à lui, est souvent lié au logiciel utilisé pour la conception des systèmes d’enregistrement, tels que FileMaker, Microsoft Access ou Microsoft Excel. Même lorsqu’il s’agit du même logiciel, on peut observer des conflits entre les différentes versions, par manque de rétrocompatibilité des itérations récentes vers les plus anciennes.

Face à ce constat, la DST a œuvré dès la fin des années 2000 pour l’harmonisation des pratiques numériques. Le concept d’harmonisation est d’ailleurs volontairement différencié de l’homogénéisation : *“l’harmonisation respecte des pratiques, prend en compte des solutions techniques préexistantes et vise à les faire communiquer alors que l’homogénéisation vise à imposer de façon unilatérale un nouveau système qui se veut une solution unique adaptée à toutes les situations”* (Koehler, Tufféry 2012, p. 234). En choisissant l’harmonisation, la DST souhaitait respecter au maximum les habitudes de travail, les outils et les méthodes de chacun, en se concentrant plutôt sur les chantiers de normalisation des formats de données et les protocoles d’échange à développer en interne.

Une étude comparative a notamment permis de mettre en évidence que malgré le vocabulaire variable et les spécificités d’enregistrement, la plupart des outils utilisaient une base de concepts communs (unité stratigraphique, fait archéologique, structure, mobilier, etc.) (Tufféry, Augry 2019). C’est pourquoi, à partir de 2015, la DSI et la DST ont conjointement lancé le développement d’EDArc (Enregistrement de Données Archéologiques), projet piloté ensuite par Ch. Tufféry. La volonté de départ était de créer un outil en accord avec les principes FAIR (Facile à trouver, Accessible, Interopérable, Réutilisable) (*cf. infra* 2.2.4. *Open data*) , qui visait à enregistrer les données de terrain essentielles, tout en affichant une modularité à même de respecter les particularités des différents contextes et pratiques de fouille. En d’autres termes, c’était à l’outil de s’adapter et de normaliser les enregistrements, quelles que soient les pratiques des utilisateurs, pour harmoniser les données. Nous avons eu l’opportunité de tester une version adaptée d’EDArc dans le cadre de ce travail doctoral, grâce à un partenariat scientifique entre l’Inrap et Bibracte EPCC. Nous reviendrons plus en détail sur l’outil dans une partie suivante, mais pour résumer ici, la version actuelle d’EDArc répond en grande partie aux ambitions techniques de départ (*cf. infra* 3.5. Outils de saisie). À partir de l’expérience du chantier des jardins de la cathédrale du Mans, Ch. Tufféry et

le responsable d'opération S. Augry notent qu'avec les formations et l'accompagnement suffisant, les automatisations de tâche apportées par le support numérique permettent de simplifier et d'améliorer la qualité et la gestion des enregistrements sur un chantier archéologique (Tufféry, Augry 2019).

À ce jour, la stratégie de l'harmonisation est toujours en application à l'Inrap, puisque EDArc n'est pas imposé et plusieurs systèmes d'enregistrement numérique de terrain continuent de coexister aux côtés de l'enregistrement papier (Seng *et al.* 2020 ; Bel *et al.* 2021 ; Tufféry 2022, p. 127-128). Après plusieurs années d'expérimentation, il s'est avéré que le support informatique était apprécié par les agents ouverts à ce type de dispositif, particulièrement si les moyens en matériel, formation et accompagnement étaient suffisants. Les bénéfices soulignés sont souvent l'automatisation de tâches répétitives, les fonctions de contrôle et de navigation dans les données, ou encore le fait de ne pas devoir numériser l'information en post-fouille (Bolo *et al.* 2014, p. 248-251 ; Augry 2018 ; Tufféry, Augry 2019). En laissant le choix de la méthode d'enregistrement à ses agents tout en déployant ses efforts sur la mise en commun des données, l'Inrap a su mettre en place une politique adaptée, capable d'évoluer avec la diversité de problématiques scientifiques, de pratiques et de contraintes techniques que rencontre une structure de cette ampleur.

Cependant, l'introduction du numérique n'a été pas sans conséquence sur les pratiques des archéologues sur le terrain (Tufféry 2017 ; Tufféry 2019 ; Tufféry *et al.* 2021). Au-delà des contraintes engendrées par le passage de l'enregistrement sur papier aux terminaux numériques, il existe aussi des aspects plus pragmatiques, comme le transport d'un nouveau dispositif tout au long de la journée sur le chantier par l'archéologue. L'usage de tablettes est d'ailleurs loin d'être majoritaire à l'Inrap (Tufféry 2022, p. 127). Plusieurs retours d'utilisateurs insistent sur la nécessité de mettre en place, sans imposer, des formations en amont des chantiers et un accompagnement au quotidien sur le terrain lorsque ce type d'équipement est proposé, afin de prévenir tout clivage pronumérique/anti-numérique au sein des équipes. Enfin, Il faut noter la durée de vie relativement courte du matériel numérique sur les fouilles. Sur les 200 tablettes acquises jusqu'à maintenant par l'Inrap, plus de la moitié sont hors services, avec une période de fonctionnement moyenne de 3 à 4 ans.

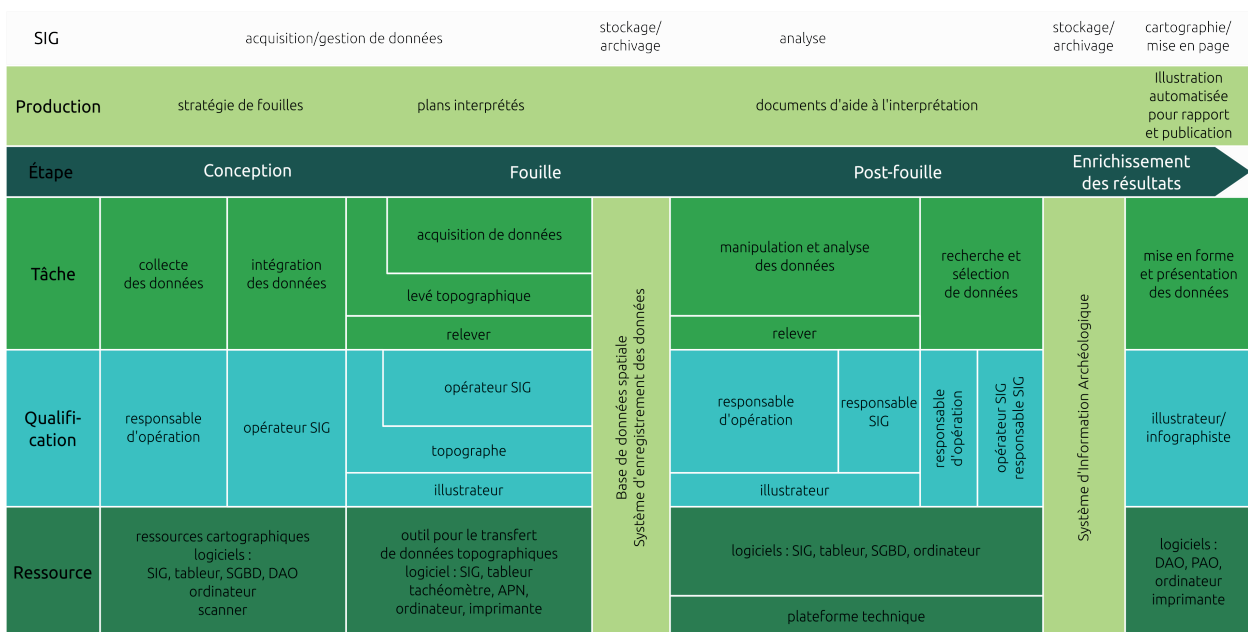
Plus récemment, l'essor rapide de l'imagerie numérique, et particulièrement de la photogrammétrie, est venu appuyer la transformation des méthodes d'enregistrement des fouilles archéologiques, à l'Inrap comme dans la discipline en général (Eusèbe 2019, p. 4-10). Les expérimentations ont débuté dans les années 2010 et ont très rapidement démontré qu'avec le

matériel et les compétences adéquates, le relevé 3D par photogrammétrie se révélait particulièrement adapté aux conditions d'enregistrement d'une opération d'archéologie préventive (Belarbi *et al.* 2012, p. 90-91 ; Gaime 2018). Avec des rendus hautement fidèles à la réalité pour une précision souvent centimétrique, la production d'orthoimages s'est ainsi largement démocratisée, parfois couplée à d'autres techniques comme la lasergrammétrie ou la tomodensitométrie (Augry 2018). En plus de fournir des illustrations inédites, ces nouvelles représentations sont principalement utilisées en tant que support, pour les relevés de tout type (plan, coupe, élévation) et pour la restitution de niveaux dans les SIG. Une flotte de 25 drones a d'ailleurs été déployée en 2016, et autant de topographes de l'Inrap ont reçu la formation de télépilote pour assurer les missions de photogrammétrie sur le terrain (Tufféry 2022, p. 163). Les approches sont aujourd'hui multiples et, comme le démontre la diversité des techniques présentées lors du séminaire scientifique et technique "Archéologie : Imagerie numérique et 3D", les archéologues de l'Inrap maîtrisent suffisamment ces outils pour adapter leur utilisation aux nombreux sujets d'étude rencontrés par l'archéologie préventive (Eusèbe *et al.* dir. 2019).

Toutefois, face à l'apparente simplicité des logiciels, Ch. Tufféry remarque que les utilisateurs ne maîtrisent pas toujours suffisamment la technique de la photogrammétrie pour produire des orthophotographies répondant aux exigences scientifiques (précision, reproductibilité, etc.) (Tufféry 2022, p. 167-168). De plus, certains auteurs, tels que S. Gaime et S. Eusèbe, alertent sur le fait qu'une image, même d'excellente qualité, ne remplace ni l'observation directe des vestiges ni l'interprétation qui peut en être faite par les archéologues (Gaime 2018 ; Eusèbe 2019, p. 10-14). D'autant plus que, selon eux, ces nouveaux modes de représentation ont tendance à accélérer l'abandon progressif des règles de sémiologie graphique archéologique, déjà amorcé avec l'arrivée de l'informatique et des logiciels de DAO et de CAO. La question se pose aussi de la représentation 3D dans le discours scientifique, car celle-ci reste pour le moment sous exploitée par les archéologues, par manque de compétence ou d'outils adéquats. Enfin, ces mêmes auteurs s'interrogent sur la conservation de ce type de fichier sur le long terme, particulièrement gourmand en espace de stockage.

L'implémentation des Systèmes d'Information Géographique a aussi été un large chantier, soutenu de longue date par l'institut. La réflexion a débuté en 2005, avec la commande d'une première expertise de la DST et du laboratoire CITERES-LAT UMR 6173 à Tours, pour délimiter les utilisations possibles, dans le cadre des missions de l'Inrap (Inrap 2007, p. 53 ; Rodier *et al.* 2013). Dès lors, une dynamique s'est installée pour évaluer l'impact

scientifique, logistique et économique du développement de l'utilisation des SIG, à l'échelle des opérations de diagnostic ou de fouilles préventives. Un premier rapport établi en 2006 a proposé un protocole d'utilisation des SIG en 4 phases, de la création du projet archéologique à la production de documentation accompagnant la publication et la diffusion d'une opération. L'objectif était de développer l'usage du SIG au sein de l'institut, afin que les archéologues l'utilisent pour "traiter leurs données de la même façon qu'ils utilisent la pelle pour creuser" (Rodier et al. 2013). Le protocole détaillait ainsi trois domaines nécessaires à sa mise en œuvre opérationnelle : les tâches scientifiques et techniques à effectuer, le personnel qualifié et leur temps alloué et les ressources informatiques indispensables (fig. 8). Il a été testé et validé en 2009 sur 7 opérations menées par le Centre archéologique de Tours, en démontrant l'apport du SIG pour la prise de décision en contexte d'archéologie préventive. Le mouvement s'est ensuite élargi avec la création d'un programme de formation pour les agents, maintenu sur plusieurs années et toujours en cours (Inrap 2010, p. 62 ; Inrap 2012, p. 40 ; Inrap 2014, p. 50). Ainsi, en 2015, plus de 600 agents étaient formés à l'utilisation des SIG, même s'ils restaient principalement utilisés en post-fouille, pour l'analyse et les besoins de production cartographique (Inrap 2016, p. 77).



8. Chaîne opératoire analytique pour le déploiement des SIG à l'Inrap (d'après Rodier et al. 2011)

Si l’Inrap a encouragé la production de données spatiales, il était logique de prévoir une solution pour centraliser et archiver les informations manipulées par les agents. C’est ainsi qu’a été conçu CAVIAR (Catalogue de visualisation de l’information archéologique) en 2017, un système de gestion des données spatiales de l’Inrap sur l’ensemble du territoire (Trébuchet *et al.* 2017 ; Marx *et al.* 2021 ; Tufféry 2022, p. 112-113). La base de données enregistre les emprises des opérations archéologiques de l’Inrap et les données descriptives associées, qui s’organisent en 6 grands thèmes (“prescription”, “ouverture”, “unité d’observation”, “axe”, “geo” et “us”). Elle est consultable à l’aide d’un logiciel de SIG par un flux WMS ou en ligne via l’application Iliad pour le personnel de l’institut.

Avec le SIG, l’Inrap a ainsi su capitaliser sur une technologie pertinente et mettre en place une stratégie nationale sur le long terme pour la formation de son personnel et le développement d’une infrastructure normalisée. L’établissement maîtrise maintenant une base de données spatiale à l’échelle nationale, actualisée en flux continu par du personnel qualifié, selon le rythme des activités de terrain. La production de cette entrée cartographique généralisée se révèle particulièrement adaptée aux pratiques des archéologues en contexte préventif, tant par l’exploitation des données des opérations passées que pour la contextualisation des futurs travaux.

En plus de CAVIAR, le système d’information de l’Inrap est complété par deux autres outils de centralisation et de partage des données. Le premier, DOLIA, est un catalogue numérique qui assure le signalement, le référencement et la localisation et l’hébergement de la documentation scientifique conservés par les différents centres de l’établissement (Bryas *et al.* 2013, p. 75) (fig. 9). Il a été continuellement alimenté depuis sa mise en ligne en 2009, et les agents de l’Inrap pouvaient accéder en 2021 à plus de 42 000 références de littérature grise et d’ouvrages scientifiques, ainsi qu’aux fichiers lorsqu’ils étaient disponibles au format numérique (Marx *et al.* 2021). Pour les utilisateurs extérieurs à l’Inrap, les notices sont consultables en accès libre depuis 2012, ainsi qu’une sélection de rapports disponibles au téléchargement. Le second outil est COMODO, qui assure la gestion et la localisation des collections archéologiques et de la documentation associée.

Avec DOLIA, COMODO et CAVIAR, l’Inrap a développé une infrastructure numérique de gestion des productions scientifiques qui présente plusieurs avantages. Elle est spécialisée, car chaque outil est construit pour s’adapter aux caractéristiques spécifiques des différents types de données (bibliographie, collection archéologique et information spatiale). Tous ces outils communiquent entre eux, d’abord pour limiter les erreurs liées aux doubles entrées,

The screenshot shows the Dolia digital catalog interface. The header includes the Dolia logo and the text 'Catalogue des fonds documentaires et bibliothèque numérique de l'Inrap'. The search results are displayed in a table with columns for 'Type document', 'Titre propre', 'Sous titre', 'Titre partie', 'Auteur', 'Auteur moral', 'Editeur', and 'Date'. The search criteria are 'Question : Titre, auteur, sujet, résumé = [bibracte%]'. The results list 12 items related to Bibracte, including archaeological sites, publications, and reports.

Type document	Titre propre	Sous titre	Titre partie	Auteur	Auteur moral	Editeur	Date
1	Vestiges archéologiques en milieu extrême					Monum, Editions du Patrimoine	2003
2	Situlae, images d'un monde disparu	[Livret de l'exposition du Mus		Kern Anton, Directeur du département de		Bibracte	2008
3	Les processus d'urbanisation à l'âge du Fer	Eisenzeitliche Urbanisationspr		Guichard Vincent, Directeur général de B		Centre archéologique européen du Mont Be	2000
4	Diagnostic archéologique préalable à l'extension du parking du Musée d	rapport de diagnostic		Bataille Gérard		Inrap GES	2008
5	Bibracte	Archéologie d'une ville gauloi		Romero Anne-Marie		Centre archéologique européen	2006
6	Saint-Léger-sous-Beuvray « La Croix du Rebut » (Saône-et-Loire)	extension du parking du musée		Bataille Gérard		Inrap GES	2010
7	Les monnaies gauloises et romaines de Toppidum de Bibracte			Gruel Katherine		Bibracte	2007
8	La Gaule Lyonnaise			Ferdière Alain		Editions Picard	2011
9	Gallia Lugdunensis	Eine Römische Provinz im Herze		Ferdière Alain		Verlag Philipp von Zabern	2011
10	Le forum en Gaule et dans les régions voisines			Bouet Alain		Ausonius Editions	2012
11	Reconditionnement, inventaire et conservation sélective du mobilier a	Mâlain, La Boussière (Côte-d'O		Labonier Emmanuel		Inrap GES	2017
12	Revue archéologique de l'Est	volume annuel. Tome 65				Société archéologique de l'Est	2016

9. Exemple de recherche du sujet "bibracte" sur le catalogue Dolia

mais surtout pour naviguer et exploiter une importante base de connaissances de façon naturelle, pratique pour l'humain ; et automatisée, adaptée aux machines. Enfin, l'infrastructure ne dépend pas de service ou de technologie externe en ce qui concerne sa maintenance ou son développement. Il en résulte une infrastructure scientifique numérique à la fois souple et robuste. Sa souplesse est liée à sa capacité à évoluer au rythme des progrès technologiques, mais aussi des problématiques scientifiques et des pratiques méthodologiques des chercheurs. Elle est en même temps robuste, car les compétences sont présentes en interne pour régler les problèmes d'ordre technique. Ces deux derniers points sont évidemment dépendants des décisions futures de l'Inrap pour maintenir le savoir-faire informatique et les moyens logistiques nécessaires en interne sur le long terme.

En parallèle de la construction de son infrastructure numérique, l’Inrap a progressivement intégré plusieurs programmes nationaux et internationaux qui œuvrent pour la gestion et la diffusion de données patrimoniales numériques. Elle rejoint d’abord en 2014 le programme ARIADNE, le projet d’infrastructure européenne pour la gestion des données archéologiques numériques piloté par l’ADS (Inrap 2015, p. 60). Par la nature de ses activités, l’Inrap apportait son expertise en archéologie préventive et la gestion des données spécifiques à ce champ disciplinaire. Ensuite, en tant que seule partenaire française, elle avait pour rôle de définir les besoins des utilisateurs en matière de portail européen sur le territoire. Enfin, toujours dans le cadre d’ARIADNE, plusieurs tests d’appariement ont été menés entre EDArc et l’extension CRM*archaeo* du CIDOC-CRM, une ontologie pour décrire des données archéologiques (Tufféry *et al.* 2017). L’un des objectifs de cette ontologie est de créer un métalangage capable de faire communiquer des bases de données archéologiques qui ne présentent pas les mêmes modèles conceptuels de référence, en associant les champs de chaque base à des classes normalisées définies par le CIDOC-CRM. L’Inrap a également participé à ARIADNEplus (2019-2022), la suite du programme ARIADNE, avec notamment la publication sur le portail en ligne de toutes les notices de rapports hébergées sur DOLIA, en respectant les normes du CIDOC-CRM (Tufféry 2022, p. 114-118). Enfin, elle fait partie des membres principaux qui dirigent l’action COST SEADDA¹⁰ (*Saving European Archaeology from the Digital Dark Age*) qui, face à la numérisation massive des données archéologiques en Europe et ailleurs, explore des solutions informatiques pour les conserver sur le long terme.

Au niveau national, l’institut a d’abord intégré en 2016 le consortium Couperin.org, un regroupement d’établissements de l’enseignement supérieur et de la recherche française, qui s’appuie sur les compétences des spécialistes de l’information (documentalistes, archivistes, etc.) pour construire un réseau de partage des publications scientifiques en accès libre (Inrap 2017, p. 77). Elle a également rejoint la même année le consortium MASA (Mémoire des Archéologues et des Sciences de l’Archéologie), le groupement de spécialistes des données numériques archéologiques soutenu par l’Infrastructure de Recherche “étoile” (IR*) Huma-Num. Dans ce cadre, l’Inrap a poursuivi les expérimentations d’appariement entre EDArc et le CIDOC-CRM*archaeo*, ainsi que le partage de son expérience sur les Plans de Gestion de Données pour produire le premier modèle national adapté aux données archéologiques¹¹ (*cf. infra* 2.3. PGD) (Tufféry, Augry 2019). Enfin, des agents de l’Inrap participent

10 <https://www.seadda.eu/> ; consulté le 20/10/2023

11 https://dmp.opidor.fr/public_templates?search=MASA ; consulté le 20/10/2023

au consortium 3D-SHS, lui aussi soutenu par l'IR* Huma-Num, qui se fixe pour objectif de définir des standards et des recommandations pour l'usage des données 3D, ainsi que pour mettre en place des outils et plateformes répondant aux besoins des chercheurs en SHS (Granier *et al.* 2019). Ainsi, par ses multiples implications, l'Inrap est aujourd'hui reconnu par la communauté archéologique internationale comme l'interlocuteur principal sur la question des pratiques numériques en France.

L'IR* Huma-Num

Depuis quelques années, l'Infrastructure de Recherche "étoile" Huma-Num¹² joue un rôle prépondérant dans le développement des pratiques numériques en Sciences Humaines et Sociales (SHS) à l'échelle française et européenne. Son nom signifie *Humanités Numériques*, la traduction littérale du concept de *Digital Humanities*. Un collectif de chercheurs réunis à Paris en 2010 dans le cadre d'une THATCamp (*The Humanities and Technology Camp*) le définissait ainsi (Mounier 2010) :

1. *"Le tournant numérique pris par la société modifie et interroge les conditions de production et de diffusion des savoirs.*
2. *Pour nous, les digital humanities concernent l'ensemble des Sciences humaines et sociales, des Arts et des Lettres. Les digital humanities ne font pas table rase du passé. Elles s'appuient, au contraire, sur l'ensemble des paradigmes, savoir-faire et connaissances propres à ces disciplines, tout en mobilisant les outils et les perspectives singulières du champ du numérique.*
3. *Les digital humanities désignent une transdiscipline, porteuse des méthodes, des dispositifs et des perspectives heuristiques liés au numérique dans le domaine des Sciences humaines et sociales."*

Depuis une dizaine d'année, la France et l'Europe encourage l'ouverture des résultats de la recherche, dans le cadre du mouvement de la science ouverte (*cf. infra 2.2. Open*). Ainsi, l'IR* Huma-Num s'est donnée pour mission en 2013 d'accompagner les SHS dans le tournant numérique en leur proposant une infrastructure technique d'envergure internationale¹³, afin de faciliter la diffusion des résultats de la recherche. Elle fournit donc un ensemble de

12 <https://www.huma-num.fr/> <https://www.huma-num.fr/>

13 <https://www.huma-num.fr/quest-ce-que-l-ir-huma-num/> ; consulté le 21/10/2023

solution techniques permettant de développer, de réaliser et de préserver sur le long terme les programmes de recherche, en préservant leurs données et leurs outils. De manière générale, les services proposés par Huma-Num incitent au respect des principes FAIR (*Findable, Accessible, Interoperable, Reusable*, ou Facile à trouver, Accessible, Interopérable et Réutilisable) pour l'ouverture de la donnée scientifique (*cf. infra 2.2.4. Open data*)

En plus des données et des outils, l'IR* œuvre pour la définition de pratiques numériques communes au sein des disciplines, afin d'anticiper le caractère ouvert des données dès les étapes de conception des projets scientifiques. Pour ce faire, Huma-Num a considéré qu'imposer des standards communs ne permettait pas d'apprécier les spécificités propres à chaque discipline. Elle a au contraire préféré offrir des moyens aux différentes communautés SHS afin qu'elles s'organisent en consortiums pour développer des outils et des pratiques adaptés aux problématiques de leurs disciplines et en accord avec la science ouverte.

Deux consortiums Huma-Num sont à prendre en compte pour l'archéologie. Le premier, le plus spécifique, est le consortium Mémoire de Archéologues et des Sciences de l'Archéologie (MASA). Créé à la fin de l'année 2012, il réunit en ensemble d'acteurs spécialistes de la gestion documentaire archéologique. Dans le cadre des principes FAIR, ils supportent le développement de nombreux outils¹⁴ pour favoriser la diffusion et la réutilisation des données de la recherche archéologique. Il comprennent des solutions pour l'alignement de vocabulaire et la mise en relation des jeux de données archéologiques (OpenArcheo, Opentheso, Pactols, CIDOC CRM Game), leur structuration (LogisticWriter), leur diffusion (Modèle de plan de gestion de données MASA, ArkéoGIS) et la création de support de formations (OpenGuide). Le collectif est également à l'origine d'un Livre Blanc qui définit un ensemble de bonnes pratiques pour maîtriser le cycle de vie de la donnée archéologique numérique (Marlet *et al.* 2022).

Le second consortium, fondé en 2014, se nomme 3D SHS. Comme son nom l'indique, il s'est donné pour mission de définir développer des outils et de définir les bonnes pratiques pour les données 3D en SHS, depuis leur création jusqu'à leur diffusion et archivage. Comme MASA, les équipes 3D-SHS ont publié un Livre Blanc en plus de livrer deux outils importants (Granier *et al.* 2019). Le premier, altag3D¹⁵, permet de documenter l'intégralité de la chaîne de production d'un modèle 3D afin de préparer son dépôt sur le second outil, le

14 <https://masa.hypotheses.org/outils> ; consulté le 21/10/2023

15 <https://altag3d.huma-num.fr/> ; consulté le 21/10/2023

Conservatoire National pour la donnée 3D en SHS¹⁶ (CND3D). Ce dernier est aujourd'hui l'entrepôt en ligne de référence pour héberger en ligne et rendre accessible de manière pérenne des objets de la recherche en SHS modélisés en 3D.

État des lieux sur les usages des technologies numériques en archéologies

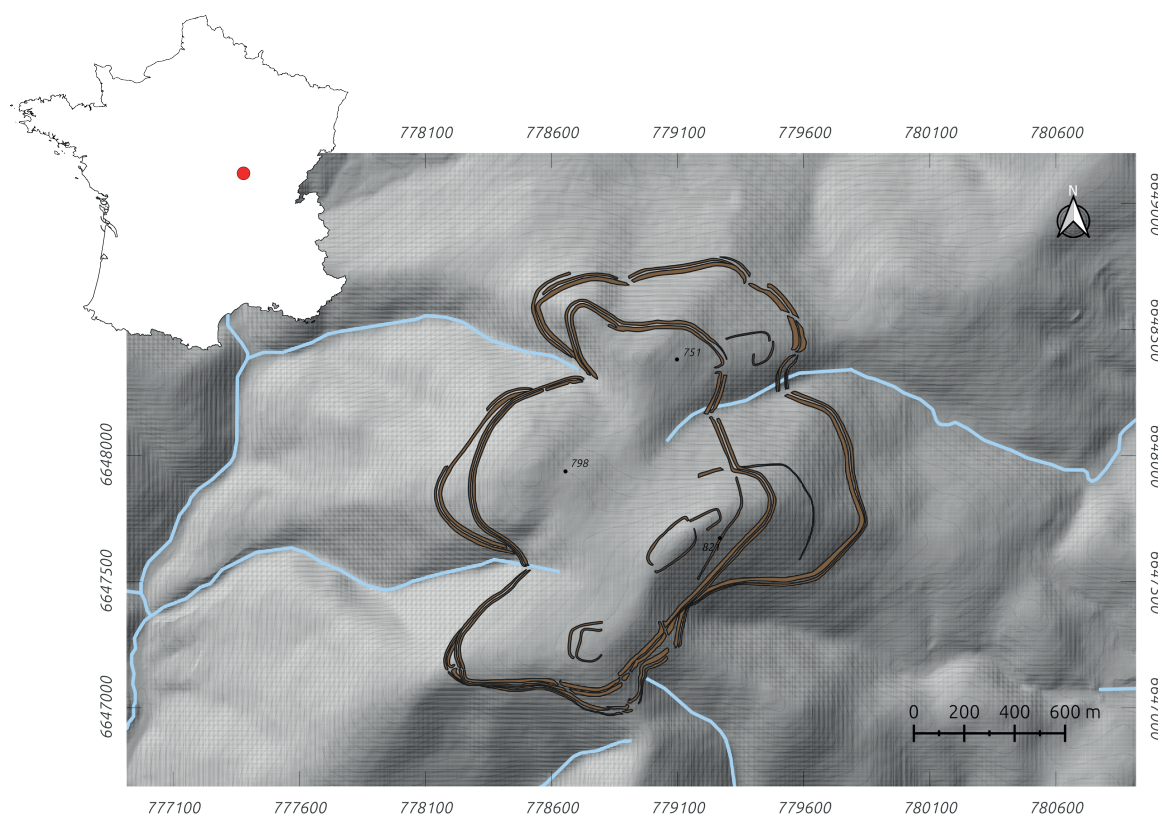
Bien que les pratiques numériques en archéologie remontent aux années 1960, le tournant opéré dans les années 1980 avec l'apparition de la micro-informatique s'est largement concrétisé durant les trois décennies suivantes. Ce développement rapide des technologies numériques dans les méthodes archéologiques a conduit à des adaptations majeures, bien souvent plus en réaction qu'en anticipant les usages. Ils ont ainsi progressivement élargi leurs échelles de travail, depuis les discussions sur les outils jusqu'aux programmes de partages d'informations au niveau européen, pour mieux gérer la production de données et coordonner leurs pratiques numériques. L'émergence de structures comme Huma-Num en France, et le rapide engouement d'une partie des archéologues pour celles-ci, symbolisent le besoin de notre discipline de rentrer pleinement dans l'ère des *Humanités Numériques*.

Le programme de recherche européen sur l'*oppidum* de Bibracte, actif depuis le milieu des années 1980, a suivi le même cheminement. Les recherches se sont construites autour d'un véritable laboratoire méthodologique archéologique, lui aussi riche en expériences dans le domaine du numérique. Dans cette continuité, Bibracte et ses partenaires ont souhaité à leur tour s'insérer dans la dynamique de développement des outils numériques, en prenant en compte les enjeux mis en évidence par les parcours de structures comme l'ADS et l'Inrap. Grâce à son passif, depuis la mise en place du système de gestion et d'archivage de la documentation archéologique jusqu'à développement de la 3D, et en prenant en compte la multiplicité des missions assurées par le centre de recherche, Bibracte apparaît comme le lieu parfait pour penser la chaîne opératoire numérique archéologique des années à venir.

16 <https://3d.humanities.science/#> ; consulté le 21/10/2023

1.2. Le programme de recherche du mont Beuvray : Bibracte comme cadre d'expérimentation méthodologique

Depuis la redécouverte du site dans les années 1860, l'*oppidum* de Bibracte (mont Beuvray, Saône-et-Loire, Nièvre) s'est imposé comme un lieu emblématique de la méthodologie archéologique (fig. 10). À chaque époque, les chercheuses et chercheurs ont mis en œuvre des pratiques avancées pour étudier les vestiges de la capitale Éduenne au 1^{er} s. av. n. è. Le programme de recherche européen, actif depuis le milieu des années 1980, s'est inscrit dans la continuité des travaux anciens en créant autour du site un laboratoire méthodologique archéologique international. Comme pour l'ADS et l'Inrap, les nouvelles pratiques explorées à Bibracte ont impliqué la manipulation de nombreux outils numériques, qui s'est constitué une riche expérience dans le domaine. Nous revenons ici sur cette longue tradition de recherche, afin de comprendre en quoi Bibracte apparaît comme le lieu adapté pour penser la chaîne opératoire archéologique numérique des années à venir.



10. Localisation et topographie générale du mont Beuvray (Saône-et-Loire, Nièvre, France)

1.2.1. De J.-G. Bulliot aux fouilles récentes

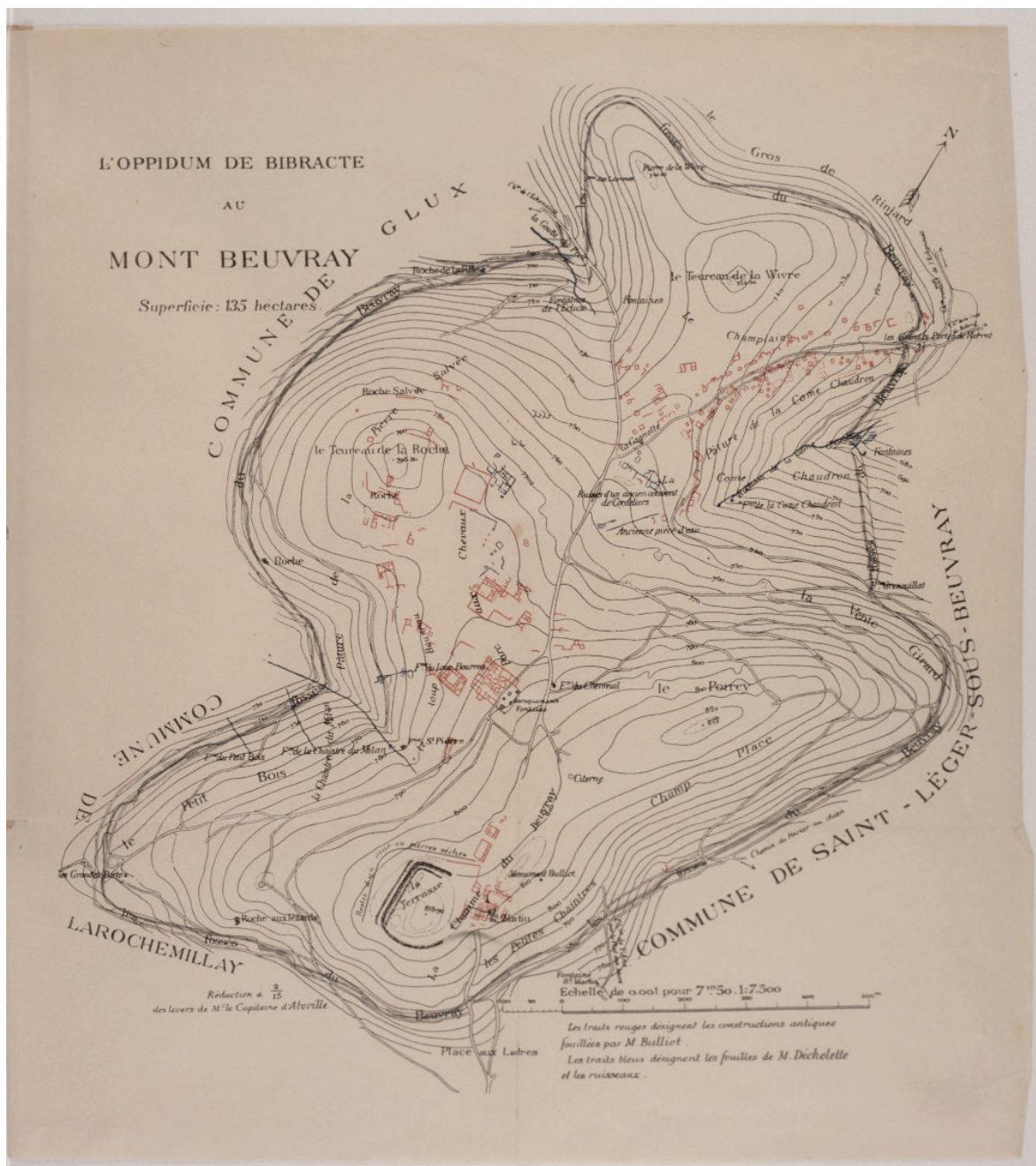
les fouilles anciennes

L'empereur Napoléon III, qui était fasciné par la figure historique de Jules César, a commandité dans les années 1860 une série de campagnes de fouilles pour identifier les principaux lieux évoqués dans les Commentaires sur la guerre des Gaules (Rafowicz 2017). Des moyens financiers exceptionnels pour l'époque ont été attribués afin d'explorer les sites des batailles césariennes comme le plateau de Gergovie, Alise-Sainte-Reine ou le mont Beuvray. Sur ce dernier, une première série de travaux exploratoires ont été lancés en 1864, avant que le président de la Société Éduenne des Lettres, Sciences et Arts, J.-G. Bulliot, soit officiellement mandaté en 1867 pour l'exploration systématique de la montagne morvandelle (Garenne 1867 ; Fochesato 2020, p. 55). Considéré aujourd'hui comme l'inventeur du site, lui et ses équipes sont intervenus pendant 30 ans, jusqu'en 1895, sur la quasi-totalité des secteurs qui caractérisent l'*oppidum*.

À l'époque, les fouilleurs étaient persuadés de se trouver sur un site archéologique homogène datant de la période césarienne. Ils ont donc mis en place une méthode de fouille adaptée pour attester de la présence de la ville gauloise de Bibracte citée dans les Commentaires sur la guerre des Gaules¹⁷. Leur technique consistait à dégager les différents éléments architecturaux maçonnés, afin de recomposer la trame urbaine de l'agglomération. Ils pratiquaient donc des tranchées plus ou moins longues pour identifier les maçonneries, puis se contentaient de suivre ensuite les murs en surface pour dégager les plans de bâtiments. Des tranchées le long des murs étaient parfois réalisées pour caractériser ses éléments maçonnés. Les fouilles ont ainsi permis de sonder plus de 35 000 m² sur différents secteurs de l'*oppidum*, apportant les premiers résultats sur l'organisation générale de l'agglomération de la fin du second âge du Fer (Fochesato 2020, p. 55-57).

Ces résultats ont d'ailleurs été cartographiés, car à la même période, le capitaine H. d'Aboville, assisté de plusieurs officiers, a effectué plusieurs campagnes de relevés topographiques pour réaliser un plan complet du site (Guillaumet 1996, p. 70-71). Ce plan, de qualité militaire, a fourni les courbes de niveau équidistantes de 1m sur les 200 hectares qui composent le site, en plus de servir de base pour localiser précisément les découvertes archéologiques des fouilles anciennes (fig. 11).

17 César, *BG*, I, 23 ; VII, 55 ; VII, 63 ; VII, 90 ; VIII, 2 ; VIII, 4



11. Plan de l'oppidum de Bibracte d'après les relevés topographiques de H. d'Aboville (d'après Guillaumet 1996)

J.-G. Bulliot a également produit une documentation abondante et de qualité pour ses travaux sur le mont Beuvray. Il a d'abord consigné les résultats dans des carnets de fouille, avant de les publier quasi annuellement dans les *Revue archéologiques*, puis dans les "*Mémoires de la Société éduenne*" (par exemple Bulliot 1870 ; Bulliot 1883). Ces publications ont participé à l'introduction des standards de nos rapports d'opérations actuels, car ils présentaient déjà les résultats des campagnes de fouille par secteurs, sous forme de textes, de

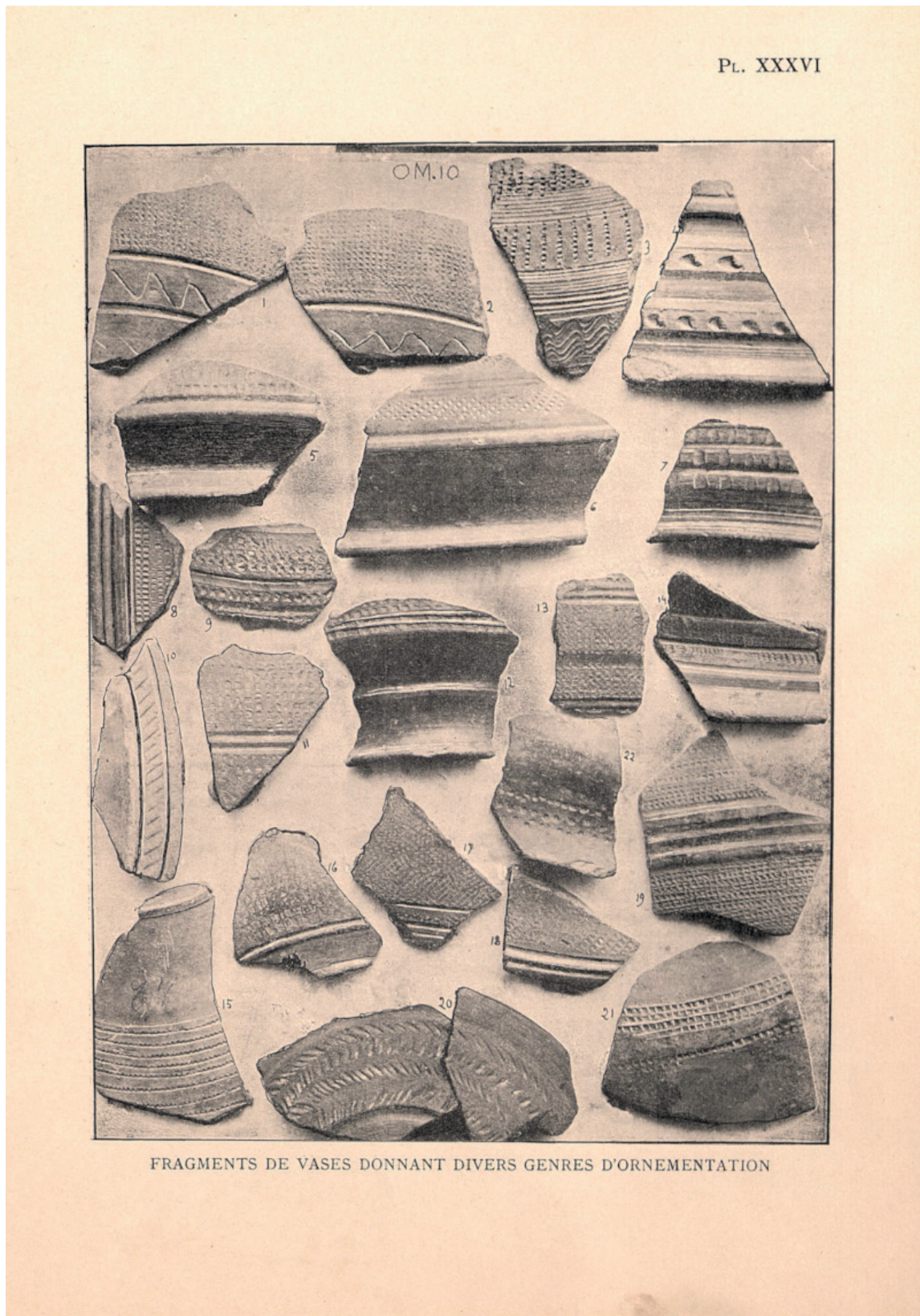
plans ou d'illustrations. Enfin, en 1899, J.-G. Bulliot a proposé une synthèse globale des 30 années de fouille, accompagné de plans généraux et de 61 planches d'objets photographiés (Bulliot 1899 ; Bulliot *et al.* 1899) (fig. 12).

Le neveu de J.-G. Bulliot, J. Déchelette, a poursuivi les fouilles jusqu'en 1907, mais en portant une attention plus spécifique à l'étude du mobilier archéologique. C'est d'ailleurs sur cette base qu'il a fondé son concept de "civilisation des *oppida*", en utilisant la typomorphologie pour démontrer les liens entre les cultures matérielles de différents grands sites européens datés du Ier s. av. n. è. Malheureusement, J. Déchelette a brutalement disparu pendant la Première Guerre mondiale, sonnant la fin de la première période d'exploration du mont Beuvray. Cependant, l'ampleur des fouilles, la pluralité des méthodes mises en œuvre et la qualité des publications laissées par les deux savants ont permis d'inscrire l'oppidum de Bibracte comme un site de référence pour l'étude de la fin de l'âge du Fer en Europe durant tout le XXe. s. (Goudineau 2007, p. 109).

Le programme de recherche européen

Lors de la reprise des fouilles modernes, en 1984, le site de Bibracte s'est affiché comme un lieu de rencontre de chercheurs de l'Europe entière. Ces équipes universitaires se sont organisées autour d'un programme de recherche commun pour étudier l'organisation de la société de la fin de l'âge du Fer à travers l'exemple de Bibracte, mais également "*l'occupation du sol, l'évolution des paysages et de l'agriculture de la région*" de l'*oppidum* (Guillaumet 1990, p. 11). Pour répondre à ces objectifs ambitieux et transdisciplinaires, le site accueillait chaque année plusieurs campagnes de fouilles en simultané, intervenant sur différents secteurs du mont Beuvray et ses alentours. Ces rencontres ont naturellement dynamisé les échanges autour des pratiques et méthodes à l'échelle internationale, tout en renforçant l'expérience de Bibracte dans la mise en commun et la conservation des résultats : "*Toute l'année arrivent des collègues français et étrangers, chercheurs ou professeurs intéressés par les résultats déjà acquis et par la façon dont la fouille de ce site, réputée difficile, a été abordée scientifiquement et techniquement.*" (Guillaumet 1990, p. 8).

Cette dynamique a été renforcée par la construction à partir de 1989 du Musée et du Centre archéologique Européen de Bibracte. Les deux édifices ont été inaugurés 6 ans plus tard, en 1995, par le Président de la République F. Mitterrand. Le centre a accueilli dès sa création des aménagements spécifiques aux besoins des archéologues, tels que des réserves pour le mobilier, un atelier de restauration et un atelier graphique avec son studio de déve-



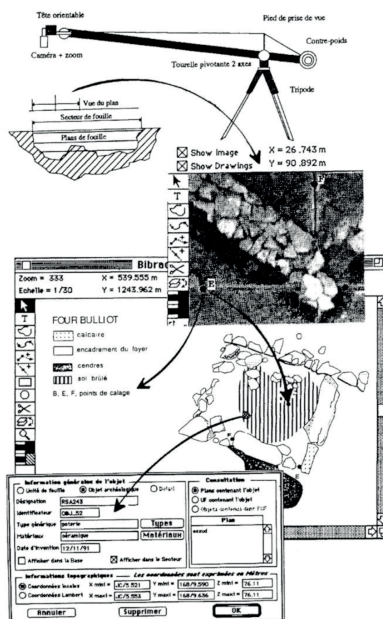
12. Exemple de planche d'objets photographiés issue de la publication de synthèse des travaux archéologiques de J.-G. Bulliot sur le mont Beuvray (d'après Bulliot et al. 1899)

loppement photographique. Un espace a également été prévu pour installer une bibliothèque spécialisée sur la protohistoire européenne, ainsi qu'une maison d'édition, dont le premier volume de la collection Bibracte est paru en 1996 (Buchsenschutz, Richard 1996). Avec ces infrastructures, Bibracte a été un des premiers l'un des premiers espaces spécialement pensés et conçus pour les archéologues. En regroupant un site archéologique, un centre de recherche et un musée, l'établissement a ainsi concentré la majorité des étapes de la chaîne opératoire archéologique en un seul lieu.

Grâce à l'émulation scientifique, la diversité des métiers représentés et la qualité du site et des aménagements, Bibracte est donc devenu un lieu de prédilection pour l'expérimentation méthodologique en archéologie dès la fin des années 1980. Cette période correspond aussi à la démocratisation de l'informatique dans le domaine, dont plusieurs expériences notables ont eu lieu à Bibracte.

Tout d'abord, en 1994, le mont Beuvray a été sélectionné pour expérimenter Arkéoplan, une solution pionnière pour le relevé par outils informatiques directement sur le terrain (Gruel, Buchsenschutz 1994). Elle se composait d'une caméra placée sur un bras articulé et reliée à un micro-ordinateur installé en bordure du chantier. Le système permettait de capturer une image des vestiges, de la redresser grâce à des marqueurs topographiques et de relever les éléments avec des outils de DAO (fig. 13). Chaque élément vectorisé se voyait automatiquement attribuer un identifiant et des coordonnées, qui pouvait être complété de descripteurs qualitatifs, afin de constituer une base de données graphique. Une base Arkéoplan pouvait ainsi gérer l'ensemble de la documentation graphique d'une campagne de fouille, l'imprimer ou l'exporter vers d'autres logiciels (tableurs ou systèmes de gestion de base de données).

À Bibracte, l'utilisation d'Arkéoplan a permis le relevé d'une portion de mur qui menaçait de s'effondrer juste après son dégagement. L'élévation a pu être photographiée et étayée, pendant que le relevé était effectué sur Arkéoplan. Plusieurs couches d'interprétation ont été relevées à partir de la même image, comme autant de calques permettant de comparer la nature géologique des blocs et de leurs liants. Finalement, avec le relevé numérique d'éléments archéologiques en lien avec une base de données à référence spatiale, Arkéoplan était capable de construire sur de véritables bases graphiques à références spatiales en chantier archéologique, à une époque où les logiciels pour la gestion de Systèmes d'Informations Géographiques étaient encore peu développés en archéologie.

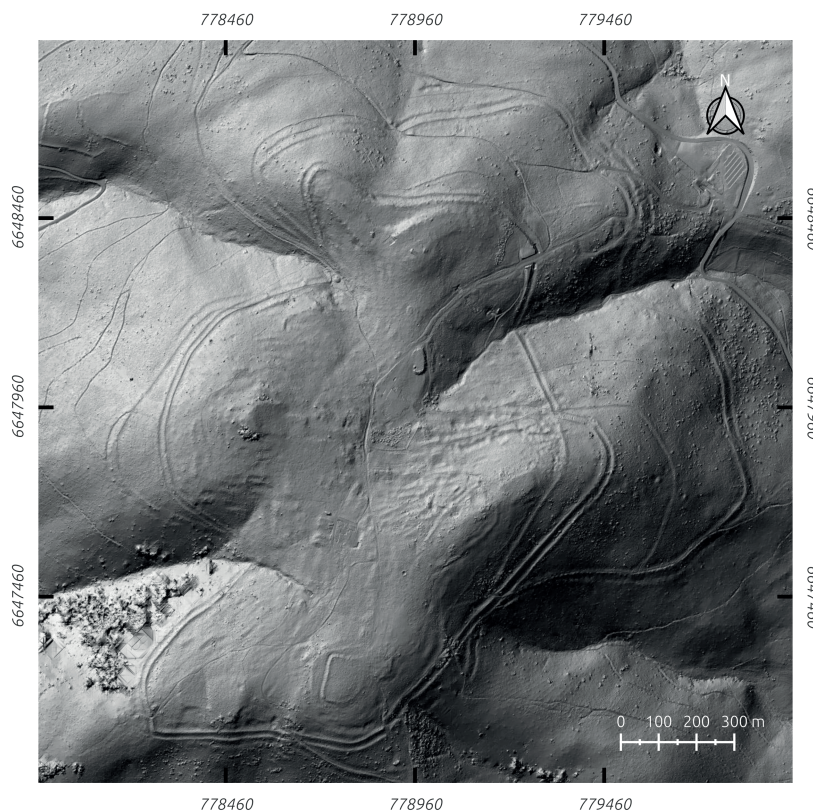


13. Schéma du système Arkéoplan (d'après Buchsenchutz et Gruel 1994)

Toujours sur les images numériques, il faut aussi mentionner à la même période l'utilisation précoce de la photogrammétrie pour actualiser le plan général du site. Le levé a été effectué en 1993 par le cabinet privé Régnier-Léveillé, sur commande du centre de recherche (Barral *et al.* 1995, p. 224). L'objectif était de préciser et de localiser dans un espace de référence le précédent plan de H. d'Aboville, datant des années 1870. En effet, la couverture forestière importante du mont Beuvray ne permettait pas de réaliser une relecture fine de la microphotographie par la photogrammétrie. En revanche, le levé a été calé à partir d'une trentaine de points positionnés grâce à la technologie satellite GPS (*Global Positioning System*), encore relativement peu utilisée à l'époque, afin de positionner avec précision les résultats dans un Système de Coordonnées de Référence (Lambert II). Les équipes du centre n'ont pas récupéré d'ortho-image ou de modèle 3D, mais un plan topographique exploitable sous le logiciel AutoCad, couvrant une surface d'environ 5 x 4 km.

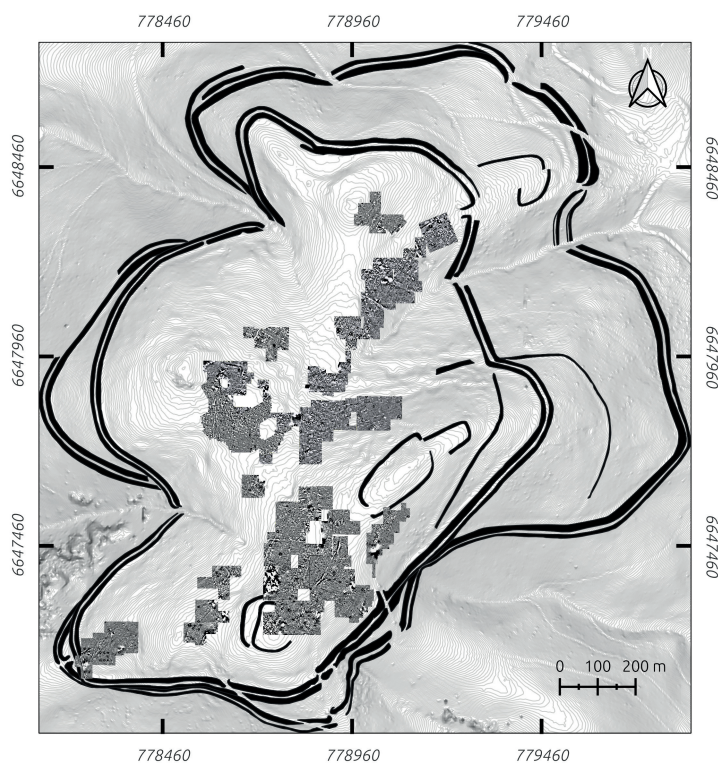
Le programme GéoTopoCart (Géodésie, Topographie et Cartographie du mont Beuvray) a été lancé en 1995 sous la coordination de F. Schubert, avec le même objectif d'améliorer la cartographie du mont Beuvray. Le but était d'utiliser les technologies modernes pour produire une “*documentation microtopographique précise et complète de l'ensemble de l'oppidum de Bibracte*”, coordonné par F. Schubert (Schubert 1997, p. 43). De ce programme est né le projet du “Grand Atlas Topographique du mont Beuvray”, qui ambitionnait de compiler les résultats topographiques et archéologiques du programme GéoTopoCart sous forme

de plans livrés à la fois dans une publication papier et sur support informatique (Schubert 2000, p. 48-49). C'est dans ce cadre qu'en 2007 le site et ses alentours ont fait l'objet d'un levé LiDAR (*Light Detection and Ranging*) aéroporté (Schubert *et al.* 2007) (fig. 14). Cette technologie permet en effet de couvrir de très large surface en avion pour lever le relief terrestre avec une très haute définition à l'aide d'un capteur laser. Elle offre aussi la possibilité de supprimer des éléments en élévation, tels que la végétation ou les lignes électriques, pour ne conserver que les informations au sol. Son utilisation à Bibracte représentait l'un des premiers cas d'application français en contexte archéologique. Deux secteurs ont été levés, l'un centré sur le mont Beuvray et l'autre sur le massif voisin au nord-ouest, autour des sources de l'Yonne. Chacun représente un carré d'environ 3 km de côté, soit près de 20 km², pour une précision en moyenne d'un point au sol tous les 10 cm. Ce premier levé LiDAR a permis de mettre en évidence de nombreux aménagements archéologiques sous le couvert forestier du mont Beuvray, mais également d'améliorer l'approche intégrée de la gestion et de la valorisation du site dans son environnement proche, à l'image du projet européen REFIT (*Resituating Europe's First Towns*) (Moore *et al.* 2020). Un nouveau levé LiDAR aéroporté a été réalisé en 2020 sur l'ensemble du Parc Régional Naturel du Morvan, dont les résultats ont un niveau de précision encore plus élevé.



14. Levé LiDAR du mont Beuvray, réalisé en 2007 (Bibracte EPCC)

Depuis la reprise des fouilles modernes, le mont Beuvray est également le terrain d'interventions de prospections géophysiques régulières. Les travaux ont débuté dès 1985, sous l'impulsion de G. Ducomet et A. Hess (Ducomet, Hesse 1985). Depuis, de multiples intervenants ont eu l'occasion de mettre en œuvre diverses méthodes pour documenter les indices naturels et anthropiques présents dans le sous-sol, faisant de Bibracte un haut lieu de l'évolution technique et scientifique des sciences géophysiques appliquées à l'archéologie. Plusieurs synthèses font état de ses recherches, et des travaux ont été entrepris pour réunir les résultats sur la plateforme web SIG de Chronocarto¹⁸ (Dabas *et al.* 1997 ; Alix 2000 ; Sanchez *et al.* 2019) (fig. 15).



15. Exemple de synthèse cartographique des prospections magnétiques menées par l'université de Brno de 2011 à 2019 sur le mont Beuvray (d'après Sanchez *et al.* 2019)

Au milieu des années 2010, une convergence s'est opérée entre les nouvelles problématiques d'exploration du mont Beuvray et le développement de plusieurs technologies, conduisant à une évolution rapide des pratiques sur les chantiers de fouille. En effet, les progrès techniques récents dans les domaines des capteurs photographiques numériques, des

18 <http://www.chronocarto.eu/spip.php?article76&lang=fr> ; consulté le 21/10/2023

drones aéroportés et de la puissance de calcul des ordinateurs ont favorisé le développement de la photogrammétrie, rendant son processus accessible au plus grand nombre. Dans le même temps, les équipes des universités de Franche-Comté et de Paris-Sorbonne mettaient au jour un bâtiment monumental sur le secteur de la terrasse PC15, requérant l'ouverture de grandes surfaces de fouilles. Cette conjoncture a conduit à plusieurs tests de dispositifs numériques, avec pour objectif de déterminer la pertinence de faire évoluer nos méthodes de fouille.

En 2013 et 2014, de premières acquisitions photogrammétriques au drone ont été effectuées par B.-N. Chagny de la société Kap-archéo, à la fin de la campagne de fouille de la terrasse PC15. Les clichés n'ont pas été particulièrement exploités, les équipes ayant déjà effectué les plans généraux à l'alidade durant la campagne. Il est d'ailleurs intéressant de noter que les orthophotographies présentaient une précision similaire aux plans manuels réalisés avec la lunette d'artilleur (fig. 16). Par contre, la vitesse d'exécution et la création d'images orthonormées détaillées offraient des perspectives intéressantes pour les années suivantes. Les mêmes années, les portions dégagées des murs de terrasses PC14 et PC15 ont également fait l'objet d'acquisitions et de traitements photogrammétriques au sol. Cependant, les résultats n'ont jamais été exploités, car il s'agissait plus de tests pour explorer les possibilités offertes par la méthode que d'une véritable volonté de poursuivre une problématique archéologique.

À partir de 2015, les expérimentations méthodologiques autour de la photogrammétrie se sont accélérées. Cette année-là, les fouilleurs des terrasses PC14 et PC15 sont venus avec le premier drone de la MSHE C.-N. Ledoux (UAR 3124), dans l'objectif de mettre en place un protocole de levé photogrammétrique qui remplacerait les plans à l'alidade. Pour les levés en cours de fouille, un élément important à prendre en compte à Bibracte est la présence de grands abris sur les chantiers (fig. 17). Ils sont volontairement arrimés au sol le temps du chantier, afin d'endurer les conditions météorologiques parfois rudes à 800m d'altitude sur le mont Beuvray. Dans ce contexte, il était impossible d'envisager l'acquisition photogrammétrique par drone sur les surfaces protégées. Ainsi, tous les niveaux de sol des espaces sous abris ont fait l'objet d'acquisitions photogrammétriques à la main, couche par couche (fig. 18). Toutefois, ce protocole n'a jamais été reproduit, car il s'est rapidement avéré que les acquisitions étaient beaucoup trop chronophages et la masse de données trop importante à traiter en fin de campagne. Des tests ont également été réalisés avec un scanner pour lever les vestiges en 3D, mais là encore les abris se sont montrés trop contraignants pour envisager le déploiement de cette solution.



16. Comparaison entre l'orthophotographie de fin de chantier et les relevés numérisés sur le chantier PC14/PC15 à Bibracte en 2013
(cliché B.-N. Chagny, Kap-archéo)

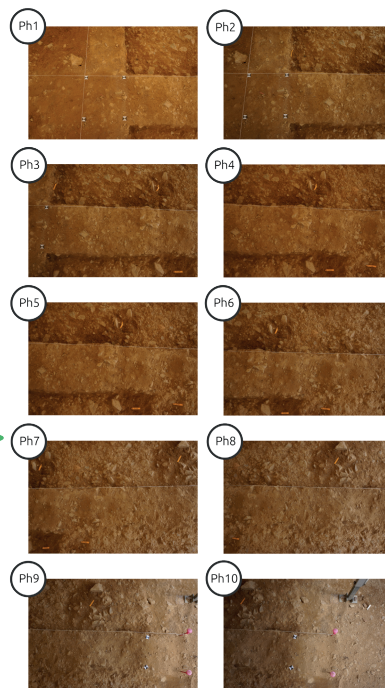


17. La nécessité d'abriter les chantiers archéologiques sur le mont Beuvray
(cliché M. Thivet)

Fichiers acquis pour traitement
photogrammétrique

- zone O.files
- Zone O1 - 28.05.15
- Zone O2 - 29.05.15
- Zone O3 - 03.06.15
- Zone O3 - 28.05.15
- zone P.files
- Zone P1 - 04.06.15
- Zone P1 - 27.05.15
- Zone P2 - 28.05.15
- Zone P3 - 27.05.15
- Zone P3 - 28.05.15
- zone Q_1.files
- zone Q_2.files
- Zone Q1 - 03.06.15
- Zone Q1 - 26.05.15
- Zone Q1 - 27.05.15
- Zone Q2-3 - 27.05.15
- Zone Q2-3 - 28.05.15
- Zone R1 - 03.06.15
- Zone R2 - 01.06.15
- Zone R3 - 01.06.15
- Zone S -T - 12.06.15
- Zone U-V - 24.06.15

Extrait de l'acquisition photo
de la zone Q1 le 27/05/15



Volume total des fichiers

Taille : 110 Go (118 706 609 505 octets)
 Taille sur le disque : 110 Go (118 720 815 104 octets)
 Contenu : 10 993 Fichiers, 901 Dossiers

18. Résultats non concluant des premiers tests pour un protocole d'acquisition photogrammétrique
sous abri

En revanche, le levé effectué au drone en fin de fouille, une fois les abris démontés, a permis de construire une image détaillée et orthorectifiée qui a été utilisée pour dresser plusieurs plans de synthèse dès la phase de post-fouille, facilitant par la même occasion la rédaction du rapport et la construction des illustrations (fig. 19). Le protocole a été amélioré et validé l'année suivante en pratiquant une acquisition au drone du chantier à chaque grande étape de la fouille, selon un rythme plus ou moins hebdomadaire. La photogrammétrie a ainsi permis de produire une documentation inédite, offrant la possibilité de revenir sur des points de détails comme de présenter une vue d'ensemble des différentes phases du chantier. Les images ont par exemple permis de réaliser le premier SIG intra-site du mont Beuvray, compilant l'intégralité des couches explorées sur les terrasses PC14 et PC15 durant les campagnes 2012 à 2017 (Hantrais 2017). Ces premiers tests concluants ont mené Bibracte à s'équiper de son propre drone en 2017.



19. Orthophotographie aéroportée de la terrasse PC15 (cliché M. Thivet, Q. Verriez)

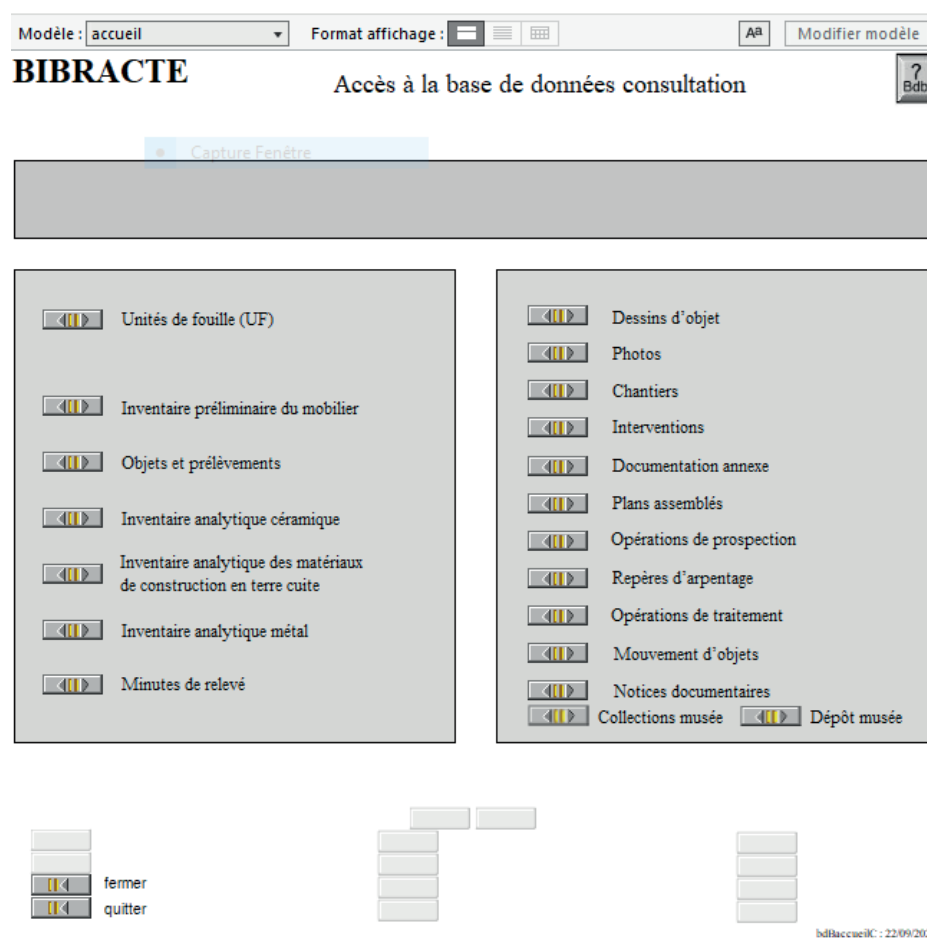
Plus récemment, le centre a soutenu les travaux de J. Wilczek sur la reconnaissance automatique de formes céramiques. Ce dernier a en effet mis en place un outil capable d'orienter, de dessiner et classer automatiquement la vaisselle céramique du second âge du Fer en se basant sur l'analyse morphométrique de modèles 3D d'ensembles mobiliers provenant de plusieurs sites européens, dont Bibracte (Wilczek 2017 ; Wilczek *et al.* 2018). L'objectif était d'éprouver les possibilités apportées par les développements récents dans les sciences mathématiques, statistiques et l'imagerie 2D/3D pour l'analyse des céramiques, dont les classifications sont parfois jugées subjectives.

Les pratiques d'enregistrement "classiques" se sont aussi rapidement numérisées : DAO, photographie numérique, inventaires sous forme de tableurs, *etc.* À travers les expérimentations méthodologiques et les résultats des campagnes de fouilles, les données archéologiques numériques produites à Bibracte se sont révélées riches et variées dès les premières années des fouilles modernes. Pour mettre en commun les résultats, il était indispensable de concevoir un système d'administration des données centralisé, capable de s'adapter aux évolutions scientifiques et méthodologiques du programme de recherche du mont Beuvray, tout en respectant et en conservant la richesse des données anciennes. Une première base de données informatisée a ainsi été mise en place, dès 1986, à partir du système CENTRAR (Buchsenschutz *et al.* 1986 ; Chaillou 2003, p. 86-87). Cette base a connu une première évolution majeure dix ans plus tard, lorsque V. Guichard a fait évoluer son modèle conceptuel en définissant les trois notions fondatrices de l'enregistrement archéologique à Bibracte :

- **L'intervention** : elle correspond à une action de recherche spécifique, quelle que soit sa nature (fouille, prospection, étude spécialisée, *etc.*), sur un temps donné, la durée d'une campagne annuelle, à condition qu'elle produise de l'information inédite.
- Le **chantier** : il s'agit d'un espace délimité qui accueille généralement une ou plusieurs entités archéologiques cohérentes.
- L'**Unité de Fouille (UF)** (Chaillou 2003, p. 126): elle représente l'unité d'enregistrement de base d'un chantier archéologique. Comme les Unités Stratigraphiques (US), elle peut être le résultat d'un phénomène naturel (une strate) ou d'une action anthropique, mais elle peut également renseigner des niveaux d'abstraction de ces US, comme dans le cas du découpage d'une couche en plusieurs unités selon un carroyage. Elles peuvent être de trois types : les UF positives, qui renseignent un volume ; les UF négatives, qui décrivent les interfaces théoriques entre deux actions ayant un effet sur

l'agencement des strates ; les UF englobantes, qui regroupent un ensemble d'UF cohérentes, quel que soit la nature de ce regroupement – les UF englobantes rassemblent sous un même terme ce que les archéologues découpent généralement en faits archéologiques, sondages, structures, *etc.* –. Chaque UF est associée à un numéro unique, et les structures en creux sont généralement définies par le numéro d'UF négative qui caractérise son creusement.

Dans le cadre de son DEA en 1997-1998, A. Chaillou s'est vu confier la mission de produire un cahier des charges pour une nouvelle mise à jour du système de documentation (Chaillou 2003, p. 87). Son objectif était d'identifier distinctement toutes les étapes d'enregistrement gérées par la base, afin de proposer une refonte clarifiée de sa structure. Ce modèle constitue depuis plus de 20 ans le socle de bdB, pour base de données Bibracte, le système de gestion et d'archivage de la documentation produite dans le cadre des activités de Bibracte EPCC (fig. 20).

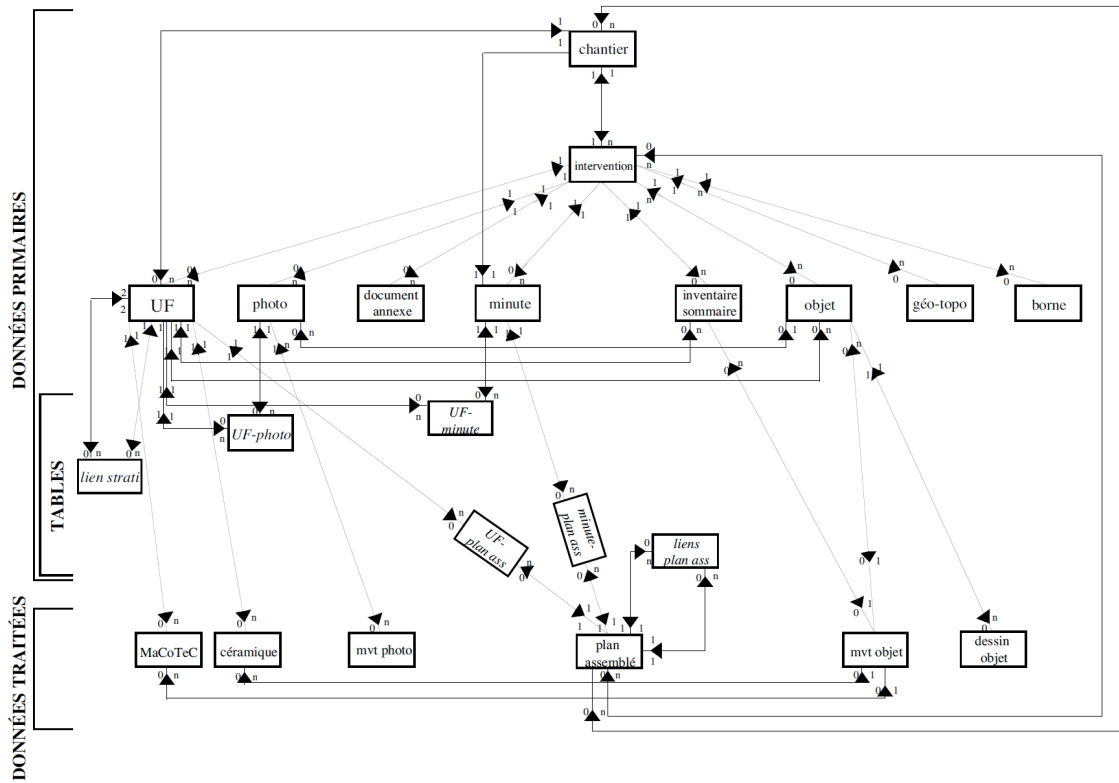


20. Page d'accueil de bdB Consultation

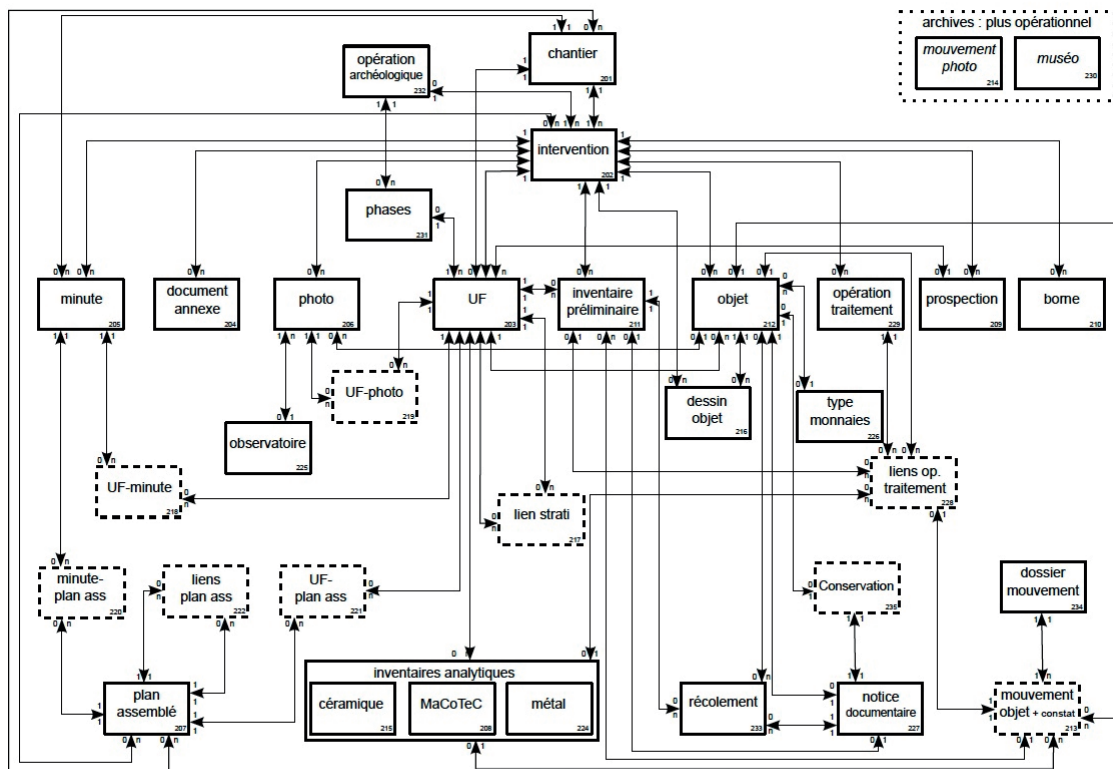
Il fonctionne depuis sa création sur le système de gestion de base de données (SGBD) FileMaker, et évolue ponctuellement avec les nouvelles versions du logiciel. Le fonctionnement de FileMaker repose sur le modèle logique relationnel. Les données sont stockées dans des tables, où chaque ligne du tableur représente une donnée et ses qualités descriptives (les attributs) correspondent aux colonnes (fig. 21). Des relations peuvent être déclarées entre les tables pour lier les données qu'elles contiennent. Le SGBD assure donc la structuration des données (classement et liens logiques), mais également le traitement des requêtes et la création et l'exploitation des interfaces graphiques de navigation et de saisie.

Dans le système bdB, l'ensemble de la documentation est homogénéisé en respectant un code d'inventaire unique par donnée, qui a été défini comme suit : *B (année)/(n° de chantier)/(n° d'ordre)* ou *B [année)/(n° de chantier)/(n° d'UF ou d'intervention)/(n° d'ordre)* (Chaillou 2004). Le caractère "B" signifie évidemment Bibracte. Ainsi, le modèle conceptuel de bdB a pu évoluer au cours du temps, au fur et à mesure des besoins exprimés par les utilisateurs, sans jamais perdre le lien avec les données déjà enregistrées (fig. 22). Il gère actuellement l'ensemble des données liées aux activités de recherche et de valorisation de Bibracte EPCC, depuis les données de fouilles (anciennes et modernes) jusqu'aux réserves du musée de France. Le système est conçu pour être accessible par tous, sans connaissance particulière en base de données ou langage de programmation, et supporte l'usage multi-utilisateur en simultané. La consultation de la base est techniquement possible par une connexion internet, mais son usage est généralement réservé au centre de recherche.

Le développement et la maintenance ont été assurés en interne, par ses principaux architectes, R. Moreau et A. Chaillou. Il en résulte un outil stable et fonctionnel, qui connaît une longévité particulièrement rare dans le monde du numérique. Sur ce point, il existe à notre connaissance seulement deux autres systèmes comparables : Syslat, le système historique développé à l'origine avec le projet de fouille du site de *Lattara* (Roure *et al.* 2021) ; la base de collections de l'ADS interrogées par *ArchSearch*, soit l'outil du leader mondial de la conservation et du partage des données archéologiques numériques (*cf. supra* 1.1.2. ADS). Pour se rendre compte de la masse de données dont elle assure la conservation, A. Chaillou évoquait quelques chiffres sur les données gérées par bdB en 2003 : 97 Mo de fichiers, 11 685 UF inventoriées, 2250 minutes, 25254 objets et 53015 photographies. En 2022, il s'agit de 15 Go de fichiers, 33 617 UF inventoriées, 7107 minutes, 87271 objets et 139852 photographies.



21. Structure de bdB en 2001 (d'après Chaillou 2003)



bdB Consult : schéma conceptuel de données 17 février 2021

— tables
 - - - - - tables de liens n ↔ n

22. Structure de bdB en 2021

Cependant, l'arrivée de nouvelles technologies dans les pratiques archéologiques a déjà mis en évidence certaines limites dans les capacités d'administration de bdB. L'expérience de Bibracte concernant le passage du support argentique au format numérique pour la photographie d'archéologie est un bon exemple de la difficulté d'intégrer de nouvelles données dans le système. Depuis la fin des années 80, l'équipe de Bibracte a comporté un archéologue-photographe professionnel, A. Mailler, et le centre a été équipé d'un studio pour la prise de vue et le développement des clichés. Jusqu'au début des années 2000, l'ensemble des photographies de fouilles étaient gérées par l'archéologue-photographe de Bibracte, depuis la prise de vue jusqu'au dépôt des clichés dans le système de gestion et d'archivage bdB. Avec l'arrivée des premiers appareils photo numériques sur le marché, les équipes de fouilles se sont progressivement procuré leur propre matériel. Ils faisaient de moins en moins appel au service d'A. Mailler, estimant leurs clichés suffisant pour documenter les campagnes de fouille. Afin de continuer d'assurer la conservation de ces photographies, il a été nécessaire de modifier le fonctionnement de bdB, qui ne prévoyait alors pas de procédure d'enregistrement des clichés par les équipes de fouilles. Malgré cela, la période intermédiaire entre l'arrivée des appareils photo numériques et la modification de bdB a représenté une perte sèche d'informations pour Bibracte. L'archéologue-photographe n'avait plus connaissance de toutes les prises de vue et les équipes de fouilles conservaient un grand nombre de photographies sur leurs ordinateurs personnels, par habitude ou parce qu'ils estimaient qu'il s'agissait de documents de travail.

Un nouveau palier technique a été atteint dans le courant des années 2000-2010. Le système de bdB est capable de gérer et d'archiver des ressources physiques (archives de fouille et mobilier) et des ressources numériques "classiques", telles que des documents textuels et des illustrations. Cependant, dans leur grande majorité, les nouvelles méthodes numériques mises en place sur les chantiers de Bibracte (LiDAR, SIG, photogrammétrie, *etc.*) ont la particularité d'être très prolifiques en matière de fichier, créant des jeux de données numériques volumineux et souvent complexes. Il s'agit dans la plupart des cas de données associant des informations textuelles, graphiques, spatiales ou temporelles. Les orthophotographies sont un bon exemple de cette situation. Plusieurs discussions ont eu lieu au sein de l'établissement pour déterminer s'il convenait de les enregistrer comme des photographies, par leur qualité de présentation des vestiges, ou comme des plans, car elles comportent des informations spatiales. De plus, dans un cas comme dans l'autre, les orthophotographies sont issues d'un processus de calcul complexe, effectué généralement à partir de plusieurs dizaines de photographies. Pour une seule acquisition, les photos et l'image finale peuvent ainsi rapide-

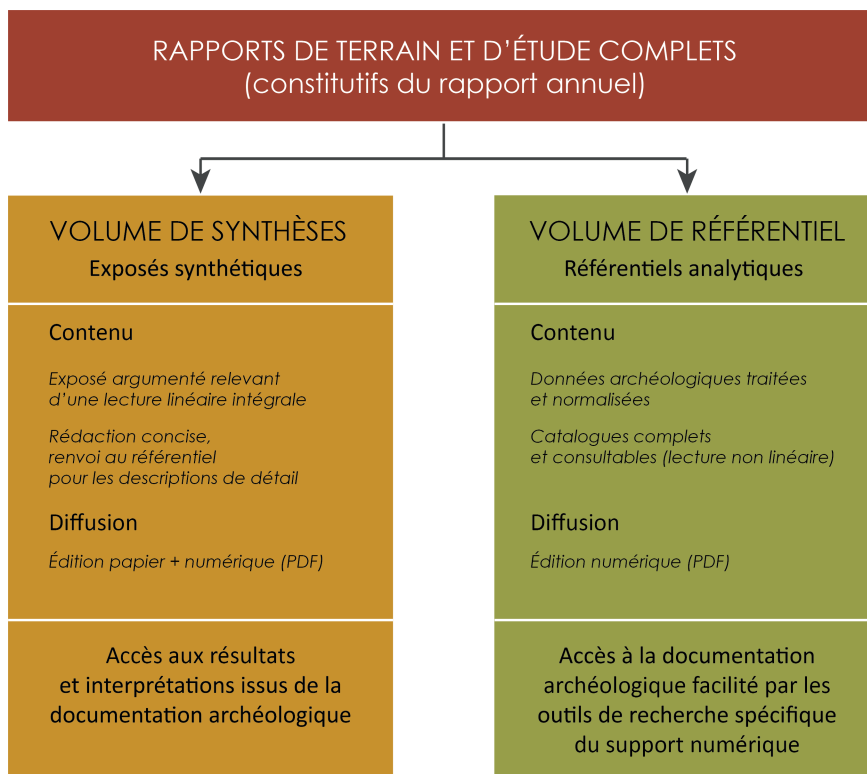
ment dépasser plusieurs Gigaoctets. Il y a encore quelques années, l'ensemble des fichiers numériques d'une campagne de fouille représentait dans la plupart des cas moins de 5 Go. Avec les nouvelles méthodes d'imagerie numérique, il s'est avéré possible de dépasser les 50 Go. La question s'est donc légitimement posée de savoir si toutes les données devaient être conservées ? Si un tri devait être effectué, quels étaient les éléments à archiver, les photographies en tant que données primaires ou l'orthophotographie comme résultat du traitement ? Aussi, la plupart des outils numériques peuvent avoir de multiples finalités, en fonction des différents traitements opérés sur les données primaires, selon les différentes problématiques des chercheurs. Pour continuer avec l'exemple de la photogrammétrie, une acquisition, selon les traitements, peut livrer une orthophotographie, un modèle 3D ou encore Modèle Numérique de Terrain (MNT). Or le système actuel n'a pas été conçu pour enregistrer cette relation entre données primaires et données traitées. Enfin, des contraintes liées à l'âge du système se font aujourd'hui ressentir. Par exemple, dans sa forme actuelle, *bdB* propose de gérer la spatialisation de la plupart des entités à partir de champs de valeur X, Y et Z, à remplir manuellement. Il a en effet été conceptualisé à une époque où la production et le traitement d'informations géographiques numériques étaient encore très peu développés en archéologie, et n'a que très peu évolué depuis sur cet aspect technique. Il ne peut donc pas interagir directement avec des solutions plus adaptées comme les SIG.

Ces différentes limites techniques ont eu pour conséquence de créer une situation similaire aux photographies numériques. Les données issues des nouvelles pratiques numériques sur les chantiers de fouilles du mont Beuvray ne sont que très rarement enregistrées dans *bdB*. Elles sont le plus souvent conservées par les auteurs sur leurs machines personnelles, parfois déposées sur les serveurs de Bibracte dans des espaces inadaptés et non signalés, ou plus rarement encore enregistrées dans la "documentation annexe" de *bdB*, un inventaire généraliste pour signaler et décrire des ressources diverses. Cette situation soulève une interrogation sur l'administration des nouvelles ressources numériques à Bibracte. Cette "documentation annexe" représente une richesse importante pour l'étude de l'oppidum, en offrant la plupart du temps des approches inédites pour répondre aux objectifs du programme de recherche de Bibracte. C'est pourquoi les expérimentations plus ou moins opportunistes se sont multipliées. Dans un processus de recherche, qui crée sans cesse de nouveaux résultats, il est parfaitement normal qu'un système de gestion de la donnée montre régulièrement des faiblesses techniques dans l'administration. Il y a quelques années en arrière, la "documentation annexe" était une solution adaptée pour enregistrer les données inédites qui sortaient

de la chaîne opératoire archéologique “classique”. Toutefois, les besoins d'enregistrement se sont aujourd'hui largement diversifiés, mettant en avant la nécessité pour Bibracte de franchir une nouvelle étape dans son processus de gestion de l'information archéologique.

Une première démarche a été mise en œuvre en 2012, lorsque l'équipe de l'établissement a procédé à une refonte du modèle de conception des rapports scientifiques annuels (Desachy *et al.* 2012). Chaque année, les multiples interventions menées dans le cadre du programme de recherche font l'objet d'une seule décision de fouille, attribué au directeur scientifique de l'établissement. Ce modèle implique la production d'un unique rapport, qui compile l'ensemble des résultats, classés par interventions. Afin d'améliorer la clarté des comptes-rendus, le choix a été fait de scinder ce rapport en 2 publications distinctes et complémentaires : le volume de synthèse et le volume de référentiel (fig. 23). Le premier volume propose un exposé synthétique des opérations, en livrant en priorités les interprétations et hypothèses (chronologiques, fonctionnelles, spatiales, *etc.*) découlant des résultats. Le deuxième volume, quant à lui, contient tous les éléments descriptifs issus des données primaires des fouilles et des études spécialisées de l'année courante. Il est présenté sous la forme d'un catalogue analytique au format numérique édité automatiquement à partir des éléments enregistrés par les différents chercheurs dans bdB. Pour compléter ce modèle, le centre de recherche a mis en place en 2016 un espace dédié sur l'archive ouverte HAL¹⁹ pour la diffusion de l'intégralité des rapports scientifiques depuis 1984. Cette méthode a permis d'améliorer significativement la capacité de partage de données archéologiques de Bibracte. Cependant, la synthèse comme le référentiel reste des documents pensés pour une diffusion au format papier. Les données sont structurées, lisibles et accessibles, mais le format PDF limite très fortement les possibilités offertes par les technologies numériques pour le traitement automatique d'informations (requête, tri automatique, import-export, *etc.*). De plus, ces publications ne prennent en compte que les données textuelles et graphiques enregistrées dans bdB, alors que comme nous venons de l'évoquer, il existe certaines nouvelles données au fort potentiel de diffusion qui ne sont pas enregistrées dans le système.

19 <https://hal.science/BIBRACTE> ; consulté le 21/10/2023



23. Structuration du rapport annuel du programme de recherche du mont Beuvray depuis 2012
(d'après Desachy et al. 2012)

En définitive, Bibracte, en tant que site majeur de l'archéologie européenne et laboratoire méthodologique de renom, occupe une place cruciale dans la production, la consolidation, la gestion et la diffusion des données numériques archéologiques. Au fil des années, cette position de pionnier amené le centre à traiter des volumes de données toujours plus importants, devenant un véritable réservoir d'informations archéologiques. Cependant, avec l'accroissement de la complexité et du volume de ces données numériques, l'outil historique de gestion et d'archivage informatique, bdB, montre désormais des signes de faiblesse. Cette prise de conscience a incité le centre archéologique européen à concevoir un projet ambitieux visant à repenser l'intégralité de l'environnement numérique de Bibracte, depuis les phases de fouille jusqu'à la publication des données. Ce projet d'envergure représente une étape cruciale dans l'adaptation de Bibracte aux défis actuels de l'archéologie numérique et de la gestion des données, garantissant ainsi la pérennité de ce site emblématique en tant que source d'informations pour la recherche archéologique européenne.

1.2.2. Le programme Bibracte Numérique

Présentation du projet

L'apparition de nouvelles données a mis en évidence le besoin de faire évoluer le système de gestion et d'archivage bdB, afin de ne pas perdre d'informations essentielles par manque de cadre d'enregistrement. Toutefois, les démarches de l'Inrap et de l'ADS ont démontré que les institutions d'archéologie n'échappent pas à la nature systémique du numérique. Chaque modification d'outil ou de processus a des répercussions significatives sur l'ensemble du système de production, de traitement, de partage et de conservation des données archéologiques. Cette réalité était d'autant plus prégnante pour Bibracte, car bdB joue un rôle central dans l'infrastructure numérique de l'établissement en assurant la centralisation de la documentation. Ainsi, l'évolution de ce système exigeait une réflexion globale sur l'ensemble de la chaîne opératoire numérique pour l'archéologie à Bibracte.

C'est de ce constat qu'est né en 2017 le programme Bibracte Numérique, qui s'est donné pour ambition de préparer la refonte complète de l'ensemble des processus opérationnels numériques de l'EPCC (Ayache *et al.* 2017). L'objectif était de répondre aux nouveaux besoins techniques en matière d'enregistrement, mais également d'accélérer la circulation des données archéologiques en exploitant pleinement le potentiel des sciences de l'information et de la communication. Pour cela, les acteurs du programme se sont appliqués à repenser et à transformer chaque étape de la chaîne opératoire, afin de garantir une intégration harmonieuse des dispositifs et des méthodes numériques tout au long du cycle de vie des données archéologiques. En ce sens, le programme rejoignait la démarche de la l'IR* Huma-Num, mais en l'appliquant à l'échelle de l'EPCC. Les outils et services de l'IR* Huma-Num devaient ainsi garder une position centrale dans les futurs développements.

Le programme a été conçu en se basant sur quatre sous-projets collaboratifs, chacun correspondant à une étape ou un point clé du processus archéologique numérique.

Projet 1 : Diffuser la connaissance par une médiation archéologique en s'appuyant sur le potentiel du numérique

Le premier projet se donnait pour but de poursuivre le développement de méthodes de médiation innovantes au musée de Bibracte, sur le mont Beuvray et sur le site internet. À l'aide de dispositifs numériques mis à jour au rythme des avancées du programme de recherche. Il était par exemple question d'explorer les possibilités offertes par la réalité virtuelle sur smartphone et tablette pour découvrir le site archéologique, ou le partage de ressources documentaires sur le web.

Projet 2 : Développer l'équipement numérique des chantiers de fouille au bénéfice des archéologues et des visiteurs

Ce projet, qui a été le cadre principal de ce travail doctoral, est né d'un double constat :

- L'utilisation d'outils numériques pour l'acquisition de données par les équipes de fouilles s'est largement développée durant la dernière décennie, mais peu d'entre eux s'intègrent dans une démarche intégrée aboutie.
- Les visiteurs du site ont la possibilité d'explorer des chantiers archéologiques en activité, cependant, la documentation produite lors des fouilles n'est pas disponible pour aider à la présentation des observations des archéologues en temps réel.

Ainsi, le projet 2 visait à améliorer l'utilisation des outils numériques sur les chantiers de fouille archéologique à Bibracte, dans le but de bénéficier à la fois aux archéologues et aux visiteurs. L'objectif était d'évaluer les capacités des équipements et des technologies numériques pour faciliter la collecte, la documentation et l'analyse des données archéologiques en temps réel sur le terrain. Le projet devait déterminer si l'introduction raisonnée d'outils numériques permettait aux archéologues de travailler de manière plus efficace et d'accéder rapidement aux informations pertinentes, tout en offrant aux visiteurs une expérience immersive et interactive lors de leur visite des chantiers de fouille.

Projet 3 : Construire une chaîne de production de la connaissance archéologique qui s'appuie sur le potentiel du numérique et organiser la connaissance pour une diffusion auprès des différents publics

Le projet 3 visait l'identification concrète d'outils et de protocoles pour gagner en efficacité, en rapidité de diffusion des résultats de la recherche, en diversité des publics atteints et en sécurité d'archivage.

Projet 4 : Mettre en place une infrastructure numérique adaptée

Comme son nom l'indique, le projet 4 devait offrir les solutions techniques pour mettre en œuvre la stratégie numérique élaborée dans le cadre du programme Bibracte Numérique. Dans cette optique, une des priorités était d'améliorer la connectivité locale et au réseau internet, relativement faible à Bibracte.

Dès son lancement, le programme Bibracte Numérique a réuni un large groupe de travail, constitué de partenaires scientifiques historiques du centre de recherche, mais également d'autres acteurs de l'archéologie française (collectivités territoriales, laboratoires, archéologie préventive et universités) et des humanités numériques en général. Cette dynamique de groupe a entraîné la création de plusieurs projets thématiques associés, ayant pour point commun de profiter de la richesse des communautés représentées à Bibracte et des données numériques qu'elles génèrent (fig. 24). De son côté, le programme Bibracte Numérique pouvait profiter des retours d'expérience de ce collectif pour nourrir sa réflexion globale sur l'évolution de l'écosystème archéologique numérique de l'établissement.

En tant que responsable du projet à la base du processus archéologique, notre recherche doctorale devait par conséquent prendre en compte les résultats des projets associés afin de définir comment produire de l'information archéologique numérique capable de s'intégrer dans un vaste réseau de connaissances, lui-même supporté par une infrastructure logicielle diversifiée et en constante évolution. De plus, le programme Bibracte Numérique et cette thèse devaient s'inscrire dans les principes de la science ouverte, afin de s'accorder avec la démarche d'Huma-Num et des principaux acteurs de la diffusion de la donnée archéologique.

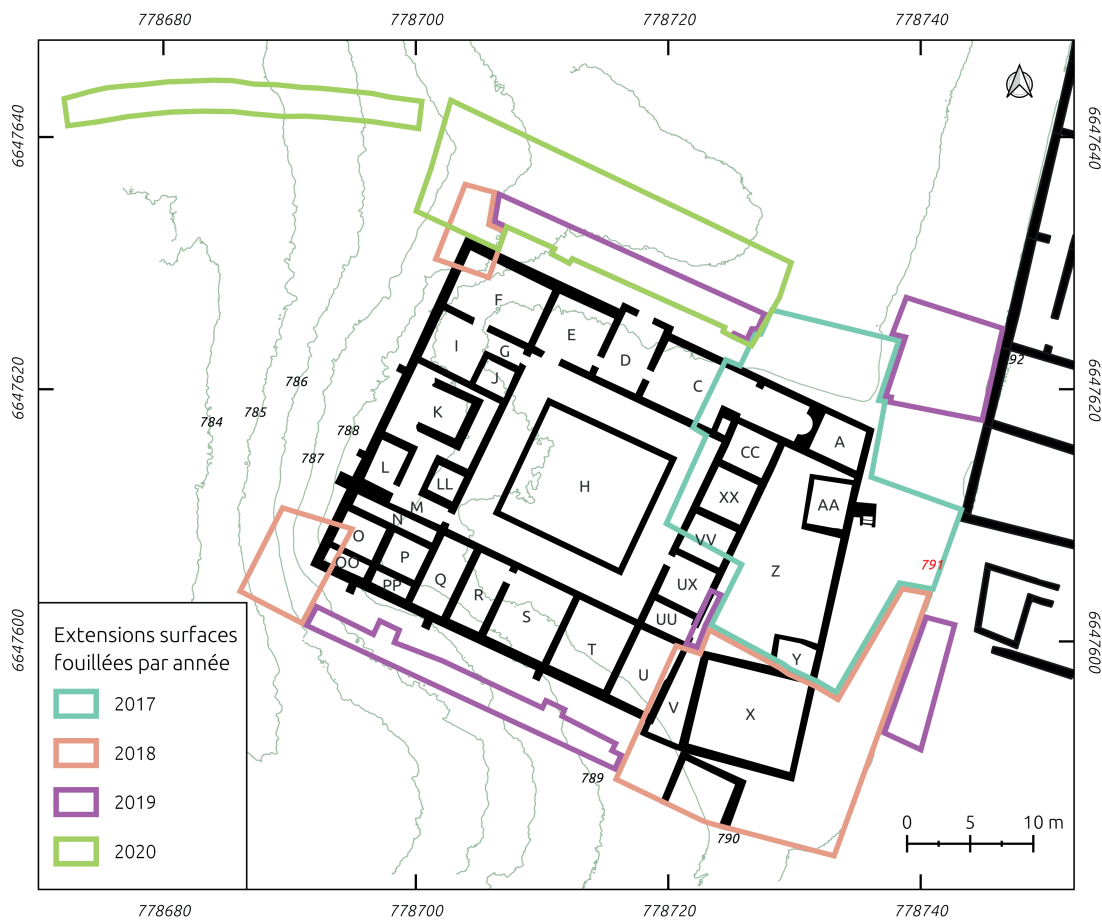
Le projet de la thèse a été construit autour d'une Convention industrielle de formation par la recherche (Cifre) entre Bibracte EPCC, le laboratoire ArTeHiS (UMR 6298) et l'École doctorale SEPT de l'Université de Bourgogne-Franche-Comté. Ce type de thèse encourage la conception, la mise en application et la valorisation de produits ou de processus innovants à travers le partenariat de laboratoire de recherche et d'entreprises sur le sol français. C'est pourquoi la commande de départ était de construire un protocole d'enregistrement numérique pour les chantiers archéologiques du mont Beuvray, à partir d'expérimentations conduites sur l'un de ses nombreux sites : la *domus* PC2.

Nom	Période	Porteur	Partenaires	Objet
HyperThésau	2018-2020	UMR 5133 Archéorient	EA 3083 ERIC ; UMR 5138 ArAr ; FR 3747 MOM J. Pouilloux ; UMS 3602 Persée ; UMR 7204 CESCO- Museum national d'Histoire naturelle ; Universitat Autònoma de Barcelona ; Bibracte EPCC ; SAVL ; musées d'Archéologie de Catalogne – site d'Ullastret ; Archéodunum	Constitution de micro-thésaurus pour l'archéologie, issus des pratiques-métiers des chercheurs, alignés sur les grands référentiels du web sémantique (Library of Congress Subject Headings, data.bnf.fr, IdREF, etc.)
ArteBib	2019-2020	UMR 5133 Archéorient	Bibracte EPCC ; Ville d'Autun (musée Rolin) ; Musée d'Archéologie nationale (MAN) ; FR 3747 MOM J. Pouilloux	Numérisation des objets archéologiques issus des fouilles anciennes et récentes réalisées sur l'oppidum, mise à disposition sur la base de données en ligne collaborative Artefacts et création d'un thésaurus propre à Artefacts.
Bulliot, Bibracte et moi	2019-2021	UMR 5133 Archéorient	Bibracte EPCC ; UMS 3602 Persée ; musée des Beaux-Arts et d'Archéologie Joseph- Déchelette (ville de Roanne) ; agence Réciproque ; Société éduenne des lettres, sciences et arts	Tanscrire et publier les carnet de fouille manuscrits de J.-G. Bulliot, en associant une méthodologie de « science participative » et un moteur de reconnaissance/transcription mobilisant l'intelligence artificielle : Transkribus
HisArc-RDF	2020-2021	UMR 5133 Archéorient	UMR 5190 LARHRA ; FR 3747 MOM J. Pouilloux ; Bibracte EPCC ; UMR 6249 Chrono-Environnement ; UMR 7044 Archimède ; UMR 8546 AOrOc ; réseau SPATIO des MSH ; EA 3083 ERIC ; ABES ; Archéodunum	Interconnecter un thésaurus (sur la plateforme Opentheso), un environnement de gestion d'ontologies (OntoME) et des notices d'autorités d'IdRef (ABES) afin de décrire des jeux-tests de données archéologiques selon les principes « FAIR » (des données trouvables (findable), accessibles, interopérables et réutilisables)
Bibracte ville ouverte	2021-2022	UAR 3124 MSHE Ledoux	Bibracte EPCC ; FR 3747 MOM J. Pouilloux ; ABES	Publication du catalogue de la bibliothèque de Bibracte sur Frantiq

24. Les projets associés à Bibracte Numérique

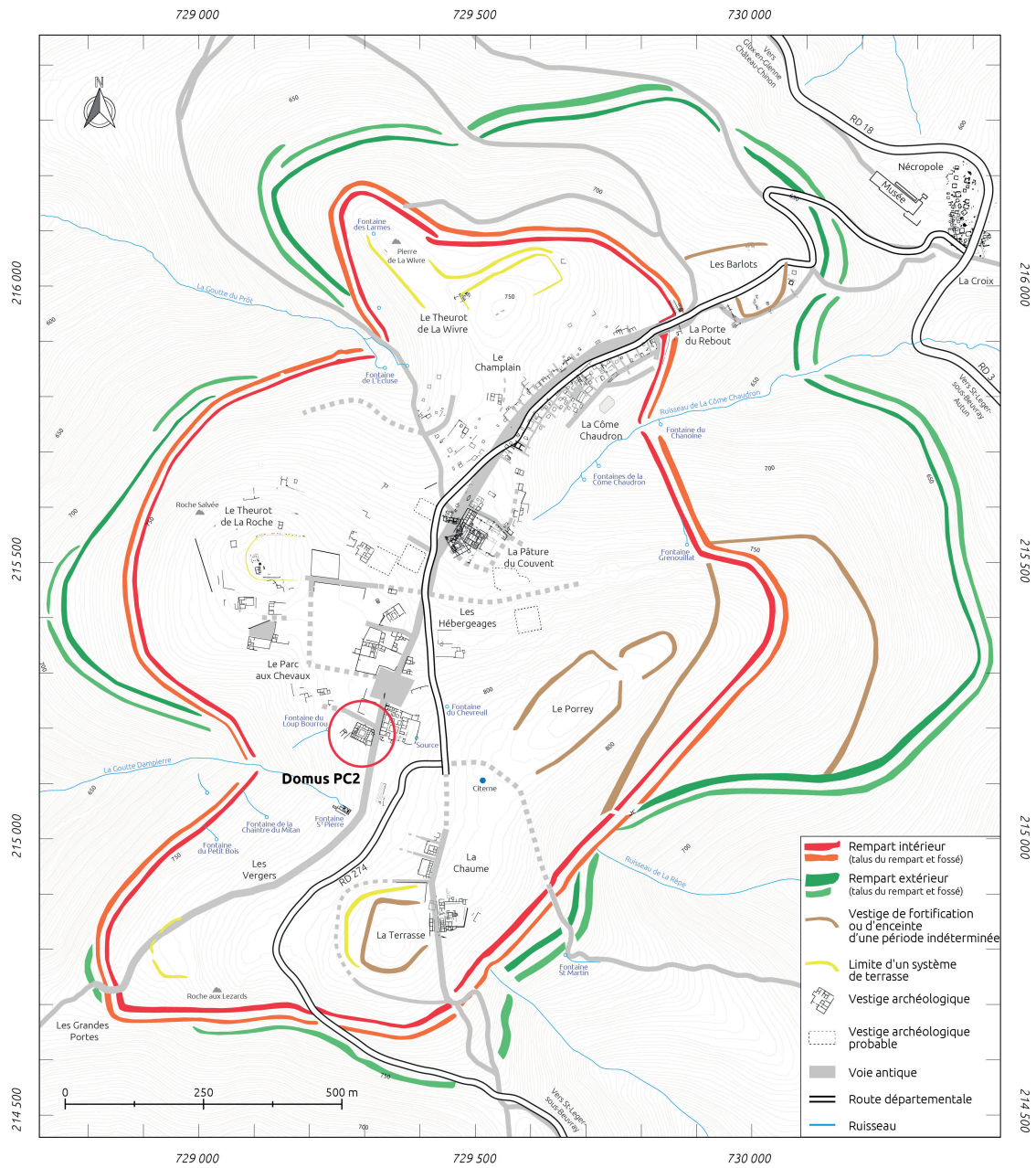
Le cadre d'expérimentation du projet 2 : la domus PC2

Les expérimentations se sont déroulées sur le chantier de la *domus* PC2, dont nous avons assuré la responsabilité scientifique des opérations de terrain pour les campagnes 2017 à 2020 (Verriez, Martini 2017 ; Verriez, Martini-Picot 2018 ; Verriez 2020 ; Defer 2021) (fig. 25). Le choix s'est porté sur ce secteur de l'*oppidum*, car il présentait plusieurs avantages au démarrage du projet.



25. Plan de la domus PC2

Pour comprendre l'intérêt du site, il est nécessaire de présenter la nature et l'historique des travaux sur le bâtiment. La *domus* PC2 est localisée dans le secteur du Parc-aux-chevaux, un plateau au centre de l'*oppidum*, délimité par le sommet du Theurot de la Roche au nord-ouest, par une rupture de pente au nord-est et par les pentes des sommets de la Chaume et du Porrey au sud (fig. 26). Le replat s'étend sur environ 50 hectares pour une altitude moyenne de 785 m NGF. J.-G. Bulliot et ses équipes ont exploré le secteur à plusieurs occasions entre 1866 et 1887. Chaque vestige a été identifié par le suffixe "PC", pour Parc-aux-Chevaux, suivi d'un numéro compris entre de 1 et 39, en suivant l'ordre de découverte (fig. 27).



26. Plan de localisation de la domus PC2 (DAO A. Meunier, Q. Verriez)



27. Plan d'ensemble du Parc-aux-Chevaux (d'après Guillaumet 1996)

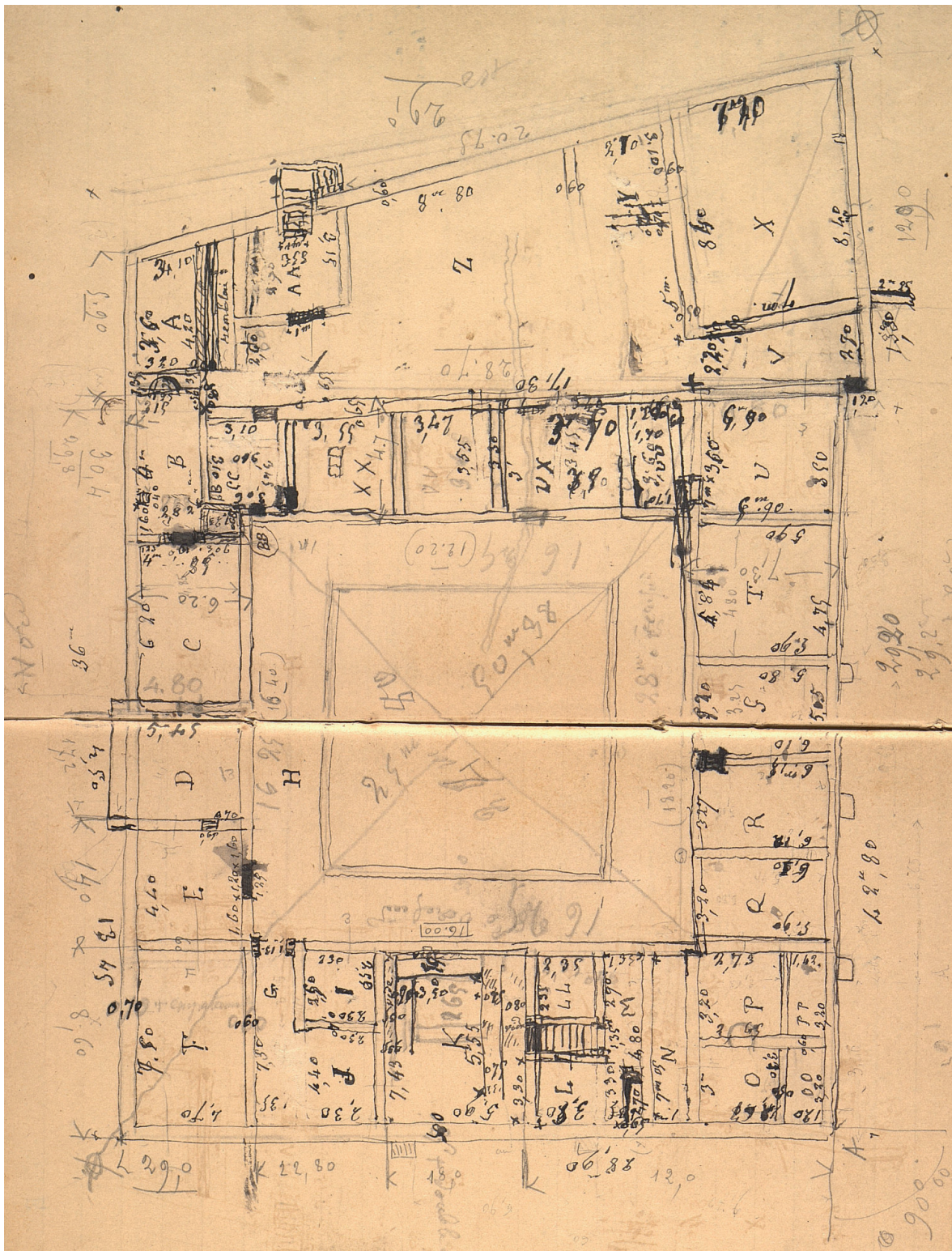
La *domus* PC2 se situe à l'extrémité sud du *Parc-aux-Chevaux*, sur un promontoire d'environ 1 500 m² qui fait face à la grande *domus* PC1 (fig. 28-29). L'entièreté du bâtiment a été fouillée en une seule campagne, durant l'été 1882 (Bulliot 1883). Suivant les méthodes alors en vigueur, seuls les murs et quelques pièces ont été entièrement mis au jour, afin d'identifier les tracés des maçonneries. Ces premières fouilles ont révélé un bâtiment quadrangulaire de 40 par 30 mètres de côtés, installé sur une terrasse artificielle qui surplombe la pente naturelle vers l'ouest. Dans son carnet de fouille, J.-G. Bulliot a dressé un schéma particulièrement bien documenté de l'édifice, qui a ensuite été mis au propre avec la publication des résultats (fig. 30-31).



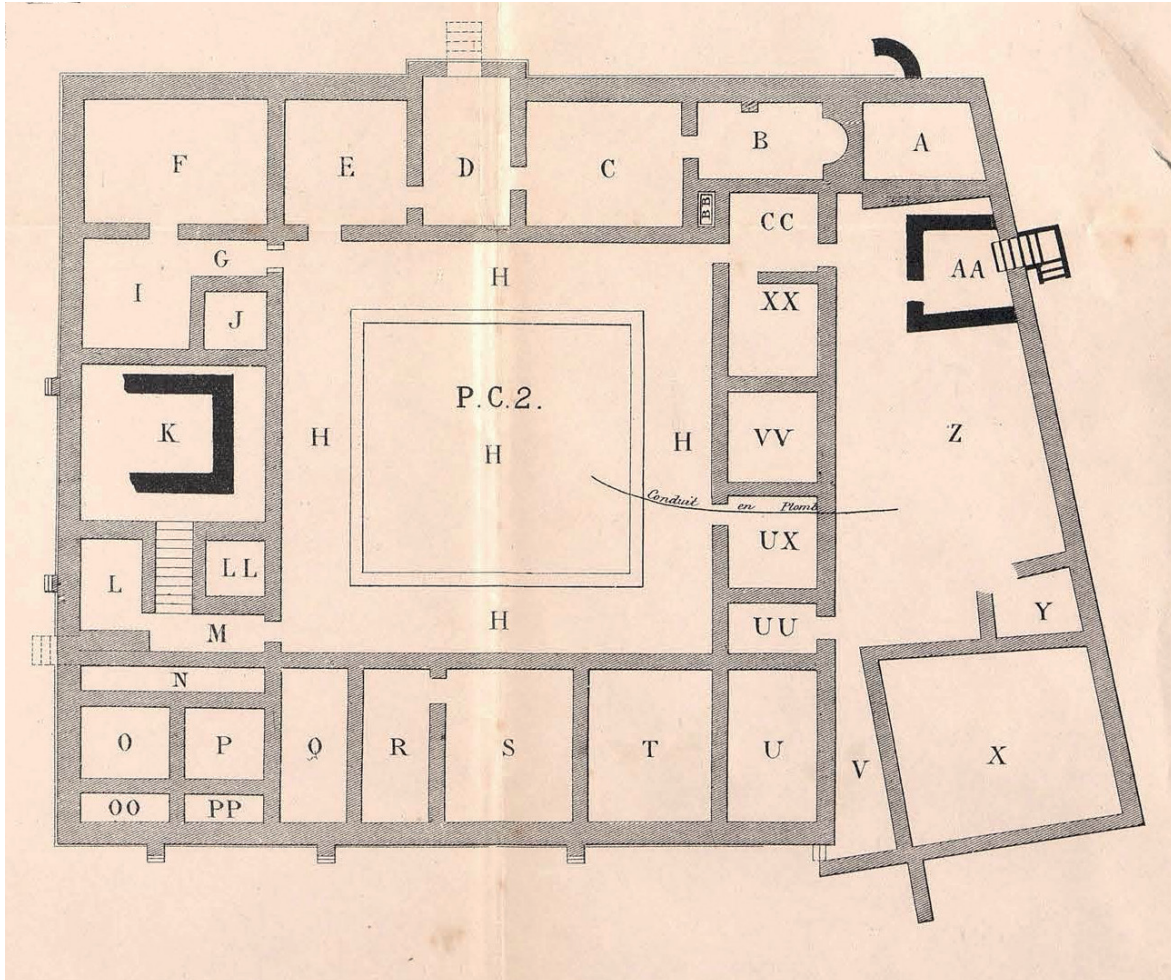
28. Début de la fouille de la *domus* PC2 en 2017, face à l'angle sud-ouest de la *domus* PC1



29. La terrasse de la *domus* PC2, vue depuis le sud-ouest

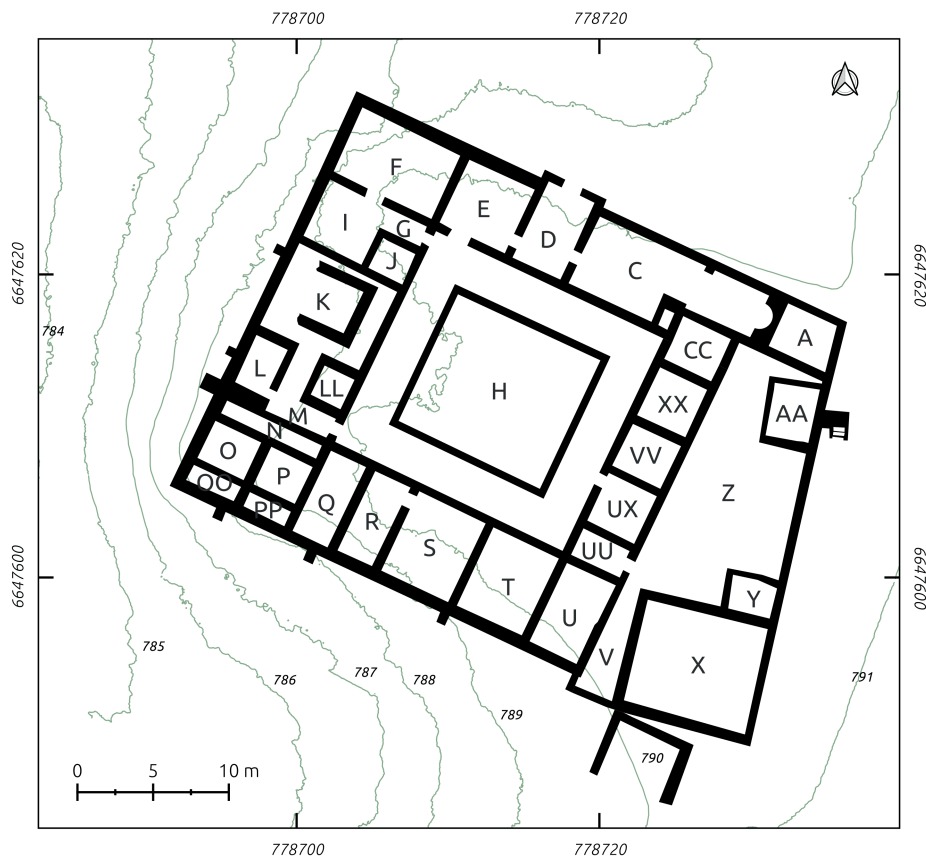


30. Schéma de la domus PC2 réalisé par J.-G. Bulliot dans son carnet de fouille 1882



31. Plan de la domus PC2 après les fouilles de 1882 (d'après Bulliot 1899)

Le bâtiment s'organise en deux parties : un édifice de plan carré d'environ 30 m de côté, organisé autour d'une cour centrale, sur lequel vient s'accoler une aile occidentale en forme de trapèze (fig. 32). L'ensemble comporte 35 pièces et espaces ouverts que J.-B. Bulliot a nommés de A à XX lors de la découverte. Il a également rapporté que certaines maçonneries possédaient jusqu'à 3 m d'élévations conservées, notamment au niveau de la façade méridionale et de la cave AA. L'intégralité du bâtiment a été remblayée à la suite de la campagne de 1882, et seule la cave AA a été brièvement rouverte par J. Déchelette en 1907 (Déchelette 1908, p. 58).



32. Plan de la domus PC2 actualisé

Avec la reprise des fouilles modernes, les abords de l'angle nord-est de bâtiment ont été explorés à l'occasion d'un sondage réalisé en 1994-1995 sur la voie qui sépare les *domus* PC1 et PC2 (Dufay 1995). Cette fouille, sous la direction d'A. Dufay, a représenté la première approche de la *domus* PC2 par le chantier-école de Bibracte, un chantier de fouille réservé à des adolescents souhaitant se former aux métiers de l'archéologie. En 2005, F. Meylan a compilé pour sa thèse les données de chaque structure archéologique connue du mont Beuvray, dont la *domus* PC2, afin de préciser les modalités architecturales employées et l'organisation spatiale de l'*oppidum* (Meylan 2005, vol. 3, p. 119-128). Enfin, en 2014, des mesures électromagnétiques et radars ont été réalisées sur la surface complète du bâtiment dans le cadre des interventions géophysiques à grande échelle dirigées par P. Milo, permettant d'attester de la qualité des plans anciens (Milo 2014).

Les fouilles ont repris sur la *domus* PC2 en 2016, toujours dans le cadre du chantier-école de Bibracte (Martini, Hamon 2016). Après 20 ans passés à sonder la *domus* PC1, l'objectif était de faire le lien des deux édifices adjacents, de manière à déterminer la chronologie du secteur. Dans la *domus* PC2, les fouilles de 2016 ont portées sur les espaces A, B, CC, XX, AA et Z, constatant de nouveau la précision des relevés et des observations effectués par J.-G. Bulliot en 1882.

Lors de la construction du programme Bibracte Numérique en 2017, la fouille de la *domus* PC2 présentait donc de nombreux avantages pour le test d'outils numériques sur un chantier archéologique. Comme nous venons de le voir, le site était particulièrement bien documenté par les fouilles anciennes et les travaux récents. La connaissance précise du plan et de la nature de certains vestiges, plutôt rares pour un site archéologique enfoui, permettait de définir en amont une stratégie de fouille adaptée aux différents objectifs du futur chantier. Tout d'abord, la conservation des vestiges permettait de tester l'enregistrement archéologique numérique dans de multiples situations : des structures maçonnées ou non, en creux et en élévations, dans les espaces restreints des pièces et les espaces plus ouverts des cours ou de la périphérie du bâtiment. De plus, par sa nature de chantier-école, la fouille de PC2 devait être extrêmement didactique et maintenir un niveau d'enregistrement des informations d'une qualité exemplaire. Comme tous les autres chantiers de Bibracte, il fait l'objet d'un compte-rendu scientifique dans le rapport annuel du programme de recherche, évalué par le ministère de la Culture. Enfin, Bibracte soutient et organise ce chantier depuis plus de 30 ans, sans interruption. Le chantier-école peut donc adapter son rythme de travail en fonction de l'évolution des objectifs scientifiques et pédagogiques définis par le conseil scientifique et la direction de l'établissement. Ainsi, l'excellente conservation des vestiges associés à la fouille méticuleuse de la *domus* PC2 constituait un terrain d'essai idéal pour expérimenter de nouveaux protocoles et technologies *in situ*. À la fois pour les possibilités offertes pour les tests et pour la qualité des vestiges à restituer aux visiteurs.

S'il était toutefois question au départ de construire un protocole, il est rapidement apparu dans nos recherches qu'il ne suffisait pas de proposer de nouveaux outils numériques pour que ces derniers soient adoptés par les archéologues. En particulier dans une communauté aussi riche et diversifiée que celle de Bibracte. C'est pourquoi le sujet de la thèse a rapidement évolué vers une réflexion plus globale, cherchant à comprendre les implications de l'utilisation du numérique en archéologie de terrain.

1.2.3. La nécessité de rationaliser les données numériques archéologiques de terrain

Le rapide tour d'horizon des années 1960 à aujourd'hui a démontré que le numérique en archéologie n'a rien d'une révolution. Il s'agit au contraire d'une longue tradition méthodologique, construite petit à petit. Ch. Tufféry, qui emprunte la notion à C.-L. Strauss, appelle cette manière de faire le "bricolage" : faire avec "*sa propre structure cognitive, par exemple en utilisant les seuls moyens à sa disposition pour atteindre son objectif [...]. Le "bricolage", comme processus de négociation, permet donc d'agencer autrement ce qui est déjà disponible et maîtrisé par les uns et les autres, que ce soit des matériaux, des dispositifs, des savoirs, des savoir-faire*" (Tufféry 2022, p. 139).

Une explication de cette situation peut se trouver dans la raison même de l'archéologie : appréhender les sociétés passées à partir d'indices matériels anthropiques et naturels. Ainsi, l'archéologie représente un champ d'études très vaste, et ses objets d'étude sont par définition extrêmement variés et fragmentaires. Pour répondre à cette problématique, les archéologues ont développé des méthodologies qui empruntent des savoir-faire techniques à de nombreux corps de métiers. Il en résulte des corpus de données volumineux et complexes, où chaque élément trouve sa signification à travers les multiples liens qu'il entretient avec les autres.

Les chercheurs ont vu dans les technologies numériques des outils pertinents pour traiter ces jeux de données. Cependant, face à la masse importante de données, leur niveau d'intrication et le manque de moyens humains, les archéologues n'ont pas pu ou pas su déléguer la mise en place d'outils numériques adaptés à leurs besoins. Ils les ont donc pour la plupart construits à partir des technologies qu'ils maîtrisaient déjà, quitte parfois à en détourner les usages.

De fait, les outils numériques – puisque le numérique ne représente bien qu'un ensemble d'outils matériels ou logiciels – sont aujourd'hui intégrés dans tous les domaines que constitue notre métier, à toutes les étapes de création de la connaissance archéologique. Ces outils, et par extensions les pratiques et les données qui en découlent, se multiplient ces dernières années de manière exponentielle, compte tenu des progrès techniques et ergonomiques récents, ainsi que de la baisse constante du coût du matériel. Ils sont présents à toutes les échelles (individuelle, collective, institutionnelle, nationale et au-delà), dans des états plus ou moins aboutis (du prototype au standard international). Par conséquent, plutôt qu'une révolution numérique en archéologie, nous préférons parler d'une intensification des pratiques archéologiques numériques.

Cette intensification est encouragée par le mouvement de la science ouverte, lui-même rendu possible par le développement massif d'internet depuis les années 2000. La mise en réseau des ordinateurs a permis d'envisager le partage des résultats de la recherche, les synthèses et les données primaires, à une échelle encore jamais égalée. Un tel exercice est théoriquement vertueux pour l'archéologie, comme pour les sciences en générale, car il encourage la critique et l'émulation scientifique, en plus d'accélérer la transmission des connaissances vers les autres publics. Toutefois, cette mise en commun sans précédent impose de construire des standards d'échange à l'échelle de la communauté archéologique dans son ensemble. Sans initiatives comme celles de l'ADS ou d'Huma-Num pour accompagner la diffusion du savoir archéologique, le partage de données à l'ère du numérique et des réseaux a plus de chance d'entraîner un "bruit inaudible" que de favoriser une véritable émulation.

Aussi, la systématisation des outils numériques en archéologie entraîne quelques dérives. Avant l'introduction du numérique, les archéologues étaient contraints d'identifier et d'échantillonner les informations pertinentes en fonction de problématiques préalablement définies. Aujourd'hui, la démarche a parfois tendance à s'inverser : puisque l'outil informatique peut théoriquement gérer d'énormes volumes de données, il est nécessaire de formater l'information en fonction de cet outil pour qu'elle puisse être traitée. Cela revient à adapter la démarche scientifique aux contraintes techniques imposées par les outils. Cette remarque n'est en réalité probablement pas à imputer au numérique uniquement. Des problèmes similaires devaient sans doute déjà exister avec les outils analogiques. Toutefois, le risque de conditionner la pratique archéologique aux possibilités des outils est largement démultiplié par les capacités de traitement, d'automatisation et de partage des dispositifs numériques récents. D'autant plus quand ils masquent leurs complexités derrière des interfaces graphiques ergonomiques qui permettent aux utilisateurs d'effectuer des tâches sans en maîtriser les fondements.

L'évolution des outils numériques peut d'ailleurs s'avérer très rapide et induire des blocages humains dans la chaîne opératoire archéologique. Sur ce point, le cas du développement du SIG à l'Inrap est plutôt explicite (*cf. supra* 1.1.2. Inrap). Même si l'outil s'est révélé techniquement adapté aux travaux archéologiques, sa mise en place s'est étalée sur plusieurs années. L'Inrap a pris le temps d'accomplir des études de faisabilité, d'organiser des formations, de prendre en compte les retours des utilisateurs pour améliorer le système, *etc.* Autrement dit, le déploiement d'un outil a des effets sur l'infrastructure technique d'une organisation, mais également sur les habitudes de travail des utilisateurs. L'intégration d'un nouvel

outil peut s'avérer difficile si ces deux aspects ne sont pas gérés conjointement, voire mener à son rejet complet. Ainsi, réfléchir aux usages du numérique en archéologie revient donc à prendre en compte les utilisateurs autant que les outils. Dans le même temps, de nombreux experts, comme le GIEC, commencent à alerter sur les effets directs et indirects des technologies sur l'environnement : *“Les technologies numériques peuvent favoriser d'importantes améliorations de l'efficacité énergétique grâce à la coordination et à la réorientation de l'économie vers les services, mais elles peuvent également accroître considérablement la demande d'énergie en raison de l'énergie utilisée par les appareils numériques”* ; *“Les appareils numériques, y compris les serveurs, augmentent la pression sur l'environnement en raison de la demande de métaux rares et du traitement des déchets en fin de vie”* (Shukla et al. dir. 2022, p. 140).

La question qui s'est alors posée à Bibracte était de savoir comment gérer la systématisation massive des usages numériques en archéologie, accélérer par la tendance à la science ouverte, sans passer par une nouvelle période d'ingérence dans la chaîne opératoire ? Pour y répondre, le programme Bibracte Numérique a découpé les étapes du processus archéologique en plusieurs projets, et ce travail s'intéresse à la première d'entre elles : la fouille. Notre objectif est alors le suivant :

Comment améliorer le traitement, la gestion et la conservation des données archéologiques numériques de terrain dans un lieu comme Bibracte, qui regroupe autant de profils de chercheurs différents, avec leurs objectifs, leurs méthodes et leurs compétences propres ? Dans ce sens, nous tenterons de définir dans quelle mesure la maîtrise des technologies numériques à la disposition des archéologues permet-elle de fluidifier le cycle de vie des données archéologiques de terrain ? Dans le même temps, nous nous interrogerons sur la capacité de la communauté archéologique à mesurer et à intégrer dans leurs démarches numériques les aspects sociaux et environnementaux induits par le déploiement de ces technologies ?

Ainsi, nous proposons dans ce manuscrit plusieurs solutions pour rationaliser l'utilisation des outils numériques en chantier archéologique, en explorant à la fois les limites techniques et, dans une moindre mesure, humaines et environnementales d'une telle pratique. L'objectif n'est pas de présenter un nouvel outil de pointe. Nous souhaitons au contraire profiter de notre regard d'archéologue pour interroger la place et l'efficacité de dispositifs et de logiciels déjà couramment utilisés par cette communauté, parfois depuis plusieurs dizaines d'années. Notre but est d'identifier les avantages et les limites de l'utilisation d'outils numériques sur le terrain, afin de dégager les points essentiels à prendre en compte pour leur implémentation

dans une démarche de fouille archéologique. Ce travail se base sur notre propre savoir-faire, à partir de l'expérimentation conduite sur le chantier de la *domus* PC2, mais aussi sur d'autres fouilles auxquelles nous avons participé. Nous avons néanmoins systématiquement tenté d'inscrire nos observations dans un champ de pratiques plus larges, en prenant en compte l'historique des outils numériques questionnés dans la discipline archéologique.

Toutefois, avant d'aller plus loin avec les outils, il nous semble important de revenir un peu en amont du cheminement scientifique. En archéologie comme dans toutes autres sciences, les outils sont mis en œuvre pour collecter des données, dans l'intention de répondre à une problématique. Présenté par certains comme le nouveau pétrole²⁰, le concept de données est très lié à l'essor du numérique, particulièrement dans le cas de la science ouverte. C'est pourquoi nous souhaitons à présent définir ces deux notions essentielles.

20 La paternité de cette idée est généralement attribuée au mathématicien Clive Humby, qui aurait déclaré en 2006 "*data is the new oil*" : https://en.wikipedia.org/wiki/Clive_Humby ; consulté le 21/10/2023

DEUXIÈME PARTIE

Les éléments constitutants du numérique en archéologie : données, science ouverte et Plan de Gestion de Données

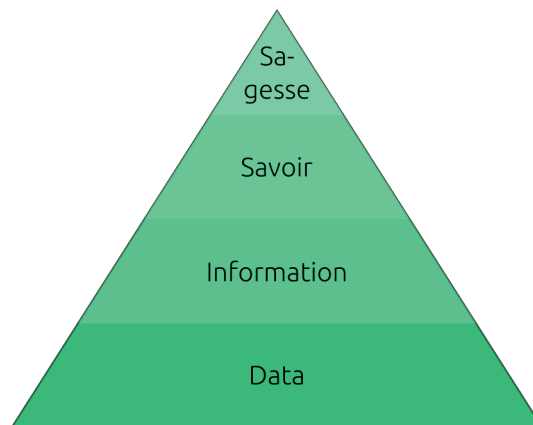
2.1. Les données numériques en archéologie

2.1.1. La donnée : définition

Un des modèles utilisés par les spécialistes de l'information pour définir ce qu'est une donnée est la hiérarchie, ou pyramide, Donnée-Information-Savoir-Sagesse (*data-information-knowledge-wisdom* ou *DIKW hierarchy* en anglais) (fig. 33). Dans un article de 2007, J. Rowley propose une relecture de ce modèle à la lumière des publications récentes sur le sujet (Rowley 2007). Elle y explique que le modèle est traditionnellement attribué à R. Ackoff, qui le présente dans un article de 1989 intitulé "*From data to wisdom*" (Ackoff 1989). J. Rowley résume les définitions et l'organisation de ce dernier ainsi (Rowley 2007, p. 166) :

- *“Les données sont définies comme des symboles qui représentent les propriétés des objets, des événements et de leur environnement. Elles sont les produits de l’observation. Mais elles ne sont d’aucune utilité tant qu’elles ne sont pas sous une forme utilisable (c’est-à-dire pertinente). La différence entre les données et les informations est fonctionnelle et non structurelle.*
- *Les informations sont contenues dans des descriptions, des réponses à des questions qui commencent par des mots tels que qui, quoi, quand et combien. Les systèmes d’information génèrent, stockent, récupèrent et traitent les données. Les informations sont déduites des données.*
- *La connaissance est un savoir-faire, c’est ce qui permet de transformer l’information en instructions. La connaissance peut être obtenue soit par transmission d’un autre qui la possède, soit par instruction, soit par extraction de l’expérience.*
- *L’intelligence est la capacité d’accroître l’efficacité.*
- *La sagesse est la capacité d’accroître l’efficacité. La sagesse ajoute de la valeur, ce qui requiert la capacité mentale que nous appelons jugement. Les valeurs éthiques et esthétiques que cela implique sont inhérentes à l’acteur et sont uniques et personnelles.”*

Le concept de donnée se construit par opposition avec ceux d'information ou de connaissances, selon un lien hiérarchique fort : sans observation, impossible de créer de l'information. En archéologie par exemple, selon le modèle DIKW, l'unité stratigraphique est une information produite par la mise en commun des propriétés (les données) d'une couche,



33. Organisation de la connaissance selon le modèle DIKW (d'après Rowley 2007)

telles que la couleur et la nature du sédiment. Les données sont donc les éléments à la base de toute construction intellectuelle issue de l'observation d'un individu.

J. Rowley précise la définition de chaque élément du modèle en croisant 16 publications de référence sur les systèmes d'information et la gestion des connaissances du début des années 2000. Voici ce qu'il en ressort pour les données (Rowley 2007, p. 170) :

- *“Les données n'ont pas de sens ni de valeur, car elles sont dépourvues de contexte et d'interprétation.*
- *Les données sont des faits ou des faits discrets et objectifs, non organisés et non traités, qui ne véhiculent aucune signification spécifique.*
- *Les données sont une description élémentaire et enregistrée de choses, d'événements, d'activités et de transactions”.*

Il apparaît clairement que, pour les spécialistes, les données n'ont pas de réalité tangible. Elles n'existent que par l'action d'enregistrer le fruit d'une observation “donnée à voir” par un phénomène qui présente des propriétés physiques. Elles ne trouvent d'ailleurs de sens qu'après compilation et analyse, ce qui a pour effet de créer de l'information et de la

connaissance. Par exemple, les données “hétérogènes”, “fragments de terre cuite orangés” et “12 cm” prisent individuellement ne signifient pas grand-chose. Mais en les associant, un archéologue cherche plus d’indices pour déterminer s’il se trouve en présence d’un niveau d’effondrement de toiture d’une douzaine de centimètres d’épaisseur.

Voici pour le cadre théorique. Cependant, comme souvent, la réalité est bien plus complexe. Dans la pratique, *“les données sont sujettes à interprétation ; leur statut de faits ou de preuves est déterminé par les personnes qui produisent, gèrent et utilisent ces données”* (Borgman 2007, p. 121). L’Organisation de coopération et de développement économique (OECD) va d’ailleurs dans ce sens, en proposant une définition des données de la recherche beaucoup plus large :

“[les données de la recherche sont des] enregistrements factuels (chiffres, textes, images, son, vidéo...) utilisés comme sources primaires pour la recherche et qui sont habituellement acceptés par la communauté scientifique comme étant nécessaires pour valider les résultats de la recherche” (OECD 2007, p. 13 ; MESRI 2018, p. 10).

En archéologie particulièrement, la frontière entre objet d’étude et observation n’est pas tranchée. Dans la littérature spécialisée, il est courant d’employer le terme “donnée” pour signifier les propriétés d’un sujet, d’un objet ou d’un vestige, mais aussi le sujet en lui-même :

“À la base de toute étude archéologique, les données sont par définition des vestiges ou ensemble de vestiges matériels, avec les informations dont ils sont porteurs. Ces vestiges font l’objet d’une analyse, d’un traitement, d’un commentaire, en bref d’une construction intellectuelle, mobilisant un certain “savoir” (faits et arguments plus ou moins explicites), et débouchant sur une interprétation” (Desachy 1990, p. 21).

Pour l’archéologue, les données sont donc à la fois des *data*, ce qui est donné à être observé, et des *capta*, ce qui est capturé par des instruments de mesure, selon le sens défini par R. Kitchin et M. Dodge :

“les capta sont des unités qui ont été sélectionnées et récoltées à partir de la somme de toutes les data potentielles, où les data sont la somme totale des faits relatifs à une entité ; en d’autres termes, en ce qui concerne une personne, les data sont tout ce qu’il est possible de savoir sur cette personne, les capta sont ce qui est capturé de manière sélective par la mesure” (Kitchin, Dodge 2011, p. 5).

Il est possible d'aller encore plus loin et d'utiliser le résultat du processus scientifique, l'interprétation, comme une donnée à part entière :

“La donnée est tout élément que l'archéologue trouve dans la fouille, mobilier, objet, couches, ainsi que les liens entre ces éléments, des observations et des interprétations qu'il peut faire et aussi tout document permettant de garder une trace visuelle de ces objets, de ces couches et des liens comme des photographies ou des dessins. Une donnée pourra aussi être le résultat de l'étude de ces premières données, études des mobiliers, [etc.]” (Chaillou 2003, p. 29).

Finalement, les archéologues utilisent le terme de donnée pour définir tout ce qui peut être observé, mesuré et interprété pour retracer l'évolution des sociétés passées. Il peut s'agir d'observations directes (propriétés physiques) ou indirectes (mesurées par des instruments), d'éléments matériels (vestiges mobiliers et immobiliers) ou de déductions structurées ou non (hypothèses, interprétations, représentations). Par la nature même des indices observés, les données archéologiques sont donc nécessairement très variées, hétérogènes et fragmentaires.

Si cette définition s'éloigne du concept de donnée défini par les sciences de l'information et de la communication, elle n'est pourtant ni erronée ni abusive. En effet, comme le souligne Ch. Borgman, ce qui prime dans la démarche scientifique, c'est avant tout le regard critique que peuvent porter les pairs sur ce qui est présenté comme une donnée. Autrement dit, en archéologie, comme dans d'autres disciplines, est accepté comme donnée scientifique tout élément à la base d'une construction intellectuelle, tant qu'il est reconnu et validé comme tel par la communauté qui anime la discipline.

Enfin, le processus d'abstraction de la réalité en valeurs normalisées implique forcément une perte d'information. Ce constat n'est pas pour autant négatif, puisqu'il relève directement de la démarche scientifique : sélectionner les éléments essentiels et pertinents pour répondre à une problématique.

2.1.2. Les types de données

Physiquement, les données peuvent se présenter selon deux types de signal : analogique et numérique. Les données analogiques sont des représentations d'une réalité, d'une grandeur physique par une autre, comme lors d'un raisonnement par analogie. La grandeur physique cible est généralement normée pour faciliter l'exploitation de la donnée (Taillet *et al.* 2018, p. 27). Par exemple, un relevé technique sur papier millimétré est une représentation normalisée de l'organisation des couches observées sur le terrain. Pour les archéologues, de nos jours, cela correspond essentiellement à la documentation enregistrée sur papier, telle que les fiches d'enregistrement, les notes, les schémas ou les relevés. La majorité du mobilier archéologique, lorsqu'il est considéré comme de la donnée, et aussi quasi systématiquement traduit en valeurs physiques (dessin, mesures, notes) avant d'être conservé en réserve.

Les données numériques, qui nous intéressent plus particulièrement dans le cadre de cette recherche doctorale, correspondent à des signaux qui discréditent une réalité physique en une valeur numérique à base de 0 et de 1, le langage binaire à la base de l'informatique. Il s'agit de toutes les ressources numérisées, qu'elles le soient de manière native ou non (Cardon 2019, p. 22). Elles représentent la très grande majorité des données qui circulent actuellement en archéologie, et s'organisent en plusieurs catégories.

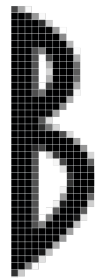
Les données numériques simples

Les images matricielles (raster)

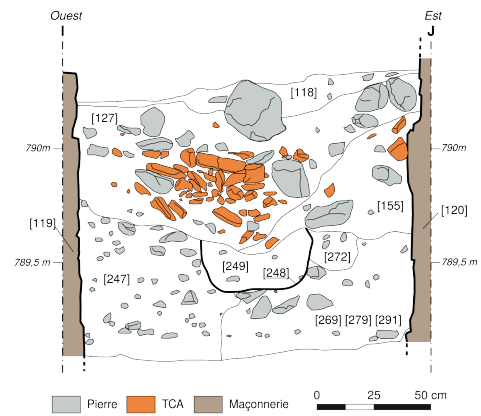
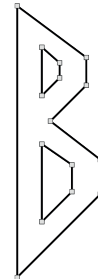
Un raster se compose d'un ensemble de pixels, disposés selon une grille régulière, ou matrice (Bernier *et al.* 2014, p. 20-21) (fig. 34). Le pixel, en tant qu'unité de base, correspond à une cellule de cette grille, généralement représenté sous la forme d'un carré. Chaque pixel possède une valeur interprétée par les programmes informatiques, comme un code couleur (photographie), une altitude (Modèle Numérique de Terrain ou d'Élévation) ou encore une densité (carte de chaleur). C'est l'association de millions de pixels qui crée l'image. Plus la définition d'une image raster est haute, plus les pixels et les informations qu'ils contiennent sont nombreux, donc plus le fichier est volumineux. Dans les faits, ce type de données correspond aux photographies (.jpg, .tiff, *etc.*), aux images d'illustration (.png, .bmp, *etc.*) et à certains documents associant du texte et des images (.pdf).

Image matricielle
(raster)

Image vectorielle
(vecteur)



B I B R A C T E



34. Différence entre les images matricielles et les images vectorielles

Les images vectorielles (vecteur)

Le vecteur résulte de la disposition de points dans l'espace, des nœuds ou sommets, reliés par des arcs calculés qui forment une géométrie (Bernier *et al.* 2014, p. 21-22). Le vecteur est un objet purement géométrique, composé de nœuds avec des coordonnées associées à des formules mathématiques. Les programmes informatiques (visualiseur, logiciel de DAO, PAO, SIG, etc.) interprètent ensuite les formules pour calculer une image. Ainsi, quelle que soit la taille d'un vecteur, le fichier reste léger, car il n'enregistre que les équations, pas des millions de pixels. Trois types de géométrie de base composent le vecteur : le point (un seul nœud), la ligne (des nœuds reliés par des traits) et le polygone (des nœuds reliés entre eux pour constituer une forme fermée) (fig. 35). En plus de cette composante graphique, un vecteur possède également une composante attributaire, soit des informations descriptives qui lui sont propres.



35. Les types de géométrie vectorielle

Les données textuelles

Les données textuelles sont composées de valeurs alphanumériques structurées ou non. Il s'agit la plupart du temps de documents texte, au format Text (.txt), *Microsoft Word* (.doc et .docx) ou *OpenDocument Text* (.odt), ou de données organisées en tableau, de type *Comma-separated values* (.csv), *Microsoft Excel* (.xls ou .xlsx) ou *OpenDocument Spreadsheet* (.ods). Contrairement aux images, les données textuelles sont éditables à l'aide de logiciels de traitement de textes ou de tableau. Les bases de données sont généralement comprises dans cette catégorie, car il s'agit avant tout d'informations alphanumériques structurées. Le langage de base de données utilisé pour renseigner les relations et interroger les données intervient comme une surcouche.

Les données numériques composées

Les données spatiales

On définit comme donnée spatiale “*toutes les données qui peuvent être localisées d’une manière ou d’une autre*” (Deweirdt 2010, p. 8). Pour cela, les données sont associées à des coordonnées qui s’exprime soit dans un espace local, soit dans un Système de Coordonnées de Référence (SCR), tel que le Réseau Géodésique Français 1993 (RGF93) pour la France. Toutes les données simples évoquées précédemment (image matricielle, vecteur, textuelle) peuvent être spatiales. Par nature, la très grande majorité des données archéologiques sont spatiales (Rodier *et al.* 2011, p. 19-20). Chaque élément découvert est normalement rattaché à un lieu : une unité stratigraphique, une unité de prospections ou plus largement un lieu-dit ou une région. L’espace est donc un facteur essentiel de l’archéologie, tant comme composante contextuelle que comme facteur d’analyse (organisation, dynamique, *etc.*).

Les données 3D

Les données 3D sont à l’origine constituées de points qui comportent des valeurs ou coordonnées dans les trois dimensions (souvent notées X, Y et Z) permettant de les positionner dans l’espace (Granier *et al.* 2019, p. 107-110). L’ensemble des points 3D composent un nuage de point, qui modélise plus ou moins finement un objet réel ou virtuel. Les points peuvent être reliés entre eux par des vecteurs pour former des polygones 3D. Le maillage ainsi créé donne une impression de surface à l’objet modélisé. Ce type de données peut être produits par plusieurs techniques qui se développent rapidement ces dernières années en archéologie, tels que les LiDAR, les scanners 3D ou la photogrammétrie.

Le cycle des données

En plus du signal, les données s’organisent en différentes catégories en archéologie, qui correspondent principalement à l’état d’avancement des traitements et des analyses de ces dernières. Dans sa thèse portant sur la gestion informatisée des données en archéologie, A. Chaillou met en garde sur la notion de “*donnée brute*” couramment utilisée pour définir les données provenant directement de la fouille (Chaillou 2003, p. 30-31). S’il n’est pas explicité, cet intitulé peut suggérer qu’il s’agit de données objectives. Or, comme le démontre l’auteur, aucune donnée n’est objective à partir du moment où elle a été décrite par un individu qui utilise des couches d’abstractions telles que des instruments, mais aussi et sur-

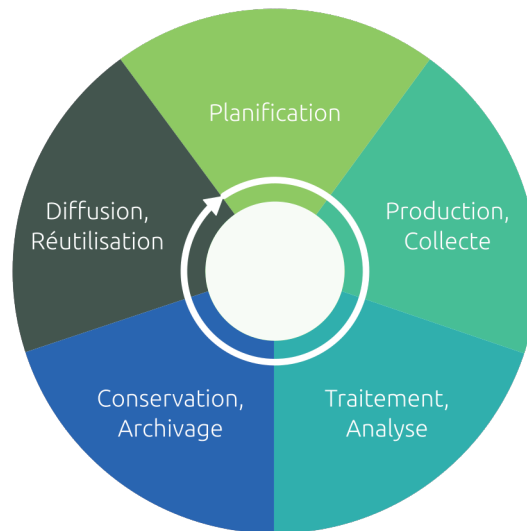
tout, sa culture scientifique. Soit une somme de connaissances acquises dans le cadre d'une formation spécifique et d'expériences qui lui sont propres. Comme aucune observation ne peut être enregistrée de manière rigoureusement identique par tous les potentiels observateurs, A. Chaillou préfère parler de “*données primaires*”, qu'elle complète avec les “*données traitées*” :

- “*Les données primaires sont toutes les données qui proviennent de l'observation directe du site, de la fouille et de la documentation que l'on peut en tirer. Ces données seront donc le mobilier, les objets, les couches, les liens directs entre ces différents éléments ainsi que les enregistrements des photographies, des minutes de terrain, des journaux de fouille, des fiches de saisie sur le terrain et tout ce qui concerne la topographie du lieu (relevés, cartes, [etc.])*”. On peut aujourd'hui ajouter à cette liste les données 3D (photogrammétrie, scanner).
- Les “*données traitées*” sont tout ce qui pourrait correspondre aux étapes de l'information et de la connaissance de la hiérarchie *DIKW*, réutilisées ensuite comme source primaire pour une nouvelle étude. Soit “*la résultante d'études menées sur les données primaires (études de mobilier, [etc.]), d'assemblages de plan, de dessins d'objet ainsi que de tout ce qui touche à l'interprétation de ces données (interprétation des couches archéologiques, [etc.]), à la datation et aux regroupements de données comme les associations de couches (faits, structure, entités archéologiques, [etc.])*”.

On parle également souvent de données de travail, pour définir les données qui seraient dans un entre-deux, des données primaires en cour de traitement. Mais il s'agit la plupart du temps de données primaires ou traitées qui n'ont simplement pas vocation à être diffusées. Toutes ces catégories alimentent finalement ce que les spécialistes appellent le “cycle de vie de la donnée” (Pennock 2007). Par cycle, ils entendent qu'en partageant une donnée, celle-ci peut être réutilisée et donc commencer une nouvelle itération : collecte et création, traitement et analyse, archivage, partage, réutilisation (fig. 36).

Les métadonnées

Ces données sont parfois associées à des métadonnées. Une métadonnée, comme l'indique le préfixe grec *meta*, qui signifie ici l'autoréférence, est une donnée à propos d'une donnée. En informatique, le terme a été proposé par l'ingénieur T. Berners-Lee, considéré comme l'inventeur du *World Wild Web*, lorsqu'il était en poste au CERN (organisation eu-



36. Le cycle de vie des données

ropéenne pour la recherche nucléaire) dans les années 90. Voici comment il décrivait les métadonnées en 1997 (Berners-Lee 1997) :

“Les métadonnées sont des informations compréhensibles par les machines à propos des ressources web et d’autres choses. Les métadonnées ont été appelées “métadonnées” parce qu’elles ont commencé leur vie, et sont encore actuellement principalement des informations sur les ressources web, donc des données sur les données. À l’avenir, lorsque les langages et les moteurs de métadonnées seront plus développés, elles devraient également constituer une base solide pour un réseau d’informations compréhensibles par une machine sur n’importe quoi : sur les personnes, les choses, les concepts et les idées. Nous gardons ce fait à l’esprit lors de la conception, même si la première étape consiste à créer un système d’information sur l’information”.

Deux éléments importants sont présents dans ce passage. Le premier, c’est que les métadonnées sont principalement destinées aux machines. Aussi, les informations doivent être fortement normalisées et structurées pour être traitées par des algorithmes. Le second est déterminé par l’origine des métadonnées en informatique. À l’époque où T. Berners-Lee écrit ces lignes, le réseau internet n’a rien à voir avec ce qu’il est aujourd’hui. Il est en pleine

construction, et beaucoup de concepts ne sont pas encore définis. Les métadonnées étaient une réponse technique pour définir un objet numérique sur le web, essentiellement une page et son contenu, pour qu'il soit correctement interprété par un programme. Mais l'auteur anticipe déjà ce qu'elles sont devenues, un système de description de référence pour l'échange d'information par des machines en réseau.

Si les métadonnées sont principalement destinées aux ordinateurs, smartphones ou autres, elles sont encore largement produites par des humains, en tout cas en archéologie. Plusieurs normes de métadonnées existent, pour structurer la forme et le contenu des informations. Parmi les plus répandues, on trouve la norme Dublin Core²¹. Elle est à l'origine pensée pour la gestion de fonds sous forme bibliographique, mais elle s'ouvre maintenant beaucoup plus largement aux autres types de données. La structure des Dublin Core repose sur 15 éléments descripteurs²², plus ou moins normalisés et formalisés en langage XML (fig. 37).

2.2. Culture et savoir libre : la démarche *Open* en archéologie

Depuis les 30 dernières années, avec le développement rapide des technologies de l'information et de la communication, le flux d'informations circulant sur les réseaux mondiaux ne cesse d'augmenter. Cette dématérialisation massive pose la question des conditions d'accès et de réutilisation des données. Une des réponses du monde de la recherche est le développement de la démarche *Open*, un mouvement inscrit dans la culture universitaire anglo-saxonne, qui milite pour le partage sans contrainte de la connaissance. L'*Open* est par définition intimement lié aux pratiques scientifiques numériques actuelles, et l'archéologie ne déroge pas à cette règle. Il est donc essentiel de présenter ce courant de pensée, ainsi que ses apports techniques et juridiques, pour comprendre ses implications sur l'exercice de la discipline archéologique.

2.2.1. *Open* : le savoir libre

Historiquement, la communauté du libre trouve sa source parmi les *hackers*, des individus “*qui aime[nt] exercer [leur] ingéniosité de façon ludique*”²³, du *Massachusetts Institute of Technology* et quelques autres universités des États-Unis dans les années 1970. Il s'agit d'un

21 <https://dublincore.org/resources/metadata-basics/> ; consulté le 21/10/2023

22 <https://www.dublincore.org/specifications/dublin-core/dces/> ; consulté le 21/10/2023

23 <https://www.gnu.org/philosophy/words-to-avoid.html#Hacker> ; consulté le 21/10/2023

Élément	Élément (anglais)	Commentaire
Titre	Title	Nom donné à la ressource
Créateur	Creator	Nom de la personne, de l'organisation ou du service responsable de la création du contenu de la ressource
Sujet	Subject	Thème du contenu de la ressource (mots clés, expressions, codes de classification)
Description	Description	Présentation du contenu de la ressource (résumé, table des matières, représentation graphique du contenu, texte libre)
Éditeur	Publisher	Nom de la personne, de l'organisation ou du service responsable de la mise à disposition ou de la diffusion de la ressource
Contributeur	Contributor	Nom de la personne, de l'organisation ou du service responsable de contributions au contenu de la ressource
Date	Date	Date de création ou de mise à disposition de la ressource
Type	Type	Nature ou genre de la ressource (catégories, fonctions, genres généraux, niveaux d'agrégation du contenu)
Format	Format	Manifestation physique ou numérique de la ressource
Identifiant de la ressource	Identifier	Référence univoque à la ressource dans un contexte donné (URI, ISBN)
Source	Source	Référence à une ressource dont la ressource décrite est dérivée (URI)
Langue	Language	Langue du contenu intellectuel de la ressource
Relation	Relation	Référence à une ressource apparentée
Couverture	Coverage	Couverture spatio-temporelle de la ressource (domaine d'application)
Gestion de droits	Rights	Informations sur les droits associés à la ressource (IPR, copyright, etc.)

mouvement social qui prône le respect de la liberté des utilisateurs de logiciels et qui s’officialise à l’initiative de R. Stallman en 1985, lors de la création de la Fondation pour le Logiciel Libre²⁴ (*FSF – Free Software Foundation*). Avec la popularisation d’internet à la fin des années 1990, les possibilités de partager de l’information explosent et la notion d’*Open* apparaît, par opposition aux modèles propriétaires qui verrouillent les possibilités de réutilisation et de distribution de ces informations (Machado 2015, p. 190-192). L’*Open Knowledge Foundation* (OKF), une association internationale créée en 2004, la décrit comme suit (OKF s. d.) :

“Open means anyone can freely access, use, modify, and share for any purpose (subject, at most, to requirements that preserve provenance and openness)”²⁵.

Nous utilisons volontairement cette définition dans sa langue d’origine pour insister sur un point qui peut porter à confusion. Le mot “*freely*” est ici entendu comme “librement”, et non pas comme “gratuitement”, ce qui est une erreur couramment faite pour décrire la démarche *Open* (*FSF 2019*). Même si ce n’est pas recommandé, la monétisation ne va pas à l’encontre de la philosophie du libre. Ni la provenance d’ailleurs, car la source peut être citée et, s’il y a réutilisation ou transformation, ce qui est libre doit le rester.

Tous ces aspects sont précisés dans la version détaillée de la définition de l’OKF²⁶. Pour les besoins de notre étude, nous proposons en annexe une traduction de la version la plus aboutie de ce document, la version 2.1 (Annexe 1). N’étant toutefois pas particulièrement familier avec les traductions de textes juridiques, nous renvoyons le lecteur vers la version originale du document, en anglais, en cas de doute ou d’incompréhension.

Cette définition, par la notion d’œuvre, souhaite englober le maximum de productions intellectuelles. Pourtant, il apparaît évident, à la lecture du document, qu’elle touche en premier lieu les productions numériques. Dans le préambule d’abord, quand elle fait appel à des définitions provenant des communautés du logiciel libre, puis dans le texte ensuite, quand elle formalise des principes tels que l’accès par internet ou l’interopérabilité de l’œuvre.

Selon l’Association Francophone des Utilisateurs de Logiciels Libres (AFUL) et le Référentiel Général d’Interopérabilité (RGI), “*l’interopérabilité est la capacité que possède un*

24 <https://www.fsf.org/history/> ; consulté le 21/10/2023

25 <http://opendefinition.org/> ; consulté le 21/10/2023

26 <https://opendefinition.org/od/2.1/en/> ; <https://opendefinition.org/od/2.1/en/>

*produit ou un système, dont les interfaces sont intégralement connues, à fonctionner avec d'autres produits ou systèmes existants ou futurs et ce sans restriction d'accès ou de mise en œuvre*²⁷ (DINUM 2015, p. 7). Il s'agit d'une vision très orientée vers la technique, qui décrit la capacité de données numériques et d'algorithmes à fonctionner ensemble et/ou entre eux. Mais le sens de l'interopérabilité peut également être élargi, tel que l'entend la Commission Européenne : *“l'interopérabilité est l'aptitude d'organisations disparates et diverses à interagir en vue de la réalisation d'objectifs communs mutuellement avantageux, arrêtés d'un commun accord, impliquant le partage d'informations et de connaissances entre ces organisations à travers les processus métiers qu'elles prennent en charge, grâce à l'échange de données entre leurs systèmes de TIC [(technologies de l'information et de la communication)] respectifs”* (JOUE 2009, p. 23).

Finalement, l'interopérabilité définit la capacité d'entités, sociales ou techniques, à fonctionner ensemble par le partage d'informations. Cela signifie que les individus ou groupes doivent être ouverts et formés au partage, et que les machines, algorithmes ou données doivent nativement pouvoir fonctionner en dehors d'interactions prédéfinies. On comprend alors l'importance que revêt cette notion pour la démarche du savoir libre.

2.2.2. *Open science* : la science ouverte

Quelle est la valeur du fruit d'une recherche s'il reste inconnu ? Contrairement aux milieux artistiques ou industriels, qui se rémunèrent la plupart du temps de la vente directe de leur production, l'objectif premier du milieu scientifique est, généralement, de diffuser le plus largement possible leurs productions. Pour authentifier la qualité de la recherche par les pairs d'abord, et pour contribuer à l'avancée de la discipline ensuite. La science ouverte est donc l'idée de partager les méthodes et les résultats scientifiques, et plus largement la connaissance, au plus grand nombre.

Un objectif aussi vaste suscite forcément beaucoup d'interprétations. C'est pourquoi deux chercheurs de l'Université de Valencia, R. Vicente-Saez et C. Martinez-Fuentes, ont mené une étude en 2017 compilant des centaines de mentions d'*Open Science*. Ils ont retenu 75 publications qui en offraient une définition ou affichaient des données qui soutiennent son interprétation. L'objectif était de livrer les textes à une analyse sémantique pour faire ressortir une définition universelle du concept d'*Open Science*, qu'il propose en fin d'article : *“La*

27 <https://aful.org/gdt/interop> ; consulté le 21/10/2023

science ouverte est [défini par] la connaissance transparente et accessible qui est partagée et développée par des réseaux de collaboration” (Vicente-Saez, Martinez-Fuentes 2018, p. 7). Cette définition, bien que succincte, à l’avantage de faire ressortir tous les concepts en lien avec la science ouverte : la transparence, le partage et le travail collaboratif.

Pour une définition plus longue, mais plus explicite, nous pouvons aussi citer celle proposée par le projet européen FOSTER (*Fostering the practical implementation of Open Science in Horizon 2020 and beyond*) : “*La science ouverte est la pratique de la science de manière à ce que chacun puisse collaborer et contribuer, où les données de recherche, les notes de laboratoire et les autres processus de recherche sont librement accessibles, dans des conditions qui permettent la réutilisation, la redistribution et la reproduction de la recherche et de ses données et méthodes intrinsèques*”²⁸.

La science ouverte est donc un mouvement étroitement lié à l’essor d’internet, qui prône la mise en réseau des acteurs et le partage massif des savoirs pour développer l’émulation scientifique. Dans la pratique, elle s’articule autour de trois piliers essentiels à la préservation de la connaissance : l’*Open Access*, l’*Open Data* et l’*Open Source* (Di Cosmo, Zacchiroli 2017, p. 4).

2.2.3. *Open access* : le libre accès, ou accès ouvert, aux publications scientifiques

Par accès ouvert, on entend la mise à disposition du produit final d’une recherche : la publication. Encore une fois, c’est le passage de l’imprimé au numérique qui a permis de faire évoluer le modèle économique du monde de l’édition. L’accès ouvert se base sur 3 déclarations publiques du début des années 2000 : l’*Open Access Initiative* de Budapest²⁹, la Déclaration de Bethesda sur la publication en accès ouvert³⁰ et la Déclaration de Berlin sur l’accès ouvert aux connaissances dans les sciences et les humanités³¹. Le professeur de philosophie P. Suber, rédacteur principal de la première et co-auteur de la seconde, propose une “*définition BBB*” (Budapest, Bethesda, Berlin) synthétique : “*La littérature en accès ouvert est numérique, en ligne, gratuite et exempte de la plupart des restrictions en matière de droits d’auteur et de licences.*” (Suber 2012, p. 4-8).

28 <https://www.fosteropenscience.eu/taxonomy/term/100> ; consulté le 21/10/2023

29 <https://www.budapestopenaccessinitiative.org/read/> ; consulté le 21/10/2023

30 <http://legacy.earlham.edu/~peters/fos/bethesda.htm> ; consulté le 21/10/2023

31 <https://openaccess.mpg.de/Berliner-Erklaerung> ; consulté le 21/10/2023

Par “*la plupart des restrictions*”, l’auteur indique que la filiation d’un document n’est pas contradictoire avec l’accès ouvert. Il est possible de le partager en accès ouvert tout en gardant le contrôle de l’intégrité des données et le droit d’être cité comme auteur. Les logiciels antiplagiat fonctionnent d’ailleurs mieux avec les publications en libre accès sur le web (INSA 2021, p. 13). Une autre croyance voudrait que l’accès ouvert soit un moyen de se passer des comités de lecture pour l’évaluation d’une publication. Pourtant, le libre accès aux publications et le processus de validation sont deux sujets bien distincts (Suber 2012, p. 99-104). De nombreuses revues de qualité proposent aujourd’hui leur contenu en accès ouvert, tout en conservant un système de validation par les pairs. En archéologie notamment, on retrouve les revues *Journal of Archeological Science*³², *Gallia*³³ ou encore la *Revue archéologique de l’Est*³⁴.

L’accès ouvert est avant tout un moyen de garantir l’accès à la connaissance en s’affranchissant de barrières financières ou de conditions d’accès restrictives. La pratique est d’ailleurs largement répandue dans le milieu académique, comme le démontre le travail de bibliographie de Jr. Charles W. Bailey. En 2013, il référençait déjà plus de 1100 travaux universitaires sur le mouvement de l’accès ouvert³⁵.

En donnant accès au résultat d’une recherche, l’accès ouvert est paradoxalement la première étape de l’ouverture de la connaissance. Pourtant, si une publication permet d’authentifier et d’évaluer l’existence d’une expérience ou d’une hypothèse, un simple texte accompagné de figures peut se révéler complexe à réexploiter. Une relecture de l’initiative de Budapest en février 2017 insiste d’ailleurs sur l’intérêt pour les publications d’offrir “*la possibilité de les transmettre sous forme de données à un logiciel ou les utiliser à toute autre fin licite*” (Alperin *et al.* 2017). La deuxième étape consiste donc à donner l’accès aux données primaires.

32 <https://www.sciencedirect.com/journal/journal-of-archaeological-science> ; consulté le 21/10/2023

33 <https://journals.openedition.org/gallia/> ; consulté le 21/10/2023

34 <https://journals.openedition.org/rae/> ; consulté le 21/10/2023

35 <http://digital-scholarship.org/tsp/w/tsp.html> ; consulté le 21/10/2023

2.2.4. *Open data* : les données ouvertes

Le partage de toutes les données

Dans un article de 2015, qui fait le point sur la science ouverte et les données ouvertes, J. Machado résume parfaitement la situation : “*Le débat actuel sur l'accès ouvert à la connaissance scientifique se croise avec celui sur les données ouvertes. De nos jours, les protocoles, les formats et les plateformes qui favorisent l'interopérabilité, le traitement, le croisement et la réutilisation des informations sont très répandus. L'accès ouvert traditionnel, tel qu'il a été discuté au cours des 10 ou 15 dernières années, est aujourd'hui dépassé et insuffisant au regard des progrès réalisés dans l'utilisation et la réutilisation de bases de données ouvertes*” (Machado 2015, p. 190).

Au départ, les données ouvertes reposent sur 8 principes fondateurs, qui ont été établis en 2007 à l'attention des dirigeants politiques américains, pour souligner l'importance de la transparence des données gouvernementales dans une démocratie³⁶ (Annexe 2). Les auteurs entendent par données “*toutes informations ou enregistrements conservés électroniquement*”. C'est une définition nécessairement très vaste, puisque nous rappelons qu'elle concerne toutes les données publiques que peut potentiellement manipuler un gouvernement, issus de domaines variés tels que le social, l'éducation, ou encore l'économie. Le lecteur comprend aussi clairement que l'essor rapide des technologies de l'information et de la communication est à la base de ces déclarations. L'objectif est autant de faire prendre conscience aux gouvernements de la richesse des données numériques que de prévenir d'éventuelles dérives concernant leurs utilisations. Pour autant, les données ouvertes ne sont pas au-dessus des lois : des restrictions de sécurité, d'embargo ou liées à la protection de libertés individuelles peuvent intervenir en fonction des législations.

Les principes Panton³⁷, rédigés en 2010, se concentrent, quant à eux, plus spécifiquement sur les données ouvertes dans le domaine scientifique. Les auteurs soulignent la nécessité de mettre à disposition les données de la recherche pour pratiquer une science “*fondée sur la construction, la réutilisation et la critique ouverte de l'ensemble des connaissances scientifiques publiées*”. En ce sens, ils insistent sur les conditions de partage, les plus libres et explicites possibles, et ils conseillent un certain nombre de licences adaptées à la diffusion

36 https://public.resource.org/8_principles.html ; consulté le 21/10/2023

37 <https://pantonprinciples.okfn.org/index.html> ; consulté le 21/10/2023

des données. Ils préconisent même que les données scientifiques entrent, autant que possible, dans le domaine public, c'est-à-dire sans aucune contrainte liée à la propriété intellectuelle.

Toujours selon ces principes, est entendue comme donnée ce *“qui est librement disponible sur internet, permettant à tout utilisateur de le télécharger, de le copier, de l'analyser, de le retraiter, de le transmettre à un logiciel ou de l'utiliser à toute autre fin sans obstacle financiers, juridiques ou techniques autres que ceux indissociables de l'accès à l'internet lui-même”*. La définition découle directement de celle du savoir libre de l'OKF évoquée plus tôt, en tenant compte des aspects techniques propres aux données. On comprend alors que pour la démarche des données ouvertes, l'objectif n'est donc pas de définir ce qu'est à proprement parlé une donnée, mais plutôt d'encourager le partage libre de celle-ci. Par conséquent, les données archéologiques, telles que nous les avons définies précédemment, peuvent tout à fait être ouvertes.

De la théorie à la mise en oeuvre : les principes FAIR

Si les principes de la donnée ouverte existent, un groupe de 53 représentants des mondes académique, de l'industrie, des agences de financement et de l'édition scientifique ont constaté que leur mise en pratique n'était pas si aisée (Wilkinson *et al.* 2016). Selon eux, la difficulté à réutiliser des données de la recherche, même ouvertes, ne vient pas d'un verrou technologique, mais d'un manque d'attention des producteurs au moment de la production et de l'archivage des données. Ce constat est pour eux d'autant plus dommageable que *“la qualité [d'une] publication - et, plus important encore, l'impact de la publication - est fonction de sa capacité à être trouvée, réutilisée et citée de manière précise et appropriée au fil du temps, par toutes les parties prenantes, tant humaines que mécaniques”*. Or cette capacité prend tout son sens quand elle est appliquée aux données en lien avec la publication.

Les auteurs identifient plusieurs barrières à l'exploitation des données, en opposant l'humain et la machine. L'humain possède un sens intuitif de la sémantique d'un objet numérique, il est capable d'identifier et d'interpréter une large variété d'éléments constitutifs ou de mise en forme pour comprendre la signification ou l'intention de cet objet. Il possède ainsi une bonne aptitude pour constituer des corpus logiques à partir de données hétérogènes. Cependant, il *“est incapable de fonctionner à la portée, à l'échelle et à la vitesse requises par l'ampleur et la complexité des données scientifiques numériques contemporaines”*. Au contraire, la machine est, en théorie, capable de gérer toutes les dimensions de ces données, en relative autonomie. Mais pour cela, les données doivent présenter un niveau de détail nor-

malisé suffisant, sans quoi elles ne peuvent pas être identifiées, échantillonnées et utilisées par les algorithmes.

L'objectif du groupe était donc de minimiser l'investissement humain au départ, en améliorant la capacité des données à être découvertes et traitées par des machines, pour maximiser le potentiel de réutilisation et d'analyse ensuite. Cela s'est traduit par la création de principes FAIR³⁸ (pour *Findable, Accessible, Interoperable, Reusable*), que l'on peut résumer ainsi :

- **Facile à trouver** : la première étape de la (ré)utilisation des données consiste à les trouver. Les métadonnées et les données doivent être faciles à trouver, tant pour les humains que pour les ordinateurs. Les métadonnées lisibles par machine sont essentielles pour la découverte automatique des ensembles de données et des services.
- **Accessibilité** : une fois que l'utilisateur a trouvé les données requises, il doit savoir comment y accéder, éventuellement sous réserve d'authentification et d'autorisation.
- **Interopérable** : les données doivent généralement être intégrées à d'autres données. En outre, les données doivent être interopérables avec des applications ou des flux de travail pour l'analyse, le stockage et le traitement.
- **Réutilisable** : les métadonnées et les données doivent être bien décrites afin qu'elles puissent être reproduites et/ou combinées dans différents contextes.

La mise en œuvre des données FAIR (*FAIR data*) repose sur deux notions indispensables : les métadonnées et l'interopérabilité. Les métadonnées doivent être suffisamment structurées pour permettre un traitement rapide par la machine, et rester intelligibles pour l'humain, même les non spécialistes des sciences de l'information et de la communication. Le groupe définit l'interopérabilité selon des termes très proches de ce que nous avons déjà évoqué : *“la capacité des données ou des outils provenant de ressources non coopératives à s'intégrer ou à fonctionner ensemble avec un effort minimal”* (Wilkinson *et al.* 2016). Dans la pratique, il s'agit essentiellement d'utiliser des formats de données et des structures de logiciels standardisés, simples et ouverts. Si les recommandations semblent simples, elles restent, dans les faits, complexes à appliquer. Surtout dans les disciplines qui n'ont pas développé de pratiques généralisées sur l'harmonisation et le partage des données, comme c'est le cas en archéologie.

38 <https://www.go-fair.org/fair-principles/> ; consulté le 21/10/2023

2.2.5. *Open Source* et *Free software* : deux visions du logiciel libre

Lorsque l'information est disponible par une publication (accès ouvert) et que les données primaires sont documentées et partagées (données ouvertes), il reste une dernière étape à franchir pour maximiser les chances de réutilisation du corpus : le libre accès au logiciel d'exploitation. *Software Heritage*, une initiative internationale pour la conservation du code de programmes libres, soutenue par le Plan national pour la Science Ouverte, considère que le logiciel est un élément indissociable de la donnée, indispensable pour l'ouverture de la connaissance (Di Cosmo, Zacchiroli 2017, p. 1 ; MESRI 2018, p. 6).

Comme son nom l'indique, le *Free Software* est le premier-né du mouvement social du logiciel libre dans les années 1970. La *Free Software Foundation* propose dès 1986 une première définition de ce qu'est le logiciel libre, que la communauté a fait évoluer par la suite à travers de nombreuses versions. La définition plus récente date de 2019 et présente les logiciels libres comme des logiciels qui respectent la liberté des utilisateurs, soit "*la liberté d'exécuter, copier, distribuer, étudier, modifier et améliorer ces logiciels*³⁹". Il faut donc bien entendre ici "libre" (*free* en anglais) au sens de la liberté offerte par le logiciel, pas de sa gratuité. Quatre "*libertés essentielles*", qui s'adressent à l'utilisateur du programme, accompagnent la définition :

- "*la liberté de faire fonctionner le programme comme vous voulez, pour n'importe quel usage (liberté 0) ;*
- *la liberté d'étudier le fonctionnement du programme, et de le modifier pour qu'il effectue vos tâches informatiques comme vous le souhaitez (liberté 1) ; l'accès au code source est une condition nécessaire ;*
- *la liberté de redistribuer des copies, donc d'aider les autres (liberté 2) ;*
- *la liberté de distribuer aux autres des copies de vos versions modifiées (liberté 3) ; en faisant cela, vous donnez à toute la communauté une possibilité de profiter de vos changements ; l'accès au code source est une condition nécessaire*".

La conception d'un logiciel libre tient donc autant de normes de développement que d'une action politique et sociale pour la défense des libertés collectives et individuelles.

39 <https://www.gnu.org/philosophy/free-sw.html> ; consulté le 21/10/2023

Un autre groupe de la communauté du libre, souhaitant une approche plus “*pragmatique et accessible*” qui relève plutôt de “*la persuasion discrète que de l’activisme public*”, a fondé en 1998 l’*Open Source Initiative*⁴⁰. Là encore, la définition de l’*Open Source*, parfois traduit par code source ouvert ou code libre, repose sur 10 principes⁴¹, inspirés par le Contrat Social du système d’exploitation Debian de 1997⁴², qui ont évolué jusqu’en 2007 (Annexe 3).

La différence entre *Free Software* et *Open Source* réside surtout dans la philosophie des concepteurs de chacune des notions⁴³. Pour la première, il s’agit avant tout d’un mouvement qui œuvre pour une société libre, éthique et juste à travers la diffusion d’un modèle ouvert dans un monde géré de plus en plus par le numérique. Pour la seconde, les concepteurs proposent une réponse pratique pour améliorer l’accès aux logiciels, outils indispensables pour le développement des sociétés actuelles. Les personnes qui portent les deux philosophies emploient les termes de FOSS, pour Free and Open Source Software, ou comme le préfère la communauté du logiciel libre, FLOSS, pour Free/Libre and Open Source Software⁴⁴.

2.2.6. Le libre et la loi : protéger les œuvres de l’esprit et leurs auteurs

En France, une des premières références officielles à la communauté du libre est la loi de 2004 “*pour la confiance dans l’économie numérique*” (JORF 2004). L’article 4 définit le standard ouvert par “*protocole de communication, d’interconnexion ou d’échange et tout format de données interopérable et dont les spécifications techniques sont publiques et sans restriction d’accès ni de mise en œuvre*”. Il faut ensuite attendre la loi de 2016 “*pour une République numérique*”, dont la première partie porte sur “*la circulation des données et du savoir*”, pour inscrire des principes du savoir libre dans la législation française (JORF 2016). Elle permet notamment l’ouverture des données produites par les administrations publiques, ou la possibilité pour les chercheurs de publier rapidement et librement, malgré les restrictions pratiquées par certains éditeurs.

Deux ans plus tard, en 2018, le Ministère de l’Enseignement Supérieur, de la Recherche et de l’Innovation (MESRI) crée le Plan national pour la Science Ouverte (PnSO), un appel aux institutions de la recherche publique à s’inscrire dans le mouvement de la science ouverte.

40 <https://opensource.org/history/> ; consulté le 21/10/2023

41 <https://opensource.org/osd/> ; consulté le 21/10/2023

42 https://www.debian.org/social_contract#guidelines ; consulté le 21/10/2023

43 <https://www.gnu.org/philosophy/open-source-misses-the-point.fr.html> ; consulté le 21/10/2023

44 <https://www.gnu.org/philosophy/floss-and-foss.en.html> ; consulté le 21/10/2023

Il a permis, entre autres, de créer un fonds national dédié, de soutenir les institutions déjà en place et de développer la pratique de la science ouverte sur le territoire et à l'international (MESRI 2021a). Le Plan a d'ailleurs été reconduit sur la période 2021-2024 (MESRI 2021b).

Dans la suite directe du PnSO, le CNRS a publié une *“feuille de route [...] pour la science ouverte”*, qui vise à garantir l'accès libre à la totalité des productions scientifiques de l'institution en agissant pour la formation des acteurs et le développement des infrastructures nécessaires (CNRS 2019).













Enfin, un groupe composé des principales organisations de financement européennes, dont l'Agence nationale de la recherche (ANR) fait partie, s'est organisé pour produire le Plan S (cOAlition S 2021). Ce dernier, qui a pris effet au 1er janvier 2021, impose que *“toutes les publications académiques sur les résultats de la recherche financée par des subventions publiques ou privées accordées par des conseils de recherche et des organismes de financement nationaux, régionaux et internationaux doivent être publiées dans des revues en libre accès, sur des plateformes en libre accès ou mises à disposition immédiatement et sans embargo par le biais de dépôts en libre accès”*.

Pour résumer, soutenir l'ouverture des résultats et des données de la recherche est dorénavant une volonté nationale et européenne, car les gouvernements ont bien compris les enjeux scientifiques, politiques et financiers qui se jouent à travers la diffusion libre de la connaissance.

Du point de vue juridique, le statut ouvert d'une publication, d'une donnée ou d'un logiciel est défini par la licence de partage, qui détermine les droits de propriété intellectuelle de l'objet diffusé. Plusieurs initiatives se sont organisées pour protéger l'objet numérique libre (ce qui est libre doit le rester), ainsi que ses auteurs. Ces licences libres s'inspirent du logiciel libre et respectent 4 libertés : le libre usage, la libre copie, la libre modification et la libre diffusion des modifications (Benabou, Farchy 2007, p. 7). Certaines licences ne respectant qu'une ou plusieurs de ces libertés peuvent parfois être qualifiées de licences ouvertes, plutôt que licences libres.

Beaucoup de licences libres et ouvertes coexistent aujourd'hui. Nous ne citerons ici que les plus couramment utilisées dans notre discipline sur le territoire national. Tout d'abord, depuis le décret du 27 avril 2017, les administrations publiques françaises sont incitées à dif-

fuser les informations publiques sous la Licence Ouverte/Open Licence⁴⁵ (LO/OL) d'Etalab ou l'Open Database Licence⁴⁶ (OdbL) de l'OKF (Etalab 2017 ; JORF 2017). Il est également courant de rencontrer la licence CeCILL⁴⁷ du CNRS, adaptée à la diffusion de logiciels libres. Enfin, depuis plusieurs années maintenant se développe l'utilisation généralisée des *Creative Commons*⁴⁸ (CC), un ensemble de 7 licences créées par l'association internationale du même nom, fondée en 2001 (fig. 38).

Termes	Licences	Sigles	Licences libres
 L'auteur initial a dédié l'œuvre au domaine public.		CC0 1.0	✓
 Attribution. Vous devez créditer l'auteur initial.		CC BY 4.0	✓
 Pas d'utilisation commerciale. Vous n'êtes pas autorisé à faire un usage commercial de l'oeuvre.		CC BY-SA 4.0	✓
 Partage dans les mêmes conditions. Vous devez diffuser l'oeuvre modifiée avec la même licence.		CC BY-NC 4.0	
 Pas de modifications. Vous n'êtes pas autorisé à distribuer ou mettre à disposition l'oeuvre modifiée.		CC BY-ND 4.0	
		CC BY-NC-SA 4.0	
		CC BY-NC-ND 4.0	

38. Les licences Creative Commons

45 <https://www.etalab.gouv.fr/licence-ouverte-open-licence/> ; consulté le 21/10/2023

46 <https://opendatacommons.org/licenses/odbl/1-0/> ; consulté le 21/10/2023

47 <http://www.cecill.info/index.fr.html> ; consulté le 21/10/2023

48 <https://creativecommons.org/share-your-work/> ; consulté le 21/10/2023

On l'aura compris, toutes les branches du mouvement du libre sont en constante évolution, en lien direct avec les technologies de partage qui se développent massivement depuis une trentaine d'années. Toutes tendent d'ailleurs, au fur et à mesure des modifications, vers une définition commune. On y retrouve des valeurs comme la transparence de l'information, favoriser la réutilisation des données ou encore encourager le travail collaboratif. Ce mouvement est donc par nature très lié au monde de la recherche, à la base de la production du savoir.

En résumé, la science ouverte représente aujourd'hui un mouvement essentiel, qui a pour objectif d'assurer la pérennité et le partage universel de la connaissance, sans barrières cognitives, techniques ou financières, et ce afin de maintenir une recherche de qualité et de promouvoir l'innovation.

2.2.7. L'archéologie : une science ouverte ?

Des initiatives multiples pour une archéologie libre

L'un des pionniers de l'ouverture des données archéologiques est, comme nous l'avons vu précédemment, l'*Archaeology Data Service* (ADS). Toutes les données hébergées sur l'ADS sont libres. Le service a même créé en 2017 des "*termes d'utilisation et d'accès aux données*⁴⁹", libres par défaut, pour les jeux de données ne citant pas les *Creatives Commons*. Il en va de même avec son homologue nord-américain, *the Digital Archaeological Record* (tDAR) du *Center for Digital Antiquity* (*Arizona State University*), qui stipule, dans l'accord des contributeurs que, sans mention spécifique, les fichiers sont diffusés sous licence libre CC BY 3.0⁵⁰.

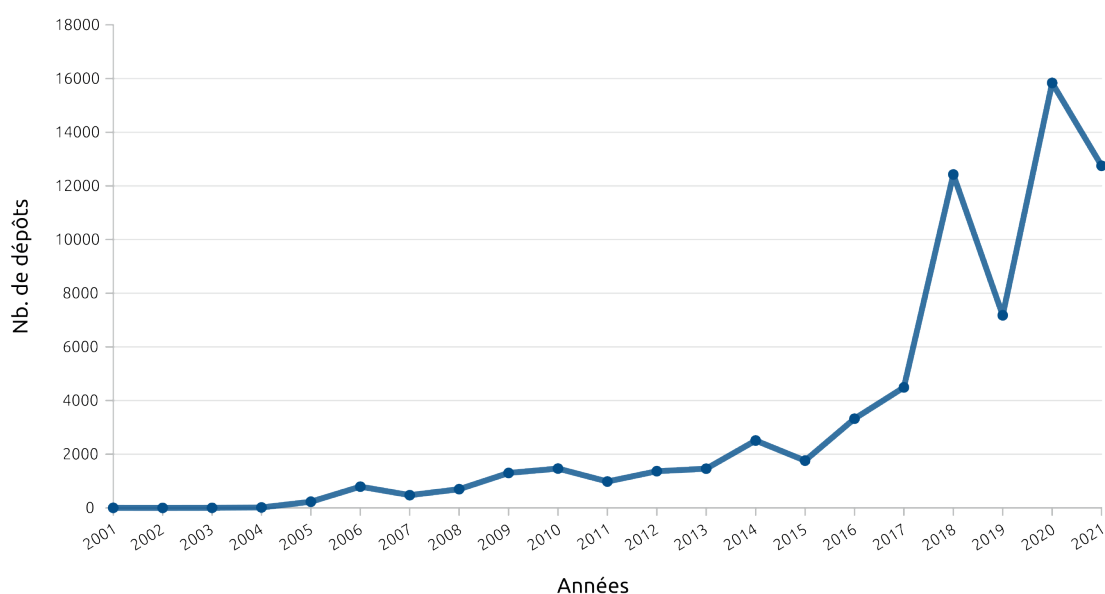
Sur notre territoire, à l'heure actuelle, le modèle est différent. La communauté archéologique française utilise plusieurs outils pour diffuser les données et les résultats librement. Ces outils ne sont d'ailleurs pas nécessairement destinés au seul public des archéologues. Comme souvent, les plus connus sont orientés vers l'*open access*. On retrouve tout d'abord l'environnement de diffusion scientifique libre HAL⁵¹. Créé en 2001 par le CNRS, il est aujourd'hui soutenu par le PnSO et accueille plus de 130 portails thématiques et institution-

49 <https://archaeologydataservice.ac.uk/advice/termsOfUseAndAccess.xhtml> ; consulté le 21/10/2023

50 Article B. <https://www.tdar.org/about/policies/contributors-agreement/> ; consulté le 21/10/2023

51 <https://hal.science/> ; consulté le 21/10/2023

nels⁵², tous interrogeables depuis le site principal. Sans entrer dans les détails, il faut noter que si le nombre de ressources déposées (notices et publications) mentionnant le terme "archéologie", dans le titre ou en mot clé, n'a cessé d'augmenter depuis 2001, la plateforme n'est devenue un lieu de dépôt privilégié de la discipline qu'à partir de 2018 (fig. 39).



39. Nombre de ressources déposées sur HAL contenant le mot clé "archéologie"

Persee.fr, de l'UMS 3602 Persée, est un autre portail couramment utilisé en archéologie pour partager des publications⁵³. Sa particularité repose sur le fait qu'il héberge de nombreux documents anciens, pour la plupart traités avec un système de reconnaissance automatique de caractères (*Optical Character Recognition* ou OCR), les rendant exploitables par des outils d'édition informatique.

52 <https://www.ccsd.cnrs.fr/portails-hal/> ; consulté le 21/10/2023

53 <https://www.persee.fr/a-propos> ; consulté le 21/10/2023

L'infrastructure de recherche nationale OpenEdition, portée par l'UAR 2504 OpenEdition Center, propose également un environnement complet d'édition numérique libre en SHS, qui accueille à ce jour 23 revues d'archéologie d'envergure régionale et nationale⁵⁴. Des initiatives institutionnelles existent aussi, à l'image d'Éveha OpenArchive⁵⁵, un portail qui centralise depuis 2015 une partie de la production scientifique la société d'archéologie préventive du même nom. Malheureusement le site être inactif depuis quelques temps. Dans le même esprit, le catalogue Dolia permet de consulter les notices bibliographiques de l'Inrap, mais seuls les agents de l'institut ont accès aux données.⁵⁶

Plusieurs logiciels libres (FLOSS) spécifiquement conçus pour les archéologues existent également, même s'ils sont peu nombreux. En France, le plus connu est probablement Syslat⁵⁷, un environnement numérique multiplateforme pour l'enregistrement, l'exploitation et l'analyse des données de fouilles. Le logiciel Ishtar, de la société d'informatique libre Igdrasil, se concentre essentiellement sur l'après-chantier, avec la gestion des archives et du mobilier de fouilles⁵⁸. Pour la phase d'analyse, on peut citer Chronomodel, un outil d'aide à la construction de chronologies à partir d'analyses statistiques⁵⁹. Quelques applications collaboratives en ligne existent également, à l'image des bases de données Artefacts⁶⁰, une encyclopédie collaborative pour l'étude du petit mobilier archéologique, ou Dicocer⁶¹, pour Dictionnaire des céramiques anciennes.

Il nous semble utile de mentionner les applications développées par B. Desachy, même s'il ne s'agit pas strictement d'outils libres. Le premier, le Stratifiant⁶², permet de produire automatiquement un diagramme de Harris et de contrôler les erreurs logiques de l'enregistrement des relations US. Le second, le sériographe EPPM, offre la possibilité de présenter graphiquement sous la forme de tableur des associations significantes à plus de deux variables (par exemple entre plusieurs types de mobiliers selon une chronologie) (Desachy 2004). Pour

54 <https://www.openedition.org/catalogue-journals?searchdomain=catalogue-journals&q=arch%C3%A9ologie> ; consulté le 21/10/2023

55 <https://openarchive.eveha.fr> ; consulté le 21/10/2023

56 <https://dolia.inrap.fr/> ; consulté le 21/10/2023

57 <http://www.syslat.fr/> ; consulté le 21/10/2023

58 <https://ishtar-archeo.net/> ; consulté le 21/10/2023

59 <https://chronomodel.com/> ; consulté le 21/10/2023

60 <https://artefacts.mom.fr/fr/fr/home.php> ; consulté le 21/10/2023

61 <http://dicocer.cnrs.fr/> ; consulté le 21/10/2023

62 <https://abp.hypotheses.org/3965> ; consulté le 21/10/2023

les deux outils, l'auteur a fait le choix de créer des macro VBA exploitables avec le logiciel propriétaire Microsoft Excel. Il considérait qu'utiliser ce logiciel déjà largement répandu chez les archéologues faciliterait la prise en main par les utilisateurs (Desachy 2008, vol. 2, p. 4-6).

Enfin, de nombreux autres logiciels libres sont couramment utilisés par les archéologues, même s'ils n'ont pas été développés spécifiquement pour eux. Parmi les plus célèbres, on retrouve le Système d'Information Géographique QGIS⁶³, la suite bureautique Libre Office⁶⁴ ou Apache Open Office⁶⁵, le gestionnaire de bibliographie Zotero⁶⁶, ou plus récemment le logiciel de dessin vectoriel Inkscape⁶⁷.

Sur le partage de données ouvertes (*open data*), les solutions numériques sont aussi multiples en France. Le service de référence est Nakala⁶⁸, un entrepôt de données scientifiques en ligne pour les SHS. Mis en place depuis 2015 par l'IR* Huma-Num, il a pour objectif d'assurer la publication et la réutilisation des données en fournissant un lien unique et pérenne vers chaque élément hébergé. D'autres outils de partage des données sont plus orientés vers l'archéologie, tels que Chronocarto⁶⁹, la plateforme de partage de l'UMR 8546 AorOc et ArkeoGIS⁷⁰, une application pour le partage de bases de données des sciences du passé par une entrée spatialisée (Dabas, Riahi 2019) (*cf. infra* 4.3. Humanités Numériques). Si les données partagées sur ces deux outils sont majoritairement libres, ils restreignent toutefois l'accès à leurs portails selon des conditions particulières (création de comptes, accès réservé en fonction d'un statut ou d'une institution d'appartenance). Ce genre de pratiques est néanmoins relativement courant dans la discipline, car le patrimoine archéologique est malheureusement fréquemment soumis au pillage.

63 <https://www.qgis.org/fr/site/> ; consulté le 21/10/2023

64 <https://www.libreoffice.org/> ; consulté le 21/10/2023

65 <https://www.openoffice.org/fr/> ; consulté le 21/10/2023

66 <https://www.zotero.org/> ; consulté le 21/10/2023

67 <https://inkscape.org> ; consulté le 21/10/2023

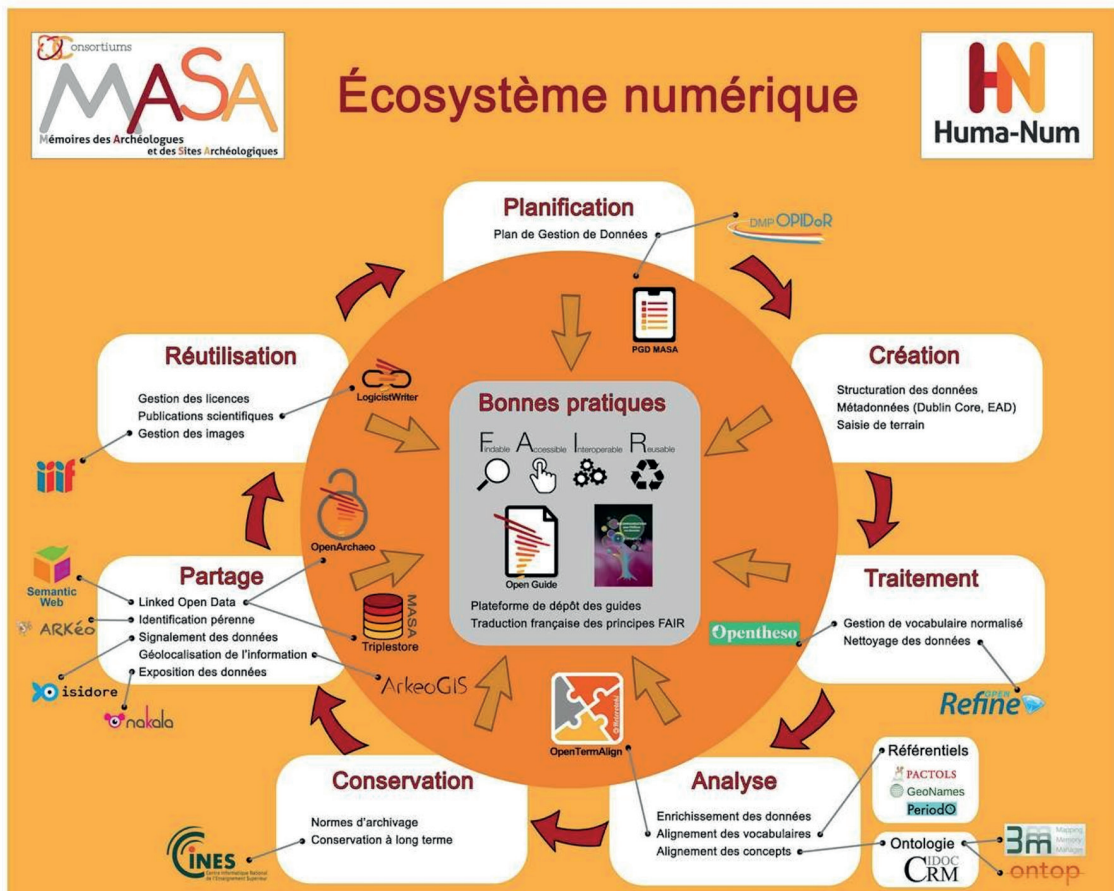
68 <https://nakala.fr/> ; consulté le 21/10/2023

69 <https://www.chronocarto.eu/> ; consulté le 21/10/2023

70 <https://arkeogis.org/> ; consulté le 21/10/2023

L'influence de l'IR* Huma-Num

Pour l'organisation générale du mouvement de la science ouverte en archéologie, l'IR* Huma-Num a délégué la coordination à des consortiums thématiques. Pour notre discipline, il s'agit du consortium Mémoires des archéologues et des sites archéologiques (MASA), lancé en 2012 (Marlet *et al.* 2022, p. 5-6). Le collectif composé de 6 Maisons des Sciences de l'Homme, le réseau Frantiq, le Musée d'Archéologie National, l'Inrap et le Réseau des Ecoles Françaises à l'Étranger. Il s'est donné pour mission de gérer la vaste masse de données numériques produites par les archéologues. Les participants ont ainsi créé un "écosystème numérique fondé sur le cycle de vie des données et centré sur la mise en œuvre des principes FAIR, s'inscrivant ainsi résolument dans la dynamique de la science ouverte" ⁴⁰ : l'écosystème numérique MASA pour les données archéologiques (extrait de Marlet *et al.* 2022) (fig. 40). Le consortium MASA, qui a été récemment renouvelé sous le sigle MASA+, est aujourd'hui reconnu internationalement pour son action et ses outils, et se place parmi les spécialistes mondiaux des bonnes pratiques numériques en archéologie.



40. L'écosystème numérique MASA pour les données archéologiques (extrait de Marlet *et al.* 2022)

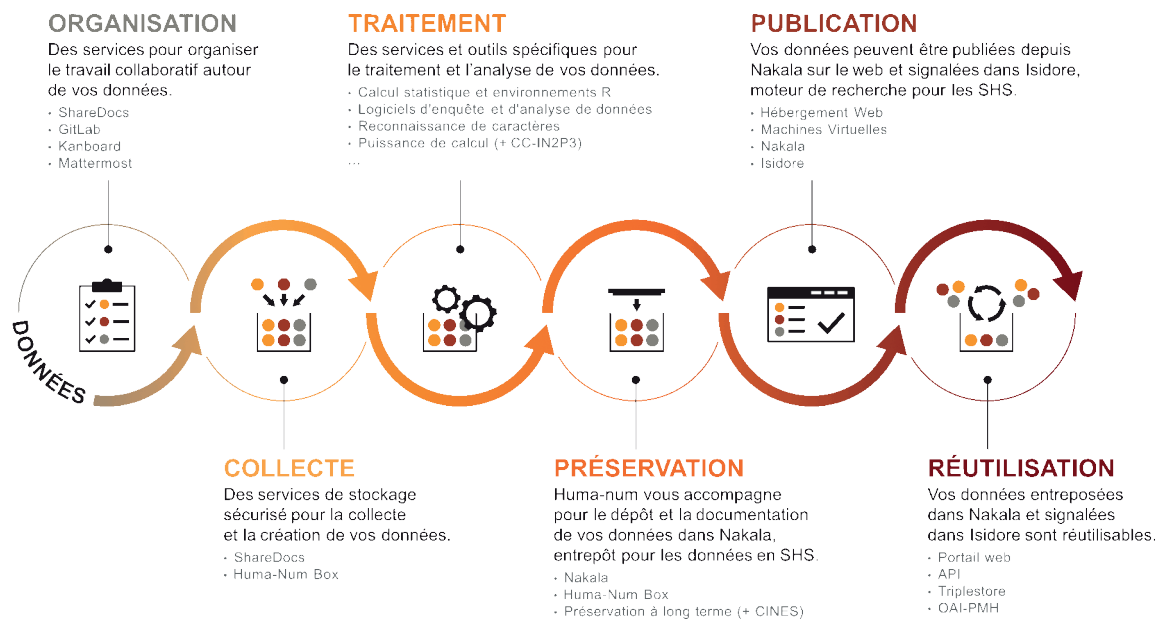
MASA n'est pas le seul groupement de ce genre, il existe aussi, depuis 2014, le consortium 3D SHS. Si, comme son nom l'indique, il s'intéresse à l'ensemble des données 3D produites en SHS, l'équipe à la base du projet est celle d'Archéovision (UMR 6034 Archéosciences Bordeaux), qui réunit des spécialistes des technologies 3D appliquées au patrimoine et à l'archéologie, et la majorité des acteurs qui l'anime proviennent d'institutions d'archéologie (Granier *et al.* 2019, p. 31). 3D SHS s'est attelé à toutes les difficultés liées à la production, l'exploitation, la gestion, la diffusion et l'archivage des données 3D en SHS, et principalement des vestiges archéologiques. Ils souhaitent offrir à chacun la possibilité de maîtriser ce cycle de vie, par la définition de normes et de bonnes pratiques, et la création d'outils adaptés. Notamment le Conservatoire National des Données 3D⁷¹ (CND3D), un service qui propose l'entrepôt et la visualisation pour le partage de modèles 3D issus de projets SHS. Il vient aujourd'hui compléter les nombreuses solutions libres qui existent pour partager et exploiter de la donnée archéologique numérique en France, de la simple publication à des objets plus complexes comme les données spatialisées ou les modèles 3D.

Ces deux consortiums, comme plusieurs des outils mentionnés ci-dessus, sont supportés par une structure centrale pour la science ouverte en SHS : l'Infrastructure de Recherche Étoile Huma-Num (IR* Huma-Num). L'IR* construit depuis 2013, en appuis avec le réseau des MSH, un environnement numérique complet pour les SHS, en pourvoyant à tous les aspects techniques envisageables, qu'il s'agisse d'outils ou d'espaces de stockages⁷². Les nombreux services qui en résultent reprennent quasiment tous le cycle de vie des données, étape par étape, de la gestion à la publication et la réutilisation de celles-ci (fig. 41). Le mouvement du savoir libre a grandement influencé cette structure, car parmi ses principes fondateurs on retrouve l'ouverture des données, la mise à disposition de services maîtrisés par les scientifiques et la pérennité des données sur le long terme. Ainsi, l'action de l'IR* Huma-Num est déterminante pour la communauté archéologique qui, en s'appropriant les outils, évolue progressivement vers la science ouverte. Cette appropriation est réelle, car outre le support des consortiums, par exemple, sur les 325 sites hébergés par Huma-Num, 55 revendiquent une thématique archéologique, soit presque 17 % du total⁷³.

71 <https://3d.humanities.science/> ; consulté le 21/10/2023

72 <https://www.huma-num.fr/quest-ce-que-l-ir-huma-num/> ; consulté le 21/10/2023

73 <https://www.huma-num.fr/annuaire-des-sites-web/> ; données relevées le 23 juin 2022



41. L'écosystème numérique de l'IR* Huma-Num (image d'après huma-num.fr)

Le cas de Bibracte

Comme nous l'avons déjà évoqué, Bibracte s'intègre aussi dans cette mouvance, notamment avec la refonte de son système de diffusion des résultats scientifiques en 2012 (*cf. supra* 1.2.1. Prog. de recherche européen) (fig. 23). La démarche consistait à clarifier et structurer les informations archéologiques pour accélérer le partage en libre accès sur le portail dédié de la plateforme HAL (Desachy *et al.* 2012) (fig. 42). Le projet Bibracte Numérique intervient dans la continuité direct de cette expérience de diffusion numérique. C'est donc naturellement que le projet s'est donné pour objectif de respecter, à chaque étape des réflexions, les principes de la science ouverte.

42. Le portail institutionnel de Bibracte sur HAL

Le bilan encourageant de la science ouverte en archéologie

En définitif, le mouvement de la science ouverte s'est d'abord manifesté en archéologie par l'ouverture des publications, avant de progressivement mettre en œuvre des solutions pour le partage des données. Actuellement, l'utilisation de logiciels libres reste avant tout opportuniste, principalement par manque de moyens financiers, mais nous pouvons espérer que des initiatives comme celles des consortiums MASA et 3D-SHS permettent à des outils spécialisés et largement partagés de voir le jour.

Après avoir clarifié les notions qui gravitent autour de ce mouvement, il nous semble important d'introduire un document en lien avec les questions d'ouverture des données scientifiques : le Plan de Gestion de Données. Il s'agit d'un type de document conçu pour renseigner le cycle de vie complet d'un jeu de données, dans l'optique de faciliter sa réutilisation. Il apparaît donc comme un dispositif déterminant pour la mise en place d'une démarche numérique ouverte et intégrée telle que celle ambitionnée par la communauté archéologique au sens large et au projet Bibracte Numérique en particulier.

2.3. Plan de Gestion de Données

2.3.1. État de l'art

Dans un article très complet qui retrace l'histoire des Plans de Gestion de Données (PGD), des chercheurs australiens démontrent que l'outil n'est pas nouveau (Smale *et al.* 2018). Il remonte même au début de l'informatique, dans les années 1960. Sans grande surprise, les premiers utilisateurs de PGD sont les sciences de l'aéronautique et de l'ingénierie, qui produisent très tôt des données numériques complexes en grande quantité. Gérer et exploiter ces données demande de l'organisation, particulièrement à une époque où les résultats de calculs informatiques sont enregistrés sur des cartes perforées. Les premiers PGD sont donc produits par des chercheurs qui maîtrisent parfaitement leur objet d'étude. Ils mettent alors en place des routines pour optimiser l'acquisition, le traitement et le stockage de données à court terme, afin de faciliter l'exploitation des résultats. Ainsi, les auteurs définissent les années 1960-2000 comme une période de "développement organique" du PGD : il est informel, personnel et n'a pas vocation à être partagé.

Cette période prend fin dans les années 1990 avec le développement d'internet et la prise de conscience générale du potentiel scientifique et économique des données de la recherche. Le cas du *National Environment Research Council* (NERC), un organisme de financement du Royaume-Uni, illustre bien le phénomène. La première version de leur guide pour le déploiement de politiques de gestion de données, publiée en 1996, résume la situation ainsi :

"[...] la science en général et la science environnementale en particulier impliquent la collecte de données [...]. Si les données sont effectivement manipulées par le chercheur pour fournir du matériel à publier, elles constituent une ressource à part entière. Correctement gérées et préservées, elles peuvent potentiellement être utilisées et réutilisées par de futurs chercheurs, et exploitées à des fins commerciales ou éducatives" (NERC 1996, p. 1-2).

Ainsi, le NERC propose l'un des premiers modèles de PGD "générique", soit une série de 9 questions pour qualifier la valeur scientifique et économique des données, et anticiper les besoins logistiques pour leur exploitation sur le long terme (NERC 1996, p. 10-11). L'objectif du NERC, avec son modèle, n'est plus celui des chercheurs des années 60-80. Les "*universitaires environnementalistes*" ne sont d'ailleurs pas la cible de cette première version, ils ne seront rajoutés à la liste des destinataires que deux ans plus tard (NERC 1998, p. 1).

Le but initial est plutôt de documenter un corpus de données, pour en faciliter la diffusion et la réutilisation, afin de gagner en visibilité et de s'assurer le contrôle des retombées commerciales (brevets, produits dérivés, etc.). Le même phénomène est visible aux États-Unis. Par exemple, la déclaration sur le partage des données de la recherche du National Institut of Health, au début des années 2000⁷⁴, impose que tout porteur de projet qui demande des subventions supérieures à 500 000 \$ par an produise un "*plan pour le partage des données*".

Très rapidement, les instances gouvernementales et internationales réalisent aussi les enjeux de la gestion des données de la recherche, en même temps que le manque d'infrastructures et de moyens pour appréhender ces questions (Lord, Macdonald 2003, p. 71-72). Un symbole fort de cette étape est la déclaration de l'OCDE pour l'ouverture des données (OECD 2004), qui marque le début des grandes politiques nationales pour la science ouverte. La réponse des gouvernements est de deux types.

D'abord sur la partie théorique, plusieurs grands audits sont lancés au Royaume-Uni et aux États-Unis, entre 2005 et 2009, afin de dresser un état des lieux sur la gestion et le partage des données de la recherche. Tous se donnent pour objectif d'harmoniser les pratiques des institutions au niveau national. Parmi les recommandations on retrouve bien évidemment le PGD. Sans donner de modèle formalisé, les rapports insistent sur le fait que les organisations de financement doivent dorénavant exiger systématiquement des PGD pour les projets candidats aux subventions (NSB 2005, p. 47 ; DAC *et al.* 2005, p. 12 ; Lyon 2007, p. 6 ; Jones 2009, p. 30 ; IWGDD 2009, p. 15).

Puis sur le manque d'infrastructures, plusieurs universités créent des services spécialisés dans la gestion des données de la recherche, à partir de services d'archives déjà en place. C'est le cas notamment, dès 2004, du Data Curation Center (DCC) en Écosse⁷⁵ et de l'UC Libraries Digital Preservation Repository Program (CDL 2004) – qui deviendra le University of California Curation Center (UC3) – ou plus tard, en 2012, du UK Data Archive qui fonde le UK Data Service⁷⁶.

À la fin des années 2000, les plus grands organismes de financement suivent ces recommandations et réclament des PGD dans les dossiers de candidature (Smale *et al.* 2018,

74 <https://grants.nih.gov/grants/guide/notice-files/NOT-OD-02-035.html> ; <https://grants.nih.gov/grants/guide/notice-files/NOT-OD-03-032.html> ; consulté le 21/10/2023

75 <https://www.dcc.ac.uk/about/history-dcc> ; consulté le 21/10/2023

76 <https://www.data-archive.ac.uk/about/> ; consulté le 21/10/2023

p. 6-8). Par exemple, une équipe a recensé qu'aux États-Unis, en 2012, parmi les 26 principales agences de financements de la recherche, 17 requerraient maintenant un PGD (Dietrich *et al.* 2012). Pourtant, si le document devient un prérequis pour obtenir des financements, les chercheurs non-initiés manquent encore d'informations sur la forme et le contenu à donner à ce dernier (Jones 2009, p. 27). Les nouvelles infrastructures de gestion vont répondre à ce manque en produisant des guides pour préciser la nature des PGD et, surtout, en expliquer l'intérêt aux chercheurs (Jones 2011, p. 2 ; Van den Eynden *et al.* 2011, p. 3). On y retrouve des valeurs comme la gestion et le partage simplifié des résultats de la recherche, la collaboration et l'émulation du débat scientifique.

Ils reprennent et standardisent aussi le concept du modèle, une suite de questions censée accompagner un auteur dans la description de son jeu de données. Le plus complet est la *checklist* du DCC, qui développe les mêmes notions que le modèle du NERC quinze ans plus tôt, mais en complétant largement les questions pour décrire l'ensemble du cycle de vie de la donnée (Donnelly, Jones 2009).

2.3.2. Création des plateformes de Plan de Gestion de Données

Parallèlement, en 2010, le Data Curation Center (DCC) lance la première plateforme d'assistance de création et de partage de PGD, nommée DMPonline⁷⁷ (Getler *et al.* 2014, p. 194). Pour la création, le site web héberge des modèles de PGD interactifs. Ainsi, il est possible de choisir son modèle, qui se présente comme un formulaire avec des onglets en fonction des thèmes abordés par le document. Des exemples et/ou recommandations sont donnés aux auteurs pour répondre à chaque question. Une fois le formulaire complété, DMPonline propose de télécharger le PGD finalisé, mais aussi de l'héberger et de le rendre accessible en tant que PGD public. Par conséquent, en centralisant et en facilitant la diffusion des PGD, le DCC souhaite offrir aux chercheurs un moyen de partager les modes d'emploi pour la réutilisation de données partagées.

La recette va fonctionner, car seulement deux années plus tard l'équipe du DCC assiste l'UC3 pour lancer DMPTool⁷⁸, un outil similaire à DMPonline pour les États-Unis (CDL s. d.). Ils vont ensuite s'associer et construire DMPRoadmap⁷⁹, une interface de programmation (API) ouverte et collaborative pour inciter d'autres institutions à reproduire l'outil. C'est

77 <https://dmponline.dcc.ac.uk/> ; consulté le 21/10/2023

78 <https://dmptool.org/> ; consulté le 21/10/2023

79 <https://github.com/DMPRoadmap/roadmap> ; consulté le 21/10/2023

ainsi que vont naître, toujours sur le même modèle, des outils comme DMP OPIDoR⁸⁰ en France et DMP Assistant⁸¹ au Canada. Attention toutefois, ces plateformes sont avant tout des outils de création. Par exemple, DMP OPIDoR ne se charge pas de l'archivage. Les PGD déposés sont conservés jusqu'à 5 ans après la dernière connexion d'un utilisateur⁸². Il est donc conseillé de les télécharger et de les héberger sur d'autres portails pour en assurer une diffusion à plus long terme.

2.3.3. Diffusion du Plan de Gestion de Données en Europe

La pratique du PGD s'étend au reste de l'Europe entre 2014 et 2016, là encore sous l'impulsion d'un financement, le programme Horizon 2020 (H2020) de la Commission Européenne. En 2014, au lancement du programme, la Commission crée l'*Open Research Data Pilot* (ORD pilot), qui se fixe les mêmes objectifs que les agences du Royaume-Uni et des USA :

"Le projet ORD pilot vise à améliorer et à maximiser l'accès et la réutilisation des données de recherche générées par les projets Horizon 2020 et tient compte de la nécessité de trouver un équilibre entre l'ouverture et la protection des informations scientifiques, la commercialisation et les droits de propriété intellectuelle, les préoccupations en matière de respect de la vie privée, la sécurité ainsi que les questions relatives à la gestion et à la conservation des données" (European Commission 2016, p. 3).

Si l'*ORD pilot* était à la base réservé à certains secteurs, il s'étend à toutes les subventions du programme H2020 à partir de 2017. Dès lors, toutes les demandes doivent être accompagnées d'un PGD, que la Commission définit comme suit (European Commission 2016, p. 4) :

80 <https://dmp.opidor.fr/> ; consulté le 21/10/2023

81 <https://dmp-pgd.ca/> ; consulté le 21/10/2023

82 <https://opidor.fr/category/dmp-faq/> ; consulté le 21/10/2023. Voir la question " combien de temps mes plans sont-ils conservés dans DMP OPIDoR ?"

"Les plans de gestion des données (PGD) sont un élément clé d'une bonne gestion des données. Un DMP décrit le cycle de vie de la gestion des données pour les données à collecter, traiter et/ou générer par un projet Horizon 2020. Dans le cadre de l'objectif de rendre les données de recherche trouvables, accessibles, interopérables et réutilisables (FAIR), un DMP doit inclure des informations sur :

- *le traitement des données de recherche pendant et après la fin du projet ;*
- *les données qui seront collectées, traitées et/ou générées ;*
- *la méthodologie et les normes qui seront appliquées ;*
- *si les données seront partagées/en libre accès ;*
- *comment les données seront-elles conservées et préservées (y compris après la fin du projet)"*

La nouveauté repose surtout sur l'intégration des principes FAIR, publiés en 2016 (Wilkinson *et al.* 2016), directement dans la définition et dans le modèle de PGD proposé en annexe du document.

Enfin, le guide de l'H2020 met en avant une notion importante du PGD : la temporalité (European Commission 2016, p. 5). D'abord, le PGD doit être soumis au maximum dans les 6 mois qui suivent le début du projet. Ensuite, le PGD doit suivre l'évolution du projet et être mis à jour à chaque changement significatif (nouvelles données, changement dans les règles de gouvernance ou la composition de l'équipe). Enfin, le PGD doit être actualisé périodiquement, selon un rythme préalablement déterminé (biannuel, annuel, etc.).

L'extension de l'*ORD pilot* en 2017 comprend aussi, par défaut, le programme de l'*European Research Council* (ERC). L'organisme publie donc l'année même un modèle relativement succinct, avec seulement 6 questions qui s'inspirent directement des principes FAIR ("*rendre les données trouvables*", "*rendre les données ouvertement accessibles*", etc.) (ERC 2017). Les concepts évoqués dans le modèle sont développés l'année suivante dans un guide sur la gestion des données ouvertes dans le cadre d'un projet ERC (ERC Scientific Council 2018).

2.3.4. Diffusion du Plan de Gestion de Données en France

En France, début 2018, une équipe de chercheurs d'universités parisiennes produit, elle aussi, un guide très complet pour répondre aux exigences de la Commission Européenne sur les PGD (Reymonet *et al.* 2018). Le document comprend une première partie qui explique en détail la nature du PGD et son intégration dans le mouvement de l'*Open* et les principes FAIR. Les auteurs y fournissent notamment une définition plus synthétique du PGD (Reymonet *et al.* 2018, p. 4) :

"Un Data Management Plan (DMP) ou plan de gestion de données est un document formalisé explicitant la manière dont seront obtenues, documentées, analysées, disséminées et archivées les données produites au cours et à l'issue d'un processus ou d'un projet de recherche".

C'est la définition que nous retiendrons dans le cadre de notre travail, car selon nous, elle met en évidence l'aspect normé du document, tout en détaillant de manière concise la notion de cycle de vie de données de la recherche. La phrase qui suit est également intéressante :

"Ce guide est conçu pour répondre aux exigences des financeurs, mais il est aussi un outil pour gérer les données tout au long du projet en intégrant la notion de cycle de vie".

Ici, les auteurs sont conscients que le PGD est une initiative qui émane des financeurs, mais ils soutiennent l'hypothèse que la communauté scientifique peut se l'approprier pour en faire un outil de gestion efficace, afin d'améliorer la qualité de la recherche.

La seconde partie est beaucoup plus pratique, car elle est construite comme un "*guide de saisie*". Il s'agit d'un modèle complet qui accompagne les chercheurs dans la rédaction d'un PGD pour les projets supportés par le programme H2020. Il est découpé en trois phases : contextualisation, gestion en cours de projet, archivage et diffusion à l'issue du projet. Pour chaque phase, les auteurs listent les points essentiels à prendre en compte, les explicitent et livrent des recommandations pour les remplir.

Également en 2018, le MESRI lance le Plan National pour la Science Ouverte. Comme pour les audits de la deuxième moitié des années 2000, ce plan invite à généraliser l'utilisation du PGD dans les appels à projets (MESRI 2018, p. 4-5). L'Agence Nationale de la Recherche (ANR) suit cette recommandation dès l'année suivante en imposant le rendu d'un PGD dans les mêmes conditions que le programme H2020 (ANR 2019b, p. 12). Elle dépose

en même temps un modèle de PGD sur DMP OPIDoR, très centré sur les moyens techniques et logistiques pour gérer le cycle de vie de la donnée, accompagné de recommandations forcément larges, mais explicites (ANR 2019a).

Il est intéressant de noter que dans les quatre modèles européens cités (H2020, ERC, universités parisiennes et ANR), aucun ne possède de questions sur la potentielle commercialisation des données de la recherche. Les aspects économiques peuvent être abordés par les auteurs dans les conditions de réutilisation et les droits juridiques qui protègent la donnée, mais contrairement aux modèles anglo-saxons, ils n'apparaissent jamais de manière explicite.

2.3.5. Plan de Gestion de Données en archéologie

La plupart des modèles de PGD que nous avons évoqué ont un point commun : ils sont commandités par des financeurs (agences de financement et services d'État) qui soutiennent de nombreux domaines scientifiques. Ces modèles se veulent donc par définition les plus génériques possibles. Les questions, comme les recommandations, sont le plus souvent très larges, pour englober le plus de types de données et de domaines de connaissances possibles.

C'est en tout cas le constat que nous avons fait en publiant le premier PGD en archéologie sur DMP OPIDoR en 2020, du PCR "*La confluence Saône-Doubs à l'âge du Fer (VIe s. av. J.-C. au I^{er} s. ap.)*" (Verriez 2020a). Initialement, une première ébauche de ce plan a été rédigée en utilisant le modèle de l'ANR, car le nombre élevé de questions semblait indiquer un modèle bien détaillé. Cependant, il est rapidement apparu que bon nombre de ces questions manquaient de précision et n'étaient pas adaptées à la discipline archéologique. Comme il n'existaient pas à ce moment là de modèle spécifique à l'archéologie, nous avons finalement préféré le modèle du programme H2020. Ce dernier présentait des questions moins nombreuses et plus ouvertes, pour un résultat plus souple d'utilisation.

Le consortium MASA a dressé le même constat bien avant nous. C'est pourquoi ses membres ont œuvré dès 2018 à la création d'un modèle de PGD pour l'archéologie. Malheureusement, le modèle MASA⁸³ a été rendu public quelques jours après le dépôt du PGD du PCR Confluence. Le modèle est fondé sur une version créée par l'Inrap, combinée au modèle de l'ANR, qui propose des questions et recommandations adaptées à la discipline archéologique (Marlet *et al.* 2022, p. 10-11). Ce modèle a également été testé par des archéologues hors du consortium, afin d'en améliorer le contenu et de la faire valider auprès d'une

83 https://dmp.opidor.fr/template_export/942300517.pdf, consulté le 22/10/2023

plus grande communauté. Toutefois, la démarche reste encore méconnue ou confidentielle, car, à notre connaissance, seuls deux PGD utilisant ce modèle sont diffusés publiquement sur DMP OPIDoR⁸⁴. Enfin, des initiatives internationales autour des PGD en archéologie existent aussi, à l'image de la très récente boîte à outils PGD du programme européen ARIADNEplus, qui présente trois nouveaux modèles⁸⁵.

En résumé, les experts au sein de la communauté archéologique œuvrent pour démocratiser la pratique du PGD, mais cette dernière reste pour le moment encore très peu répandue dans notre discipline. Il semble pourtant qu'une bonne gestion des données de la recherche peut avoir de nombreuses retombées bénéfiques pour les chercheurs comme pour les institutions qui les accueillent. Particulièrement à une époque où la quantité de données générées par les dispositifs numériques dépasse les capacités de traitement d'un être humain.

En effet, l'avantage principal de la science ouverte, soit le partage en masse de connaissances, représente également sa plus grande faiblesse. Toutes les données ne sont pas appropriées pour la diffusion, et celles qui le sont doivent à tout prix respecter un ensemble de règles. Sans ces règles, le risque d'une perte de l'information, d'une donnée redondante ou de qualité insuffisante serait trop grande, alors même que les quantités de données produites sont de plus en plus volumineuses. Cette observation s'applique particulièrement à la discipline archéologique, qui se distingue par l'accumulation de corpus variés et abondants. Le risque est aussi renforcé par la nature intangible des données numériques, car ce type de documentation n'est pas "incarnée" dans une fiche papier, une minute de relevé ou un objet archéologique. Le PGD permet donc d'anticiper les règles à adopter en amont d'un projet de recherche. Il force les chercheurs à s'interroger sur l'utilité de produire, de conserver et de diffuser des données numériques. Il les accompagne également dans la sélection des standards (formats de fichiers, protocoles de partage et d'archivage) à adopter en fonction des problématiques propres à leur domaine d'expertise.

Le manque de critique sur la qualité des données diffusées est un autre risque lié à la dynamique de la science ouverte. Si des solutions existent pour les publications traditionnelles

84 https://dmp.opidor.fr/plans/17977/export.pdf?export%5Bquestion_headings%5D=true ; consulté le 21/10/2023. PGD du projet "L'apport du numérique dans l'étude des sites et des objets archéologique au Haut-Empire romain" par A. Turgis. https://dmp.opidor.fr/plans/17763/export.pdf?export%5Bquestion_headings%5D=true ; consulté le 21/10/2023. PGD du projet "Chronostratigraphie pléistocène des terrasses alluviales entre Pyrénées et Massif Central. Implications paléoenvironnementales et archéologiques pour les sites paléolithiques." par J. Vincent

85 <https://vast-lab.org/dmp/index.html> ; consulté le 21/10/2023

en libre accès, la question est tout aussi légitime pour les jeux de données publiés librement (cf. *supra* 2.2.1. *Open access*). Un PGD rédigé agit comme un véritable mode d'emploi pour un jeu de données. En décrivant les données et leur finalité, il permet également d'en évaluer la qualité. Par conséquent, le PGD représente une étape indispensable pour la validation des données par la communauté scientifique. Il apporte le formalisme nécessaire à la mise en place de *data papers*, des publications académiques à comité de lecture centrées sur la présentation des données plutôt que sur les résultats et les interprétations.

Enfin, le PGD est souvent vécu comme une contrainte imposée par les instances de financement, plutôt qu'un outil pour assurer la qualité et la pérennité des résultats d'une recherche. Le résultat dressé par l'étude australienne précédemment citée est sans appel : sur les plus de 800 PGD consultés, les 3/4 n'étaient pas complètement rédigés, les descriptions n'étaient pas claires et les données n'étaient pas accessibles (Smale *et al.* 2018, p. 16-20). Pour les auteurs, ce constat s'explique par le manque de clarté derrière la démarche des PGD. Ils mettent en avant que si les organismes de financement encouragent le partage de données, ils ne livrent pas les informations suffisantes pour le mettre en œuvre. La tâche est d'ailleurs probablement impossible, car ils visent un public trop large. Aussi, les rédacteurs ne perçoivent pas les bénéfices qu'ils peuvent tirer du document, tant les questions semblent vagues et les éléments pour y répondre peu nombreux. Pour être fonctionnel et pertinent, chaque PGD devrait au contraire différer en fonction de la discipline, de l'objet d'étude, mais aussi de l'objectif du document lui-même. En effet, les attentes ne sont pas les mêmes si le PGD doit aider à la gestion d'un projet de recherche, sa documentation en vue d'une réexploitation des résultats ou sa valorisation auprès de différents publics. Selon nous, en France, cette incompréhension peut-être aussi une conséquence de l'évolution actuelle du modèle de financement de la recherche européenne, qui s'inspire de plus en plus du modèle anglo-saxon. Malgré toutes les nobles intentions qu'on lui attribue, le PGD symbolise aussi la transformation en cours de nos institutions, qui s'orientent de plus en plus vers la recherche thématique et le modèle de financement sur appel à projet.

Pour conclure avec cette partie, nous avons souhaité démontrer ici l'intérêt de la science ouverte pour l'archéologie. Certes, la diffusion à large échelle de ce mouvement reste relativement récente, et des risques existent face à l'ouverture massive de la connaissance, en archéologie comme dans les autres domaines. Toutefois, le soutien offert par l'État et l'Europe à la science ouverte a permis d'apporter plusieurs réponses proportionnées. Elles se sont d'abord concrétisées sur les aspects techniques, avec la mise en place d'infrastructures

adaptées à la gestion d'importants flux d'informations. Puis elles ont permis de faire émerger des communautés de spécialistes capables s'emparer de problématiques autour du cycle de vie de la donnée.

En archéologie, de nombreuses questions restent à préciser sur les pratiques numériques liées à la dynamique de la science ouverte, notamment sur la définition de standards communs pour l'échange des données. Cependant, il est déjà possible de mener une recherche en s'appuyant sur des outils libres sans faire de concession sur la qualité du travail scientifique. C'est en tout cas ce que nous avons souhaité illustrer dans cette thèse. Ainsi, à part quelques outils clairement mentionnés dans ce texte, tous les logiciels utilisés de notre recherche doctorale sont libres. On parle ici du processus complet, depuis les premières expérimentations jusqu'à la rédaction de la thèse. Les rares fois où nous avons été contraint d'utiliser des outils propriétaires sont dû soit à un manque de temps de formation pour passer du propriétaire au libre, soit à l'absence d'alternatives libres. C'est par exemple le cas pour la mise en forme du manuscrit. Nous avons abandonné le logiciel de publication assistée par ordinateur (PAO) libre Scribus au profit de Adobe InDesign, car il n'est pas encore au point pour la création de publications académiques. Aussi, les publications et les données qui découlent directement de ce travail sont et seront diffusés au maximum en formats ouverts, autant que possible gratuitement, tout en recherchant à respecter une procédure de validation par les pairs.

En conséquence, la partie suivante présente un protocole d'enregistrement archéologique numérique pour les chantiers de Bibracte qui repose largement sur l'utilisation de logiciels libres. Notre volonté était d'établir des processus clairs, transparents et reproductibles, depuis l'acquisition jusqu'au partage des données, tout en proposant à chaque étape une évaluation critique des outils mis en œuvre.

TROISIÈME PARTIE

Un chantier archéologique numérique à Bibracte

3.1. Drones : nouveaux vecteurs, nouvelles échelles

3.1.1. Objectifs

Les drones sont des aéronefs pilotés à distance dont l'usage et la technologie se sont largement développés depuis les années 2010, particulièrement ceux de la catégorie des "multirotors". Ce type d'appareil est en effet capable d'embarquer des capteurs pour des vols libres et stationnaires de plusieurs dizaines de minutes. Il présentait ainsi plusieurs intérêts pour les fouilles de Bibracte, et pour les opérations archéologiques en général. En effet, un drone permet d'obtenir des clichés verticaux des sites en cours de fouille, au lieu des traditionnelles photographies obliques prises, le plus souvent, depuis une nacelle. Outre le fait d'améliorer grandement la qualité des vues d'ensembles des sites archéologiques, cette solution permet, une fois l'équipement acquis, de réduire les coûts induits par les locations récurrentes d'équipement (assez importants dans le cas des nacelles). Ensuite, un drone peut évoluer dans l'espace aérien au-dessus des chantiers, à des altitudes interdites aux avions, ouvrant les potentialités de prises de vue verticales et multiscalaires sur les sites archéologiques. Enfin, la facilité et la rapidité de mise en œuvre de ces missions, associées à la précision de navigation offerte par les drones, permettent d'envisager la réalisation de couvertures systématiques régulières des sites archéologiques fouillés. Les drones permettent ainsi de systématiser le relevé des données de terrain sous un format inédit, avec des rendues variés hautement fidèle à la réalité (orthophotographie, Modèle Numérique de Terrain, modèle 3D) et qui résulte d'opérations mathématiques qualifiables. Dans un objectif d'amélioration de la méthode scientifique archéologique, une telle approche permet d'apporter de nouveaux éléments pour justifier les observations de terrain une fois les vestiges détruits par la fouille.

3.1.2. Choix des équipements

Les premiers vols de drone effectués à Bibracte dans le cadre d'une opération archéologique ont eu lieu en 2013 et 2014, pour photographier le chantier de la terrasse PC15. Ils ont été réalisés par B.-N. Chagny, de la société Kap-archéo, après le départ des équipes de fouille. N'ayant accès qu'aux images finales, nous avons peu d'informations sur le déroulement de ces missions. Toutefois, la page web de la société livre une description détaillée du

matériel utilisé⁸⁶. Au début des années 2010, la technologie des drones multirotors n'était pas encore standardisée sur le marché européen. Les drones en circulation étaient assemblés "sur mesure" par des entreprises spécialisées ou les propriétaires eux-mêmes. Ainsi, le drone déployé sur le chantier de PC15 comportait un châssis à 8 hélices Beetec X8, contrôlé par l'électronique de vol DJI Wookong M et alimenté par une batterie au Lithium Polymère (LiPo) de 12Ah. L'ensemble embarquait une caméra de type Reflex numérique, avec un émetteur vidéo permettant de visualiser le retour vidéo, en temps réel, sur un écran porté par le pilote à l'épaule. La capture était déclenchée à distance, mais les réglages de la caméra devaient être définis avant le début de la mission, car ils ne pouvaient pas être modifiés une fois l'appareil en vol. Le drone affichait un poids total à charge de 4,5 kg et une autonomie de vol de 9 minutes. En prenant en compte le temps de navigation et le cadrage, cela offrait un temps de captation active d'environ 5 minutes, soit au maximum une vingtaine de clichés. Finalement, ces interventions ont livré deux orthophotographies de la fin du chantier, non géoréférencées et exploitées essentiellement à des fins de valorisation.

En 2015, la plateforme technologique de la MSHE C.-N. Ledoux (Besançon) s'est dotée d'un drone, grâce à une dotation de la Région Franche-Comté. Elle s'est ainsi inscrite parmi les premières institutions de recherche en France à devenir autonome pour des missions de drone multirotors en archéologie. L'objectif était d'acquérir une machine polyvalente, capable d'emporter plusieurs capteurs, dont la technologie LiDAR, réservée jusqu'alors à l'aviation. Le choix s'est porté sur l'OnyxStar FOX-C8HD, un imposant octocoptère (8 hélices) de 5 kg pouvant accueillir plusieurs types de nacelles et capable d'embarquer jusqu'à 6 kg d'instruments (fig. 43). Cette charge utile permet de supporter des APN haut de gamme avec objectif interchangeable, les plus lourds utilisés sur les chantiers de fouille. L'OnyxStar permet donc de produire des images zénithales larges et très détaillées, de moins de 5 mm par pixel, en volant à 20 m d'altitude. À ce niveau de précision, tous les indices archéologiques au sol, depuis les éléments architecturaux jusqu'aux petits objets, sont visibles sur les photos. Le drone possède également un retour vidéo, assuré par une petite caméra installée sur son châssis. Le retour sur l'écran au sol est dégradé et légèrement décalé dans le temps par rapport à la capture réelle de l'APN, mais la qualité reste largement suffisante pour effectuer la plupart des missions (fig. 44). Le signal du retour vidéo est stable sur une portée de 200 à 400 m, quand les conditions optimales sont rassemblées : un terrain plat et dégagé, sans élévation ni

86 le service d'hébergement d'Orange, sur lequel était accueilli le site en question, a été stoppé le 05 septembre 2023. Malgré nos recherches, il semblerait que l'auteur n'est pas déplacé le site vers une nouvelle solution d'hébergement (anciennement https://kap-chagny.pagesperso-orange.fr/index_fichiers/Page38.htm)



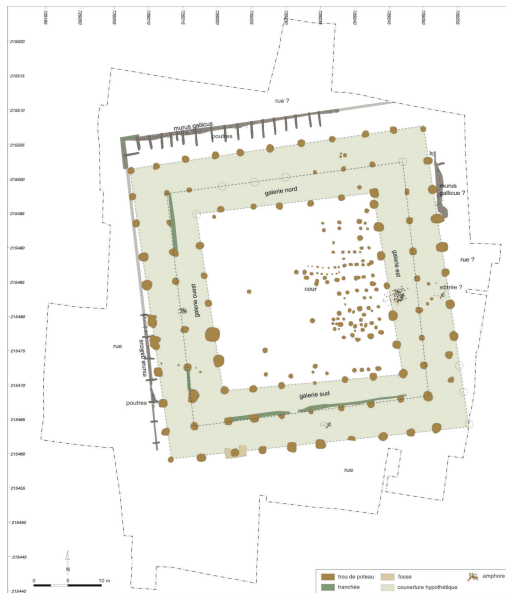
43. Le drone OnyxStar FOX-C8HD de la MSHE Ledoux (cliché A. Maillier)



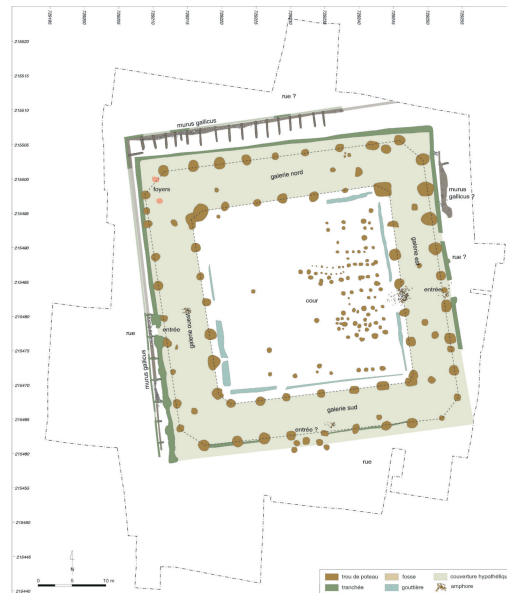
44. La télécommande et le système de retour vidéo de l'OnyxStar FOX-C8HD (cliché A. Maillier)

végétation haute, sans vent ou de faible intensité. L'appareil est aussi équipé d'une puce GPS et d'une carte électronique lui permettant de réaliser des vols autonomes. Toutefois, cette fonction requiert une programmation minutieuse du plan de vol en amont de la mission. Le matériel électronique Mikrokopter embarquée propose en effet une solution logicielle libre très complète, mais qui offre peu d'aide à l'utilisateur pour les réglages des paramètres de navigation. Il est donc nécessaire de calculer de nombreux aspects du vol, comme l'emplacement des points de navigation et de contrôle, l'altitude et la vitesse de vol du drone, ou encore le déclenchement programmé du capteur. L'alimentation est assurée par deux batteries LiPo de 10 000 mA/h, qui fournissent une autonomie de vol de 15 minutes environ, en fonction du poids embarqué et de la force du vent. Les capacités offertes par l'OnyxStar ont donc permis de réaliser les premières couvertures photographiques systématiques en haute définition de la fouille de PC15. En tout, six vols ont été effectués entre 2015 et 2016. Les missions se sont réparties selon le même programme durant les deux années : une acquisition d'images couvrant l'intégralité du chantier juste après le décapage, une seconde en milieu de campagne et la dernière en post-fouille, lorsque l'équipe avait quitté le terrain. Ces vols ont livré des données spatiales et graphiques qui retracent les étapes cruciales de la fouille avec un niveau de détail jusqu'alors inédit à l'échelle du site. Les archives ont ensuite servi de support pour la plupart des analyses chronospaciales de l'édifice monumental de PC15, permettant de produire une documentation graphique synthétique riche et détaillée pour les rapports et publications (fig. 45).

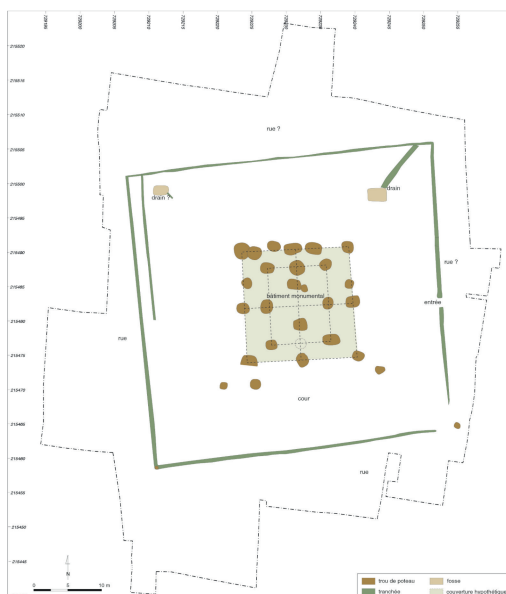
Face aux premiers résultats encourageants, L'EPCC Bibracte a également fait l'acquisition d'un drone en 2017. L'objectif prioritaire des missions de drones était la production d'images et de vidéos, pour les recherches archéologiques et la valorisation du site. Ainsi, le choix s'est porté vers une solution intégrée drone et capteur (la caméra ne peut pas être séparée du drone), avec le DJI Phantom 4 Pro (fig. 46). Il affichait alors parmi les meilleures caractéristiques techniques de sa gamme, pour un prix bien inférieur à la concurrence. Du fait de sa construction en série, seuls les accessoires propriétaires DJI sont compatibles (batterie, radiocommande, hélice, etc.) avec le drone et la gestion est assurée par les applications DJI GO et DJI GS Pro, qui ne fonctionnent que sur la tablette iPad ou le smartphone iPhone d'Apple. Cependant, l'ensemble présente une conception très ergonomique, avec une prise en main vraiment simplifiée, notamment pour le calcul de plans de vol. Le drone en lui-même est très léger (moins de 1,4 kg) et la taille de la caméra associée est, de fait, inférieure aux APN utilisés avec l'OnyxStar. Toutefois, il suffit de réduire l'altitude de vol entre 10 et 15m pour obtenir une qualité d'image clairement suffisante pour une utilisation archéologique. Le



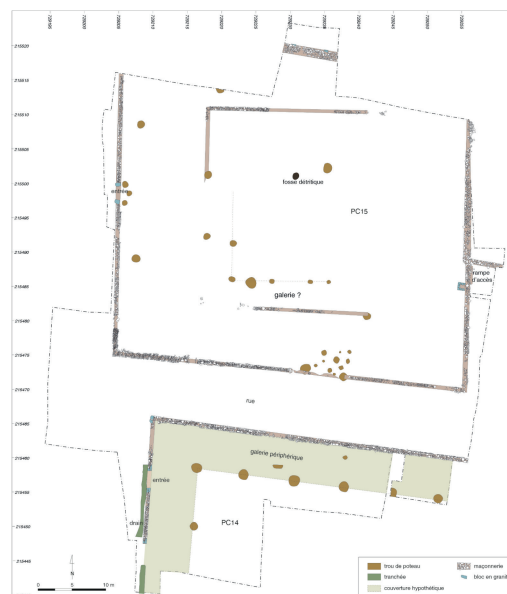
Etat I



Etat II



Etat III



Etat IV-V

45. *Évolution chronologique du bâtiment monumental de la terrasse PC15.
Plans dessinés sur l'orthomosaïque de fin de chantier de 2015 (extrait de Hantrais 2017)*



46. Le drone *Dji Phantom 4 Pro* de la *MSHE Ledoux*

Le poids permet en revanche d'offrir une bonne autonomie de vol, de 20 à 25 minutes selon les conditions de vent. Pour le pilotage, le Phantom 4 Pro est très maniable et accompagné de toute une série de capteurs infrarouges pour effectuer des vols en sécurité. Le retour vidéo, qui provient de la caméra, est intégré aux applications, et un système de fixation est prévu sur la radiocommande pour accueillir la tablette, permettant de centraliser toutes les fonctions de contrôle. Il est également possible d'agir sur les paramètres de la caméra durant le vol, et de les ajuster en fonction des conditions réelles de luminosité, de vitesse, de déplacement ou d'altitude. De plus, la qualité du positionnement GPS est remarquable, permettant à l'appareil de rester stable même en cas de rafales de vent. Enfin, le drone, les batteries, la télécommande et l'iPad tiennent dans un sac à dos. Sa facilité de transport et de mise en œuvre en a fait un instrument d'utilisation quotidienne pour la capture d'images sur les chantiers, contrairement à l'OnyxStar qui intervenait pour des missions spécifiques et planifiées.

3.1.3. Acquis des expérimentations

Lorsque le contexte le permet, les drones multirotors avec capteur embarqué sont des instruments particulièrement adaptés aux travaux archéologiques de terrain. Ils remplacent et démultiplient les possibilités autrefois offertes par les cerfs-volants, les ballons ou les nacelles, en permettant d'évoluer rapidement et avec régularité au-dessus des chantiers. Les drones participent ainsi au renouveau de la documentation archéologique de terrain, et livrent des supports graphiques et des modèles 3D à l'échelle de sites entiers, selon un niveau de précision inédit pour la discipline.

Le niveau de précision atteint par l'équipement embarqué sur l'OnyxStar s'est d'ailleurs révélé disproportionné par rapport aux besoins réels des archéologues. En effet, l'avantage principal des drones en contexte archéologique est la prise de recul rapide par rapport aux vestiges. Les prises de vue sont de deux types : les vues générales, pour mettre en évidence la distribution spatiale des niveaux archéologiques et éventuellement la structuration interne de certains aménagements (pierre à pierre) ; les prises de vues détaillées pour des structures complexes, telles que les dépôts de mobilier, les sépultures ou les mosaïques. Dans la pratique, les images produites avec l'OnyxStar associent une échelle large à un niveau de détail suffisamment précis pour réunir les deux types. Pourtant, il est apparu à l'utilisation que ces images n'étaient pas forcément pertinentes pour l'interprétation des vestiges archéologiques. Les images produites par drone servent essentiellement de support pour la création des relevés archéologiques. Or produire des plans qui associent des éléments interprétés à l'échelle du site avec des points de vue détaillés risque, dans le meilleur des cas, de mettre en parallèle des éléments qui ne sont pas comparables. Dans le pire des cas, l'excès d'informations peut nuire à l'interprétation générale des vestiges.

Par conséquent, l'OnyxStar apparaît surdimensionné pour l'enregistrement photographique systématique d'un chantier de fouille. Il fournit des archives de très grande qualité, mais scientifiquement sous-exploitées et très gourmandes en ressources informatiques, tant en matière d'espace que de puissance de calcul. En effet, le traitement des orthophotographies générales des chantiers archéologiques comme PC15 nécessite l'utilisation d'ordinateurs de calcul spécifiques, qui exploite des composants informatiques largement supérieurs à ceux présents dans les ordinateurs personnels classiques. Même dans ces conditions, les calculs peuvent prendre plusieurs heures et produire des modèles qu'il faut impérativement traiter à nouveau pour les alléger avant une quelconque exploitation. À l'inverse, le Phantom 4 Pro

offre une solution légère, ergonomique et adaptée aux besoins en matière de qualité d'enregistrement. Sa taille et sa simplicité de mise en œuvre permettent de le faire intervenir de manière quotidienne sur des acquisitions à l'échelle du site. Les modèles obtenus permettent de documenter la distribution des niveaux archéologiques, tout en conservant des tailles de fichier raisonnables. Pour les mêmes raisons, il s'avère aussi plus efficace qu'une perche, si les conditions le permettent le vol à faible altitude (moins de 10 m).

Quel que soit l'aéronef déployé, plusieurs aspects sont à prendre en compte pour opérer une mission au drone :

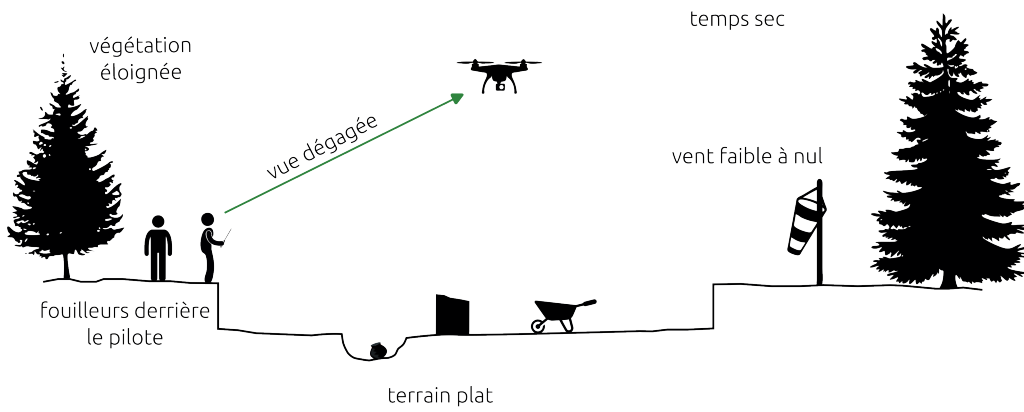
- L'utilisation des drones est soumise à des règles strictes sur le territoire français. L'opérateur doit maîtriser le pilotage d'aéronef sans personne à bord, les règles de navigation dans l'espace aérien et le scénario dans lequel s'effectue une mission, afin d'appliquer les règles de sécurité adéquates. Tous ces points sont soumis au contrôle de la Direction générale de l'aviation civile (DGAC).
- Les avancées sur les batteries LiPo ont été déterminante pour le développement des drones, car elles offrent désormais un ratio poids/puissance/autonomie suffisant pour faire décoller un appareil de plusieurs kilos pendant plusieurs dizaines de minutes. Le lithium est toutefois un élément très réactif. Il est donc important de se référer aux recommandations d'utilisation des constructeurs, afin de prévenir une détérioration irréversible du matériel et d'effectuer des vols en toute sécurité (fig. 47).
- Quel que soit le drone utilisé, son utilisation est à privilégier dans les conditions suivantes (fig. 48) : sur un terrain plat, étendu et dégagé, sans élévation ni végétation haute (ou alors avec une très bonne visibilité sur l'appareil durant tout le temps de vol), sans vent ou de faible intensité. Dans d'autres d'autres conditions, la sécurité, le temps de déroulement de la mission et la qualité des résultats ne permettent généralement pas de justifier l'utilisation de ce type d'équipement. On parle ici par exemple de vols sous couvert forestier, sous abris de chantier ou encore en milieu urbain. Une exception peut être faite pour l'archéologie du bâti, si le bâtiment étudié est isolé. Toutefois ce type de prise de vue est réservé à des pilotes expérimentés, capables de faire voler les drones en sécurité dans des espaces souvent restreints.



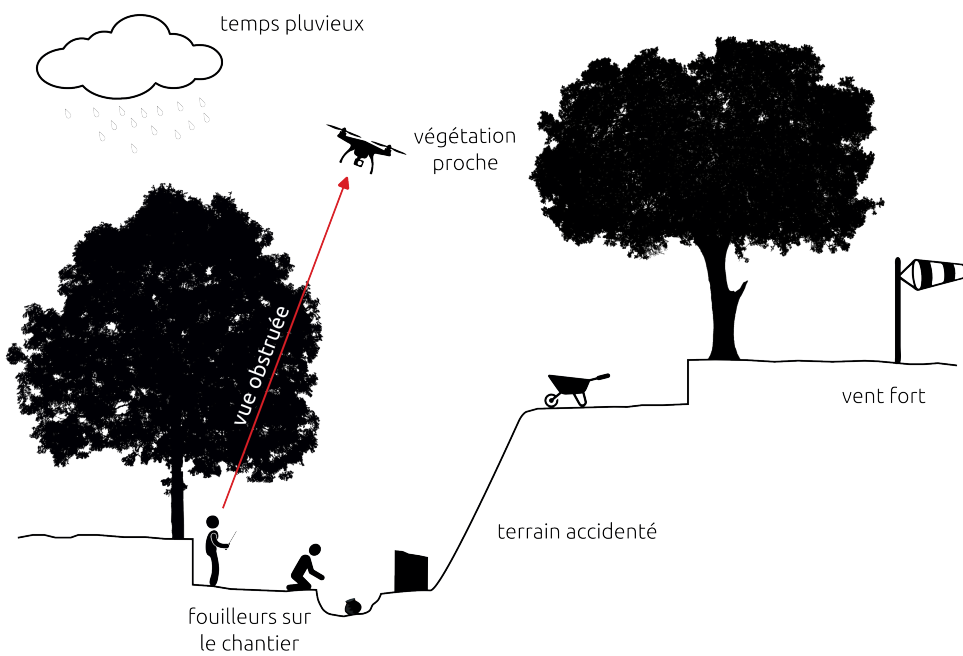
47. Conditionnement des batteries LiPo dans sacs ignifugés

Finalement, l'intégration des drones dans les chantiers archéologiques de Bibracte a considérablement augmenté le nombre d'images capturées et a ouvert de nouvelles possibilités d'utilisation. Cette abondance d'images a incité les archéologues à repenser le rôle de l'image dans leurs analyses et dans la documentation archéologique. En effet, les images aériennes offrent une perspective globale et détaillée du site, permettant de mettre en évidence des structures, des motifs et des éléments parfois complexes à appréhender à l'œil nu. De plus, les technologies de traitement d'images et de modélisation 3D permettent désormais de créer des représentations virtuelles réalistes des vestiges, offrant ainsi de nouvelles opportunités pour la conservation et la transmission des résultats. Cependant, cette profusion d'images soulève également des questions sur la sélection, l'analyse et l'interprétation appropriées de ces données visuelles. C'est pourquoi il nous a semblé essentiel de revenir aux principes fondamentaux de la photographie en archéologie, afin d'intégrer efficacement ces images numériques dans les recherches archéologiques, en tenant compte de leur potentiel, mais aussi de leurs limites en tant que source d'information.

Conditions recommandées



Conditions déconseillées



48. Conditions pour le survol de site archéologiques par drone

3.2. Appareils photo numériques : nouveaux capteurs, nouvelles résolutions

3.2.1. Objectifs

L'archéologie se distingue par le fait que la méthode d'investigation, en l'occurrence la fouille, entraîne inévitablement la destruction de son propre objet d'étude, à savoir le site archéologique. Une part importante du travail de l'archéologue sur le terrain consiste donc à collecter les objets et enregistrer ses observations, afin de pouvoir revenir sur les différents éléments durant la phase d'interprétation, d'analyse et de synthèse. Cette phase intervient généralement *a posteriori*, avec la rédaction du rapport d'opération. La photographie tient donc une place de choix parmi la documentation produite, car elle permet de conserver une trace visuelle des différentes étapes du chantier.

La photographie est "*l'ensemble des techniques permettant d'obtenir des images permanentes grâce à un dispositif optique produisant une image réelle sur une surface photosensible*⁸⁷". Par extension, elle désigne aussi les images produites à partir de ce procédé. Le lien entre archéologie et photographie a été formalisé très tôt, dès son invention à la fin des années 1830. Dans son discours de présentation de la photographie à l'Académie des Sciences le 19 août 1839, F. Arago démontre l'utilité de ce procédé à travers l'exemple de l'étude des hiéroglyphes découverts sur les monuments de l'Égypte antique (Chéné *et al.* 1999, p. 8). Selon lui, la photographie allait permettre la copie de "*millions de hiéroglyphes*" par une seule personne, là où auparavant la tâche aurait nécessité "*des vingtaines d'années et des légions de dessinateurs*". Dès lors, Ph. Collet, photographe de l'École Française d'Athènes (EFA), relate les différentes expéditions qui ont eu lieu dans les années 1850, tout autour de la méditerranée, pour photographier des monuments antiques (Collet 1996, p. 325-326). Si cela semble anecdotique aujourd'hui, il faut bien comprendre qu'à l'époque la photographie n'était ni parfaitement maîtrisée, ni simple à mettre en œuvre : entre les optiques et la chimie nécessaire, les bagages pouvaient dépasser la tonne de matériel à transporter. Comme le souligne Ph. Collet, ce n'était donc pas l'aspect pratique de la photographie qui a "*séduit les premiers archéologues, mais sa précision et sa fidélité scientifique*", contrairement aux autres méthodes de représentation alors en vigueur, le dessin alors jugé trop "*subjectif*" et la peinture "*approximative*".

Plus d'un siècle et demi plus tard, la photographie est devenue une pratique courante en archéologie, toujours vantée pour ses capacités à reproduire de manière factuelle la réalité. Pourtant, la plupart des spécialistes de la photographie d'archéologie s'accordent à dire que l'image n'est pas neutre. Dans sa préface du manuel de M. B. Cookson, M. Wheeler parlait même de la technique photographique comme d'une "*affreuse menteuse*" qui à travers un "*œil inhumain [...] falsifie le plus simple des sujets*" (Cookson 1954). Pour Ph. Collet, "*la photographie normalisée [...] est une illusion*", car chaque image s'intègre dans la culture scientifique et "*l'environnement d'image*" propre à son époque (Collet 1996, p. 337).

L'archéologie n'est pas la seule discipline scientifique à avoir envisager la photographie comme un système d'enregistrement objectif du monde. L. Daston et P. Galison ont démontré, à travers l'évolution des atlas scientifiques, que l'objectivité est une construction intellectuelle qui a émergé en même temps que l'invention de la photographie, particulièrement autour des sciences naturalistes, dans la deuxième moitié du XVIII^{ème} s. (Daston, Galison 2012, p. 26-27). Avant cette période, les moyens techniques à la disposition des scientifiques pour représenter leurs observations étaient largement manuels : le dessin et la peinture. La rigueur scientifique d'une image reposait alors sur la capacité de l'auteur, et de son illustrateur le cas échéant, à représenter les qualités essentielles de son sujet (nombre, forme, proportion, position), sans se laisser perturber par les caractéristiques accidentelles pouvant varier d'une itération à l'autre (comme la couleur) (Daston, Galison 2012, p. 69-78). Seul le scientifique, par son expérience et ses observations assidues, était capable de déceler ces qualités constantes, certaines. Il intervenait ainsi directement dans la production de l'image pour créer des représentations "*typiques, caractéristiques, idéales ou moyennes*", parfois soutenues par des modèles mathématiques, comme le célèbre Homme de Vitruve de L. da Vinci. Les auteurs nomment cette pratique "active" de la production d'images scientifiques la "*vérité d'après nature*".

L'invention de la photographie s'accompagne alors d'un nouveau courant de pensée. Certains savants ont commencé à considérer la part faillible de l'être humain, car il pouvait commettre des erreurs ou appliquer ses propres projections à ses observations. Si les scientifiques n'étaient pas capables de retranscrire sans simplifier, théoriser ou même embellir, les machines devaient prendre le relais pour créer mécaniquement. C'est ce que L. Daston et P. Galison décrivent comme "*l'objectivité mécanique*", soit le "*désir impérieux de réprimer toute intervention volontaire de l'artiste-auteur en mettant en place des méthodes capables d'imprimer la nature sur la page suivant un protocole strict, voir automatique*" (Daston,

Galison 2012, p. 137-148). Ce concept sera d'ailleurs élargi à la photographie, puisqu'un débat aura lieu dans les années 1860, qui opposera la photographie mécanique, scientifique et objective, à la photographie esthétique, artistique et subjective (Daston, Galison 2012, p. 157-163). Le terme mécanique sous-entend deux significations. Il renvoie d'abord à la production de l'image, qui doit être automatique, sans intervention directe de l'auteur sur celle-ci. Mais il renseigne également la capacité de reproduction de la photo, avec la multiplication automatique et à l'identique d'une même image. À l'inverse de la vérité d'après nature, le principe de l'objectivité mécanique était d'effacer la volonté du scientifique dans la création d'images en employant des instruments à la fois observateurs et artistes, ces derniers ne comportant pas de biais interprétatifs ou ethnocentristes.

Pourtant les auteurs remarquent que le débat entre photographie scientifique et esthétique était surtout d'ordre rhétorique. La plupart des utilisateurs de la photographie de la fin du XVIII^e s. étaient conscients de l'importance du savoir-faire et du jugement dans la création d'une photo, en plus de la possibilité de trafiquer une photo (Daston, Galison 2012, p. 157). Pour illustrer ce propos, le même exemple revient à plus de 50 ans d'intervalle chez M. Wheeler et les auteurs de l'ouvrage *La pratique de la photographie en archéologie* (Cookson 1954 ; Chéné *et al.* 1999, p. 16). Pour le même sujet, avec le même niveau de compétence et le même équipement, chaque photographe produira une image différente. Il est alors difficile de dissocier la technique de l'humain, et de qualifier la photographie d'objective et de mécanique. Les auteurs du second manuel avertissent d'ailleurs le lecteur : la photographie peut devenir "*un instrument dangereux, car il est d'autant plus falsificateur qu'on ne le soupçonne pas d'induire en erreur*". L'objectivité mécanique est donc plutôt un "*idéal régulateur*" qu'une finalité en soi, "*un point de repère inatteignable, mais qui tend à avoir une pratique éthique de la science*" en réduisant l'interprétation au minimum (Daston, Galison 2012, p. 144-145).

Pour A. Chéné *et al.*, cette recherche de l'objectivité à travers les photographies d'archéologie passe par deux points essentiels : la maîtrise technique de la photographie et la connaissance de la destination de l'image. Pour la maîtrise technique, ils défendent à juste titre le fait qu'il est impossible de s'approcher d'une quelconque objectivité sans être conscient des limites du procédé photographique (Chéné *et al.* 1999, p. 16-20). Cette remarque, qui dépasse largement le champ de la photographie pour s'appliquer à l'ensemble des techniques,

est détaillée ici selon trois aspects. Le premier est le choix du point de vue et la nature du sujet à photographier, qui détermine le choix de l'objectif à employer avec l'appareil photo, afin de limiter les déformations optiques (*cf. infra* 3.3. Photogrammétrie).

Le second aspect, et non des moindres, est le choix de l'éclairage. Le terme photographie signifie littéralement "écrire avec la lumière", ce qui laisse entendre les enjeux de la bonne gestion de l'éclairage pour retranscrire une image fidèle. Les auteurs dévoilent un exemple assez parlant pour expliquer le rôle de la lumière en photographie : dans un portrait réalisé avec un éclairage en contre-plongé, nous avons tellement l'habitude des visages éclairés par le haut que la personne est généralement méconnaissable. Il en va de même pour les vestiges archéologiques, pour lesquels les reliefs et les couleurs peuvent radicalement changer d'une photo à l'autre, en fonction de l'angle et de l'intensité de la source de lumière. Il est vrai que sur le terrain, les archéologues ont rarement le choix, puisque que la luminosité est bien souvent liée aux conditions météorologiques. La qualité du photographe ne s'exprime pas alors dans son choix de l'orientation et l'intensité de l'éclairage, mais plutôt dans sa capacité à saisir la lumière optimale en fonction du sujet, en portant une attention particulière à la position du soleil et à la couverture nuageuses (Cookson 1954, p. 13).

Enfin, A. Chéné *et al.* rappellent dans un troisième point que les mires couramment employées en photographie d'archéologie sont uniquement là pour indiquer un ordre de grandeur. En effet, la photographie produit des perspectives, et non des projections en plan. Ainsi, d'après eux, la taille de certains objets peut varier de près de 20 % sur l'image, selon leur position dans l'espace par rapport à l'échelle. Il est intéressant de noter que si à l'époque de la publication les auteurs employaient des appareils photo argentiques, tous les points évoqués restent valables avec les appareils photo numériques actuels (Djindjian 2017, p. 169-170).

Sur la question de la finalité de la photo de terrain, la plupart des guides méthodologiques en archéologie s'accordent à dire que tous les éléments jugés pertinents doivent être photographiés (Barker 1993, p. 128-134 ; Greene, Moore 2010, p. 137 ; Djindjian 2017, p. 169-170). Pour juger, il est donc indispensable que le photographe possède une culture archéologique poussée, afin de saisir les problématiques qui justifient la prise de vue. Puisqu'il ne s'agit pas d'enregistrer factuellement des éléments de la réalité, le principal objectif de la photographie d'archéologie est de "*servir un discours qui se veut scientifique, celui qui naît et s'élabore au cours de la recherche archéologique*" (Chéné *et al.* 1999, p. 12-14). Les photos se classent alors selon trois catégories fonctionnelles, qui peuvent se recouper :

- la "*pièce à conviction*" (montrer) : quelle que soit la destination des photos, toutes ont "*en commun un rôle de témoignage de l'existence d'un objet (ou d'un groupe d'objets)*", ou d'une situation "*à un moment donnée*". Ce rôle est indispensable en archéologie puisque que les originaux sont par nature éphémères (processus de fouille, érosion, restauration ou destruction des élévations, problème de conservation des objets, etc.) ;
- la "*photo démonstrative*" (démontrer) : elle est construite comme un éléments du discours scientifique et appuie les observations et interprétations des archéologues. C'est ce rôle particulièrement qui requiert une bonne connaissance de la finalité de l'image. Ce pose alors le même défi qu'à l'époque de la vérité d'après nature : si l'archéologue n'est pas le photographe, la création du cliché doit se faire à partir d'un véritable dialogue entre les deux individus ;
- la "*photo d'illustration*" : certaines images sont principalement dans une démarche esthétique, pour alléger visuellement le contenu des publications scientifiques ou dans une démarche de valorisation ("photo de couverture", poster, etc.).

Avant les années 2000, lorsque la technologie argentique prévalaient, le photographe devait maîtriser les moyens techniques (choix équipement, réglages) à employer en fonction des problématiques liés aux vestiges étudiés, de leur nature (conservation, superficie, etc.) et des conditions d'acquisition, tout en modérant le nombre de clichés. Un tel savoir-faire permettait d'identifier et comprendre les sujets à photographier, afin de produire une documentation visuelle "parlante" et "lisible" par la communauté archéologique, mais également de limiter les coûts et le temps de développement des images, relativement élevés à l'époque. C'est pourquoi les photographes capables de réaliser des photographies d'archéologie étaient, plus naturellement qu'aujourd'hui, reconnus comme des spécialistes, tels que qu'A. Maillier, l'archéologue-photographe de Bibracte EPCC, ou Ph. Collet de l'EFA, déjà évoqué plus haut.

Mais à la fin du XX^e s., début du XIX^e s., la pratique de la photographie d'archéologie à progressivement évoluée, comme nous l'avons déjà mentionné pour le cas de Bibracte (*cf. supra* 1.2.1. Prog. de recherche européen). Cette période correspond à une large démocratisation de photographie, d'abord avec l'apparition des appareils photographiques jetables, ou prêt-à-photographier, dans les années 1980. Mais c'est surtout avec l'avènement de la photographie numérique dans les années 2000 que cette tendance s'est concrétisée, pour deux raisons. Premièrement, le développement des photographies s'est retrouvé grandement facilité. En effet, une fois le matériel acquis – un APN, des cartes mémoires et un ordinateur pour

la visualisation –, et tant que les images restent au format numérique, le développement est quasi-instantané et sans frais supplémentaire. Deuxièmement, les APN ont aussi proposés les premiers modes de réglage automatiques ou semi-automatiques. Jusqu'alors, les appareils photo haut de gamme, généralement utilisés en photographie d'archéologie, étaient réservés à des utilisateurs confirmés qui maîtrisaient les nombreux paramètres à régler manuellement. Grâce à l'ajout de cartes mères, les APN se sont progressivement dotés de logiciels internes capables de déterminer automatiquement les réglages adaptés en fonction de l'exposition de la scène. Cette fonction a permis d'élargir la base d'utilisateurs des APN haut de gamme, même si elle a tendance à opacifier la gestion des paramètres essentiels de la photographie.

La facilité de production apportée par les APN a donc entraînée un changement de pratique généralisé, dont la photographie d'archéologie de terrain n'a pas été exempté. Avec le numérique, il devenait possible de multiplier les clichés pour enregistrer visuellement, en série, chaque étape, chaque scène du chantier sans se soucier des contraintes de création, de développement ou de conservation. La fonction de l'image en archéologie de terrain s'est donc progressivement transformée, en même temps que les contraintes techniques semblaient se lever. C'est en tout cas ce qui ressort de l'étude menée par un groupe de recherche sur les pratiques photographiques en chantier archéologique (Desprès-Lonnet 2013). Ils ont notamment travaillé pendant trois ans sur les photographies d'archéologie produites à Bibracte, en lien direct avec les archéologues. D'après l'auteur de l'article, M. Desprès-Lonnet, les chercheurs avec qui le groupe a eu l'occasion d'échanger classent eux-mêmes leurs clichés selon deux catégories, là encore en fonction de leur destination :

- Les "*photos de travail*" : relevant d'une pratique plus "amateur", il s'agit de photographies prises sans porter une attention particulière à la technique, qui s'apparente à des "*prise de note visuelle*". Il peut s'agir de simples pense-bêtes, dont seul l'auteur possède les éléments pour comprendre le contenu de la photo. Leur utilité est généralement pensée sur le court terme, pour se souvenir d'éléments particuliers jusqu'à la rédaction du rapport par exemple. Dans ce cas là, l'avantage de la photographie est de réaliser une prise de vue instantanée, sans effet sur le déroulement du chantier. Les photos de travail peuvent aussi comprendre des images plus abouties, comme dans le cadre d'un enregistrement visuelle systématique des US, où les mêmes éléments normatifs sont repris pour chaque unité. L'objectif est ici d'inventorier et de conserver une trace des éléments fouillés.

- Les "*bonnes photos*" : ces images sont pensées dès leur création comme des documents visuels, qui répondent à un cadre normatif et affichent les attributs nécessaires pour être comprises par un public spécialiste. Les archéologues consultés indiquent qu'elles ont pour objectif de "*montrer*", "*d'informer*", mais on retrouve en réalité le rôle de "*démontrer*" déjà défini par A. Chéné *et al.* . Il est question ici d'anticiper l'image et d'identifier les paramètres adaptés pour montrer ce que l'on voit, ou plus exactement, ce que l'on souhaite que l'Autre, le lecteur, voit. Pour le public interrogé, les bonnes photos doivent comme toujours conserver par l'image des éléments voués à disparaître, principalement pour l'Autre qui ne pourra donc pas revenir "*à la source*". Mais tous les éléments présents sur le chantier ne peuvent pas ou ne requiert pas de faire l'objet d'un tel traitement photographique. Cette prise de décision est à la discrétion du fouilleur, qui connaît les problématiques qui motive l'opération de terrain.

Dans l'utilisation de l'image, l'autrice note que la frontière entre les deux catégories reste floue. Par exemple, des photos produites dans le cadre d'un enregistrement systématique peuvent très bien être considérées comme de bonnes photos. De même, une photo de travail, même si elle n'est pas d'une grande qualité technique, peut finalement s'avérer être le témoin le plus représentatif d'un élément que l'archéologue aurait compris *a posteriori*, et devenir une démonstration type "*la preuve par l'image*" dans son discours.

Ainsi, avec l'introduction du numérique, les usages des photos se sont diversifiés et intensifiés sur les chantiers archéologiques. Cependant, quelles que soient la technique ou la pratique envisagée, il apparaît que la photographie d'archéologie n'est pas une action d'enregistrement neutre et automatique. Elle ne fabrique pas des reproductions véridiques de scènes de fouilles archéologiques en série. Il s'agit au contraire d'une démarche scientifique, qui repose sur la capacité du photographe à questionner une situation observée pour en rendre compte à travers une image. Dans les faits, même s'il est souvent admis que les photos doivent être réalisées sous une lumière diffuse, sans ombre portée, centrées sur le sujet, avec un environnement propre sans éléments perturbateurs, sur un axe vertical pour les photos en plan et un axe horizontal pour les photos d'élévation ou de coupe, en réalité chaque archéologue est conscient qu'une photo de terrain est avant tout affaire de compromis. La capacité du photographe est donc intimement liée à sa capacité à lire le terrain, en fonction de sa culture scientifique, et à retranscrire ce qu'il comprend de la scène à travers une image. Il n'est d'ailleurs pas rare de prendre plusieurs clichés d'une scène avant de sélectionner le "*bon*", ce qui correspond plus à une expérimentation pour produire un résultat qu'à l'enregistrement d'un fait objectif.

Par conséquent, nous considérons que la photographie d'archéologie de terrain est bien un témoignage, dans le sens où une photo ne se contente pas d'attester, mais plutôt de transmettre ce que l'auteur perçoit du site qu'il est entrain de fouiller. L'archéologue construit une archive de fouille. L'image photographique est donc un élément constitutif du discours scientifique du chercheur qui interprète le site fouillé. Le passage aux technologies numériques permet techniquement d'améliorer la qualité de ces archives, car le développement instantané et immatériel facilite grandement le contrôle des images. Pourtant, dans le cas de Bibracte, l'évolution des pratiques vers les équipements numériques a entraîné une baisse de qualité de la majorité des photographies d'archéologie. Elles étaient auparavant effectuées par un professionnel de la photographie avec un équipement haut de gamme qui maîtrisait les problématiques de la prise de vue en contexte de chantier archéologique. Depuis, les images sont dans leur grande majorité effectuées par des personnes peu au fait des enjeux de la photographie d'archéologie, qui travaillent avec du matériel grand public ne répondant pas forcément aux standards de l'image archéologique. C'est pourquoi il nous semble important de revenir sur les notions techniques essentielles de la photographie et son utilisation en archéologie de terrain.

3.2.2. Choix des équipements

Les éléments décrits par la suite sont issus de nos lectures, ainsi que d'expérimentations sur trois capteurs différents : le Sony Alpha 7, le Canon EOS 6D et le capteur intégré du drone *Dji Phantom 4 Pro*.

En photographie comme dans beaucoup d'autres techniques, le choix de l'équipement est déterminé par le résultat souhaité. Hors nous l'avons vu, les usages de la photographie se sont multipliés en archéologie avec l'introduction de la photographie numérique. De plus, ce qui relève de la "bonne" photographie d'archéologie de terrain est plutôt variable, et dépend essentiellement des problématiques qui ont motivées la fouille du site et de l'aptitude des archéologues photographes à illustrer un cheminement scientifique à travers une image. Dans ces conditions, il semble difficile de recommander l'usage d'un type de matériel en particulier.

Il est toutefois possible de considérer que quels que soient les choix de l'auteur, une photographie d'archéologie de terrain doit toujours chercher à produire, sur le plan technique, des images nettes et bien exposées, afin de maximiser leur lisibilité. Dans les faits, avec les

progrès réalisés ces dernières années sur les capteurs et les systèmes optiques, il existe aujourd'hui tout une panoplie d'équipements capable de produire ce type de résultats, allant des APN professionnels haut de gamme jusqu'à certains smartphones.

La plupart de ces appareils permettent également de modifier les paramètres d'acquisition manuellement, avec des modes prédéfinis ou complètement automatisés. Les paramètres principaux à prendre en compte sont la sensibilité du capteur, l'ouverture du diaphragme de l'objectif, la vitesse d'obturation du capteur et la balance des blancs (*cf. infra* 3.3.1. Contraintes photographiques). Comme nous l'avons déjà évoqué, il est impossible de normaliser l'usage de ces paramètres, car les conditions de prise de vue changent à chaque situation. N'étant pas nous même photographe de formation, nous renvoyons le lecteur vers les publications spécialisées pour appréhender les caractéristiques techniques à maîtriser en fonction du sujet à photographier, de son environnement et de la finalité de l'image (Cookson 1954 ; Matthews 1968 ; Simmons 1969 ; Conlon 1973 ; Chéné *et al.* 1999). Parmi nos lectures, une publication se démarque par la clarté et le niveau de détail du discours : le chapitre de G. Verhoeven dans l'ouvrage *3D Recording, Documentation and Management of Cultural Heritage*, intitulé *Basics of Photography for Cultural Heritage Imaging* (Verhoeven 2016). En une centaine de pages, l'auteur reprend toutes les bases de la technique photographique pour la prise de vue de sujets patrimoniaux. Il passe en revue les procédés physiques théoriques, les caractéristiques matérielles des différents équipements et jusqu'aux formats d'export des photos numériques. L'objectif n'est pas d'anticiper une situation, mais au contraire de décrire de la manière la plus exhaustive possible tous les mécanismes de la photographie, ainsi que leurs limites, afin d'exploiter pleinement le potentiel de cette technique.

La formation prime donc sur le choix de l'équipement. Dans la pratique, un spécialiste de la photographie d'archéologie de terrain équipé d'un APN compact milieu de gamme obtiendra probablement de meilleurs résultats qu'un archéologue ayant une pratique "amateur" de la photographie avec une APN reflex haut de gamme (fig. 49). En poussant cet argumentaire plus loin, même le choix de la technologie n'est pas réellement contraignant, puisque que déjà à la fin des années 1990, il pouvait être difficile de faire la différence entre une image nativement numérique ou issue d'une pellicule argentique numérisée (Collet 1996, p. 339). Nous souhaitons cependant apporter ici une nuance pragmatique : si la maîtrise du matériel permet de produire des images de grande qualité, le mode automatique des APN actuels permet dans la plupart des cas de réaliser des photographies lisibles et convenables pour une exploitation scientifique.



Compact

Sony DSC-HX99



Hybride

Sony Alpha 7 II



Reflex

Canon Eos 6D

49. Types d'appareils photo numérique

Maintenant que ces aspects sont évoqués, nous pouvons néanmoins fournir quelques conseils. L'ambition est simplement ici de partager un savoir empirique, basé sur notre propre expérience de terrain ou celle de collègues proches.

Dans le cadre d'une opération de fouille archéologique, un des aspects les plus importants à prendre en compte est la polyvalence de l'APN. Les archéologues sont amenés à réaliser des clichés de sujets très diverses, qui vont de la photo de détail d'un objet de quelques centimètres aux images générales du site. Pour résumer grossièrement, les APN d'entrée de gamme, tous types confondus, vont généralement se retrouver rapidement limités pour les photos de sujets très proches, très éloignés ou en basse lumière. Ils peuvent aussi parfois afficher une définition trop basse et produire des images "pixelisées", même si les développements récents des capteurs tendent à infirmer ce dernier point. À l'inverse, les APN haut de gamme sont beaucoup plus polyvalents, particulièrement ceux avec un objectif interchangeable, afin d'adapter la prise de vue à toutes les situations.

Si les possibilités de prise de vue offertes par le changement d'objectif sont démultipliées, les objectifs à focale variable standards permettent de réaliser la grande majorité des photos de terrain (sur les types d'objectif, cf. *infra* 3.3.1. Objectif photographique). Un modèle classique, tel qu'un objectif zoom 24-70mm équivalent 35mm, produit en général des clichés de sujets centimétriques à plusieurs dizaines de mètres, sans entraîner d'effets optiques très marqués, comme l'effet *fish-eye* des objectifs ultra grand-angles (fig. 50). De plus, le fait d'emporter un seul objectif peut être un avantage dans le cas de missions où le transport est contraignant, sur des sites difficiles d'accès ou lors de décapages extensifs très étendus par exemple. En effet, le poids d'un objectif peut varier d'une centaine de grammes à plus d'un



50. Exemple d'objectif interchangeable polyvalent : le Canon EF 24-70mm f/4L IS USM

kilo. Sachant qu'il s'agit d'un équipement fragile, à conserver de préférence dans un contenant adapté, à l'abri des chocs et de la poussière, le matériel photographique peut rapidement devenir lourd et/ou volumineux, donc problématique à déplacer.

La possibilité de changer d'objectif reste malgré tout très appréciable, car les APN sont utilisés dans de nombreuses autres tâches que la photographie de terrain en archéologie, chacune avec leurs besoins techniques spécifiques. Il s'agit par exemple de l'étude de détails centimétriques à subcentimétriques sur des objets, à l'aide d'objectifs macro, ou encore d'acquisitions photogrammétriques, pour lesquelles un objectif à focale fixe est conseillé (*cf. infra* 3.3.1. Objectif photographique). Selon la qualité de l'appareil, les objectifs interchangeables permettent donc de faire évoluer un APN pour répondre à tous les besoins des archéologues en terme de photographie, sans faire de compromis sur la qualité des images.

Enfin, un dernier point évident, mais qu'il nous semble utile de rappeler tant il peut être source de frustration, est le cas des batteries. Nous conseillons de privilégier les APN avec batterie indépendante, et de posséder plusieurs batteries de rechange, afin de ne jamais se retrouver bloqué par une batterie épuisée ou hors d'usage.

Pour toutes les raisons évoquées et selon notre expérience, notre préférence pour la photographie d'archéologie de terrain se porte sur les APN hybrides milieu ou haut de gamme, équipés d'un objectif à focale variable standard (24-70mm équivalent 35mm). Ce type de kit associe un appareil compact et performant avec un objectif permettant de prendre des photos de détail comme des vues d'ensemble. Lorsque le transport n'est pas un problème, les APN

reflex, plus larges, offrent aussi un véritable confort de prise de vue, avec une prise en main ergonomique et la possibilité de choisir entre le viseur optique ou le retour numérique pour cadrer la photo.

3.2.3. Acquis des expérimentations

Finalement, la photographie d'archéologie de terrain relève autant de la maîtrise technique de l'équipement que de la sensibilité du photographe face à la nature des vestiges et le discours qu'il souhaite en dégager. Il est nécessaire que ce dernier maîtrise les problématiques archéologiques à l'origine de la fouille pour :

- identifier quels éléments nécessitent d'être photographiés et selon quels paramètres afin de transmettre, par l'image, une vision fidèle à celle développée par les chercheurs sur le terrain ;
- intervenir en fonction de l'avancée des travaux, être capable d'évaluer à quelle étape du processus de fouille la photo sera la plus "parlante", tout en limitant l'impact de la prise de vue sur le déroulement du chantier ;
- repérer les conditions idéales d'acquisition, naturelles (luminosité et hydrométrie) et logistiques (accessibilité du site).

C'est en étant capable de "calculer" le compromis optimal entre le quoi, le comment et le quand que l'archéologue photographe produit une image pertinente et de qualité scientifique.

D'après notre expérience, plusieurs points méritent d'être abordés pour faciliter la recherche de ce compromis. Avec les automatismes apportés par les APN, la prise de vue est techniquement facilitée. Ils devraient en toute logique conduire à l'amélioration globale des clichés. Pourtant, les photos réalisées par les équipes de fouilles ont tendance à être de qualités très variables. C'est en tout cas ce que nous avons pu constater sur les chantiers auxquels nous avons participé durant la période de cette recherche doctorale. Par exemple, sur la fouille de la *domus* PC2, malgré un temps de formation accordé à tous les fouilleurs sur l'importance de l'image pour l'enregistrement archéologique, plusieurs photos sont à la limite de l'exploitable. Certaines images sont floues ou mal cadrées, d'autres affichent un mauvais éclairage ou la présence d'éléments perturbateurs (des outils, des tas de terre ou les pieds du photographe).

Selon nous, deux raisons principales peuvent être envisagées pour expliquer cet état de fait. La première peut-être lié à l'apparente facilité de production des images depuis l'arrivée des APN sur les chantiers archéologiques. Les appareils ont atteint depuis quelques années un niveau de performance et d'automatisme élevé, qui a largement contribué à démocratiser l'usage de la photographie. Mais cette recherche de maîtrise "intuitive" des aspects techniques de la part des constructeurs a dans le même temps limité les interrogations des photographes amateurs sur les modalités de création d'une photo. Pour la génération d'archéologues formée depuis la fin des années 2000 – dont nous faisons partie –, il suffit d'appuyer sur un bouton pour produire des dizaines d'image, sans se soucier de leur développement, de leur intérêt ou de leur conservation. Le phénomène s'est possiblement renforcé avec l'intégration d'APN directement dans les smartphones, permettant à une grande majorité de la population de manipuler la photographie quotidiennement.

Il est probable que pour cette génération, le statut de l'image ait évolué avec les progrès techniques récents, ce qui a entraîné des répercussions sur la qualité des photographies d'archéologie de terrain. Les personnes de cette génération sont prédisposées à photographier chaque unité stratigraphique, chaque détail du chantier, simplement parce qu'il en ont la possibilité. Mais la multiplication des images semble s'accompagner d'une perte d'intérêt pour des questions essentielles, que les archéologues se posaient peut être plus naturellement à l'époque de la photographie argentique : quel est le sujet de la photo, pourquoi le photographier et comment ? La photographie d'archéologie est ainsi passée d'un objet réfléchi à une production de masse, qui s'accompagne généralement d'une baisse de qualité du résultat final et traduit peut-être une peur de ne plus se souvenir, de ne pas comprendre. Plus largement, ces questions traduisent aussi un problème de réflexion scientifique, avec l'incapacité à savoir quand produire une archive destinée à la communauté archéologique. L'argument de la fracture générationnelle autour des nouvelles technologies est aussi à relativiser : il est probable que les archéologues non formés aux enjeux de la photographie à l'époque des appareils argentiques soient tombés dans les mêmes dérives méthodologique lorsque que les APN ont fait leur apparition.

De plus, il nous semble que cette production en série tend à renforcer l'idée de la photographie de terrain comme un système d'enregistrement visuelle "objectif" de la réalité. Si l'archéologue n'est ni sensibilisé par sa formation, ni contraint par le matériel à s'interroger sur les raisons de la prise de vue, il est alors tenté d'agir par automatisme méthodologique (prendre un cliché pour prendre un cliché). Pourtant, il est pourtant nécessaire de questionner

en amont l'action elle-même pour respecter le caractère de témoignage de l'image : qu'est-ce que je veux "montrer", ou plutôt "démontrer", avec cette photo ? Une telle démarche conduit à auto-évaluer le résultat, à considérer si les détails visibles sur la photo sont conformes à notre propre compréhension de la fouille. D'un point de vue pratique, il est d'ailleurs conseillé de vérifier le plus rapidement possible les clichés sur un grand écran, car très souvent les défauts sont peu visibles sur l'écran de l'APN. C'est particulièrement le cas de l'auto-focus, qui a tendance à créer des légers flous difficiles à visualiser sur l'écran de l'APN, mais qui gênent effectivement la lecture de l'image en plus grand format.

La recherche d'objectivité à travers la photographie d'archéologie de terrain comporte selon nous une autre limite. Comme l'ont démontrés L. Daston et P. Galison avec l'objectivité mécanique, la régularité avec laquelle sont effectuées les prises de vue sur les chantiers appuie l'impression d'une démarche systématique (Daston, Galison 2012, p. 137-148). L'automatisme de la pratique serait gage de sa scientificité. Mais, paradoxalement, elle restreint aussi la capacité de mémoire visuelle de la photographie sur un chantier. Souvent, tous ce qui n'est pas une photo de couches archéologiques ou d'objets est considéré comme une photo d'ambiance, non scientifique. Pourtant, la documentation par l'image des méthodes et des instruments mis en œuvre lors de la fouille a une véritable valeur scientifique, car elle permet aux chercheurs extérieurs de comprendre comment ont été produites les données. On parle ici d'images qui illustrent des outils ou des techniques de fouille spécifiques. Elles sont particulièrement utiles, notamment pour accompagner des parties méthodologiques dans les rapports ou les publications. Ce type de prise de vue, en plus de constituer des archives précieuses sur nos pratiques, encourage la critique et la reproductibilité des processus de fouille, et renforce donc leur caractère scientifique.

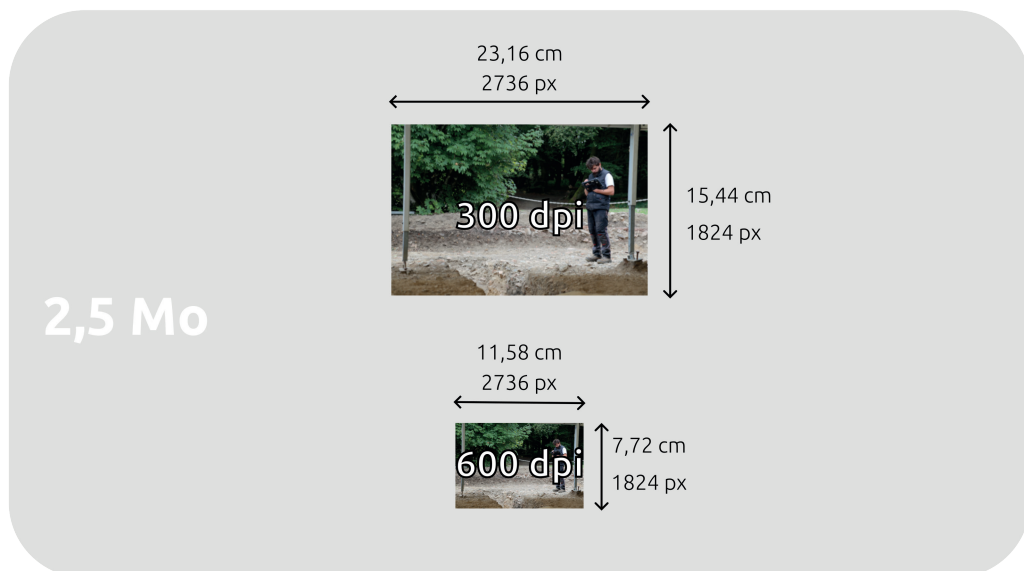
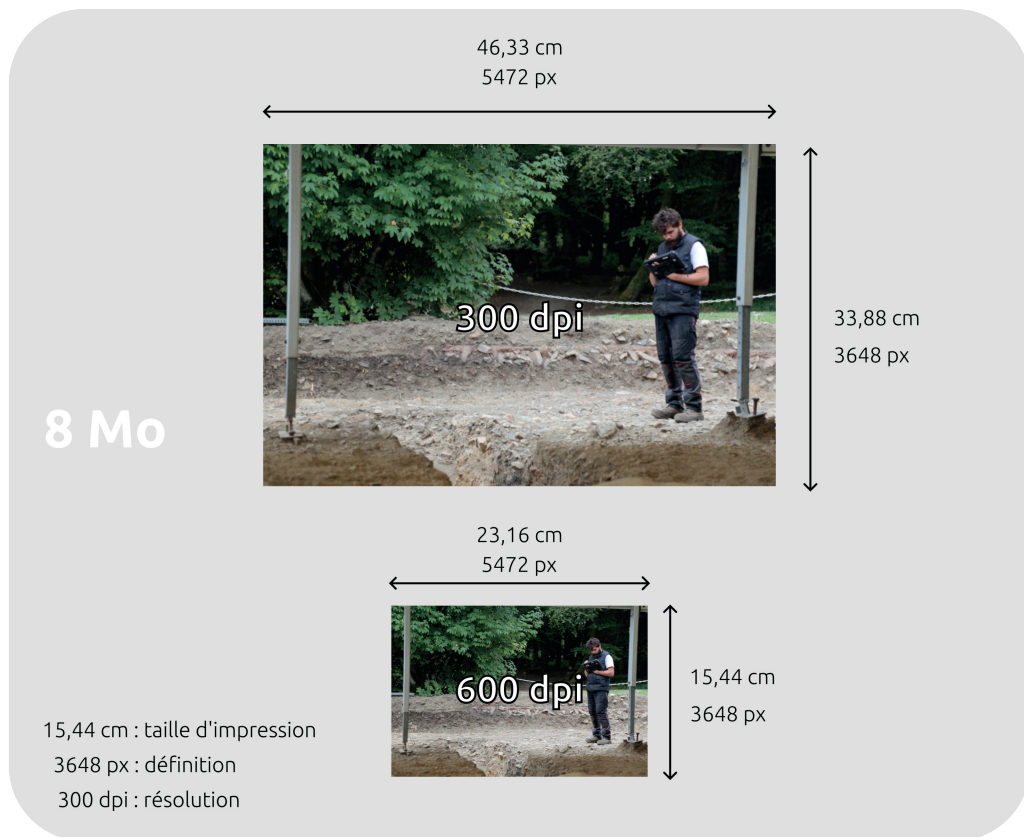
D'une manière plus générale, nous insistons sur l'importance de réaliser des photos sur un chantier, même si d'autres moyens de représentation visuelle sont mis en œuvre (dessins, photogrammétrie, lasergrammétrie, etc.). Lorsque les archéologues sont formés à l'utilisation du matériel et aux objectifs de la photographie d'archéologie de terrain, l'image reste un moyen rapide, efficace et particulièrement accessible pour retranscrire une perception du site.

Attention toutefois à ne pas tomber dans l'extrême opposé. La photographie d'archéologie revêt un caractère plus subjectif que ce que les archéologues souhaitent communément lui attribuer. Mais accepter ce caractère subjectif ne constitue pas pour autant un argument suffisant pour se passer des normes existantes dans la discipline. Si la fonction principale

de la photographie est de transmettre une vision du site, la personne réceptrice de l'image doit pouvoir comprendre et interpréter ce regard. Le respect des codes et des attributs photographiques classiquement utilisés par les archéologues constitue la base d'un langage visuel commun, sans lequel l'échange d'informations serait grandement complexifiés. Par codes, nous comprenons par exemple le fait de réaliser des prises de vue à la verticale pour les éléments en plan, à l'horizontale pour les éléments en coupe ou en élévation, de centrer la photo sur le sujet et son contexte proche, d'évacuer toute personne, objets ou matériaux qui pourrait obstruer la lecture des couches stratigraphiques et de limiter au maximum les ombres portées. Les attributs traditionnellement utilisés sont le panneau lettreur, pour renseigner le contenu de l'image, les échelles photographiques et la flèche qui indique la direction du Nord pour les photos en plan.

Pour les aspects plus techniques, propres à la photographie numérique, nous conseillons tout d'abord d'être attentif au réglage de la définition des photos. Si l'on prend l'exemple de l'APN Canon EOS 6D utilisé sur le chantier de la *domus* PC2, il capture à sa définition maximum des clichés au format JPEG de 5472 x 3648 pixels, d'un poids moyen de 8 Mo (fig. 51). En considérant la résolution minimale généralement recommandée par les éditeurs, soit 300 dpi, la taille recommandée pour l'impression de la photo est de 46,33 x 30,88 cm. À 600 dpi, la taille d'impression est logiquement divisée par deux, soit 23,16 x 15,44 cm. Lorsque l'on prend les mêmes clichés, mais avec une définition divisée par deux, soit 2736 x 1824 pixels, on obtient à 300 dpi une image imprimable de 23,16 x 15,44 cm et à 600 dpi de 11,58 x 7,72 cm, d'un poids de 2 à 3 Mo.

La question est alors de savoir s'il est utile de produire et de conserver toutes les photos dans un format plus grand qu'un A3. Dans notre exemple, en diminuant la définition par 2, le fichier reste dans les normes recommandées aux auteurs de nombreuses publications, mais permet de conserver un fichier presque trois fois plus léger. Ce type de définition est en général suffisant pour les photos d'inventaire, réalisées avant et après fouille d'une couche. Nous incitons donc à maîtriser la définition en fonction de l'usage de la photo, et à réserver les hautes définitions aux images destinées à de l'impression sur des formats au-delà du A3. Si la modification des réglages de l'APN se révèle problématique en cours de fouille, il est possible de prendre toutes les photos à la définition maximum, puis d'effectuer un traitement par lot pour les sous-dimensionnées en fin de campagne, avec le plugin BIMP du logiciel GIMP



51. La différence entre la définition, la résolution et la taille d'impression d'une photographie

par exemple (Francesconi 2023). Cette aspect technique n'est d'ailleurs pas suffisant. Le tri n'est pas neutre. Le choix des images à conserver participe à la construction du discours scientifique, et s'en abstraire revient à construire un "bruit" dans les informations archéologiques.

À propos des formats d'export, certains appareils permettent de créer les images dans un format numérique brut, généralement désigné par le terme générique *RAW*. Il s'agit de fichiers au format propriétaire, contenant les informations brutes du capteur. Ils sont lourds et nécessitent une étape de développement numérique pour être exploités, à l'aide de logiciels de traitement d'images numériques brutes, type Darktable. Il n'est donc pas utile de les produire sans justification préalable. Le format le plus adapté à la création et à la manipulation d'image reste le JPEG, un format standard et libre et internationalement reconnu depuis les années 1990. Mais le JPEG n'est pas qu'un format de fichier. C'est également une norme de compression d'image qui peut mener à une dégradation de sa qualité. C'est pourquoi il est recommandé d'opérer une sélection sur les photos à conserver en fin de campagne et de les convertir en format TIFF, un autre standard libre plus lourd, mais stable et adapté à l'archivage numérique

Parmi les nouvelles méthodes d'imagerie récemment déployer sur les chantiers archéologiques figure une technique en apparence proche de la photographie d'archéologie de terrain : la photogrammétrie. Cette dernière utilise des photos comme données primaires pour produire des images plus complexes, mais se supplante pas la photographie pour autant. Il s'agit de deux méthodes complémentaires. À la différence de la photographie qui retranscrit la perception du photographe, la photogrammétrie permet de reproduire une surface en 3D à partir d'images 2D. Le résultat est issu d'un calcul mathématique, reproductible, dont les produits peuvent être jugés de manière logique et systématique. Elle offre donc une nouvelle opportunité de produire les "enregistrements visuels objectifs" tant désirée par la communauté archéologique. Mais comme pour toute technique, la photogrammétrie nécessite de maîtriser un certain nombre de notions et d'appréhender leurs limites, sans quoi, comme pour la photographie, la valeur scientifique du résultat peut rapidement être remise en question.

3.3. Photogrammétrie : nouvelle technique, nouveaux (re)levés

3.3.1. Objectifs

Le relevé, ou levé, est une notion qui vient des techniques de la topographie et de la topométrie. Selon l'École Nationale Supérieure des Travaux Publics (ENSTP), elle comprend d'abord "*l'ensemble des opérations destinées à recueillir sur le terrain les éléments nécessaires à l'établissement d'un plan ou d'une carte*" (Ngouajio M. 2015, p. 5). Par extension, le relevé désigne également le support créé par l'action de relever, soit le plan lui-même.

En archéologie, plusieurs types de relevés sont produits, classés selon la technique employée, tels que les relevés manuels, les relevés topographiques ou plus récemment les relevés LiDAR ou photogrammétriques (Esnault 2019). Toutes ces techniques aspirent au même objectif : prendre les mesures de points dans l'espace, en 2 ou 3 dimensions (X, Y ou X, Y, Z), pour représenter les indices archéologiques (mobilier, couches, élévations), selon des niveaux de détail plus ou moins élevés (de la silhouette à la représentation fidèle) et des échelles variables (de l'objet au territoire).

Selon la technique employée, relever peut revenir à la fois à enregistrer et interpréter les vestiges. Dans le cas d'un relevé manuel par exemple, le but est de retranscrire, à une étape de la fouille, la localisation et l'organisation des couches archéologiques selon un cadre graphique normalisé. Mais pour représenter les différents éléments, l'archéologue fait appel à sa capacité à lire les indices archéologiques. Cette capacité est nécessairement influencée par sa formation initiale, son expérience professionnelle et ses usages méthodologiques en archéologie. Il n'est d'ailleurs pas rare que l'existence ou l'emplacement exact de limites de couche soit sujet à discussion durant l'élaboration d'une minute. Avec le relevé manuel, l'action consiste donc autant à enregistrer factuellement une trace archéologique qu'à en proposer une première interprétation. C'est également le cas des relevés topographiques, car l'emplacement de la mesure d'un point est à la discrétion de l'archéologue.

À l'inverse, dans certaines situations l'enregistrement et l'interprétation sont séparés. Il s'agit des relevés réalisés à l'aide de scanner tridimensionnels ou de la photogrammétrie. Dans ces cas là, quel que soit la méthode, l'instrument opère une multitude de mesures sur une surface donnée. Il en résulte un ensemble de points qui modélisent la surface en 3D, appelé communément nuage de points. Divers traitements peuvent ensuite être appliqués sur

les nuages de points, pour produire par exemples des profils, des modèles de profondeur ou des images orthonormées. C'est en conjuguant leurs observations directes et l'analyse de ce type de support, que les archéologues peuvent formuler des interprétations.

Nous proposons alors de faire une distinction, pour l'archéologie, entre les concepts de lever et relever. Le relevé correspond à la première situation, lorsque que les mesures sont opérées conjointement à l'interprétation. Le levé, quand à lui, regroupe l'ensemble des actions où les mesures et les interprétations sont formalisées sur des temps différents. Autrement dit, nous parlons de relevé manuel et topographique, et de levé LiDAR, au scanner 3D ou photogrammétrique.

Contrairement aux autres techniques de levé qui exploitent des instrument de mesure directe, la photogrammétrie a recourt aux photographies en 2D pour reconstituer des scènes en 3D (Kasser, Egels 2001, p. 19). Si la technique est connue et utilisée en archéologie depuis la fin du XIX^{ème} s., c'est grâce au développement simultané de la photographie numérique et de la puissance de calcul des ordinateurs personnels qu'elle a pu récemment se démocratiser parmi les archéologues (Ebrahim 2003, p. 4-6 ; Egels 2011, p. 42 ; Djindjian 2016, p. 4-5). Cette situation a mené au remplacement progressif de tout ou partie de nos relevés de terrain traditionnels (à la main, sur support papier millimétré) par des levés produits par photogrammétrie. Les images de vestiges archéologiques deviennent alors de véritables supports aux relevés techniques et scientifiques. Or la bonne pratique de ce procédé implique la maîtrise d'un ensemble de connaissances et de technique lors des phases d'acquisition et de traitement que de nombreux utilisateurs ont tendance à négliger, tant les interfaces des logiciels actuels en simplifient l'usage. Cette impression, qui n'en est clairement qu'une, conduit même parfois à caricaturer son usage en une démarche "presse boutons" alors qu'en réalité sa mise en œuvre relève de processus pourtant hautement scientifiques.

Si les applications de la photogrammétrie sont nombreuses et variées, l'emploi qu'en fait l'archéologie est aujourd'hui bien standardisé dans la grande majorité de ces usages. Il s'agit principalement de remplacer les traditionnels relevés en plan, en coupe et en élévation des vestiges archéologiques, par la production d'ortho-images, dites orthomosaïques, et de modèles numériques de surfaces.

C'est pourquoi la partie qui suit est issue d'un guide de photogrammétrie à destination de la communauté archéologique, que nous avons récemment publié avec nos collègues A. Tomasinelli et M. Thivet (Verriez *et al.* 2023). L'idée de cette publication a été formulée après

huit années d'expérimentation et de développement méthodologique au cours desquelles nous avons pu tester de multiples logiciels et équipements dédiés, des plus coûteux à la solution libre retenue, en passant par les systèmes les plus couramment utilisés dans notre discipline. Une grande partie de ces développements méthodologiques ont été réalisés dans le cadre des projets de recherches archéologiques de l'UMR 6249 Chrono-environnement, et tout particulièrement celles menées sur le site de Bibracte et nos travaux de thèse. Toutes ces années d'expérimentation et d'observations ont conduits à plusieurs constats : une grande diversité dans la mise en œuvre de ces techniques sur le terrain (des attendus du levé au choix des équipements) ; la formation autodidacte de très nombreux archéologues ; le recours quasi-systématique à des logiciels qui privilégient la simplicité d'utilisation à la reproductibilité des résultats ; l'hétérogénéité notable dans la façon de qualifier (qualité, précision) les données photogrammétriques publiées.

Par conséquent, le premier objectif de ce guide est de permettre à tous d'avoir recours à ces outils, dans la pratique quotidienne de sa discipline, par le recours à des équipements commun et par l'usage d'un logiciel libre et gratuit, issus des recherches scientifiques menées par les principales unités de recherches françaises en photogrammétrie. Le second objectif est de proposer un processus complet et ciblé aboutissant à la production d'orthomosaïque à partir de plusieurs photographies numériques. En effet, les principes de la photogrammétrie permettent également la production de nuages de points décrivant la scène capturée sous une forme tridimensionnelle, mais cet aspect de la méthode et son utilisation ne seront pas ou quasiment pas abordés dans cet ouvrage.

Nous proposons donc ici un protocole de photogrammétrie complet, basé sur des exemples concrets et éprouvés, directement issus de contextes archéologiques et respectant les recommandations récentes du Consortium 3D SHS (Huma-Num) (Granier *et al.* 2019). L'objectif est de fournir aux archéologues les informations essentielles pour produire des orthomosaïques pertinentes et critiquables par la communauté scientifique.

L'imagerie verticale en archéologie : définitions

Avant tout, il est important de faire un point sur les différentes notions qui cohabitent lorsque l'on parle d'images verticales en archéologie.

Photographie zénithale

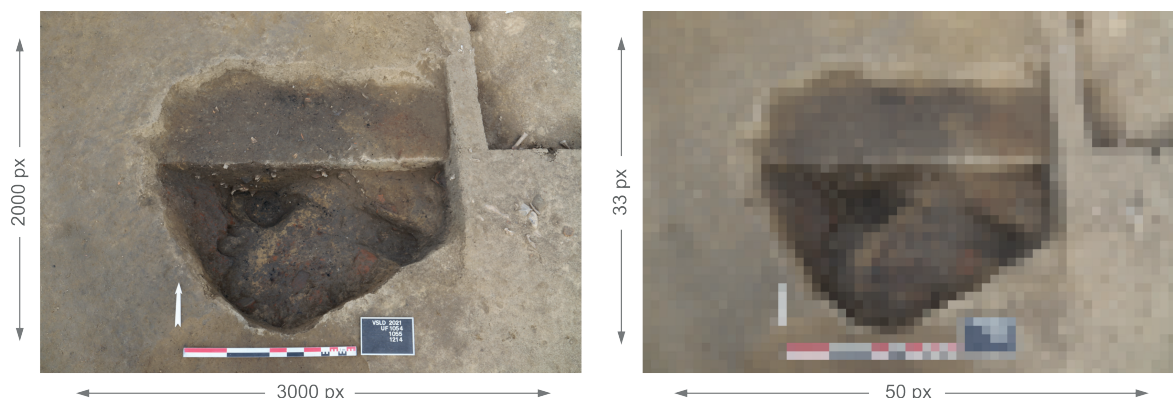
Il s'agit d'une photo dont le plan de l'image est positionné le plus parallèlement possible du plan à relever, pour limiter les effets de déformation liés à l'axe de prise de vue. Elle se veut complémentaire des relevés techniques : la photographie offre une vision "objective" d'une surface, alors que le dessin propose déjà une première interprétation. Mais si les photos numériques sont précieuses pour contextualiser les observations du fouilleur, elles comportent malgré tout plusieurs contraintes techniques pour une exploitation scientifique poussée.

La définition

La définition correspond aux dimensions d'une photo numérique, exprimées en pixels (*cf. supra* 2.1.2. Images matricielles). En photographie, il est classiquement représenté par un petit carré associé à une couleur. C'est l'association de millions de pixels de couleurs différentes qui crée une image. Plus la définition d'une image est élevée, plus elle est composée de pixels "52. une photo archéologique à deux définitions différentes. Verdun-sur-le-Doubs, le Petit Chauvort, Saône-et-Loire, 2021 (cliché E. Hammon)" (fig. 52). Par exemple, une photo d'une définition de 4 000 x 3 000 possède 4 000 pixels de large pour 3 000 pixels de haut. La définition dépend donc d'une part de la dimension du capteur de l'APN, mais également du nombre de pixels (ou photosites) que le capteur comporte. Dans l'exemple présenté, la photo correspond bien à la dimension du capteur, on parle alors d'un APN 12 Méga pixels, car $4\,000 \times 3\,000 = 12\,000\,000$.

Pour la plupart des APN destinés aux amateurs, le couple boîtier (donc capteur) et objectif (qui détermine la longueur focale) ne permet que rarement de prendre une photo d'un sujet très large (plusieurs dizaines de mètres) avec une définition suffisante pour en apprécier les détails centimétriques, qui apparaissent alors trop "pixelisés".

Attention aussi à ne pas confondre la définition et la résolution. La résolution indique la densité de pixels dans l'image, notée en Points Par Pouce (DPI, *Dots Per Inch*) ou Pixels Par Pouce (PPP), soit le nombre de pixels par 2,54 cm². En effet la taille d'un pixel n'est pas



52. Une photo archéologique à deux définitions différentes. Verdun-sur-le-Doubs, le Petit Chauvort, Saône-et-Loire, 2021 (cliché E. Hamon)

fixe. Leur dimension varie en fonction de la taille d'impression de l'image. Plus l'image est agrandie, plus les pixels sont larges et susceptibles d'être visibles à l'œil nu.

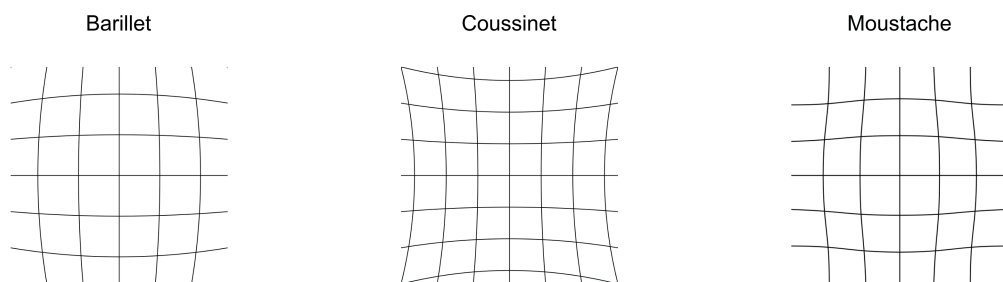
Les déformations d'images ou aberrations géométriques

Les principales déformations rencontrées lors des prises de vue en archéologie proviennent de la difficulté pour les appareils photos à restituer la totalité des détails de l'image dans toute sa profondeur. Dans la pratique cela se traduit par du flou ou par des distorsions optiques par exemple. Nous présentons ici les deux principales aberrations géométriques rencontrées lors de captations d'images destinées à la photogrammétrie en contexte archéologique.

Les distorsions optiques se traduisent par des lignes normalement droites apparaissant légèrement courbées sur les clichés (fig. 53). Tous les objectifs, même les plus hauts de gamme, présentent cette déformation. Elles peuvent être de trois types (Egels 2011, p. 44-45) :

- Distorsion en **barillet** : la photo semble être posée sur une surface convexe, une sphère ou un baril. Les lignes directrices se courbent du centre vers l'extérieur de la photo. Elle est caractéristique des objectifs grand angle ;
- Distorsion en **coussinet** : La photo semble être posée sur une surface concave, comme enfoncée dans un creux. Les lignes directrices se resserrent vers le centre de l'image. Elle est particulièrement visible avec les objectifs de longue focale, ou téléobjectifs ;
- Distorsion en **moustache** : identique à la distorsion en barillet, sauf que les lignes se redressent légèrement sur les bords de la photo.

La déformation de volume, ou anamorphose, est caractérisée par un effet d'étirement des sujets au bord de la photo. Elle est particulièrement visible avec les objectifs grand angle et lorsque la distance entre l'APN et le sujet est courte. Sur les photographies aériennes, les bâtiments semblent parfois "couchés" à cause de l'anamorphose. On parle alors d'effet de déplacement, ou de dévers (Smith 1995).



Exemple de barillet



Exemple de coussinet



53. Les déformations optiques courantes. Bibracte, domus PC2, Nièvre, 2017

Photographie redressée

Pour réaliser ce type d'image, le fouilleur installe des cibles (de 3 à n , en fonction de la précision souhaitée) réparties de manière régulière sur la zone à photographier. Les coordonnées géographiques des cibles sont relevées par un outil de mesure topographique (GNSS, tachéomètre, station totale, etc.) ou positionnées dans un système relatif orthonormé (carroyage). Le principe consiste ensuite à faire coïncider manuellement sur la photo le pixel du centre de chaque cible avec son repère topographique à l'aide d'un logiciel type SIG. La photo est dite redressée, car elle est contrainte (ou transformée) par certains points dans un

Système de Coordonnées de Référence (SCR). Le résultat offre une photo avec un positionnement dans l'espace relativement précis, mais qui conserve néanmoins plusieurs défauts (fig. 54) :



54. Photos redressées de la pièce XX de domus PC2, à Bibracte. Bibracte, la domus PC2, Nièvre, 2017

- Cette méthode ne corrige pas les déformations optiques. Au contraire, si le type de transformation le permet, la manipulation va de nouveau déformer l'image pour compenser les distorsions. De telle manière que certaines lignes droites peuvent apparaître courbes ou en "S" ;
- Il est généralement compliqué d'assembler plusieurs photos redressées. Il faut ajouter aux deux points précédents les problèmes de colorimétrie, de luminosité changeante et la superposition des clichés. L'image finale peut se révéler inutilisable à cause de toutes ces discontinuités entre les photos voisines.

- La photo redressée peut donc être un bon outil de travail pour visualiser et contextualiser rapidement des vestiges archéologiques sur un SIG. Mais, pour les raisons évoquées, elle ne peut pas être considérée comme un support de mesures géométriques fiable et en aucun cas elle ne saurait remplacer un plan.

Photogrammétrie et orthophotographie

Depuis le passage à la photo numérique, les boîtiers se sont rapidement perfectionnés, rendant accessible au plus grand nombre la production d'image en haute définition (Egels 2011, p. 42). Il en va de même pour la micro-informatique, avec l'augmentation exponentielle de la puissance de calcul des ordinateurs personnels. Ces deux innovations technologiques sont à l'origine de la démocratisation de la photogrammétrie, une technique qui requerrait auparavant soit des compétences poussées, soit du matériel complexe et onéreux. Le phénomène a d'ailleurs profité de la mise en réseau de connaissances dans le domaine universitaire pour créer des outils libres et complets, à la portée de la communauté archéologique.

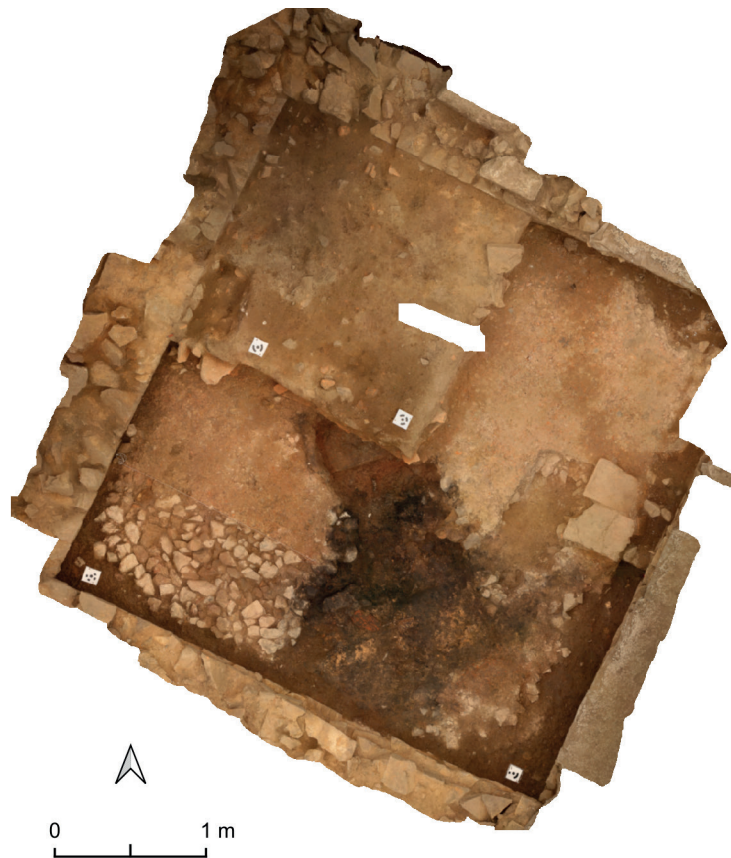
La photogrammétrie est une technique pour modéliser la forme et la position d'un objet ou d'une surface dans un espace tridimensionnel à partir de données en deux dimensions (Paul *et al.* 1997, p. 244). Les données 2D sont le plus souvent des images photographiques, mais elles peuvent être aussi radars, radiographiques, microscopiques ou provenant d'un appareil à balayage. Fondamentalement, "*le processus de photogrammétrie est indépendant du type d'image*" (Kasser, Egels 2001, p. 19). Pour la modélisation d'un objet ou d'une surface, deux méthodes existent (Egels 2011, p. 42) :

Avec une seule image. Il est alors nécessaire de l'associer avec un fichier qui renseigne la forme de l'objet ou de la surface (type MNT ou MNE) ;

Si le fichier de forme n'existe pas, ce qui est le plus souvent le cas en archéologie, il est possible d'utiliser plusieurs images comme plusieurs points de vue sur l'objet ou la surface pour calculer la profondeur. C'est le même principe que la vision stéréoscopique humaine, d'où son nom de stéréophotogrammétrie.

Si la pratique de la photogrammétrie s'accroît ces dernières années, son invention date pourtant du milieu du XIX^e siècle. Elle est utilisée pour la première fois en archéologie une vingtaine d'années plus tard seulement, sur les ruines du site de Persépolis, en Iran (Ebrahim 2003, p. 4 ; Djindjian 2016, p. 4-5). Depuis lors, elle n'a cessé d'être employée pour l'enre-

gistrement et la conservation du patrimoine culturel, et plus particulièrement le patrimoine architectural. Si la photogrammétrie a trouvé un terrain d'application privilégié dans ce domaine, c'est parce qu'elle permet de produire, entre autres, des orthophotographies (fig. 55).



55. Orthophotographie de la pièce XX de domus PC2, à Bibracte. *Bibracte, la domus PC2, Nièvre, 2017*

Une orthophotographie, ou ortho-image, est le résultat du traitement mathématique d'une photographie aérienne (Boldo 2003). Lors du processus de production, l'image de base est recomposée pixel par pixel et présente à la fin les caractéristiques suivantes :

- Les distorsions optiques de l'objectif sont corrigées ;
- Tous les pixels de l'image respectent un axe de vue à la verticale parfaite (comme au centre d'une photo classique) ;
- Chaque pixel est associé à des coordonnées (système local à mondial) ;
- Le relief du sol naturel et la courbure de la terre sont pris en compte.

Autrement dit, l'image est rectifiée pour être exploitée comme une carte plane, tout en affichant les caractéristiques réelles du terrain (Smith 1995). On peut alors calculer les distances au sol et modifier le facteur d'échelle du cliché.

Comme chaque orthophotographie est le résultat d'une opération mathématique, il est possible de relancer un calcul pour les assembler. L'image issue de cette fusion, une orthomosaïque (ou orthophotomosaïque), présente de nombreux avantages pour l'archéologue :

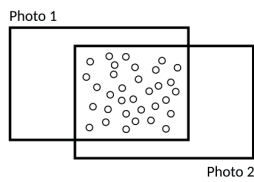
- Elle est entièrement rectifiée. Tous les éléments visibles sont mesurables. Il est alors possible de l'utiliser comme support de dessin technique ou pour mesurer des distances ;
- Elle est géolocalisée bien plus précisément qu'avec les solutions de positionnement GNSS fournies par les APN. De plus elle est nativement exploitable par des outils de cartographie, tels que les SIG ;
- Malgré sa grande taille, elle est en très haute définition, car elle conserve la définition native de chaque image assemblée. On peut ainsi naviguer du vestige de quelques centimètres à l'ensemble du site avec la même image.

Orthomosaïque : processus théorique de production

Nous présentons ici les grandes étapes du processus photogrammétrique pour produire une orthomosaïque (fig. 56). L'ordre de ces étapes peut varier et dépend du logiciel de traitement utilisé. La majorité des informations ici présentées sont synthétisées des thèses de L. Girod et M. Samaan, adaptées pour un public d'archéologues (Samaan 2016, p. 29-70 ; Girod 2018, p. 23-41).

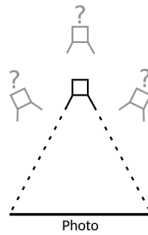
Acquisition des photographies

Sur le principe, tous les éléments visibles sur les images d'un set de photos se retrouveront dans la scène 3D. Mais il faut pour cela respecter quelques notions élémentaires lors de l'acquisition des photos. D'abord, la qualité de la surface et la précision du modèle 3D dépendent directement de la qualité des photos utilisées. Ainsi, une photographie pour un usage photogrammétrique doit répondre à deux conditions, être nette et bien exposée. De plus, les photos doivent recouvrir l'ensemble de la scène et tous les angles de vues souhaités. Enfin, il faut toujours prévoir un recouvrement important entre les photos voisines, pour retrouver les éléments communs qui permettront de les assembler.



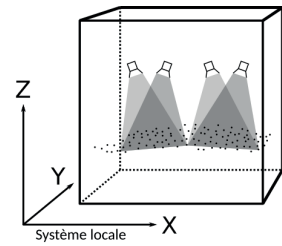
1 - Découverte des points d'attache

SIFT



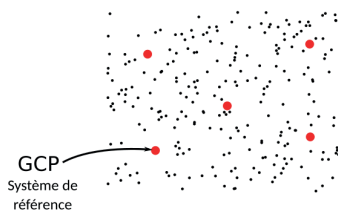
2 - Calibration et orientation des caméras

SfM



3 - Calcul des points homologues (nuage peu dense)

Stéréoscopie



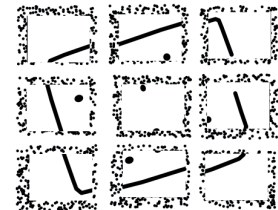
4 - Géolocalisation du nuage de points

Transformation



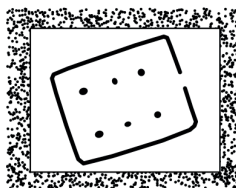
5 - Densification du nuage (nuage dense)

Stéréoscopie



6 - Calcul des orthophotographies

Orthorectification



7 - Composition de l'orthomosaïque

Légende :

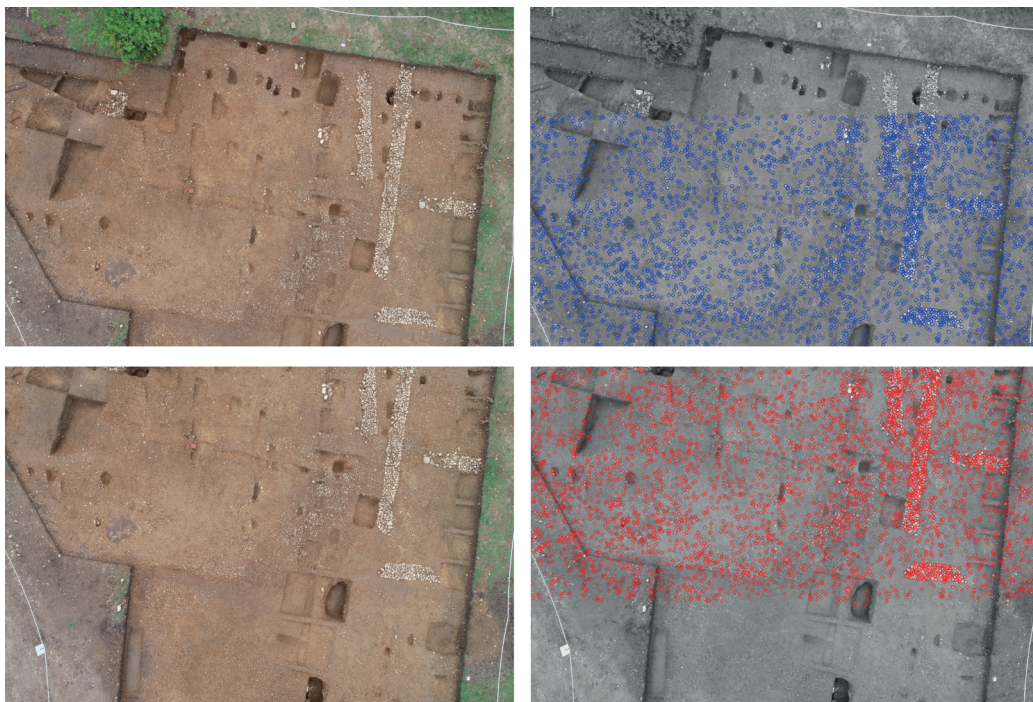
Étape du processus

Méthode

56. Processus théorique de production d'une orthomosaïque.

Points homologues (Tie point)

La première étape logicielle consiste à recomposer le puzzle général de la scène, c'est-à-dire trouver les différentes liaisons des photos les unes par rapport aux autres. Pour les relier, le programme identifie sur chaque image des points communs avec les images voisines (fig. 57). Plusieurs techniques existent pour identifier automatiquement ces points homologues. Une des plus courantes est l'algorithme SIFT, qui discrimine des points entre des couples d'images le plus indépendamment possible de la résolution, du niveau de flou, de l'échelle de rotation, du point de vue et de la luminosité de chaque photo (Lowe 2004). Le processus est appliqué à l'ensemble des images pour déterminer tous les couples et obtenir la vue complète de la scène, projetée dans un système de coordonnées local en 2D (X et Y). Les couples d'images qui présentent les meilleurs scores de corrélation sont conservés comme points homologues.



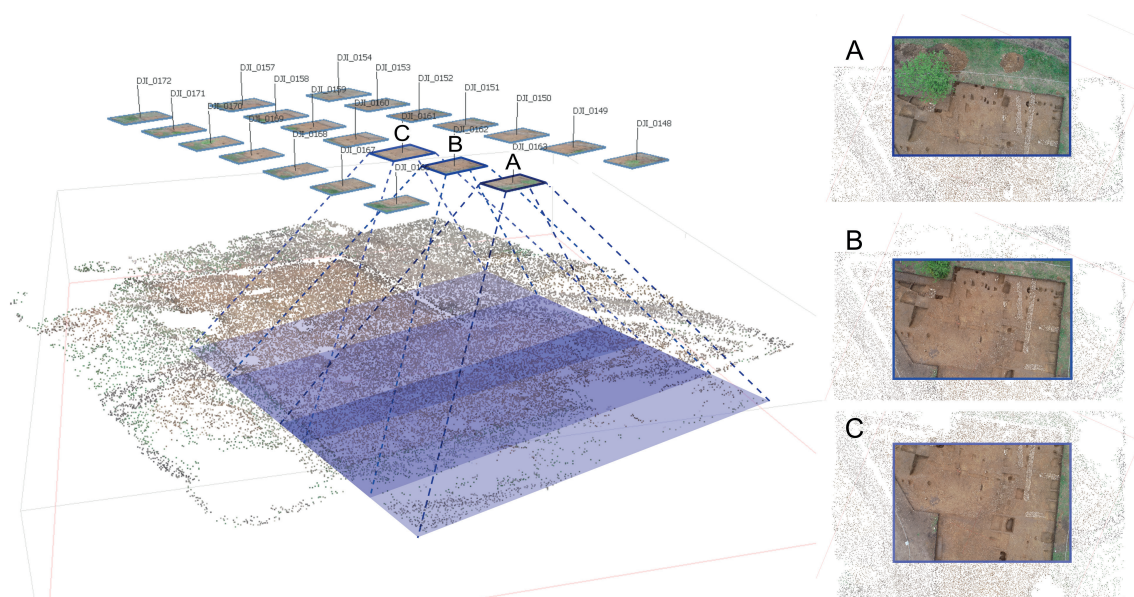
57. Points homologues entre deux images voisines. Bibracte, plateforme PC14, Nièvre, 2021

Calibration et orientation des caméras

Une fois les photos assemblées, il faut estimer la position de chacune d'entre elles pour définir l'espace 3D théorique dans lequel la scène sera projetée. Pour cela, on utilise un algorithme dit "d'ajustement de faisceau". Il détermine d'abord les caractéristiques intrinsèques de l'APN lors de l'acquisition (paramètres d'orientation interne). Puis le calcul détermine l'angle et la distance du capteur par rapport à la photo, c'est l'estimation des positions (l'orientation externe). Cette étape permet de rajouter la dimension Z au système de coordonnées local défini durant l'étape précédente. Si, dans la littérature, on retrouve principalement les termes d'orientation interne et externe, nous avons préféré, afin de limiter les possibilités de confusions, utiliser dans cet article le terme de calibration pour décrire les paramètres d'orientation interne et de limiter l'emploi du terme orientation aux paramètres d'orientation externe (Verhoeven 2016, p. 174 ; Girod 2018, p. 22-23).

Nuage de points épars

Puisque le modèle géométrique des caméras (calibration + orientation) modélisant l'emprise de la scène est maintenant connu, il est possible de projeter des points dans un espace 3D. Pour chaque point (homologue) visible sur chaque paire d'images, la position (coordonnées en X, Y et Z) de ces points est calculée et projetée dans un espace 3D (fig. 58). L'assemblage des points homologues constitue le premier nuage de points visualisable en 3D. S'il présente relativement peu de détails, le nuage dit "épars" sert avant tout de base pour vérifier la qualité de la projection 3D et l'emplacement des caméras.



58. Projection des points homologues dans un espace 3D. Bibracte, plateforme PC14, Nièvre, 2021

Géolocalisation

À ce stade, le nuage de points est toujours projeté dans un système de coordonnées local. C'est-à-dire que la géométrie qui le définit est celle tirée des caractéristiques des images, où l'unité de base est le pixel. Pour basculer dans un Système de Coordonnées de Référence (SCR), tel que le RGF93 en France, il faut renseigner les coordonnées géographiques de certains pixels. Dans la pratique, pour obtenir un modèle précis, cela se traduit par l'installation avant acquisition d'au minimum trois points de contrôle au sol (*Ground Control Point* ou GCP) sur la scène photographiée, chacun visible sur au minimum deux photos. Ils sont relevés sur le terrain dans le SCR de sortie souhaitée. Il faut ensuite les identifier sur les images pour que le logiciel transforme la position du nuage de points dans un espace géoréférencé 3D.

Ajustement de la calibration et de l'orientation des caméras

Pour obtenir la géométrie finale du modèle, une dernière étape consiste à affiner la calibration et l'orientation des caméras en fonction du SCR défini, pour améliorer la précision du nuage de points.

C'est une étape dite de compensation, correspondant à un nouvel ajustement de faisceaux, et prenant en compte les contraintes spatiales induites par les points d'appuis (GCP). In fine cela permet d'améliorer la précision du nuage. Mais on peut également dire que cette étape ajuste ou affine la géométrie de la scène (camera et nuage épars) en prenant le SCR comme contrainte.

Densification

Avec le nuage et les paramètres des caméras validés et optimisés, la construction du nuage de points dense peut commencer. Cette étape, dite de corrélation multi-stéréoscopique dense (*3D dense matching*), est basée sur la génération de cartes de profondeurs permettant de densifier le nuage selon un ensemble de résolutions allant en s'affinant (Samaan 2016, p. 39). Cette méthode, multi-échelle/multi-résolution, conduit d'une part à une meilleure robustesse dans le calcul de corrélation entre images et d'autre part à plus de rapidité.

La densification du nuage augmente le nombre de points proportionnellement à la résolution des images sources. La concentration est parfois telle qu'elle peut donner une impression de surface physique au nuage. Toutefois cette étape ne signifie pas nécessairement une amélioration de la précision du modèle.

Orthorectification

Chaque point du nuage est alors associé à des coordonnées X, Y et Z précises, qu'il est possible de projeter sur une grille. L'image n'est donc pas la photo initiale déformée, mais bien le résultat d'un nouveau calcul où chaque pixel est une case de cette grille. On supprime ainsi les distorsions optiques pour obtenir des orthophotographies.

Orthomosaïque

Il ne suffit pas de coller côte à côte les orthophotographies pour composer une orthomosaïque, sans quoi des limites grossières seraient visibles entre chaque image. La dernière étape consiste donc à fusionner les orthophotographies par un calcul d'égalisation radiométrique destiné à homogénéiser au maximum les couleurs entre les images voisines.

3.3.2. Choix et réglages des équipements

Pour réaliser une bonne acquisition, le plus important est de bien identifier l'objet d'étude et de définir en amont les besoins pour le rendu final. En effet, en photogrammétrie, l'équipement, les paramètres d'acquisitions et les propriétés de calcul varient en fonction du résultat souhaité. Nous proposons donc ici des recommandations génériques pour produire l'orthomosaïque d'un site archéologique, qu'il convient d'adapter en fonction des problématiques et du contexte d'intervention.

Matériel photographique et topographique

Un des avantages du relevé photogrammétrique réside dans le fait que le matériel nécessaire pour réaliser une campagne d'acquisition fait partie des équipements classiques d'un chantier de fouille archéologique.

Capteur

Le capteur est la partie sensible d'un APN, composée de millions de photosites qui reçoivent et transforment la lumière en autant de signaux électriques analogiques, selon l'intensité lumineuse perçue. Ces signaux sont ensuite convertis en pixels qui recomposent l'image. Sans entrer dans les détails, les progrès technologiques réalisés sur les capteurs type CMOS ont participé à une baisse importante des prix des APN à partir de 2005, les rendant plus accessibles sur le marché (Verhoeven 2016, p. 145). Les constructeurs ont ainsi continué les développements, et proposent aujourd'hui des appareils haut de gamme qui répondent aux

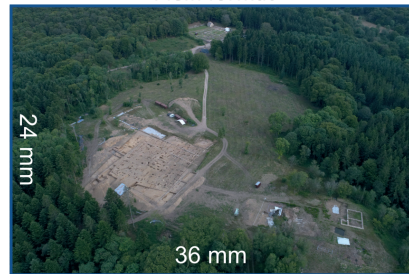
exigences techniques de la photogrammétrie. Le premier aspect à prendre en compte est la taille du capteur, et par extension la taille des photosites (*pixel size*). Logiquement, plus la surface photosensible est importante, plus elle capte de lumière sur un temps donné, ce qui influe directement sur la qualité de l'image capturée (Granier *et al.* 2019, p. 62-63). Ainsi, les grands capteurs offrent plus de souplesse pour les acquisitions, notamment dans les zones de basse luminosité. Parmi les tailles standards proposées par les constructeurs, le capteur "plein format" ou "35 mm", qui calque ses dimensions sur celle des pellicules argentiques 24 x 36 mm, est donc le plus adapté pour la photogrammétrie (fig. 59). Toutefois, un capteur plus petit peut être employé quand des contraintes de poids et de taille de l'APN existent, notamment dans le cas de la photogrammétrie aéroportée par drone. Par exemple, avec de bonnes conditions de luminosité, la caméra du *Dji Phantom 4 Pro*, avec un capteur d'un pouce (13,2 x 8,8 mm), présente une qualité suffisante pour la production d'orthomosaique de site archéologique. Une autre caractéristique technique importante est le nombre de photosites qui composent le capteur. Il correspond au niveau de détail de l'image finale, car un photosite sur le capteur est normalement égal à un pixel sur l'image, si l'APN est utilisé à sa définition maximum.

Cette valeur est généralement retranscrite en mégapixel (MP ou Mpx), équivalent à 1 million de pixels. La taille et le nombre de photosites sont donc intrinsèquement liés, car pour deux capteurs de 20 Mpx, la surface sensible de chaque photosite sera forcément plus restreinte pour un capteur APS-C (21,1 x 16,7 mm) que sur un capteur Plein format (24 x 36 mm). Ce critère peut être évalué par l'espacement en micro mètre entre le centre de deux photosites adjacents (*pixel pitch*) ou par la densité de photosites au micro mètre carré (*pixel density*). Lorsque le choix est possible, on privilégiera toujours la taille du capteur, pour une meilleure réception de la lumière, aux mégapixels, sans descendre sous la barre des 16 Mpx (Granier *et al.* 2019, p. 62-63).

Objectif photographique

Dans le choix de l'équipement de photographie, la qualité de l'objectif est primordiale. Un boîtier haut de gamme avec un objectif d'entrée de gamme livre des résultats clairement insatisfaisants (Historic England 2017, p. 28-29). Dans tous les cas, le choix de l'objectif devra être priorisé avant celui du boîtier. Ce choix détermine notamment la longueur ou distance focale (souvent raccourcie en focale) qui correspond à la distance (en millimètres) entre le centre optique de l'objectif et le capteur du boîtier quand la mise au point est faite à l'infini (De Ferrières *et al.* 2015). Elle détermine l'angle de vue de l'image, en fonction de

Plein format



Canon EOS 1D X, 5D, 6D,
Nikon D3, D4, D5, D6, D600, D700, D800
Sony α 7, α 9

APS-H (Canon)



Canon EOS 1D

APS-C / DX (Nikon)



Nikon D500, D3000, D5000, D7000

APS-C (Canon)



Canon EOS 7D, 60D, 50D

4/3'



Panasonic Lumix G

1'



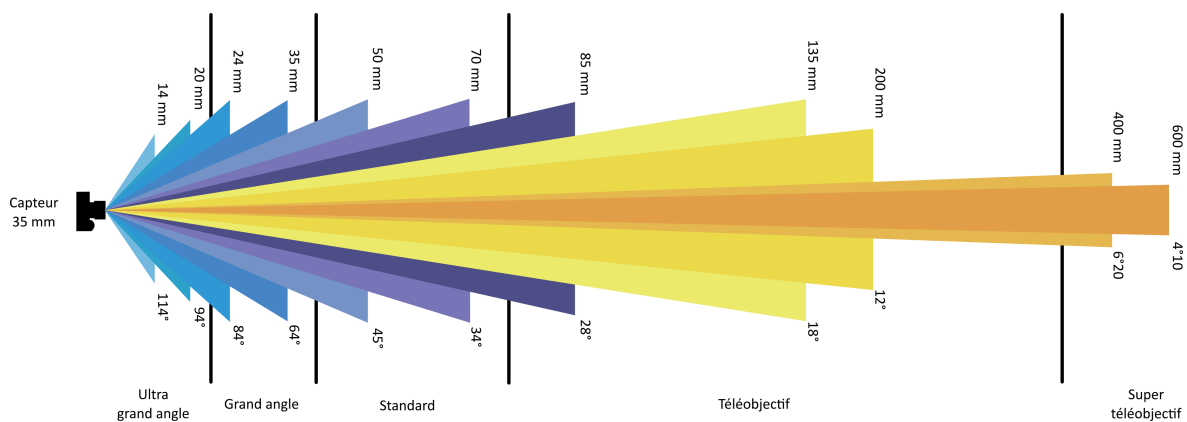
Dji Phantom 4 Pro

1/2.3'



GoPro

la taille du capteur. Plus la distance focale est élevée, plus l'objectif "voit" loin, mais avec un cadrage serré. À l'inverse, une distance focale courte permet un cadrage large, mais nécessite d'être plus proche de la scène à capturer (fig. 60). La manière d'exprimer la distance focale est issue des normes définies initialement pour les boîtiers argentiques, pérennisée aujourd'hui au travers des capteurs plein format. Avec ce type de capteur, les objectifs affichent des distances focales de 15 mm (fish-eye, ultra grand angle) à plus de 300 mm (téléobjectif). Mais à distance focale identique, un capteur plus petit cadre plus serré, avec un angle de vue plus court. Ainsi, pour obtenir le même angle de vue qu'avec un capteur plein format, il faudra choisir un objectif d'une longueur focale nominale plus petite ou égale si elle est exprimée en "équivalent 35 mm" ou "équivalent plein format". Selon les préconisations du consortium 3D SHS, il est recommandé d'utiliser en photogrammétrie une plage focale située entre 20 mm et 85 mm en équivalent plein format soit entre 14 mm et 55 mm pour un capteur de type APS-C (Granier *et al.* 2019, p. 64).



60. Schéma de relation entre l'angle de champ d'une photographie et la longueur focale de l'objectif utilisé (au format 24 x 36 mm). (d'après Pamart *et al.* 2019)

Plusieurs types d'objectifs existent, les plus couramment utilisés étant les objectifs zoom et les objectifs à focale fixe. Les objectifs zoom sont les plus polyvalents, car comme leur nom l'indique, il s'agit d'objectif à distance focale modulable, pour adapter l'angle de vue, ou le niveau d'agrandissement de la scène. Les objectifs à focale fixe, à l'inverse, sont verrouillés sur une distance focale. Il en résulte des objectifs généralement mieux optimisés et

de moindre coût, dont l'utilisation est à prioriser absolument pour une acquisition photogrammétrique. Par expérience, les objectifs à focale fixe 35 mm équivalent plein format se révèlent particulièrement adaptés pour le relevé de structures archéologiques.

Précisions également que lors de traitement photogrammétrique, les logiciels effectuent une nouvelle calibration (étape de calcul) à chaque changement de distance focale, car cela modifie les paramètres internes de la caméra. Pour éviter de multiplier les opérations de calibration, nous conseillons donc d'utiliser soit un objectif à focale fixe, soit de verrouiller la focale de l'objectif durant l'acquisition.

Type de boîtier

Les APN sont classés selon plusieurs catégories (fig. 49). Les reflex (de l'anglais *Digital Single-Lens Reflex*) sont généralement les plus adaptés pour la photogrammétrie. La spécificité de ces appareils est le système de visée, qui se fait à l'œil nu par un jeu d'optiques, comme dans des jumelles. L'objectif est d'obtenir un rendu fidèle de l'image à capturer, sans conversion numérique sur un écran. Un miroir est donc positionné devant le capteur le temps de la visée, et se déplace au moment de l'acquisition. Si le miroir ne présente pas d'intérêt en photogrammétrie, les APN reflex sont aussi par tradition des APN haut de gamme avec des capteurs plein format, des objectifs interchangeables et la possibilité de régler manuellement un grand nombre de paramètres d'acquisition (Samaan 2016, p. 33). Aussi, les progrès récents sur les APN hybrides les rendent tout aussi pertinents pour la photogrammétrie. Cette catégorie associe les attributs techniques des APN compacts et reflex : un boîtier de taille réduite combiné à des objectifs interchangeables. La réduction de la dimension du boîtier est rendue possible par l'absence de miroir devant le capteur et en direction du viseur. L'image captée est retranscrite sur le viseur ou sur l'écran après conversion numérique. Depuis quelques années, plusieurs de ces APN embarquent des capteurs plein format. Plus légers et moins encombrants que les APN reflex, ils se révèlent particulièrement adaptés pour les missions de terrain, et constituent aujourd'hui la référence en termes de photographie numérique. Enfin, une carte mémoire avec un espace suffisant, à adapter en fonction des spécifications de l'équipement utilisé, est évidemment nécessaire.

Acquisition déportée

Avec du matériel plus spécifique, il est possible de surélever la caméra pendant l'acquisition pour augmenter significativement le champ de vision de chaque image. Deux solutions existent pour cela : la perche et le drone. Pour les perches, nous recommandons une perche de 4 à 10 m de long. En dessous, le gain d'angle de vue n'est pas vraiment significatif et au-delà les perches deviennent difficilement maniables. Pour la gestion des prises de vue, de nos jours la plupart des boîtiers hybride et reflex proposent des fonctions Bluetooth ou Wifi associées à des applications sur smartphone, très pratiques pour avoir un retour d'image et un déclenchement à distance. Il est également de plus en plus courant de faire intervenir des télépilotes de drone sur les fouilles archéologiques. Nous n'entrerons pas dans les détails ici, mais il faut savoir que l'utilisation de ces outils est très réglementée par les législations française et européenne⁸⁸. En tenant compte du cadre légal et correctement paramétré, un drone présente l'avantage de réaliser une campagne d'acquisition de manière quasi automatique sans impacter le terrain. Les appareils sont souvent des solutions intégrées (aéronef + capteur) et, à ce jour, seuls les drones semi-professionnels à professionnels présentent une taille et une définition de capteur suffisante pour réaliser des orthomosaïques de qualité (ex : Dji phantom 4 pro). Chaque constructeur accompagne ses machines de solutions logicielles pour préparer et réaliser les vols (Granier *et al.* 2019, p. 70). Un ordinateur ou une tablette sont donc généralement nécessaires sur le terrain pour gérer le bon déroulement de la mission. Enfin, attention au contexte d'intervention avec ces deux outils. Si l'espace est généralement bien dégagé au niveau du sol sur un chantier archéologique, ce n'est pas toujours le cas à plusieurs mètres de hauteur (arbres, lampadaires, bâtiments, etc.).

Topographie

Pour le géoréférencement, des cibles sont requises pour représenter les GCP. En théorie, n'importe quel objet ou surface peut faire office de GCP, tant qu'il est bien visible et individualisé sur les images. Il est courant d'utiliser des cibles papier numérotées et plastifiées, maintenues par des clous. Pour relever la position des GCP, il faut être équipé d'au moins un des instruments classiques de relevé topographique (tachéomètre, station totale, GPS, etc.). Naturellement, plus la précision de l'instrument employé est bonne, plus le positionnement de l'orthomosaïque sera exact.

88 <https://www.ecologie.gouv.fr/politiques/drones-uas>

Photographie

Contraintes photographiques

Régularité

Pour éviter de créer des erreurs ou des aberrations sur le modèle 3D, il est nécessaire de conserver les mêmes paramètres d'acquisition sur l'ensemble du lot de photos. Les réglages doivent donc être effectués avant de lancer une mission d'acquisition, en cherchant à obtenir des photos nettes et suffisamment lumineuses en tous points du chantier. Les parties les plus sombres et les plus irrégulières sont souvent à privilégier pour étalonner la caméra, tout en contrôlant que les paramètres fonctionnent aussi dans les zones lumineuses. Si une exposition équilibrée ne peut pas être trouvée, il est préférable de sous-exposer les images. En effet, la sous-exposition peut être corrigée à l'aide de logiciels de traitement d'image, comme Darktable, mais pas la sur-exposition. Attention toutefois dans ce cas à bien produire et utiliser le format brut des images.

Sensibilité

L'ISO contrôle la sensibilité du capteur de l'APN à la lumière. Une valeur haute augmente avec la sensibilité et rend le capteur plus réceptif au signal lumineux. Mais elle a aussi pour conséquence de faire apparaître des particules lumineuses sur la photo, comme du "bruit" numérique. Ce phénomène est problématique en photogrammétrie, car les particules créent des artefacts uniques d'une image à l'autre, ce qui perturbe toutes les étapes du processus photogrammétrique. Il convient donc de garder l'ISO au minimum (max. 400), en jouant en priorité sur l'ouverture et la vitesse d'obturation pour régler la luminosité de l'image.

Ouverture

Les détails flous d'une photo ne peuvent pas être modélisés. Or, un objectif ne permet, en théorie, de ne faire la mise au point (focus) que sur un seul élément de la scène. La profondeur de champ détermine alors la netteté des éléments placés à l'avant et à l'arrière de la cible de mise au point (Historic England 2017, p. 24-25) (fig. 61). Il faut donc rechercher la plus grande profondeur de champ possible, pour que tous les plans à relever apparaissent nets sur l'image. Dans la pratique, cela revient à refermer le diaphragme de l'objectif. Mais une ouverture trop faible laisse passer moins de lumière jusqu'au capteur. Le but est alors de trouver le bon équilibre pour produire des images nettes et suffisamment lumineuses. La

valeur de l'ouverture, notée f/n , correspond au facteur de division de la distance focale. Plus le facteur est petit, plus l'ouverture est grande. Prenons l'exemple d'un boîtier plein format (capteur 24 x 36mm) avec un objectif d'une distance focale de 35 mm :

- Si le terrain présente peu de reliefs (terrain plat, structures en creux peu profondes, élévations de deux ou trois assises, etc.), selon les conditions de luminosité, il est possible d'ouvrir à maximum $f/4$ ($35/4 = 8,75$ mm), ce qui correspond à un diamètre d'ouverture du diaphragme de 8,75 mm. Cette configuration est lumineuse, mais induit du flou sur les arrières plans, ce qui la rend peu adaptée aux environnements accidentés ;
- Par contre, si le terrain est irrégulier (site en pente, plusieurs paliers topographiques, sondages profonds, éléments verticaux, etc.), il est préférable de ne pas descendre en dessous de $f/5.6$ ($35/5,6 = 6,25$ mm). Avec un diamètre d'ouverture réduit à 6,25 mm, on réduit la luminosité des clichés, mais on augmente la netteté des arrières plans.

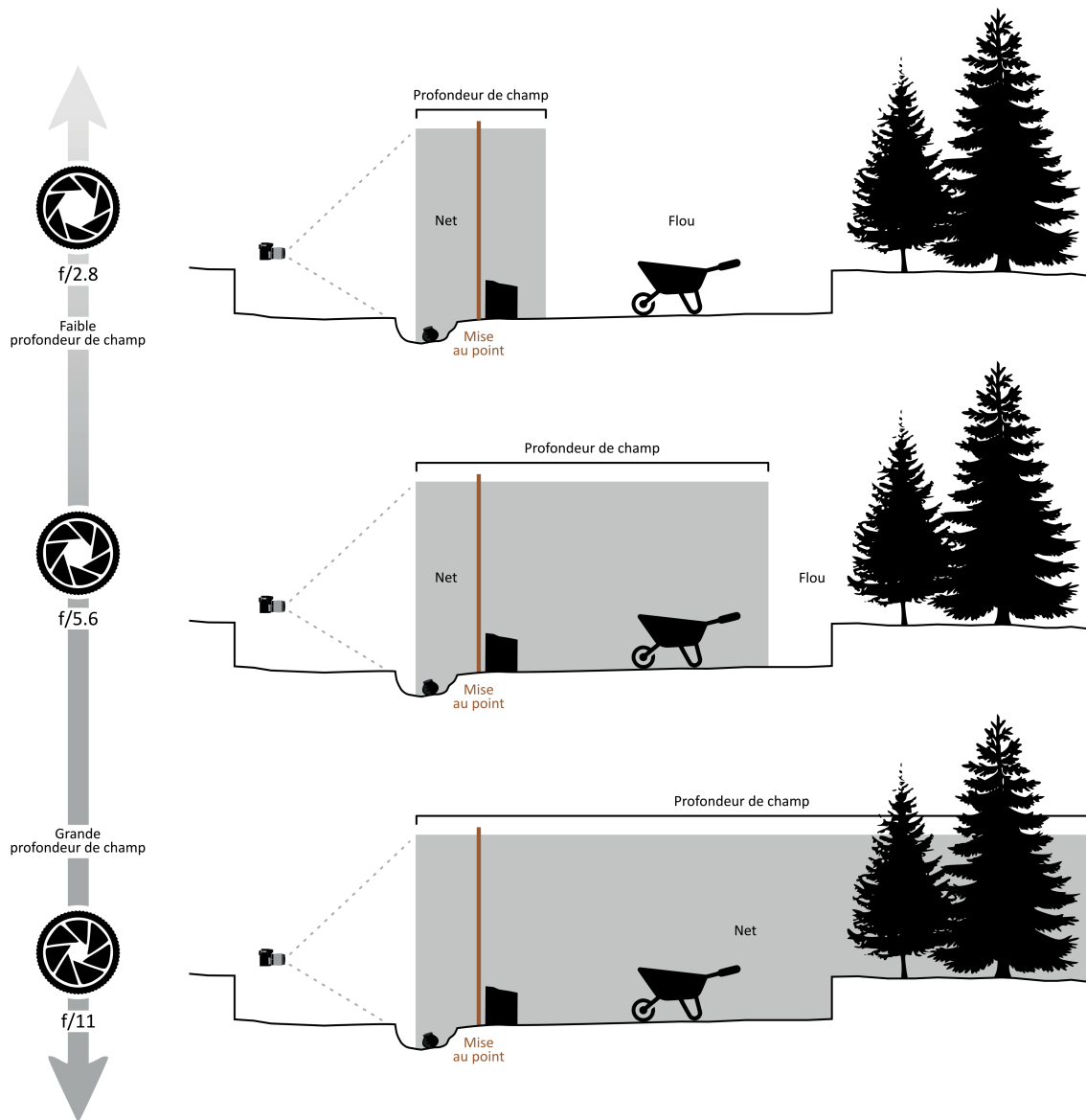
Ces exemples sont adaptés à des prises de vue manuelles nécessitant une vitesse d'obturation élevée (*cf.* ci-dessous). S'il est possible d'utiliser un trépied associé à un déclencheur à distance pour augmenter la stabilité des acquisitions, alors il est recommandé d'utiliser la plus petite ouverture possible ($f/8$ à $f/11$), afin de maximiser la netteté des images sur la totalité de leur profondeur de champ.

Vitesse d'obturation

La vitesse d'obturation représente le temps durant lequel le capteur de l'APN reste exposé pour capter la lumière. Elle se mesure en fraction de seconde. Un capteur exposé longtemps capte beaucoup de lumière. Mais il enregistre aussi tous les mouvements de l'appareil pendant cette période, ce qui crée du flou. Il faut donc privilégier une vitesse rapide pour figer une image nette, tout en captant assez de luminosité. C'est pourquoi il est préférable de ne pas descendre en dessous du centième de seconde ($1/100$) avec un APN maintenu à bout de bras ou par une perche.

Balances des blancs

La balance des blancs doit être identique sur toutes les images d'une même acquisition. Il est recommandé de sélectionner celle qui correspond le plus à l'éclairage ambiant de la scène à relever, parmi les réglages préétablis de votre appareil (nuageux, ombragé, lumière du jour, artificiel, etc.). Bien entendu le mode automatique est à proscrire (C3DC 2011, p. 14).



61. Schéma de la relation entre profondeur de champ et nombre d'ouverture

Format de fichiers

Il est de toutes circonstances, en photographie, conseillé de procéder à un enregistrement en double format, JPEG et RAW (format brut du constructeur du boîtier). Ce dernier permet, si besoin, de compenser des défauts d'exposition, sans perte, lors du développement numérique des données brutes via des logiciels spécialisés comme darktable. Pour les traitements, le format JPEG en haute définition est préférable. Le format TIFF, trop lourd, est à proscrire des acquisitions photogrammétriques.

Contraintes photogrammétriques

Recouvrement

Nous recommandons un taux de recouvrement entre les clichés de 70 % minimum. C'est-à-dire que pour deux photos prises à la suite, seul un tiers de la deuxième photo doit illustrer des éléments inédits. Ce paramètre est valable pour les clichés d'une même ligne d'acquisition ainsi qu'entre les photos de lignes adjacentes. En dessous de 70 %, il existe un risque pour que les photos voisines ne présentent pas assez d'éléments similaires pour s'aligner correctement.

Nombre de photographies

Le nombre de photos dépend de la surface à relever et de la résolution au sol visée. Pour être restitué en 3D, un point doit être vu sur un minimum de deux photos, et pour être reconstruit avec plus de précision, il doit être vu sur trois images. En pratique, nous conseillons de prendre quatre clichés de la scène au minimum. Si moins de quatre photos sont nécessaires, c'est probablement que la surface est réduite ou le niveau de détails requis faible. Dans ce cas, un relevé manuel sera sans doute plus rapide à mettre en œuvre, pour un enregistrement à un niveau de détails identique au relevé photogrammétrique. À l'inverse, si l'acquisition dépasse les 100 photos, c'est en général soit parce que la surface à relever est très importante (plusieurs hectares), soit parce que la résolution au sol dépasse largement les besoins réels.

Lot de calibration (optionnel)

Pour l'essentiel des acquisitions photogrammétriques réalisées sur les chantiers archéologiques, le lot de photos initial est suffisant pour réaliser les étapes de calibration et d'orientation des caméras. Il s'agit en effet de scènes simples, sur un même plan (une surface générale plane ou une façade), requérant un nombre raisonnable de photos (moins de 50)

pour la numérisation. Dans le cas de scènes plus complexes, qui présentent une grande profondeur et de nombreux plans, il est nécessaire de multiplier les photos et les orientations pour assurer une couverture efficace du sujet. Or le manque de convergence sur de trop nombreuses prises de vue peut entraîner des erreurs de calcul de calibration de la caméra par les logiciels. Nous rappelons toutefois au lecteur que ce type d'acquisition est généralement destiné à la production de modèle 3D, alors que cette publication se concentre sur la production d'orthomosaiques.

MicMac, le logiciel que nous proposons d'utiliser pour le traitement photogrammétrique, permet d'assurer une calibration de caméra (orientation interne) robuste en la calculant à partir d'une dizaine de prises de vue spécifiques avant d'en appliquer le résultat à l'ensemble des images de l'acquisition (Pierrot-Deseilligny *et al.* 2022, p. 53-54). Ce lot de calibration doit répondre aux caractéristiques suivantes :

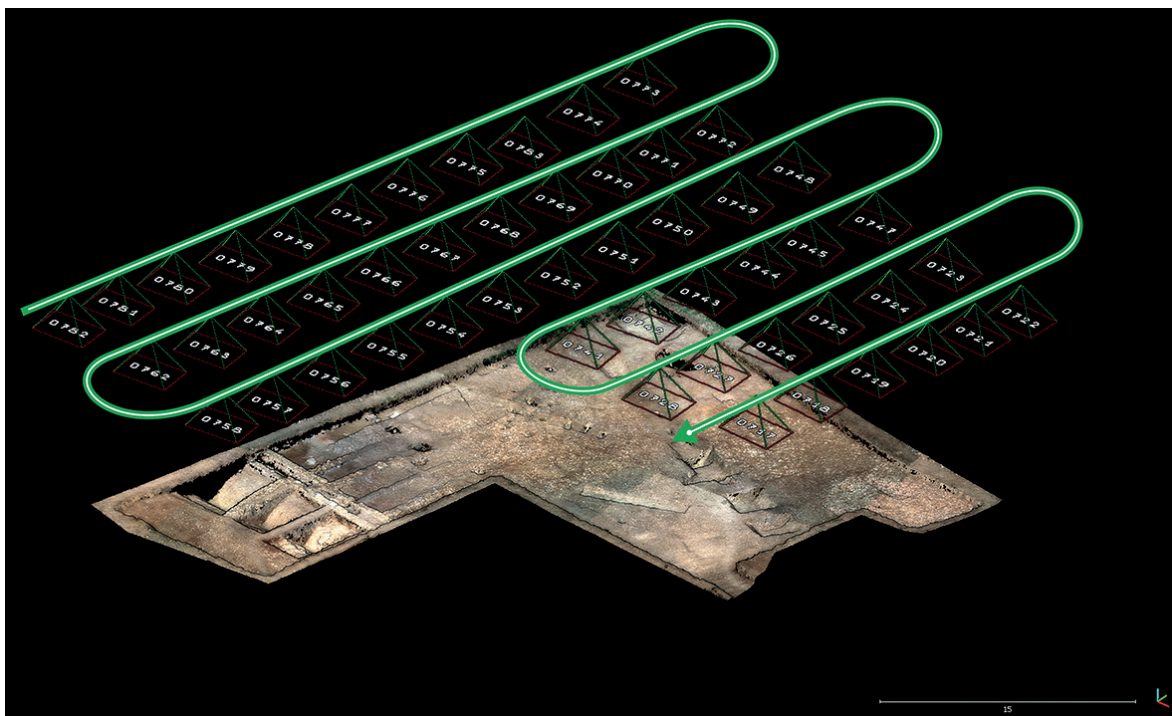
Les images doivent toutes converger vers le même point, mais en étant prises de positions et d'angles de vue différents (différences d'altitude et d'angle de rotation de la caméra), comme si l'appareil se déplaçait sur une sphère pour photographier son centre. Il doit conserver une distance à la surface photographiée toujours identique, en "tournant autour" de la portion, pour garder une image nette et bien exposée sans modifier les paramètres de distance focale ou d'ouverture de la caméra ;

- Les photos de calibration doivent évidemment être produites avec les mêmes paramètres d'acquisition que le reste des photos, et respecter les prérequis de la photogrammétrie (des photos nettes et bien exposées) ;
- Toutes les images couvrent la même portion de la scène. La surface de cette portion n'a pas besoin d'être très large, 2 m² sont suffisants ;
- Cette portion doit présenter une profondeur suffisante, avec un relief d'au minimum 20 cm. Par exemple, une portion parfaitement adaptée peut être la jonction de deux murs perpendiculaires, car elle illustre trois plans (les deux façades et le sol) ;
- Éviter les éléments trop complexes, avec beaucoup de petits détails sur des plans différents (ex : arbres), qui complexifient grandement la recherche de points homologues. Pour les mêmes raisons, les surfaces transparentes (ex : eau, vitre) ou monochromes et lisses (ex : bâche noire ou blanche) sont également à proscrire.

Contraintes contextuelles

Cheminement

Pour faciliter la recherche des couples d'images stéréoscopiques, il est conseillé de prévoir un cheminement régulier durant l'acquisition. Le plus souvent, la caméra se déplace sur un axe parallèle à la longueur du chantier, en commençant par un angle de la fouille. Une fois la ligne terminée, la caméra se déplace de quelques mètres à la perpendiculaire avant de reprendre une nouvelle ligne parallèle au premier axe (fig. 62). Ce cheminement en "S" est à répéter autant de fois que nécessaire pour couvrir la zone à modéliser. Il est aussi important de couvrir toujours un peu plus de surface que la réelle zone d'intérêt, pour éviter les manques d'information ou les erreurs géométriques sur les bords.



62. Principe d'un cheminement en "S", recommandé pour les acquisitions aéroportées (DAO. M. Thivet)

Luminosité

Le premier critère à prendre en compte lors d'une campagne d'acquisition est la luminosité ambiante du terrain. Comme sur des photos de terrains classiques, les zones d'ombre sont à proscrire, car les surfaces de pixels noirs ne présentent pas assez d'éléments remarquables pour identifier les points d'attache. Il en va de même pour une luminosité changeante durant l'acquisition, comme par une journée avec des éclaircies ponctuelles. Là encore, la recherche de points d'attache va être limitée si une même surface présente des luminosités différentes selon les photos. La luminosité optimale est donc la même qu'habituellement sur un chantier archéologique : une lumière indirecte, diffuse et homogène sur l'ensemble de la surface à acquérir, tel qu'avec un ciel couvert par temps sec.

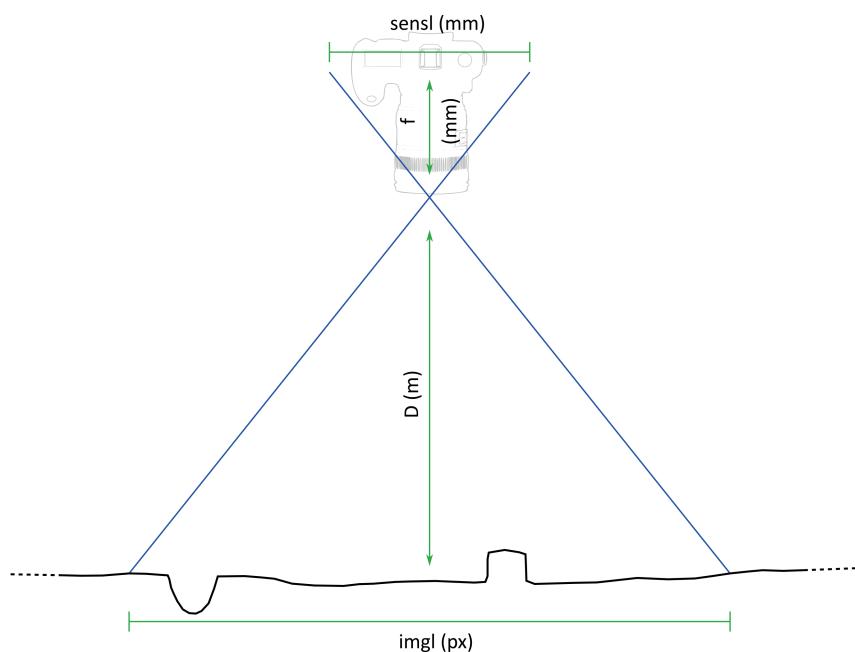
Résolution au sol

La résolution au sol est une distance réelle au sol représentée par la distance entre les centres des pixels adjacents dans une image.

La résolution au sol ou résolution spatiale (en anglais *Ground Sample Distance* ou GSD) est une valeur calculée qui vise à représenter la distance physique entre deux pixels adjacents dans une image par son équivalent en distance réelle au sol. Par exemple, pour une image qui affiche une résolution au sol de 5 cm, la distance entre les centres des pixels correspond à 5 cm réels sur le terrain. Les pixels étant carrés, cela correspond à une résolution au sol de 5 cm/px, soit une emprise de 25 cm²/px. De ce fait, les éléments en dessous de 5 cm ne seront pas identifiables sur l'image. En toute logique, plus la valeur de la résolution au sol est élevée, plus les pixels sont larges et le niveau de détails de l'image diminue. La résolution au sol influe donc directement sur la lisibilité d'une orthomosaïque et doit être estimée avant le début de la mission. Pour les photographies perpendiculaires au plan de la scène photographiée, ce facteur évolue en effet en fonction de la taille des photosites sur le capteur (*pixel size*), de la distance focale utilisée et de la distance du capteur par rapport à la surface à numériser. Dans la pratique, il est difficile de connaître la taille exacte des photosites du capteur. Aussi, il est courant d'utiliser une valeur équivalente, qui peut être obtenue par le rapport entre la largeur du capteur et la définition maximale de l'image.

Pour calculer la résolution au sol en centimètres, la formule est $(D \times \text{sensl} \times 100) / (f \times \text{imgl})$ (Neumann 2009), ce qui correspond à (fig. 63) :

- D = distance de la caméra à la surface numérisée, exprimé en mètres ;
- sensl = taille du capteur : largeur du capteur du boîtier de l'APN, exprimée en millimètres ;
- f = distance focale, exprimée en millimètres ;
- imgl = largeur de l'image, exprimée en pixels (la définition).



63. Calcul de la résolution au sol

Il est important de souligner que la recherche de la plus faible résolution au sol est lourde de conséquence pour lors des étapes de traitement photogrammétrique. En effet, plusieurs possibilités existent pour diminuer la résolution au sol d'une image : allonger la distance focale ou rapprocher l'appareil du sujet. Dans les deux cas, cela revient à diminuer l'angle de vue de la caméra, et donc multiplier les photos à l'acquisition pour couvrir une zone identique. Or augmenter le nombre de photos impacte lourdement la puissance de calcul nécessaire pour le traitement photogrammétrique par les logiciels. Selon le nombre d'images et les ressources informatiques à disposition, le temps de calcul peut varier de quelques di-

zaines de minutes à plusieurs heures, voire ne pas fonctionner du tout. Il est donc primordial de définir en amont de l'acquisition la résolution au sol souhaitée, afin d'optimiser le ratio précision/temps/objectif scientifique de l'acquisition. Par exemple, un relevé d'inhumation justifie, par un besoin de détail élevé, une résolution au sol basse (millimétrique : 0.1 cm/px) puisqu'elle ne concerne qu'une surface très restreinte. À l'inverse, le levé de la totalité d'une nécropole devra être réalisé à une résolution au sol beaucoup plus élevée (entre 1 et 5cm/px selon la surface) et adapté à la localisation des structures funéraires, à la manière d'un plan d'ensemble.

Topographie

Installation des GCP

Au minimum, 5 GCP suffisent, placer comme suit (Girod 2018, p. 31-32)(fig. 64) :

- 4 GCP aux angles qui délimitent la zone réelle à modéliser (et non pas la surface complète photographiée) ;
- 1 GCP au centre de la zone ;
- D'autres cibles peuvent être installées à l'intérieur de la zone d'intérêt pour renseigner les paliers topographiques ;
- Quelle que soit la taille de la surface, il est déconseillé de dépasser une dizaine de GCP. Cela n'améliore pas le positionnement du modèle et ralentit l'étape de saisie des points lors du traitement.

Paramètres de levé topographique

Pour faciliter l'exploitation des données, il est recommandé d'utiliser le SCR de l'aire géographique du chantier. Par exemple, pour la France métropolitaine, il s'agit du RGF93/Lambert-93 (EPSG 2154). La précision du levé dépend essentiellement de l'équipement utilisé (antenne GNSS, tachéomètre) et des conditions d'acquisitions (nombre de satellite, ouverture du milieu, paramètre humain, etc.). Attention, une orthomosaïque avec une haute résolution au sol ne signifie pas une position précise dans l'espace réel. En effet, rien n'exclut dans le processus de traitement exploité ci-dessous la production d'une orthomosaïque sans localisation ou avec une marge d'erreur importante entre chaque GCP.

Export

Pour l'export, il est nécessaire d'utiliser un fichier au format CSV qui liste les noms des points et leurs coordonnées X, Y et Z.



64. Placement optimal des points de contrôle au sol. Couan, Saint-Aubin-des-Chaumes, Nièvre, 2021

Traitement pour la production d'orthophotographies en plan avec le logiciel MicMac

Pour le traitement, nous avons fait le choix d'utiliser un logiciel libre, afin que ce protocole soit reproductible par le plus grand nombre. Plusieurs logiciels de photogrammétrie libres sont disponibles (Open-MVG, Visual SfM, COLMAP, etc.), mais étant orientés sur le traitement d'informations graphiques plutôt que vers les sciences géographiques, ils s'arrêtent pour la plupart à l'étape de la construction du nuage dense (Bartoš *et al.* 2014). Or, notre objectif est d'exploiter le modèle 3D pour produire une orthomosaique de site archéologique. Le logiciel MicMac présente plusieurs avantages. Il est tout d'abord libre de droit et gratuit, mais il est surtout l'aboutissement de nombreux travaux scientifiques portés par

une vaste communauté nationale de chercheurs en photogrammétrie. La rigueur scientifique de ces travaux et des données produites par le logiciel justifie notre choix de son utilisation comme d'autres chercheurs avant nous en Europe (voir notamment (Guimarães *et al.* 2020 ; La Salandra *et al.* 2021)).

MicMac (ou "Multi Image Correspondance, Méthodes Automatiques de Corrélacion") est un logiciel de photogrammétrie complet qui offre de nombreuses possibilités de création de modèles 3D et d'orthoimages multi-échelles. Son architecture repose sur un ensemble d'outils, des "briques" autonomes, mais interdépendantes, diffusés sous licence libre (CE-CILL-B). L'auteur historique, M. Pierrot Deseilligny, a commencé le développement du logiciel en 2003 à l'IGN. Mais depuis, le projet s'est largement diffusé et recense maintenant une douzaine de partenaires nationaux et européens (Rupnik *et al.* 2017). La documentation officielle est centralisée sur un *wiki* géré par l'ENG⁸⁹. Comme la solution est libre, son développement profite de multiples contributions. L'édition du *wiki* est également ouverte à tous. Cela permet de maintenir à jour une documentation relativement complète, même si la quantité d'informations fournies n'est pas toujours homogène d'une page à l'autre.

Une des particularités de MicMac est qu'il se gère, de base, sans interface graphique (fig. 65). Il se lance avec l'invite de commande sous Windows et le terminal sous Linux et Mac. Les instructions à opérer sont gérées directement en ligne de commande. Elles commencent systématiquement par "mm3d" (pour "MicMac 3D"), pour indiquer à l'ordinateur que la commande qui suit utilise les outils du logiciel. Vous trouverez toutes les informations nécessaires à l'installation sur la page suivante : <https://micmac.ensg.eu/index.php/Install> et dans le tutoriel vidéo suivant : <https://www.youtube.com/watch?v=pMcohhH6aJU>

Nous présentons dans le guide en ligne une chaîne de production complète⁹⁰, tirée des travaux déjà publiés par les auteurs du logiciel et adaptée pour une utilisation dans un cadre archéologique (Pierrot-Deseilligny, Pappadimitis 2006 ; Pierrot-Deseilligny, Clery 2011 ; Rupnik *et al.* 2017 ; Pierrot-Deseilligny *et al.* 2022 ; Verriez *et al.* 2023). Pour chaque commande décrite, il existe plusieurs arguments optionnels. Nous ne présentons que ceux sélectionnés sur la base de 5 ans d'expérimentations et invitons le lecteur à se référer au *wiki* pour plus d'information sur les autres possibilités⁹¹.

89 <https://micmac.ensg.eu/index.php/Accueil> ; consulté le 22/10/2023

90 <https://books.openedition.org/pufc/50723> ; consulté le 22/10/2023

91 https://micmac.ensg.eu/index.php/MicMac_tools ; consulté le 22/10/2023

```

qzv@qv-dell: ~
qzv@qv-dell:~$ mm3d
*****
**                                                                 **
**   MicMac: a free open source project for photogrammetry        **
**   hosted at Ecole Nationale des Sciences Geographiques         **
**   in Marne-la-Vallee, for IGN-France                           **
**                                                                 **
** The project is funded by :                                     **
**                                                                 **
** - Institut National de l'Information Geographique et Forestiere **
**   (IGN main funder since 2003)                                  **
** - the french FUI Project "Culture 3D Cloud" (and MAP-CNRS)     **
** - the french ANR Project "MONUMENTUM"(collaborating with MAP-CNRS)**
**                                                                 **
** Research also currently supported by :                          **
** - CNES (French Space Agency) - via TOSCA Committee (and IPGP)  **
** - Compagnie Nationale du Rhone                                  **
** - Vinci-Construction-Terrassement                             **
** - ERC Advanced Grant A.Kaeab "ICEMASS" (University of Oslo)    **
**                                                                 **
** Current Team: MP Deseilligny, D Jouin, J Belvaux, G Maillet,   **
**   L Girod, E Rupnik, JM Muller, M Daakir, TG Nguyen           **
**                                                                 **
** Contact for participating : Marc.Pierrot-Deseilligny@ensg.eu   **
**                                                                 **
** Hope you enjoy, todo list in case of any problem using MicMac : **
**   (0) Take a Pastis                                             **
**   (1) Switch your computer off and on again                     **
**   (2) Install it on Gnu-Linux (work better on)                  **
**   (3) See the wiki at http://micmac.ensg.eu/ **
**   (4) Contact the forum http://forum-micmac.forumprod.com/ **
**                                                                 **
*****

```

```

mthivet@T7910-P01: ~/pc14
mthivet@T7910-P01:~/pc14$ mm3d Tapioca MulScale ".*JPG" 500 1500

```

65. L'interface du logiciel MicMac, dans le terminal de Ubuntu, et sa première commande, pour lancer l'outil Tapioca

3.3.3. Acquis des expérimentations

En définitif, c'est plus d'une centaine d'orthomosaïques qui ont été produites dans le cadre des nos travaux de recherche doctorale, d'abord avec des logiciels propriétaires comme Agisoft Metashape, puis dernièrement avec MicMac. Le corpus principal est issu de la fouille de la *domus* PC2, où nous avons expérimenté un protocole de levé photogrammétrique systématique de chaque unité stratigraphique, avant et après fouille. L'objectif de cette démarche était double.

En premier, nous voulions éprouver les limites d'un tel enregistrement, afin de déterminer s'il s'agissait d'une méthode effective et efficace pour reconstituer la totalité d'une fouille en 3D. Permettrait-elle aux archéologues de justifier l'intégralité de leurs observations et d'enfin "*se déculpabiliser des méthodes destructrices de la fouille*" (Djindjian 2016, p. 170) ?

Les nombreux tests effectués durant la campagne de 2017 ont rapidement mis en évidence les limites pratiques et logiciels de l'enregistrement 3D. À l'acquisition, un modèle 3D requiert plus de photos qu'une ortho-image, pour couvrir tous les angles de vue du sujet. Le temps dédié aux prises de vue est de fait plus important, particulièrement pour les archéologues ayant l'habitude des clichés face au sujet. De plus, il s'est avéré qu'il n'existait pas de logiciels pour exploiter convenablement les modèles 3D, ou en tout cas, pas avec un fichiers par couche à traiter. Les problèmes rencontrés sont divers et toujours d'actualité : les modèles sont souvent lourds, la manipulation est peu fluide, le géoréférencement natif du modèle n'est pas pris en compte (voire supprimé), les possibilités d'effectuer des mesures ou de superposer les couches sont limitées.

Cependant, les couches archéologiques sont bien des objets tridimensionnels. Comme l'a déjà noté B. Desachy, nous franchirons une nouvelle étape dans l'enregistrement des données de terrain archéologiques lorsque nous serons capables de relever la stratigraphie en 3D (Desachy 2008, p. 88). Il mentionnait d'ailleurs à la fin des années 2000 plusieurs outils expérimentaux qu'il imaginait voir se développer rapidement, mais il faut bien admettre que 15 ans plus tard, les moyens techniques restent difficilement accessibles pour la communauté archéologique (Losier *et al.* 2007 ; Katsianis *et al.* 2008). Imposer un enregistrement 3D de chaque couche archéologique au détriment d'informations planimétriques, en plus d'être complexe à exploiter et lourd à conserver, peut aussi mener à une perte d'information si le public cible n'est pas formé à ce type de représentation. C'est pourquoi, à l'heure actuelle, le système de représentation en plan, horizontale ou verticale, reste le système le plus abouti

et le mieux maîtrisé par les archéologues. Ainsi, comme le reflète la partie précédente, nous nous sommes concentrés sur la production d'informations en 2D, ou plutôt en 2,5D (ou pseudo 3D). Les orthomosaïques sont des représentations planimétriques, mais elles conservent néanmoins les valeurs d'altitudes de chaque pixel.

Le second objectif de la mise en place d'un protocole de levé photogrammétrique systématique s'inscrivait dans le cadre du projet 1 de Bibracte Numérique, qui visait la diffusion des connaissances archéologiques auprès du grand public à travers des dispositifs de médiation innovants. Nous souhaitions évaluer les possibilités offertes par les modèles 3D et les orthomosaïques pour valoriser les fouilles archéologiques, et particulièrement les méthodes de travail des archéologues. Le but était d'offrir au grand public un point de vue inédit sur les vestiges détruits par les travaux en cours, ou lors des campagnes précédentes. Nous supposons que les images permettraient d'expliquer au public les choix qui ont motivé leurs destructions, ainsi que les stratégies mises en œuvre.

Malheureusement, après plusieurs échanges avec l'agence Réciproque, sélectionnée pour la conception de nouveaux services de médiation numérique, il est apparu que les données photogrammétriques produites sur le chantier étaient difficilement exploitables pour la valorisation des fouilles. Plusieurs points problématiques ont été identifiés : toutes les couches n'étaient pas forcément pertinentes à présenter au public ; opérer une sélection avant d'avoir finalisé la phase d'analyse et de synthèse de la stratigraphie s'est révélé complexe ; les images avaient pourtant besoin d'être accompagnées d'interprétations détaillées, qui sont rarement fixées dès les premières observations ; comme la précision et la lisibilité prévalait sur le caractère harmonieux des images, la qualité esthétique des rendus n'était pas homogène.

Pour toutes ces raisons, il s'est avéré complexe de mettre en place un flux de partage standardisé des levés photogrammétriques des archéologues vers les médiateurs scientifiques. Nous sommes conscient que les orthomosaïques et les modèles 3D possèdent un réel potentiel pour la valorisation des pratiques archéologiques. Contrairement au dessin technique, les scènes modélisées avec fidélité dans un espace tridimensionnel sont intuitivement intelligibles, même pour des non initiés. Cependant, le sujet mérite une réflexion à part entière, qui devra déterminer les critères de sélection et de diffusion des données vers les services de valorisation. Dans le cadre de cette recherche doctorale, nous avons préféré nous concentrer sur les apports scientifiques de la photogrammétrie.

Cette pratique "intensive" de la photogrammétrie nous a par ailleurs permis de mettre en évidence plusieurs points critiques à maîtriser pour optimiser la production et l'usage d'images orthonormées en archéologie.

Le premier aspect est relatif à la nature d'une orthomosaïque et son usage en archéologie. Comme nous l'indiquions en début de partie, ce type d'image est issu d'un levé photogramétrique et non d'un relevé, dans le sens où l'image n'est pas interprétée. Elle se "contente" d'illustrer et localiser une surface selon un système orthonormé. Dans ce sens, elle remplit un rôle de support, qui comporte plusieurs caractéristiques :

Les orthomosaïques représentent d'abord un support à l'enregistrement archéologique, qui propose un nouveau type de données graphiques, avec une précision et une fidélité de rendu jusqu'alors inédite pour les archéologues : chaque pixel de l'image est lié à des mesures, exprimées en coordonnées tridimensionnelles X, Y et Z.

Elles constituent également un support de relevé. En effet, créer une orthomosaïque revient à prélever un échantillon visuel du terrain à une étape de la fouille, alors que le dessin offre une première forme d'analyse. Par le dessin, les archéologues exercent un raisonnement interprétatif en déterminant l'agencement des éléments présents sur le site, en fonction de logiques chrono-stratigraphiques et taphonomiques. Les orthomosaïques ne permettent donc pas de supprimer l'étape du dessin, mais offrent un support idéal pour effectuer des relevés rapides et précis.

Selon la même logique, les orthomosaïques ne permettent pas non plus de remplacer des observations directes. Malgré toutes les précautions d'usage mises en œuvre dans la pratique de la photogrammétrie, une image n'est jamais le reflet exact de la réalité (*cf. supra* 3.2.1. APN). C'est pourquoi il est indispensable que les minutes de terrain réalisées à l'aide d'orthomosaïques soient créées en ayant un point de vue direct sur le sujet, sur le terrain.

Enfin, les images produites par photogrammétrie offrent un nouveau support de réflexion aux chercheurs, puisqu'il s'agit d'une donnée visuelle inédite pour revenir au terrain en phase d'analyse et de synthèse des résultats.

En second point, nous avons identifié des situations plus propices que d'autres au levé photogramétrique. En effet, après avoir théorisé, mis en œuvre et éprouvé un usage systématique, le constat est pour nous sans appel : il n'est pas scientifiquement pertinent d'opérer un enregistrement photogramétrique exhaustif des couches archéologiques mises au jour.

Une orthomosaique permet de modéliser le terrain avec fidélité et précision, mais selon la nature et l'état de conservation des vestiges, le temps investie pour l'acquisition, le traitement et le relevé ne justifie pas toujours sa mise en œuvre. Par exemple, lever un trou de poteau, produire l'orthomosaique et l'utiliser en support pour le relevé manuel prendra toujours plus de temps que la méthode classique du papier millimétré, pour une précision sensiblement comparable. Face au poids que représente les fichiers dans l'archivage des données de fouille, il est nécessaire d'évaluer les situations qui justifient l'emploi de la photogrammétrie. Notre expérience sur ce point peut se résumer ainsi :

- les relevés manuels classiques sur papier millimétré sont les plus adaptés pour les structures isolées, d'une surface de moins 5m de large et qui présentent peu de détails. Dans ces cas là, les dessins sont plus rapides à produire que les relevés à partir d'orthomosaique, pour une marge de précision acceptable (moins de 5 cm). On comprend par exemple des structures en creux (TP, fossé, etc.) ou d'aménagements simples (lambeau de sol, foyer ouvert, etc.) ;
- la photogrammétrie se révèle particulièrement adaptée pour les structures complexes et/ou larges ou les surfaces étendues présentant de nombreuses structures. Pour les surfaces restreintes et détaillées, l'orthomosaique remplace l'étape souvent fastidieuse des prises de mesures, comme c'est le cas pour les relevés d'inhumation ou certains pierre à pierre. Pour les surfaces plus larges, complexes ou non, elle permet d'éviter la mise en place de carroyages ou de limiter les erreurs dues à l'installation d'axes de références sur de grandes distances. De plus, à l'échelle d'un secteur ou du chantier, l'image offre un visuel inédit pour appréhender l'organisation spatiale des différents vestiges et facilite la prise de décision sur les stratégies de fouille à adopter.

Enfin, même si la nature du sujet est adaptée à la photogrammétrie, les conditions peuvent contraindre l'opérateur à préférer une autre méthode d'enregistrement. Une luminosité mauvaise ou changeante va par exemple compliquer la lecture de l'image. Le manque de recul entre la scène et l'APN, à cause de la configuration de la fouille, d'un abris ou de la végétation va souvent être compenser par de trop nombreuses prises de vues. Parfois encore, si le mauvais temps n'empêchera pas de réaliser une photogrammétrie, on peut s'interroger sur l'utilité d'enregistrer avec une telle fidélité des vestiges boueux et peu lisibles. Comme toujours, c'est donc à l'archéologue de maîtriser les apports et les limites de la photogrammétrie pour déterminer s'il s'agit d'une technique pertinente pour atteindre ses objectifs scientifiques.

Avec le levé photogrammétrique, il manquait dans notre démarche l'étape du relevé. Une technique relativement simple à mettre en œuvre pour interpréter l'image consiste à l'imprimer, en respectant un facteur d'échelle standard (1/20, 1/50, 1/100, etc.), puis à effectuer le relevé manuel, c'est à dire tracer les limites ou les détails des vestiges observés sur le terrain, directement sur l'image ou par l'intermédiaire d'un papier calque. La méthode de relevé manuel classique s'en retrouve facilitée, mais le passage au format papier supprime aussi tous les avantages offerts par le numérique : le changement d'échelle (agrandir un détail de l'image), le géoréférencement, la superposition précise et le jeu de transparence applicable sur les vues. C'est pourquoi nous avons imaginé un module complémentaire de dessin technique numérique réalisable sur le terrain. L'objectif était d'allier la souplesse des logiciels de dessin vectoriel avec la précision des outils cartographiques, solution que l'on retrouve dans les SIG, déjà connu des archéologues.

3.4. Système d'Information Géographique : nouveau logiciel, nouveau relevé de l'US au plan de synthèse

3.4.1. Objectifs

La partie qui suit propose une présentation synthétique des SIG en archéologie selon l'objectif initial de le rattacher à la mise en œuvre de relevé numérique à partir d'orthomosaïques. Ainsi, nous nous concentrerons plutôt sur les capacités de gestion et de modélisation des SIG que sur les fonctions d'analyse spatiale, pour le moment absentes du protocole.

Espace et SIG en archéologie

La notion d'espace est intrinsèquement liée à la discipline archéologique : "*On peut considérer qu'existent les prémices d'une démarche archéologique dès lors que cette exploration n'est plus une simple chasse à l'objet ancien recueilli (ou pillé) uniquement pour sa valeur propre, mais qu'une certaine attention est accordée au contexte, c'est-à-dire aux propriétés de l'espace dans lequel cet objet a été trouvé*" (Costa, Desachy 2020, § 4). Ainsi, en fouille comme en prospection, toutes les informations recueillies sont spatialisées. L'espace est pour les archéologues un indice, un facteur d'explication qui nourrit leurs interprétations. Comme le signale O. Barge *et al.*, il est toujours question de "*proximité ou de connectivité des objets archéologiques, de relation de ces objets avec l'espace géographique, ou [...] de répartition*" (Barge *et al.* 2004, p. 16).

Par conséquent, les représentations spatialisées ont été réalisées très tôt en archéologie, ne serait-ce que pour pointer les vestiges découverts, donnant naissance aux premières cartes archéologiques. Pour le terrain, l'évolution des méthodes de fouilles stratigraphiques dans la deuxième moitié du XX^e s. a progressivement conduit au relevé systématique des unités d'enregistrement, en plan, en coupe ou en élévation (Wheeler 1954 ; Harris 1979). Ces minutes sont des abstractions d'une réalité observée, des représentations graphiques statiques qui permettent de localiser les vestiges mobiliers et immobiliers en symbolisant les interfaces et certains détails. Elles répondent à des normes qui permettent aux archéologues de représenter graphiquement ce qu'ils observent et comprennent des sites, mais aussi de partager ces observations. Ces visuels sont précieux tant pour l'analyse que l'archivage, mais ils sont aussi limités en matière de représentation. Les unités stratigraphiques ne se résument pas à leurs interfaces. Elles possèdent également des qualités intrinsèques qui ne peuvent pas toutes figurer sur le dessin, sous peine de le rendre illisible. Des descriptions textuelles détaillées sont donc enregistrées sur des fiches papiers, les fiches "US", ou directement dans des fichiers informatiques plus ou moins structurés.

Parmi les outils informatiques capables de gérer de l'information spatiale, les Systèmes d'Information Géographique tiennent une place de choix. Comme beaucoup de logiciels, les SIG ont émergé avec le développement de la micro-informatique dans les années 80. Plusieurs travaux se sont déjà intéressés à l'histoire de cet outil en archéologie (Wheatley, Gillings 2002, p. 1-21 ; Deweirdt 2010, p. 16-20 ; Rodier *et al.* 2011a, p. 14-20 ; Costa, Desachy 2020). Nous nous concentrerons donc ici sur son utilisation en archéologie de terrain. Dans le guide "Information spatiale et Archéologie", X. Rodier cite F. Pirot et A. Varet-Vitu, qui définissent ainsi le SIG : *“Système informatisé permettant, à partir de diverses sources, de rassembler (collecter) et d’organiser, de gérer, de modéliser (modélisation spatiale, structuration spatiale), d’analyser et de combiner (analyse spatiale), de simuler (simulation spatiale, scénarii), d’élaborer et de présenter des informations localisées géographiquement afin de contribuer à la connaissance de l’espace. Un SIG est aussi un concept qui sous-tend la notion de modélisation selon une approche phénoménologique. Un système informatisé est défini comme l’ensemble des méthodes d’information et des logiciels d’application qui prennent appui sur le système informatique pour son exploitation alors qu’un système informatique est défini comme un ensemble de matériels et de logiciels.”* (Pirot, Varet-Vitu 2004, glossaire ; Rodier *et al.* 2011a, p. 14).

Cette définition est longue, mais elle a l'avantage, selon l'auteur, de mettre en évidence

deux notions distincts regroupées sous le terme de SIG. Elle désigne à la fois les aspects techniques, les outils logiciels et les méthodes, et la finalité de leur utilisation en tant que système d'information. Autrement dit, parler de SIG en archéologie peut revenir à parler des outils informatiques qui les composent ; des données collectées, hébergées, gérées, transformée et présentées ; de la pratique que les archéologues ont de ce type de logiciel (Burrough 1986, p. 6-7).

Dans sa thèse, E. Deweirdt propose une synthèse du concept d'information géographique (Deweirdt 2010, p. 8-9). Elle comprend comme information une "*sélection de données pertinentes en vue d'une application ou d'un objectif particulier*", une définition similaire à celles des spécialistes de l'information et de la communication (*cf. supra* 2.1.1. La donnée : définition). L'information géographique représente les phénomènes du monde réel à travers trois couches d'abstraction. Dans un SIG traditionnel en deux dimensions, elles correspondent à :

- la **composante géométrique**, qui renvoie à la localisation dans l'espace d'objets à partir d'un système de coordonnées de référence connu, exprimé le plus souvent sous forme alphanumérique en x,y. Elle comprend également la description morphologique surfacique de ces objets, soit la forme et la taille de chacun ;
- la **composante sémantique**, qui précise la signification et la propriété des objets. Elle renseigne les attributs qui ne sont pas liés à la position des objets, comme la couleur et la nature du sédiment d'une unité stratigraphique, ou les liens qu'ils entretiennent avec d'autres objets (liens stratigraphiques, présence d'un objet dans une US, etc.). On parle ici des données attributaires ;
- la **topologie**, qui offre la dimension graphique. Elle décrit, par des vecteurs et non par des valeurs alphanumériques, la forme, la structure et la cohérence des objets. Ces vecteurs sont calculables à partir des valeurs de la composante géométrique. La topologie renseigne ainsi la nature des relations spatiales d'un objet avec ses voisins, en illustrant la connectivité ou non des objets.

Techniquement, les SIG gèrent toutes ces dimensions puisque qu'ils associent, entres autres, les fonctionnalités des Systèmes de Gestion de Base de Données (SGBD), des solutions logicielles de cartographie automatique et de Dessin Assisté par Ordinateur (DAO) (Rodier *et al.* 2011a, p. 14-16). Pour les SIG, l'espace représente le point d'entrée et le lien principal pour afficher et comparer les données graphiques et textuelles interrogées. En

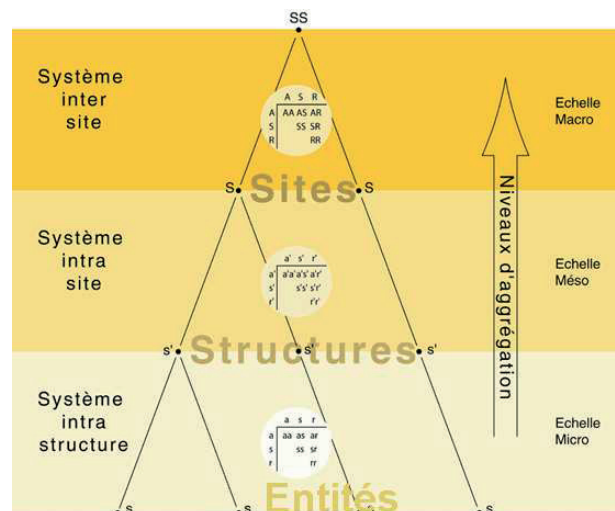
chantier, l'enregistrement archéologique est opéré de la même façon, car tous les objets d'études décrits sont *a minima* liés à une unité d'enregistrement, c'est-à-dire une surface définie et localisée. Ainsi, quelle que soit leur nature, les données archéologiques issues de fouilles répondent aux trois caractéristiques de l'information géographique. Les SIG intervenaient donc dans la continuité des plans, cartes et bases de données déjà intégrés aux méthodes des archéologues.

D. L. Clarke a défini en 1977 un système encore largement répandu pour l'étude spatiale de sites archéologiques (Clarke 1977, p. 11-15). Il comporte 3 fenêtres d'analyse, hiérarchisés selon leur résolution (fig. 66) :

- Système **inter-site**, échelle macro : il concerne l'étude des dynamiques de territoire et les interactions entre les sites archéologiques.
- Système **intra-site**, échelle méso : il s'intéresse à l'organisation spatiale interne d'un site archéologique, avec notamment la différenciation fonctionnelle, sociale et économique des différents espaces ou quartiers. F. Djindjian parle aussi de système inter-urbain (Djindjian 2017, p. 297).
- Système **intra-structure**, échelle micro : l'analyse se place au niveau des vestiges mobiliers, avec par exemple la répartition spatiale d'un type de mobilier qui permettant de définir une zone d'activité artisanale. Cette échelle permet également de documenter des micro-phénomènes, très localisés, comme l'érosion d'une partie d'une façade ou des indices tracéologiques sur des objets archéologiques (Conolly, Lake 2006, p. 12).

Les données produites par la fouille sont plus largement exploitées aux échelles méso et micro, même si certaines informations peuvent être exploitées à l'échelle macro dans des études de synthèses ultérieures.

Ce système pose aussi la question de la définition d'un site archéologique. Le site est souvent perçu comme un lieu rassemblant les vestiges mobiliers et immobiliers d'activités humaines passées (Barker 1993, p. 3 ; Renfrew, Bahn 2004, p. 54 ; Lehoërf 2009, p. 42 ; Djindjian 2017, p. 295-296). Selon cette définition, l'oppidum de Bibracte, en tant que capital du peuple Eduen au I^{er} s. av. n. e., peut évidemment être considéré comme un site archéologique, mais qu'en est-il des différents secteurs fouillés par des équipes séparées ? L'agglomération peut-elle être divisées en "sous-sites" ? Et les occupations ultérieures font-elles parties du même site ?



66. Les fenêtres d'analyse pour l'étude spatiale des sites archéologiques (extrait de Clarke 1977)

Cette ambiguïté vient du fait que, comme le souligne F. Giligny, le concept de site peut comporter plusieurs réalités spatiales, fonctionnelles et temporelles (Giligny 2009, p. 174). Dans la pratique, le site correspond d'abord à une échelle d'analyse, avant d'incarner un lieu : *"cette analyse doit être orientée par des questions précises et des objectifs à atteindre constituant la problématique archéologique, ce qui permettra ensuite de définir les sites appropriés et d'aller chercher les informations nécessaires pour répondre à cette problématique"*. Selon cette définition, la représentation spatiale de l'espace à Bibracte à l'aide d'un SIG peut ainsi être considérée comme une analyse intra-site, qu'elle intervienne à l'échelle de l'agglomération ou à celles des différents chantiers de fouille.

Les premières expérimentations françaises de SIG à l'échelle de fouilles archéologiques datent des années 1990 (Deweirdt 2010, p. 17-18 ; Rodier *et al.* 2011a, p. 16-17 ; Costa, Desachy 2020, § 18). Elles se sont essentiellement concentrées sur les fonctions d'édition cartographique et d'inventaire des SIG, pour produire des "bases de données graphiques". Il s'agit par exemple des objectifs fixés pour l'usage d'Arkéoplan à Bibracte, qui expérimentait la constitution d'une base de données à partir de dessins vectoriels référencés dans l'espace effectués sur une image numérique du terrain (Gruel, Buchsenschutz 1994).

Il est en effet possible de saisir et de rassembler sous une même interface un grand nombre de données provenant de la fouille, tels que les fiches d'enregistrement, les inventaires et les relevés manuels et topographiques. L'ensemble s'affiche sous forme de plans dynamiques,

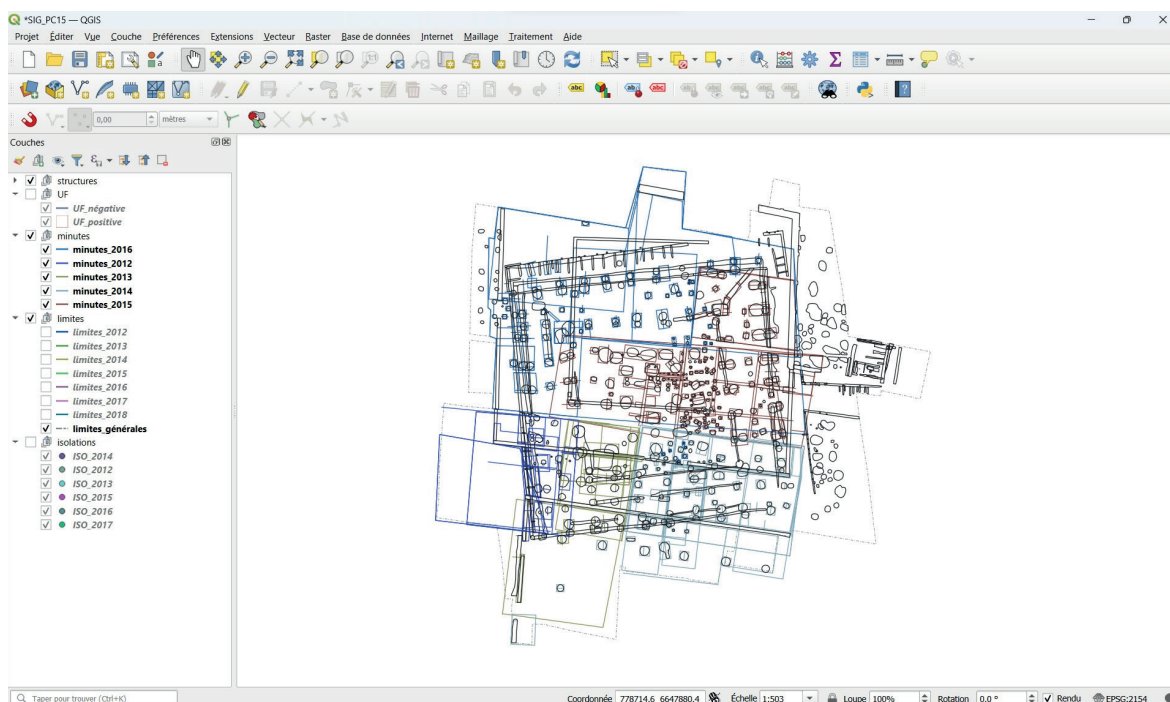
projetés dans un espace normalisé, et/ou de tableurs de données en lien avec les entités vectorielles. Toutes les informations, graphiques et textuelles, peuvent être éditées, faire l'objet de requêtes pour modifier les paramètres de visualisation et assurer les liens logiques avec des bases de données textuelles qui compilent les données archéologiques, comme le système de conservation et d'archivage bdB à Bibracte (base de données Bibracte, *cf. supra* partie 1.2.1. Prog. de recherche européen).

Les SIG ne se résument pourtant pas qu'à des outils de consultation et de gestion des ressources spatiales. Ils proposent aussi des fonctions avancées en matières d'analyse et de modélisation de l'espace pour étudier, par des traitements mathématiques ou logiques, les interactions entre les trois dimensions de l'information géographique (données attributaires, géométriques et topologiques) (Greene, Moore 2010, p. 80-82). Ils permettent ainsi d'explorer des hypothèses inédites sur l'organisation et les dynamiques spatiales qui régissaient les territoires, les sites et les objets archéologiques en confrontant les données de terrain à différentes caractéristiques spatiales (relief, distance, réseau, *etc.*).

C'est d'ailleurs en prenant en compte cette composante que les SIG vont se développer en archéologie dans les années 2000, notamment autour d'actions comme le réseau Information Spatiale et Archéologie (ISA) (Barge *et al.* 2004, p. 22-23). Dès sa création, ce réseau se donnait pour objectif, entre autres, de valoriser l'usage de la géomatique (contraction de géographie et informatique) dans la discipline archéologique. Les membres du réseau estimaient en effet que les outils, et principalement les SIG, étaient sous-exploités, alors qu'ils répondaient aux besoins des archéologues, tant en termes d'enregistrement que d'analyse.

Au milieu des années 2000, L. Costa et B. Desachy considèrent que l'usage des SIG en archéologie est entré dans une nouvelle dynamique (Costa, Desachy 2020, p. 17). Les pratiques se sont progressivement stabilisées et les SIG sont devenus des outils courants des archéologues. Néanmoins, cet usage se focalisait essentiellement sur des études inter-sites. É. Deweirdt conclue par exemple en 2008 son historiographie ainsi : "*malgré une recherche active dans le domaine du SIG en archéologie, les applications à l'échelle du site sont quasi absentes*" (Deweirdt 2010, p. 20). C'est pour cette raison que l'Inrap, qui travaille en priorité à l'échelle intra-site, a largement investi durant les années 2010 dans la formation le développement des outils, afin d'intégrer le SIG dans les pratiques archéologiques de terrain (*cf. supra* 1.1.2. Inrap). À Bibracte, la première analyse intra-site d'envergure à l'aide d'un SIG date de 2017, avec le travail de Master de J. Hantrais (Hantrais 2017). Dans ce cadre, elle a

réalisé un SIG rassemblant les données de six années de fouilles sur la terrasse PC15 et ses abords, avec pour objectif de comprendre la fonction du bâtiment et les dynamiques spatio-temporelles au sein de l'édifice monumental mis au jour sur la plateforme (fig. 67).



67. SIG de PC15 réalisé par J. Hantrais

Un SIG sur le terrain

Nous avons donc identifié les SIG comme les outils les plus adaptés pour exploiter les orthomosaïques produites systématiquement sur le terrain. Ils respectaient la position, l'orientation et l'échelle des ortho-images tout en proposant des fonctionnalités de vectorisation similaires aux logiciels de DAO couramment utilisés en archéologie. L'enregistrement des données attributaires était géré par ailleurs, avec le test de deux outils de saisie de terrain numérique complémentaires du système de gestion et d'archivage de Bibracte, *bdB* (cf. *infra* 3.5. Outils de saisie). Si cette utilisation était bien en deçà des véritables capacités des SIG, elle répondait néanmoins parfaitement à nos besoins pour finaliser le protocole de relevé à partir de données photogrammétriques. De plus, même si les SIG se sont largement démocratisés en archéologie ces dernières années, notre objectif était que l'outil déployé soit accessible au plus grand nombre. Nous avons donc cherché à mettre en place les solutions les plus simples et les plus proches des habitudes de travail des archéologues.

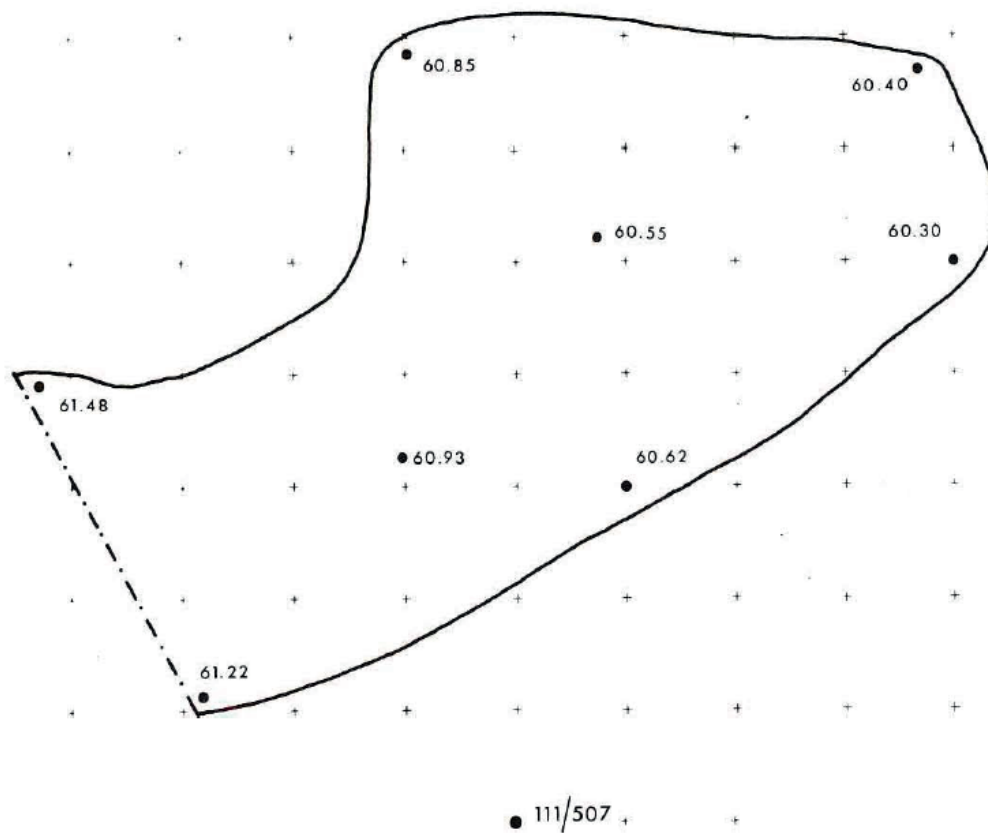
Cette expérimentation a également été rendue possible par les évolutions matérielles récentes. Au moment de l'élaboration du sujet de recherche, la MSHE C.-N. Ledoux (UAR 3124) faisait l'acquisition des premiers modèles d'ordinateurs durcis fiables au format tablette, fonctionnant avec le système d'exploitation Windows. Nous pouvions alors envisager l'emport et l'utilisation des logiciels directement sur le terrain, face aux vestiges.

En revanche, la construction de SIG intra-site en chantier devait s'inscrire dans les pratiques déjà en place à Bibracte. Depuis la reprise des fouilles modernes en 1984, le site accueille des équipes de fouille internationales qui interviennent autour d'un projet de recherche commun. Durant presque 40 ans, un effort constant a été apporté pour définir, maintenir et faire évoluer les méthodes et les outils nécessaires à l'enregistrement, au partage et à l'archivage des données archéologiques. Par conséquent, notre outil ne devait pas supplanter les techniques de relevé déjà en place. Il devait au contraire agir comme un module complémentaire, capable d'exploiter les possibilités offertes par les orthomosaïques tout en intégrant les relevés manuels traditionnels sur papier millimétré. La pratique du SIG devait également respecter la temporalité de la production et du traitement de l'information à Bibracte. Pour chaque équipe, la totalité des données primaires physiques et numériques devaient être inventoriées et archivées en fin de campagne, avant le départ des fouilleurs. La technique que nous souhaitons mettre en place ne devait donc pas alourdir la création et le traitement de l'information, ni minorer la qualité des données récoltées.

Enfin, au début de notre recherche, le mont Beuvray était aussi en grande majorité une "zone blanche", une zone géographique sans couverture réseau téléphonique ni internet. L'absence d'électricité sur le site ne permettait pas non plus d'envisager l'installation d'un réseau local. L'enregistrement sur plusieurs tablettes, utilisées simultanément par plusieurs fouilleurs, devait donc fonctionner sans la possibilité de synchroniser automatiquement les fichiers des différents terminaux sur un serveur distant. Notre démarche devait par conséquent anticiper le transfert de données par des supports physiques, type clé USB ou disque dur externe, tout en limitant au maximum les erreurs de saisie et les doublons.

Pour répondre aux enjeux scientifiques comme aux contraintes techniques, nous avons mis en place un protocole fortement inspiré de la méthode britannique du *Single-Layer Plan* (SLP), ou *Single Context Plan*. Il s'agit d'une méthode conçue dans les années 70, à partir des travaux de E. C. Harris sur l'enregistrement stratigraphique, pour documenter des fouilles en contexte urbain densément stratifiées (Harris 1989, p. 95-104 ; Desachy 2008, p. 87) (fig. 68). Les SLP sont des équivalents graphiques des fiches d'enregistrement textuelles des

Site **NEW ROAD** 1975
 Area **1** Drawn by PJO 22/8/75 Phased by Layer 464



- METHOD**
- One unit of stratification ONLY on a sheet
 - Boundary contours of the unit are drawn
 - + Some elevations placed directly on the plotted area
 - One set of co-ordinates (as sheet is pre-orientated)

SW ▲

1M Scale 1:50

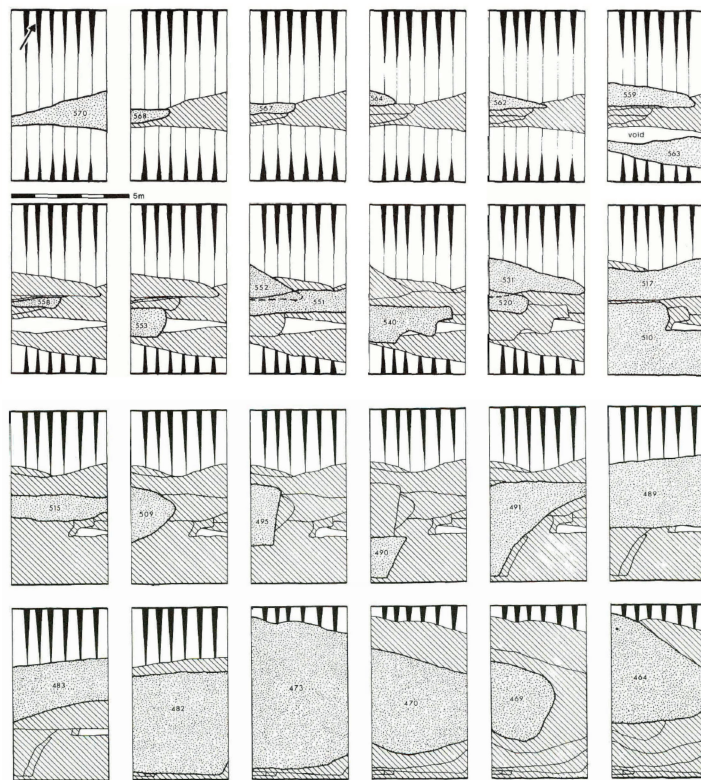
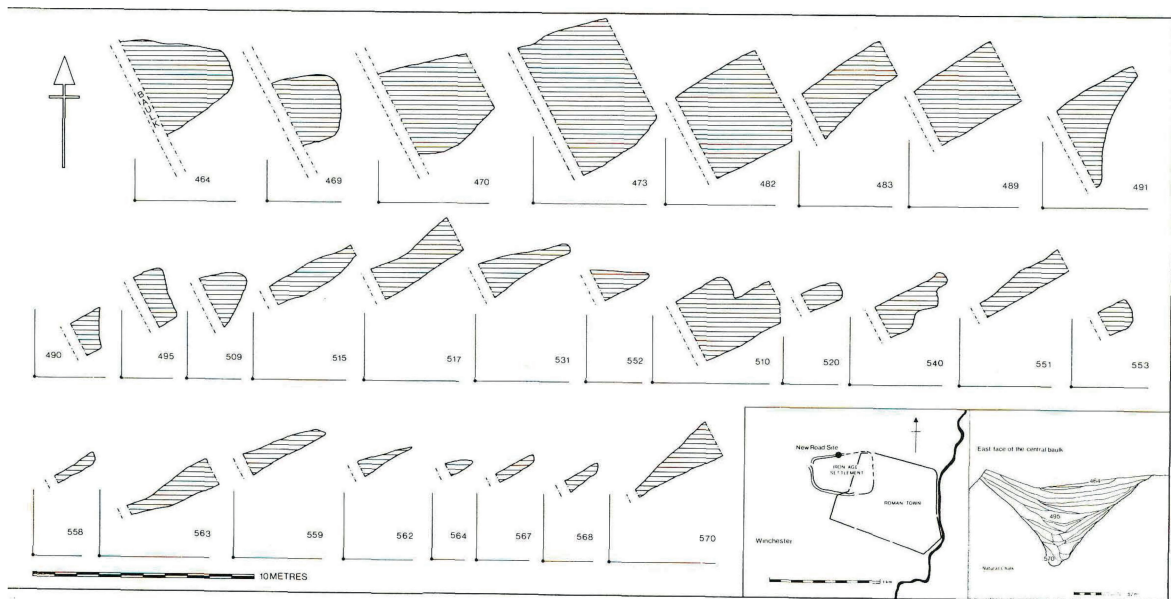
68. Principe de la méthode du Single-Layer Plan (extrait de Harris 1989)

US. Chaque US possède un SLP, et un SLP ne documente qu'une seule US. Sur cette fiche, seules les caractéristiques essentielles, répétitives et universelles de l'US sont relevées : son emprise maximale en plan, symbolisée par le tracé de son contour, localisée par des coordonnées précises et complétée d'indications d'altitudes de surface. Pour cela, le SLP est réalisé lorsque tous les éléments qui se superposent à l'US et qui masquent son étendue ont été dégagées. À l'origine, l'objectif était de pouvoir faciliter l'analyse des phases d'évolution, en recomposant manuellement en plan, couche par couche, la stratigraphie des sites étudiés (fig. 69).

Pour notre protocole, nous nous sommes fondé sur le même principe que le SLP, mais en modifiant le support d'enregistrement : "une US = une fiche " devenait "une US = un fichier de forme vectoriel (*Shapefile*)". Le but était de créer les fichiers de forme à partir des levés photogrammétriques systématiques, sur SIG, en observant directement l'US sur le terrain. En théorie, cette solution offrait plusieurs avantages pour le relever :

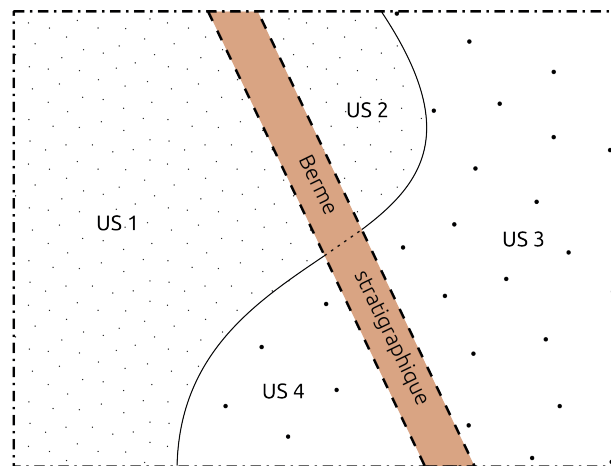
- une orthomosaïque livre des mesures topographiques beaucoup plus précises qu'avec un SLP classique (une mesure en X, Y et Z par pixel) ;
- elle permet d'archiver un visuel en haute qualité de l'emprise de l'US, qui permet au fouilleur de justifier son choix ; les fonctionnalités DAO des SIG facilitent la visualisation, la manipulation et l'édition des plans par rapport au support papier ;
- les éléments basiques d'enregistrement, tels que le numéro de l'US, le nom de l'auteur et la date de relevé, peuvent être implémentés dans la table attributaire du fichier de forme ;
- les fichiers de forme sont conçus pour une exploitation SIG, ils enregistrent les coordonnées des entités vectorielles dans un système de coordonnées de référence.

Cette méthode implique de suivre à la lettre les règles stratigraphiques de synchronie telles que définie par E. C. Harris (Harris 1979). Le principe de continuité spatiale doit être assurée sur l'ensemble de la surface de l'US (fig. 70). Deux entités de nature similaire qui ne respectent pas ce principe sont considérées comme deux US différentes, assurant un lien d'équivalence. Si avec l'avancée des travaux, la continuité physiques est finalement observée, les deux numéros d'US sont conservés et le lien de synchronie est modifié, en indiquant que les deux US sont identiques. Cette règle d'enregistrement permet d'assurer que chaque US



69. La stratigraphie d'un fossé de l'âge du fer recomposée en plan grâce aux SPL (New Road Site, Winchester 1975, extrait de Harris 1989)

est associée à un fichier de forme avec une seule entité spatiale vectorielle. Cette condition est nécessaire pour automatiser la fusion des entités en un seul fichier de forme en fin de campagne.



70. Principe de continuité spatiale : même s'il est quasi certain qu'il s'agit des mêmes couches de part et d'autre de la berme, ces dernières sont enregistrées dans des US différentes.

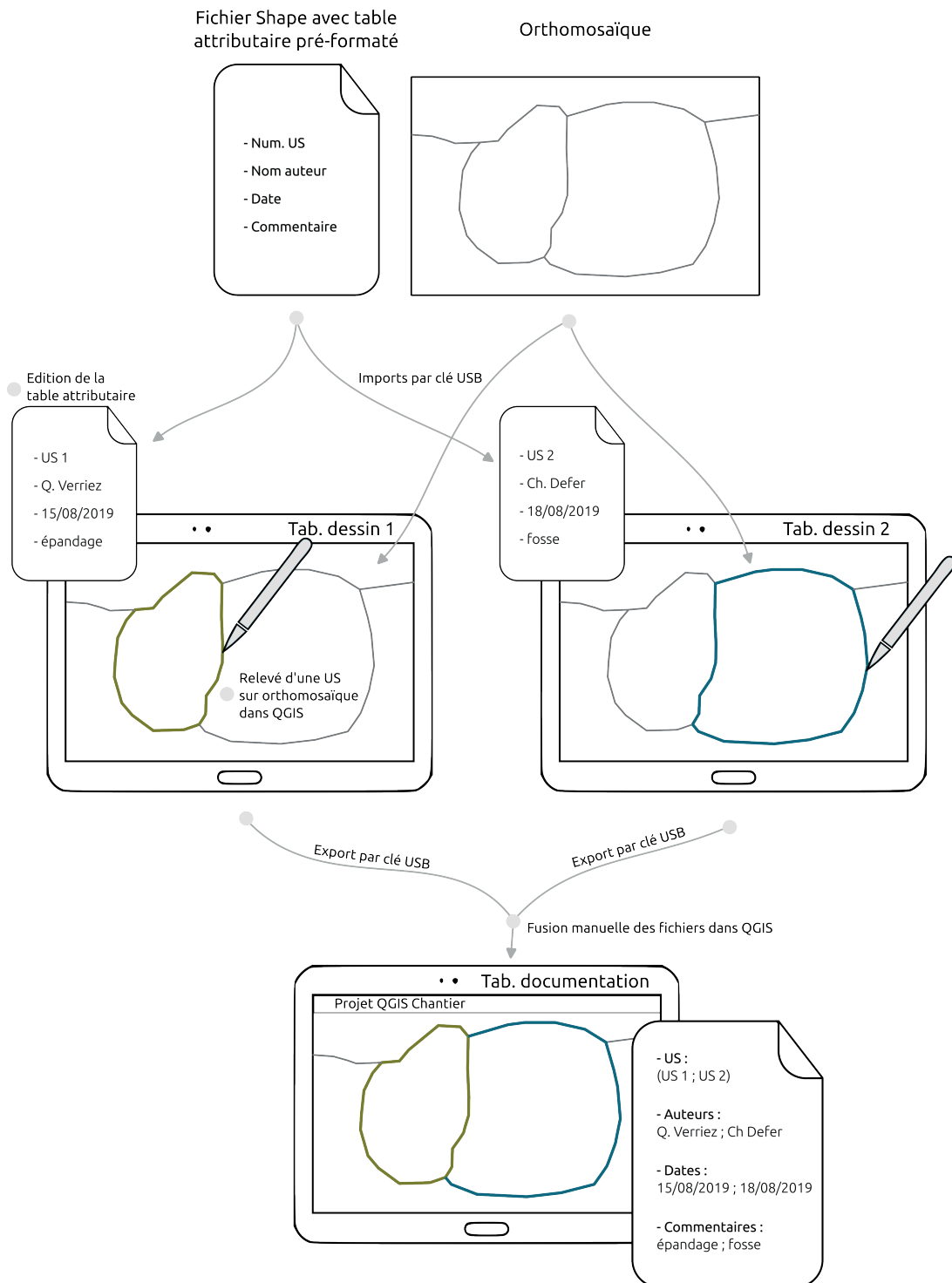
Nous avons aussi repris le principe de la fiche papier SLP, en créant un fichier de forme générique, afin de limiter les erreurs de saisie et de faciliter la gestion des fichiers pour des archéologues dans une dynamique de fouille. Ce modèle agit comme une fiche d'enregistrement vierge. Il ne comporte pas d'entité vectorielle de base, mais présente néanmoins une structure d'enregistrement normalisée. Cette normalisation repose d'abord sur une convention de nommage du fichier, qui répond à deux règles : le nom doit assurer de manière claire et intelligible le lien entre le fichier de forme et l'orthomosaïque utilisée pour le relever, qui comporte elle aussi un nom normalisé ; ce nom doit être codifié pour faciliter l'exploitation du fichier dans un système informatisé (tri automatique, requêtage), mais suffisamment explicite pour que les archéologues comprennent le contenu du fichier. La structure interne du fichier est également normalisée, car les champs attributaires sont déjà préformatés. Notre but est de faire enregistrer au fouilleur les informations essentielles pour identifier le contenu du fichier sous la forme la plus simple et la plus concise possible. Le formulaire final comporte quatre champs : le numéro de l'US documentée, le nom de l'auteur, la date de création et un commentaire libre.

Dans la pratique, sur le chantier, chaque fouilleur qui entame un relevé se munit d'une tablette destinée au dessin, du fichier de forme générique et de l'orthomosaïque qui renseigne la couche qu'il souhaite dessiner (fig. 71). Il importe les fichiers dans un projet QGIS et effectue son relevé à partir du fichier de forme pré-formaté. Lorsque l'entité vectorielle est créée, il suffit de compléter le formulaire d'attributs déjà existant dans le fichier et de l'enregistrer sous le nouveau nom qui respecte la convention de nommage. Il est aussi possible de relever plusieurs US à la fois, l'important étant d'exporter un fichier de forme par US. Une fois le relevé effectué, les fichiers de forme sont immédiatement centralisés sur un ordinateur ou une tablette "documentation", réservés au dépôt et à la gestion des données. Toute modification d'un fichier de forme se fait sur d'autres tablettes, prévues pour le dessin. Cette rigueur permet que la machine qui centralise l'information soit toujours disponible pour le dépôt de fichier, sans délai. Le SIG ainsi construit, US par US, agit comme une bibliothèque de fichier de forme, qui permet la visualisation et le contrôle en temps réel des relevés sur le terrain (fig. 72). Enfin, il est possible de fusionner tous les fichiers de forme en fin de campagne pour obtenir une véritable base de données à référence spatiale.

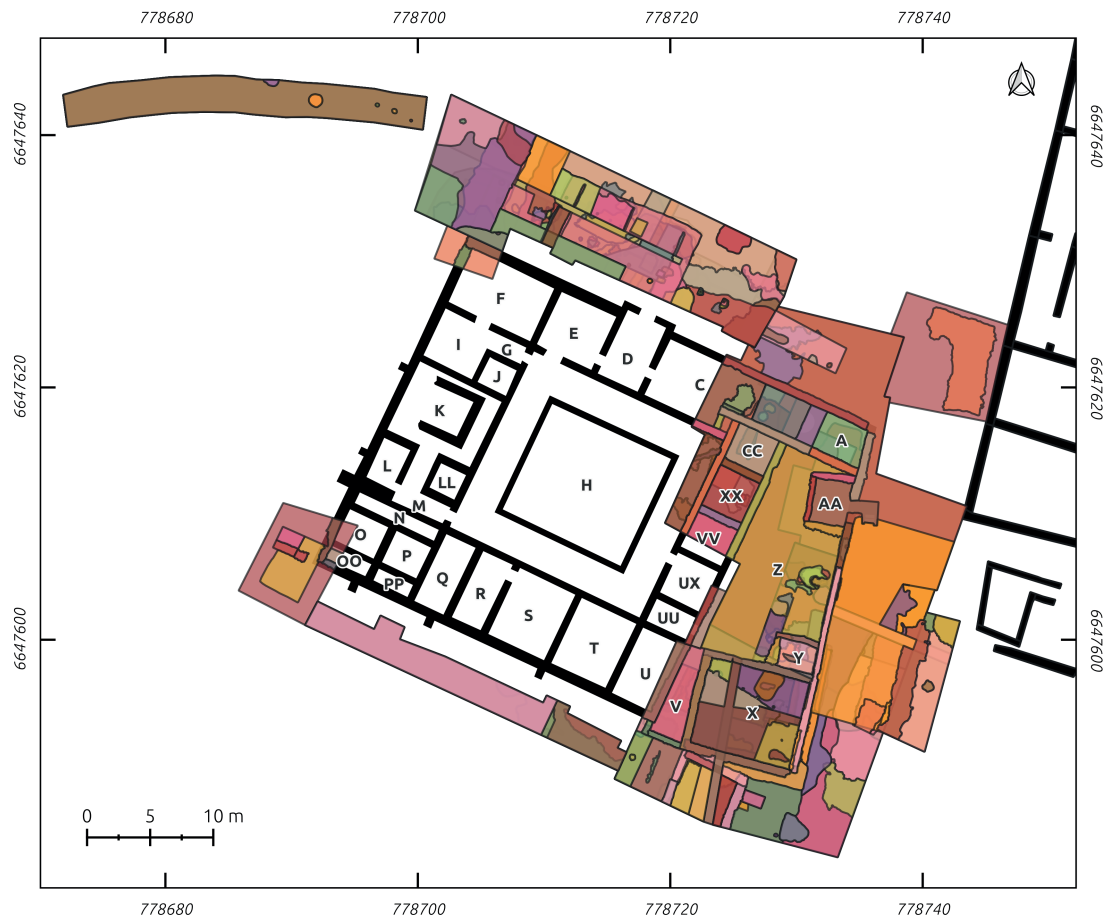
3.4.2. Choix des équipements

Pour la mise en œuvre du protocole, nous avons utilisé deux modèles de tablettes ordinateurs de la plateforme technologique SHERPA de la MSHE C.-N. Ledoux : la Xslate B10 de Xplore et la Latitude 7212 de Dell (fig. 73). L'objectif n'est pas ici de décrire en détail le matériel, mais plutôt d'identifier les avantages (et inconvénients) de ce type d'équipement pour le travail de terrain.

Le format tablette est particulièrement pratique pour l'emport de matériel informatique sur le terrain. Les fonctions tactiles et l'utilisation d'un stylet permet de retrouver les gestes classiques des archéologues pour le relevé manuel. Les équipements durcis sont aussi vivement conseillés, car ils sont prévus pour une exploitation en extérieur. Ils sont résistants à l'humidité, à la poussière et aux chocs, et ils proposent généralement une luminosité de l'écran plus forte, pour utilisation au soleil. L'écran de 11 pouces, soit environ 28 cm de diagonale, s'avère être un bon compromis entre la lisibilité et la capacité de transport des machines. Pour terminer avec les aspects matériels, comme toujours, il est important d'être attentif à la capacité des batteries. Ces équipements sont généralement prévus pour tenir une journée, mais pas forcément avec une utilisation intensive, comme dans le cas des traitements photogrammétriques. Afin d'anticiper tout problème, comme pour les appareils photo, il est préférable de sélectionner des modèles avec batterie interchangeable.



71. Processus de production des fichiers de forme selon le protocole de relevé numérique mis en place sur le chantier de la domus PC2



72. Les 566 US relevées sur le chantier de la domus PC2 durant les campagne 2017 à 2020



73. Les tablettes Xplore Xslate B10 (à gauche) et Dell Latitude 7212 (à droite) de la MSHE Ledoux

De même, le choix des composants internes se fait en fonction des capacités informatiques de l'époque. En effet, les besoins en ressource des logiciels ont tendance à augmenter conjointement avec le développement des composants informatiques, qui est en général très rapide. Pour l'exemple, lors des premières expérimentations en 2018, les tablettes Xplore Xslate B10 nous permettaient de produire et de manipuler des données SIG sans réellement de contrainte. Avec l'évolution des logiciels, nous avons rencontré quelques limitations pour effectuer ces mêmes tâches en 2021 (lenteur dans l'affichage des données sous SIG, temps de traitement photogrammétrique très long, *etc.*). D'une manière générale, l'utilisation de logiciels de SIG et de photogrammétrie requiert, au minimum, l'équivalent d'ordinateurs de bureautique haut de gamme.

À propos des logiciels, les tablettes numériques fonctionnent généralement sous des systèmes d'exploitation conçus pour les smartphones (Android de Google et iOS d'Apple), avec un environnement à base d'applications. Dans notre cas, nous avons utilisé le système d'exploitation fourni avec les tablettes Xplore et Dell, le système Windows de Microsoft, adapté à une utilisation sur ordinateur. Il ne s'agit pas d'un système d'exploitation libre, comme ceux basés sur le noyau Linux, mais il reste néanmoins le système d'exploitation le plus utilisé dans le monde (fig. 74). Il permet ainsi de dupliquer sur les tablettes l'environnement de travail informatique que la majorité des archéologues utilisent au quotidien sur leur ordinateur portable ou fixe. De plus, il facilite l'utilisation des logiciels libres, qui sont en général développés en priorité pour les systèmes d'exploitation pour ordinateur, sur Windows particulièrement.

Pour le SIG, nous avons sélectionné le logiciel libre QGIS, aujourd'hui largement répandu dans la communauté archéologique⁹². QGIS est un logiciel de SIG développé depuis 2002 par une communauté souhaitant rendre accessible gratuitement au plus grand nombre les SIG avec un projet multiplateforme (Windows, macOS et Linux) qui propose une interface intuitive⁹³. Le projet est conçu autour d'un noyau central, auquel il est possible d'ajouter des fonctionnalités par le biais d'extensions complémentaires (*plugin*), développées par des utilisateurs⁹⁴. Pour la communauté archéologique par exemple, des agents de l'Inrap ont créé les extensions ArcheoCAD⁹⁵, qui accélère le processus de relevé de structures à partir de don-

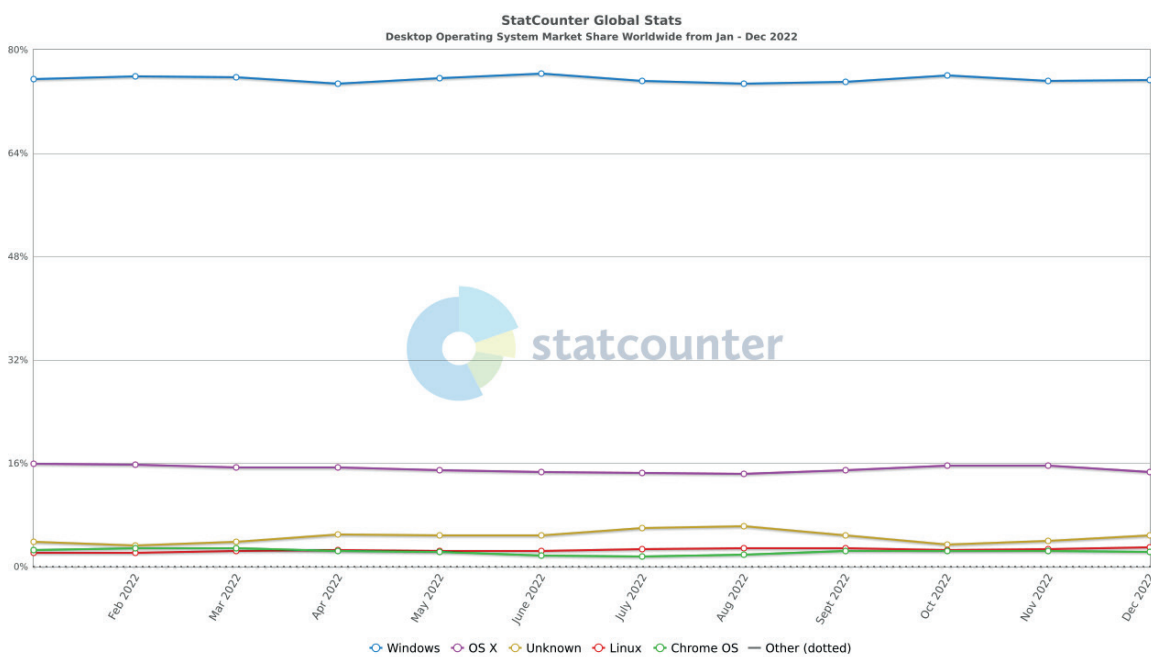
92 <https://www.qgis.org/fr/site/> ; consulté le 22/10/2023

93 https://docs.qgis.org/3.28/fr/docs/user_manual/preamble/foreword.html ; consulté le 22/10/2023

94 <https://plugins.qgis.org/plugins/> ; consulté le 22/10/2023

95 <https://github.com/narimanInrap/ArcheoCAD/blob/master/README.md> ; consulté le 22/10/2023

nées topographiques, et AGT⁹⁶ (*Archaeological Geophysics Toolbox*) pour la visualisation et le traitement de données géophysiques (Hulin *et al.* 2020). De plus, les fonctionnalités de DAO de QGIS ont été constamment améliorées depuis l'arrivée de la version 3.0 en février 2018. Il s'agit donc d'un logiciel libre parfaitement adapté pour la mise en œuvre d'un protocole de relevé intra-site en temps réel.



74. La répartition des parts de marché entre les principaux systèmes d'exploitation (extrait de statcounter : <https://gs.statcounter.com/os-market-share/desktop/worldwide/2022>)

Pour la partie cartographique, QGIS et les SIG en général prennent en charges des données de deux natures : les images numériques matricielles (*raster*) et les vecteurs (fig. 34) (*cf. supra* 2.1.2. Données simples). Les rasters correspondent la plupart du temps au fonds de carte, aux photos redressées ou dans notre cas, aux orthomosaïques. Les vecteurs correspondent aux entités géographiques numérisées sous SIG, telles que les US, les points topographiques ou l'emprise du chantier.

Comme pour la photogrammétrie, nous proposons en annexe un manuel, étape par étape, pour la réalisation d'un relevé numérique sur le logiciel QGIS à partir d'une orthomosaïque (Annexe 4). Ce document prend en compte les règles d'enregistrements propres à Bibracte, mais il peut tout à fait s'utiliser dans d'autres contextes. Le guide a été écrit avec la version 3.16.0 *Hannover* de QGIS.

3.4.3. Acquis des expérimentations

Les premières expérimentations du protocole ont eu lieu sur le chantier de la *domus* PC2 en 2018, par trois des responsables de l'équipe. Il a été validé l'année suivante et ouvert à l'ensemble des fouilleurs. La majorité des relevés en plan a été réalisée en suivant ce protocole lors des campagnes de 2019 à 2021. Il s'agit donc d'une méthode pleinement opérationnelle, qui présente plusieurs avantages significatifs.

Tout d'abord, le protocole s'intègre dans la chaîne opératoire archéologique sans impacter le déroulement du chantier. Même dans le cadre de notre expérimentation de levé photogrammétrique systématique, qui a démontré que les orthomosaïques n'étaient pas adaptées à toutes les situations, la production des relevés numériques sur SIG n'a jamais été un facteur bloquant pour l'avancée du chantier. Au contraire, pour les personnes formées, l'étape stricte du dessin est facilitée et plus rapide, car il est possible de relever seul de grandes surfaces sans prendre de mesure supplémentaire.

Sur la formation, qui est en enjeu important du chantier-école de la *domus* PC2, tous les fouilleurs ont été formés et exploitaient cette méthode de relevé quotidiennement à partir de 2019 (fig. 75). Les séjours des fouilleurs sont relativement courts sur ce chantier, de 15 jours par session. Lors des campagnes 2017 et 2018, les participants n'ont pas pu tous pratiquer le relevé manuel traditionnel, par manque de temps. Avec le relevé numérique, la prise en main des outils s'est avéré beaucoup plus rapide et chacun a pu effectuer au minimum un dessin technique. La formation en elle-même se faisait directement sur le chantier, avec une présentation de l'outil et de la méthode de relevé en condition réelle, face aux vestiges.

Il faut toutefois nuancer ce dernier point en considérant la particularité de l'équipe du chantier-école de la *domus* PC2 : les fouilleurs ont entre 13 et 17 ans et ne sont pas ou très peu expérimentés en archéologie. Contrairement à ce que les préjugés sur leur jeune âge pourraient laisser penser, ces *digital natives* ne sont pas tous à l'aise avec le numérique, et encore moins avec les ordinateurs. En revanche, nous n'avons jamais rencontré de réticence de principe face aux objets numériques. Tous se sont montrés ouverts au relevé sur SIG, quelles que soient leurs connaissances en informatique. La facilité de prise en main peut aussi venir du fait que les fouilleurs ne connaissaient pas d'autres méthodes de relevé. Le dessin sur SIG ne rentrait pas en conflit avec une autre pratique déjà façonnée par leur expérience.



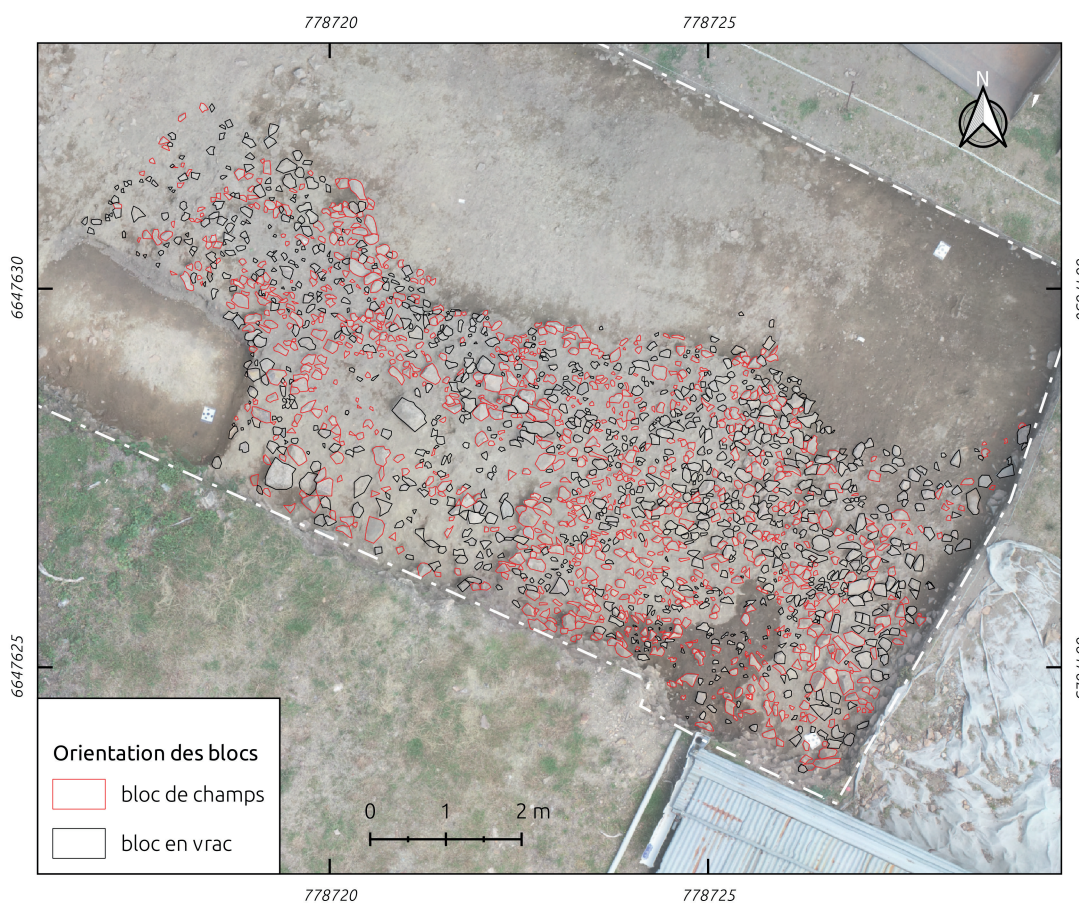
75. Une fouilleuse qui utilise une tablette comme support de relevé en 2019

Il nous semble important de préciser ici un point qui sort quelque peu du cadre de la thèse, car la remarque nous a souvent été faite par des collègues lorsque nous présentions nos résultats. Nous sommes bien conscient que former des nouveaux fouilleurs à des méthodes expérimentales avant les méthodes classiques est discutable. Il était toutefois nécessaire de trouver un équilibre entre les objectifs pédagogiques du chantier-école, les problématiques archéologiques de la fouille et nos considérations méthodologiques. C'est pourquoi nous avons systématiquement présenté aux participants la méthode de relevé manuel sur papier millimétré avant de les former au relevé sur SIG, en leur expliquant le caractère innovant, mais surtout expérimental de cette méthode.

Pour revenir à notre sujet principal, nous avons constaté une amélioration générale de la cohérence et de la qualité des informations relevées, sur le terrain comme en post-fouille. En chantier, un seul relevé par structure suffit, il n'est pas nécessaire de multiplier les dessins à des échelles différentes. En compilant l'ensemble des fichiers de forme dans un SIG, le plan d'ensemble de la fouille est mis à jour avec chaque relevé. La consultation peut passer de l'échelle de la structure à l'échelle du chantier en quelques secondes, tout en assurant une continuité spatiale et visuelle dans l'affichage des entités. Il existe également plusieurs options

pour effectuer des sélections de couches à afficher, pour isoler simplement un secteur, une entité archéologique cohérente ou une phase stratigraphique à adopter. En post-fouille, le SIG s'est révélé être une aide précieuse pour la vérification des enregistrements, en constituant un point d'entrée unique qui assure le lien entre les données textuelles et graphiques.

Dans l'état actuel de notre protocole, le tracé vectoriel enregistre uniquement la silhouette des US. L'enregistrement détaillé des éléments internes est assuré par les orthomosaïques. Cette apparente simplicité assure qu'une seule entité vectorielle soit créée par US, une condition nécessaire pour la fusion simplifiée des fichiers de forme en fin de campagne. Une expérience a toutefois été réalisée pour relever des détails avec plusieurs entités, avec la mise en évidence des pierres de maçonnerie sur le chant d'un mur effondré (fig. 76). Cette expérience démontre que le relevé de détails est possible, même si la multiplication des entités pour une seule US ne convient pas pour le moment à notre protocole de relevé. Il faudrait éventuellement prévoir dans le projet QGIS un deuxième ensemble d'entités vectorielles qui regroupent les éléments de détail plus complexes, à l'image d'un nouveau calque dans un outil de DAO.



76. Test de relevé pierre à pierre sur le mur effondré 655 en 2020

Enfin, les dessins produits avec ce protocole sont nativement au format vectoriel, mais il n'exclue pas la possibilité de réaliser les relevés sur papier millimétré. Les dessins traditionnels en plan sont également localisés dans un espace contrôlé par des marqueurs topographiques qui signalent l'axe de référence du relevé. En rajoutant un troisième marqueur sur le dessin, il est possible de numériser et de géoréférencer ces plans avec QGIS pour ensuite les reprendre en dessin vectoriel sur le SIG. Cette technique est d'ailleurs généralement plus rapide que le levé photogrammétrique pour les petites structures ou surfaces présentant peu de détails, tels que des trous de poteau, des foyers simples ou des surfaces d'épandage restreintes.

Pour la mise au propre, QGIS propose un module d'édition de cartes à partir des données gérées par le SIG. Les fonctionnalités avancées du module permettent d'atteindre une qualité de rendu graphique qui correspond aux attentes des illustrations de rapport ou de publication en archéologie⁹⁷. Ce processus peut d'ailleurs être en partie automatisé en produisant un ou plusieurs modèles de trame graphique, par exemple en fonction de l'échelle ou de la nature du sujet. Il est également possible de produire des cartes qui ne comporte que les informations essentielles (contour des couches et échelle) et de les exporter au format graphique vectoriel *Scalable Vector Graphics* (SVG). Cette étape permet ensuite d'éditer les cartes avec les logiciels de DAO mieux connus des archéologues. Si Adobe Illustrator est le plus répandu, les données SVG provenant de QGIS sont généralement mieux gérées par Inkscape, son homologue libre.

Finalement, le protocole s'est révélé particulièrement efficace pour sa rapidité de prise en main, pour l'exploitation du potentiel des orthomosaïques et pour la qualité de la création jusqu'à la diffusion de la donnée. Cependant, comme pour toute nouvelle méthode, cette solution comporte aussi certaines limitations.

Le point le plus préoccupant selon nous est la transformation du rapport au terrain qu'implique ce type d'enregistrement. Depuis la conceptualisation du protocole, nous nous sommes appliqués à mettre en place une solution numérique qui soit transportable sur le terrain. Chaque fouilleur formé a également été averti de l'importance de faire ses propres observations sur le terrain avant de les symboliser sur l'orthomosaïque. Malgré cela, nous avons constaté à plusieurs reprises que les personnes qui effectuaient les relevés procédaient

97 Exemple de cartes réalisées avec QGIS : <https://www.flickr.com/groups/qgis/pool/> ; consulté le 22/10/2023

à la démarche inverse. Même en étant face aux vestiges, elles dessinaient d'abord les couches sur les orthomosaïques avant de vérifier leur relevé en fonction de la réalité du terrain. Or l'orthomosaïque n'est qu'une abstraction de la réalité, un instantané qui respecte un point de vue sur une scène en tenant compte de paramètres d'exposition. Le capteur photographique n'est pas à même d'enregistrer toute la complexité des indices géoarchéologiques et matériels qui caractérisent une couche ou une structure archéologique. L'ortho-image n'est qu'un support précis et fidèle pour reporter facilement des observations faites sur le terrain.

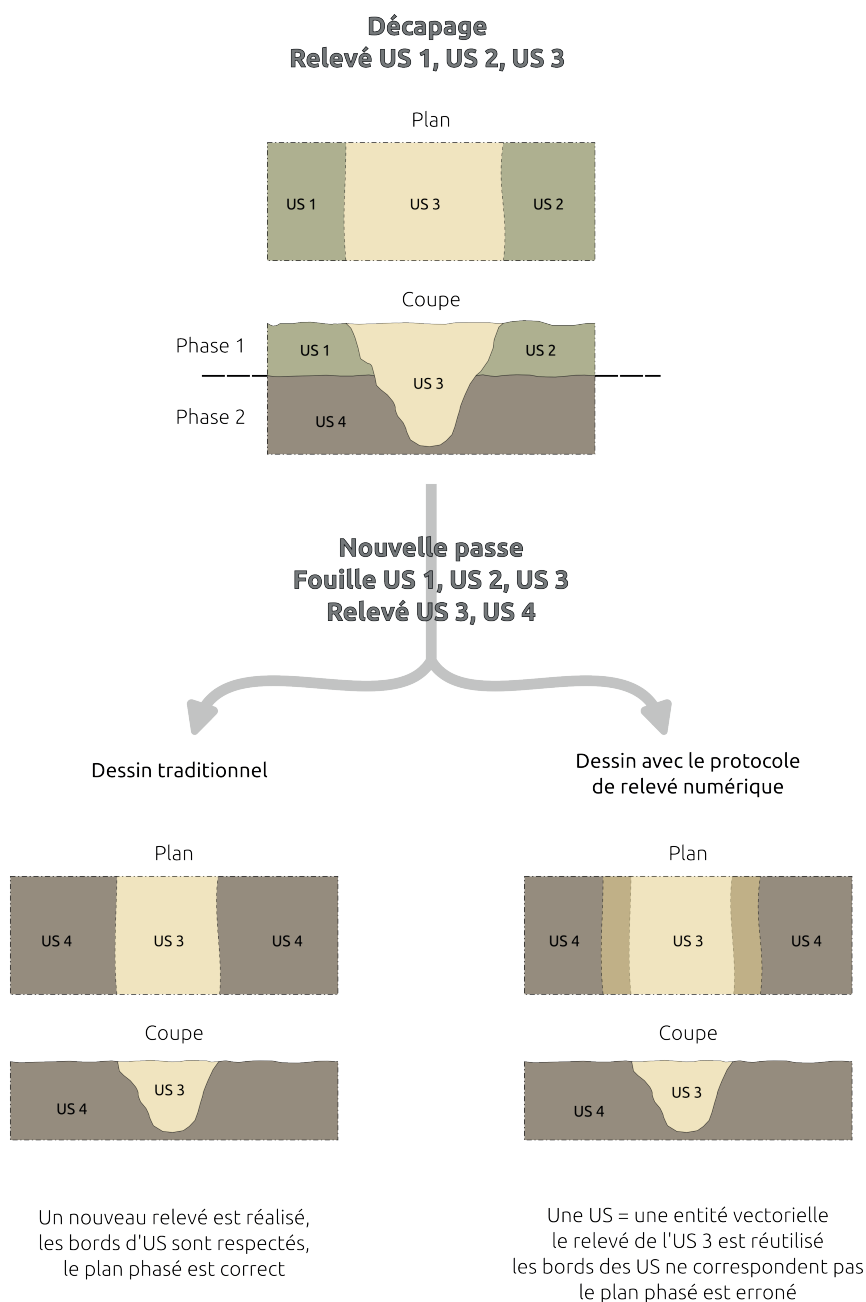
Actuellement, il n'existe pas à Bibracte de cadre d'enregistrement strictement défini dans le système de gestion et d'archivage bdB pour l'enregistrement des données d'un SIG intra-site. Avec le protocole, l'objectif initial était de constituer une documentation semblable aux minutes de terrain en exportant des plans de manière semi-automatique à partir du SIG, pour assurer la conservation des données graphiques. Pourtant, ces exports n'ont jamais été réalisés, par manque de temps, mais surtout, car la question s'est rapidement posée des données à sélectionner pour chaque export. Fallait-il reprendre la méthode des SLP et produire un plan par US ? Le but des SLP était avant tout rassembler des représentations normalisées des US en plan, dans l'objectif de simplifier la création de plans phasés. Un SIG intra-site rassemble et gère déjà ces représentations normalisées, il ne nous a donc pas semblé pertinent de doubler cette documentation en éditant plus d'une centaine de plans d'US individualisés par campagne de fouille.

De manière générale, il reste un travail important à mener sur la sémiologie graphique des relevés effectués avec QGIS. Par exemple, sur un dessin archéologique, lorsqu'une limite de couche est incertaine, elle est généralement représentée par des tirets. Cette incertitude ne s'applique pas nécessairement à l'ensemble de l'US, elle peut se limiter à une portion du tracé de la couche. Dans notre cas, chaque US est représentée par un polygone, qui affiche obligatoirement la même sémiologie graphique sur l'ensemble de tracé. Comme le démontre S. Eusèbe, ce type de contrainte a progressivement poussé les archéologues à abandonner les règles de représentation en vigueur dans notre discipline (Eusèbe 2019). Le respect de la sémiologie graphique est pourtant essentiel pour le partage de d'information à travers la communauté archéologique. Elle est le fruit d'un long héritage intellectuel, qui offre aujourd'hui un cadre normatif pour traiter et partager des informations sous forme de carte. Respecter ce cadre est gage de transmettre une information spatiale de qualité, explicite pour l'ensemble de la communauté archéologique. Pourtant, avec l'apparition des outils informatisés, ces normes sémiologiques ont tendance à disparaître, car les symboles graphiques ne sont pas

implémentés à l'origine dans les outils, ou parce que les utilisateurs manquent de formation sur les logiciels. Les solutions techniques existent pour pallier ces contraintes apparentes. Il est par exemple possible d'importer ses propres modèles de symbole dans QGIS, même si la tâche requiert des compétences particulières en graphisme. La priorité aujourd'hui pour la communauté archéologique est de s'accorder sur la transposition des règles de représentation du format papier au format numérique.

Aussi, les couches archéologiques ne sont pas des parallélépipèdes rectangles parfaitement empilés comme les briques d'une célèbre marque de pièces emboîtables. Elles n'ont pas la même emprise sur les parties supérieures et inférieures. Cette situation peut poser un problème pour la représentation des structures hautes ou profondes, comme les élévations ou leurs tranchées de fondation. Habituellement, il suffit de multiplier les minutes à différentes étapes de la fouille pour représenter un même US en plan à différents niveaux d'altitude. Mais dans notre cas, chaque US est limitée à une seule entité vectorielle, ce qui peut avoir des conséquences pour construire des plans de synthèse par phase chronologique. Il est possible que l'emprise maximale d'une tranchée de fondation soit en réalité plus haute en altitude que dans le plan actuel, ce qui peut amener à des limites de couches qui se superposent visuellement (fig. 77).

Enfin, sur des aspects très pratiques, nous avons pu constater que l'usage d'un stylet actif, c'est à dire un stylet alimenté par une source d'énergie, est quasiment indispensable pour dessiner sur SIG. Les stylets, de manière générale, ne sont pas particulièrement précis. Néanmoins, un stylet actif est reconnu par la tablette et affiche un curseur même à quelques centimètres de l'écran, ce qui permet d'améliorer leur précision en anticipant le point cliqué. Avec un stylet classique, aucun curseur n'apparaît à l'écran, le point est enregistré uniquement lors de la pression du stylet sur l'écran. De plus, malgré l'apparente légèreté des tablettes (env. 1,2 kg.), elles restent bien plus lourdes que simples feuilles de papier sur des planches à dessins. Ce poids est généralement problématique pour le relevé, qui nécessite de porter la tablette d'une main pour dessiner de l'autre. Il est donc nécessaire de prévoir un système de harnais pour maintenir l'équipement pendant la phase de relevé.



*77. Un problème rencontré avec le protocole de relevé numérique
pour la représentation de structures hautes ou profondes*

Enfin, nous avons pu achever grâce au SIG le protocole de relevé sur orthomosaïque, de l'acquisition à l'interprétation des vestiges. Cependant, comme nous l'avons déjà évoqué à plusieurs reprises, l'enregistrement archéologique ne repose pas seulement sur des éléments graphiques (plans, photographies, orthomosaïques, etc.). Il comporte également de nombreux compléments textuels sous forme d'inventaires et d'énoncés plus ou moins structurés. Cette étape nous a donc amené à tester des outils de saisie numérique pour l'enregistrement et la gestion de la documentation.

3.5. Outils de saisie : nouvelles interfaces, nouveaux relevés attributaires

3.5.1. Objectifs

Sur les chantiers de Bibracte, comme sur de nombreux autres chantiers archéologiques, l'enregistrement des observations et des interprétations préliminaires sous forme textuelle se déroule en deux phases. La première phase est caractérisée par la constitution de la documentation archéologique à l'aide de formulaires papiers normalisés sur le terrain : les fiches UF, les inventaires des photos, des minutes, des prélèvements, *etc.* Cette documentation est ensuite numérisée dans une seconde phase, soit au fil du chantier en fin de journée, soit directement à la fin de ce dernier. À Bibracte, une attention particulière a été portée à l'homogénéisation du rendu graphique des formulaires papiers et de l'interface du système de gestion bdB, afin de faciliter la phase de numérisation.

Dans ce cadre, l'apport de dispositif numérique sur le terrain pour gérer la documentation textuelle avait deux objectifs principaux. Le premier, très pragmatique, était d'accélérer et d'améliorer l'enregistrement de la documentation en supprimant la transition du support papier au support numérique. En effet, la numérisation manuel des fiches papiers est souvent laborieuse, ce qui augmente le risque d'erreurs d'inattention dans la documentation finale. Nous souhaitons donc faire évoluer l'enregistrement papier vers un format numérique plus ergonomique, sans pour autant perdre les avantages du papier : la conservation des données sur la longue durée et la souplesse dans l'enregistrement. De plus, Bibracte possède déjà un système de gestion de la documentation de fouille avec bdB. Les outils sélectionnés devaient agir comme des modules complémentaires adaptés aux travaux de terrain tout en étant capable de produire de la donnée normalisée qui puisse être par la suite reversée dans bdB.

Le second objectif s'inscrivait directement dans le programme Bibracte Numérique, avec la volonté d'utiliser l'informatique pour créer automatiquement du lien entre les informations nouvellement créées. Ce second point rejoint un des objectifs du projet 3, qui cherchaient à comprendre "*comment passer, avec les archéologues [...], de restitutions en langage vernaculaire, fluides, mais déconnectées des données, à des encodages numériques fidèles aux discours comme aux données et permettant de les relier (interopérabilité), pour finalement leur permettre de déléguer à la machine une partie de leurs capacités de mémorisation et d'exploitation (classement et calcul) ?*" (Reich *et al.* 2022, p. 37). Leur souhait était de

prendre en compte toutes les données, sans contraindre les archéologues à l'utilisation d'un outil ou d'une méthode particulière.

Nous avons donc éprouvé l'interopérabilité de deux outils de saisie et de gestion de documentation de fouille avec bdB : EDArc de l'Inrap et ODS de la société *9 heures précises*. Pour cela, nous avons doublé l'enregistrement de trois campagnes de fouille, en constituant la documentation à la fois sur les fiches papiers et sur les outils numériques.

Nous nous sommes également intéressés à Syslat⁹⁸, le Système d'Information Archéologique développé historiquement par M. Py (Roure *et al.* 2021). Ce système présente plusieurs avantages : il repose sur presque 40 ans d'expériences, il est mis à disposition gratuitement et il est optimisé pour fonctionner sur smartphone. Malheureusement les modules de saisie qui composent le système, iSyslatPro et iSyslat, ne peuvent pas fonctionner indépendamment d'une base Syslat. Toute l'architecture du système est centrée sur cette base de données pré-structurée, qui respecte ces normes d'enregistrement propre. Ors Bibracte possède déjà son propre système de gestion et d'enregistrement avec bdB. Nous ne souhaitons pas créer un doublon dans l'enregistrement sans pouvoir le réintégrer ensuite dans bdB.

3.5.2. Choix des équipements

EDArc

EDArc (Enregistrement de Données Archéologiques) a pour vocation de gérer l'enregistrement archéologique sur le terrain (fig. 78). Le projet est né d'une collaboration de la DST et la DSI en 2014, et Ch. Tufféry en a assuré le suivi à partir de 2015 (Tufféry 2022, p. 132). Nous avons eu l'occasion de tester EDArc lors des campagnes de fouille 2018 à 2020 de la *domus* PC2, sur la base d'une convention-cadre de collaboration scientifique et culturelle entre l'Inrap et Bibracte. L'enregistrement a été effectué sur une tablette Dell Latitude 7212 fonctionnant sous Windows, le même modèle que celui utilisé pour le relevé numérique.

Face à la diversité des outils alors employés à l'Inrap, l'objectif était de proposer un outil générique capable de s'appliquer aux nombreuses méthodologies en œuvre au sein de l'institut. EDArc n'a pas pour vocation de devenir le nouvel outil central capable de gérer l'intégralité de l'enregistrement, de la gestion de la documentation de fouille à l'Inrap. Il intervient

98 <http://www.syslat.fr/> ; consulté le 22/10/2023

au contraire comme un outil visant à enregistrer les données essentielles, pour répondre aux besoins communs des équipes de fouille et des spécialistes amenés à manipuler la documentation de fouille. Chaque intervenant est ensuite libre d'exporter les données produites dans EDarc vers ses propres outils, selon les domaines et les besoins. EDarc met ainsi l'accent sur l'interopérabilité plutôt que sur la gestion centralisée dans un seul logiciel.

Gestion Enregistrements Listes ?

Intervention : 921 Année : 2019 Chantier : 41 Equipe : Chantier école Auteur(s) de la fiche : Charly Chateau U

Nouvelle UF-UFÉ Générer numéro Enregistrer Supprimer

GENERAL (1/3)

Numéro d'UF : << < 378 > >> Choisir un numéro dans la liste

Type de l'UF : UF positive (couche)

Si UF englobante, choisir les UF concernées : U Si partie d'UF englobante, UF englobante : 384

LOCALISATION/DESCRIPTION (2/3)

Si UF positive :

Couleur de la matrice : Jaune orangée

Nature de la matrice : limon compact homogène Constituants grossiers : cailloutis

Si UF négative :

Forme en plan : Forme en profil : Orientation :

Longueur/Diamètre : 6,5 Largeur : 2,35 Haut./Prof./Epaiss. :

STRATIGRAPHIE/INTERPRETATION (3/3)

UF antérieures 347 379 433 437 ±

UF contemporaines 136 ± "Equivalence" (synchronisme sans continuité observable)

UF postérieures 435 ±

Datation : U Phase stratigraphique :

Contexte mobilier :

Ensemble clos UF de l'ensemble clos : U

Interprétation normalisée : PA71 - Sol construit

Interprétation et commentaires :
niveau de sol antique bordant la façade sud de la domus, équivalent à 136, observe dans la tranchée sud lors de la campagne 2019, la fouille s'est arrêtée sur ce niveau, il subit une dépression importante dans sa moitié sud, du à la présence d'un chablis ?

Nouvelle UF-UFÉ Générer numéro Enregistrer Supprimer

Photo Minute
 Photogram. Mobilier
 Enregistr. fini

Commentaires:
Fiche num. remplie par Nils Scavone.

Date d'ouverture : (AAAA-MM-JJ HH:MM)
2019-7-31 12:1

Date de fermeture : (AAAA-MM-JJ HH:MM)
2019-7-31 12:1

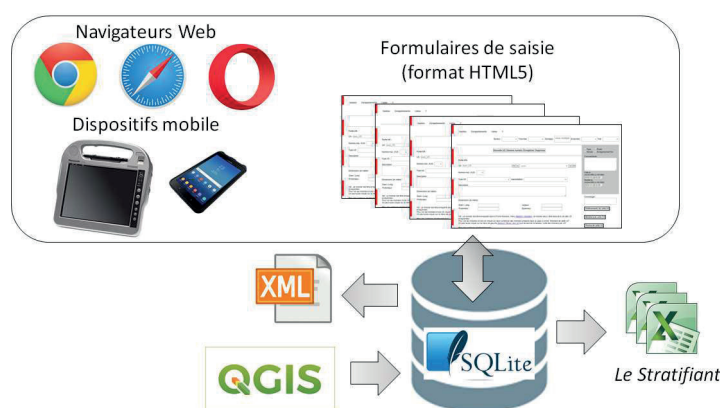
Photos de cette UF/UFÉ

Minutes de cette UF/UFÉ

78. Formulaire principal de la version EDarc pour Bibracte : l'enregistrement des UF

Dans cette perspective, EDArc possède une architecture extrêmement légère et évolutive. L'application ne requiert pas d'installation particulière, il s'agit simplement d'un dossier qui accueille plusieurs fichiers HTML5 (*HyperText Markup Language 5*). Le HTML est un des principaux langages informatiques à l'origine des pages web. En plus d'être un format ouvert et un standard mondialement reconnu, il a l'avantage d'être relativement compréhensible, même pour les personnes débutant en programmation informatique. Pour preuve, l'un des concepteurs principaux d'EDArc, Ch. Tufféry s'est auto-formé en même temps qu'il faisait évoluer le logiciel (Tufféry 2022, p. 132). Dans le cas d'EDArc, le HTML5 permet de construire les formulaires de saisie pour effectuer l'enregistrement archéologique.

Les formulaires de saisie viennent ensuite alimenter une base de données SQLite, un langage de gestion de base de données libre largement partagé à travers le monde. L'ingéniosité d'EDArc repose sur le fait d'exploiter les navigateurs internet Chrome, Chromium, Safari et Opera comme interface graphique, car ils installent tous ce type de base de données par défaut (fig. 79). L'utilisateur n'a donc pas besoin d'autre chose qu'un navigateur internet, déjà présent sur la grande majorité des machines, quel que soit leur système d'exploitation. Il suffit alors de glisser-déposer un fichier dans le navigateur pour commencer la saisie. EDArc ne requiert pas non plus de connexion particulière, le navigateur étant simplement utilisé pour son interface graphique et sa base de données SQLite.



79. Architecture générale d'EDArc (extrait de Tufféry 2022)

Il est également possible d'accéder au code de chaque formulaire, en HTML5. De ce fait, il est relativement simple de modifier le visuel d'EDArc, afin de l'adapter en fonction des besoins, sans pour autant toucher à sa structure interne. Il a donc été possible de mettre en place une interface adaptée à la logique d'enregistrement de Bibracte dès les premiers tests en 2019, et de mettre deux fois à jours cette dernière en fonction de nos retours.

L'application s'organise selon trois menus principaux : *Gestion*, *Enregistrements* et *Listes*. Le premier onglet *Gestion* permet d'atteindre toutes les fenêtres pour gérer la base de données : l'initialisation, la documentation (métadonnées), le contrôle d'état et import/export de données. Le menu *Enregistrement*, comme son nom l'indique, permet de produire et de collecter les données archéologiques. Dans le cas de la version test pour Bibracte, il s'agissait des UF, des minutes et des photos. Pour le troisième onglet *Listes*, un nombre important de champs est géré par des listes de valeurs dans EDarc. Cette fonctionnalité permet à l'utilisateur de sélectionner une valeur au lieu de la saisir manuellement, ce qui facilite l'enregistrement et limite les risques d'erreur ou de doublon. L'onglet *Listes* regroupe et permet d'éditer toutes les listes de valeur.

La base SQLite permet également de créer un lien dynamique avec QGIS, afin de lier les entités spatiales du SIG avec la base de données, mais nous n'avons malheureusement pas eu le temps d'explorer en détail cette fonctionnalité.

Lors de l'export, les données se présentent sous la forme d'un tableur XSL (*eXtensible Stylesheet Language*), exploitable par les logiciels Microsoft Excel ou LibreOffice Calc. Le fichier centralise toute l'information enregistrée, chaque fenêtre de chaque menu correspondant à un onglet dans le fichier tableur. Les données peuvent d'ailleurs être éditées dans le tableur avant d'être réimportées dans EDarc. De plus, le format XSL est en réalité une mise en forme d'un fichier XML (*eXtensible Markup Language*), un métalangage informatique à la base de nombreux formats et protocoles d'échanges ouverts, tels que le SVG (*Scalable Vector Graphics*), l'OpenDocument, le Dublin Core ou encore le RDF (*Resource Description Framework*). Enfin, EDarc propose un modèle d'export XSL pour la mise en forme des données en vue de leur exploitation dans le Stratifiant de B. Desachy⁹⁹.

Selon nous, EDarc est une solution particulièrement ingénieuse pour mettre en place un outil d'enregistrement léger et ouvert à la modification, qui respecte les méthodes de chacun tout en gardant un socle d'enregistrement robuste. Dans son état actuel, le code d'EDarc

99 <https://abp.hypotheses.org/3965> ; consulté le 22/10/2023

n'est pas encore publié en libre accès. Nous nous joignons au souhait de son concepteur principal, Ch. Tufféry, de voir un jour cet outil diffusé, car il présente tous les éléments pour fédérer une communauté autour de son développement (Tufféry 2022, p. 134).

ODS

Le concepteur à l'origine du projet, B. Bonaventure, a choisi la même stratégie que B. Desachy pour la mise en place de son outil : utilisé un logiciel propriétaire, mais largement employé par la communauté archéologique, afin de simplifier la prise en main de l'outil. ODS¹⁰⁰ s'appuie donc sur le SGBD FileMaker, disponible sur Windows, macOS et iOS (fig. 80). Nous avons utilisé ODS avec FileMaker Pro version 16, mais l'outil fonctionne déjà à partir de la version 12. Il a été installé sur la même tablette Dell Latitude 7212 avec Windows que celle utilisée pour EDArc.



80. ODS fonctionne sur de multiples supports (image 9heuresprecises.com)

Le test a été réalisé en 2021, avec l'autorisation de B. Bonaventure, sur la fouille de la terrasse PC14, dirigée par l'équipe des universités de Franche-Comté, de Bourgogne et de Toulouse. La campagne a duré 4 semaines, auxquelles s'ajoutent 2 à 3 semaines pour la mise au propre de la documentation.

ODS a été conçu en priorité pour une utilisation en contexte préventif. L'outil met donc l'accent sur une prise en main rapide et intuitive, qui passe en priorité par l'ergonomie de l'interface de saisie. Selon nous, ce point est une réussite puisque ODS se démarque par un environnement logiciel clair et aéré, adapté à son utilisation sur le terrain. L'entrée se fait par un tableau de bord, qui reprend les différentes composantes des données couramment gérées en chantier : les unités d'enregistrement, les éléments prélevés et la documentation de fouille. Chaque composante ouvre un formulaire d'enregistrement pensé pour accompagner au maximum le fouilleur dans le renseignement des informations, en fournissant des propositions logiques à sélectionner. Les liens entre les données sont également bien gérés, ce qui rend la navigation particulièrement intuitive.

Les données textuelles sont "encapsulées" dans le projet FileMaker créé pour chaque nouveau projet, mais peuvent à tout moment être exportées au format tableur sous divers formats. Des modèles d'export sont également prévus vers le Stratifiant ou les inventaires normalisés demandés par les Services Régionaux de l'Archéologie.

Pour résumer, ODS est un logiciel pragmatique et efficace, pensé en priorité pour les utilisateurs. Il utilise pleinement les possibilités du numérique pour offrir une expérience d'enregistrement la plus naturelle possible, tout en structurant et en codifiant l'information en arrière-plan. Le plus grosse contrainte de cet outil reste selon nous son fonctionnement fondé sur l'outil propriétaire Filemaker. Une version qui fonctionne sur navigateur, ODS Castor, est en cours de développement pour palier à ce problème. Nous espérons que le modèle économique de ce futur outil s'orientera vers le libre, tout en offrant à l'équipe les moyens de continuer leur activité. Il pourrait s'agir par exemple de proposer des formations payantes, des contrats de maintenance, des offres de développement spécifiques, ou encore des contrats de déploiement sur mesure d'ODS dans de larges structures.

3.5.3. Acquis des expérimentations

Lors de l'utilisation des deux outils, nous avons pu rapidement constater l'intérêt du support numérique pour la documentation archéologique. Cette étape de la fouille est caractérisée par un nombre important de tâches mécaniques et redondantes, que les outils numériques sont par nature aptes à traiter efficacement. Par exemple, dans EDArc comme dans ODS, la réciprocité d'une relation stratigraphique est déclarée automatiquement. Si l'utilisateur renseigne que l'US 1 est sur l'US 2, les deux logiciels renseigneront automatiquement l'US 2 est sous l'US 1. Selon notre expérience, ce point représente la très grande majorité des erreurs

d'enregistrement avec la documentation papier. EDArc et ODS assurent d'autres tâches automatiques comme la gestion de liens dynamiques entre les données, ou de listes de vocabulaires pour assurer l'homogénéité des informations renseignées. Une fois le logiciel maîtrisé, ce qui ne demande pas plus de temps que de maîtriser la documentation papier, EDArc et ODS facilitent tous les deux l'enregistrement archéologique, et offrent un réel confort pour la gestion et la navigation dans le corpus de données.

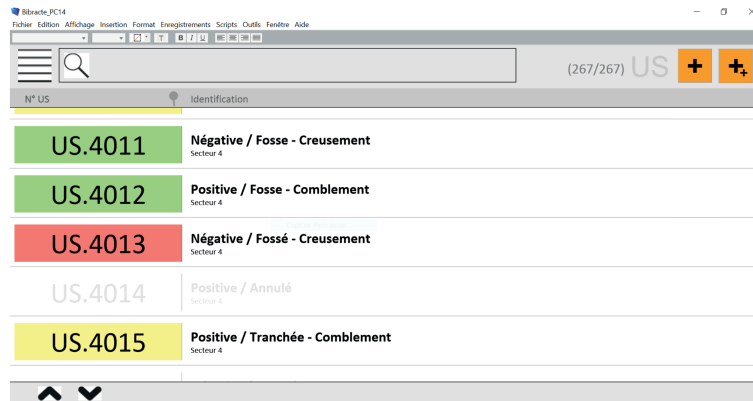
Ils proposent également tous les deux de visualiser l'état d'avancement de l'enregistrement, une fonctionnalité très pratique pour identifier les tâches qu'il reste à accomplir pour documenter chaque couche archéologique. EDArc est d'ailleurs plutôt explicite à ce propos, puisqu'il regroupe toutes ces informations dans un tableur *Reste à faire*, qui se met à jour automatiquement en fonction des informations renseignées (fig. 81). ODS gère l'état d'avancement grâce à un système de tableau de bord graphique et de code couleur, une solution claire et visuelle (fig. 82).

Reste à faire :

Liste des UF à terminer

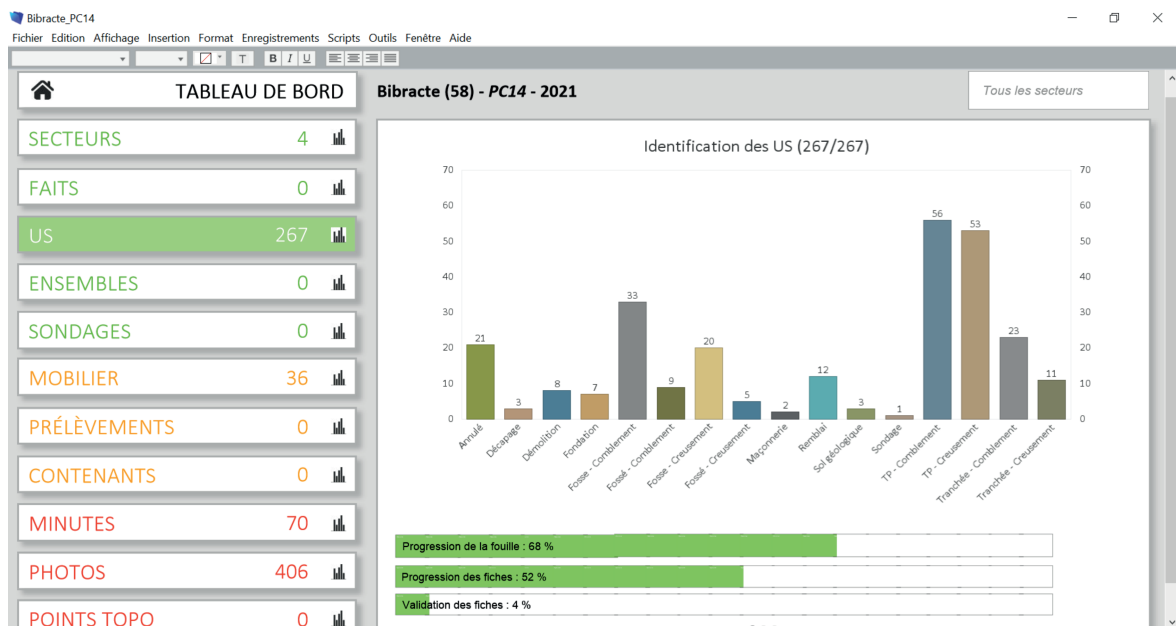
Numéro d'UF	Photo	Minute	Photogram.	Mobilier	Enregistrement fini	Voir la fiche d'UF
1	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	1
241	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	241
320	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	320
321	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	321
322	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	322
323	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	323
324	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	324
325	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	325
326	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	326
327	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	327
328	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	328
329	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	329
330	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	330
331	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	331
332	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	332
333	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	333
334	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	334
335	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	335
336	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	336

81. Exemple de gestion de l'avancement de l'enregistrement dans l'onglet "Reste à faire" de la version d'EDArc pour Bibracte



Utilisation en mode tablette

- Vert : Fiche US renseignée
- Jaune : Fiche US partiellement renseignée
- Rouge : Fiche US non renseignée
- Blanc : Fiche US annulée



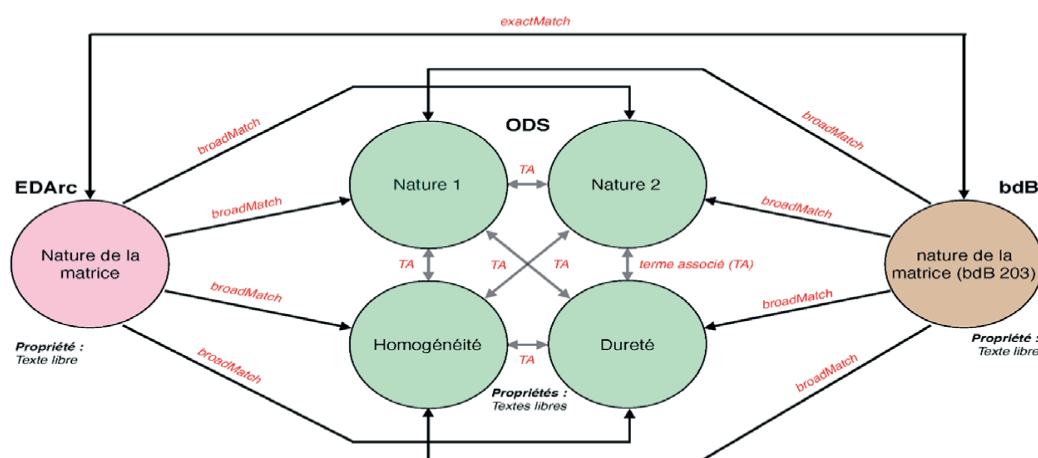
Utilisation en mode ordinateur

82. Exemple de gestion de l'avancement de l'enregistrement en mode tablette et en mode ordinateur sur ODS

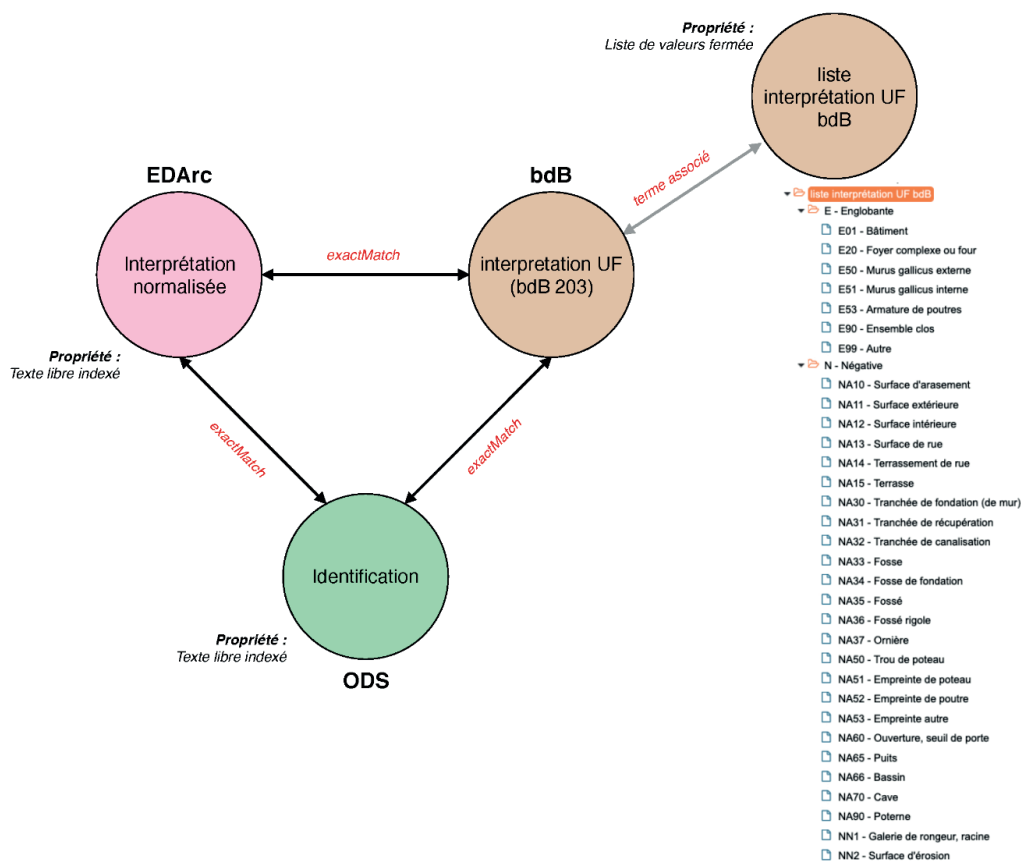
Pour autant, les outils, aussi efficaces soient-ils, ne travaillent pas à la place des archéologues. Si l'enregistrement n'est pas assidu et consciencieux, on retrouve les mêmes travers qu'avec l'enregistrement papier : des lacunes dans les informations renseignées, des doublons, des relations non déclarées, etc.

Une fois le chantier terminé, les outils de saisie ont permis d'éviter l'étape laborieuse de la numérisation manuelle des données textuelles. Toutefois, les données devaient être importée dans bdB, le système en place à Bibracte depuis plus de 30 ans. Grâce à ce travail, il a été possible de mesurer le niveau d'interopérabilité entre EDARc, ODS et bdB (Reich *et al.* 2022). Après analyse, il s'est avéré que les outils partagent une base commune très large, autant dans leurs logiques d'enregistrement que dans les similitudes du vocabulaire employé dans la description des champs d'enregistrement. Cependant, les divergences dans les modèles logiques de chacun, même mineures, sont restées complexes à compenser manuellement, rendant la démarche d'import/export des données fastidieuse.

Ce type d'expérience a donc mis en évidence le besoin de normaliser les échanges de données numériques entre outils. Pour cela, Bibracte a fait le choix de normaliser la manière de documenter les logiques d'enregistrement de chaque outil, plutôt que de standardiser les outils en eux-mêmes. L'équipe réunit autour du projet 3 de Bibracte Numérique (la gestion documentaire, *cf. supra* partie 1.2.2. Bibracte Numérique) a ainsi proposé une méthode de documentation et d'alignement de modèles logiques d'outils de gestion de données archéologiques, en exploitant la structure de la norme des thésaurus polyhiérarchiques (fig. 83). Selon cette méthode, chacun est libre d'utiliser l'outil qu'il considère le plus adapté à ses besoins. L'important est que les producteurs de données renseignent le modèle logique de l'outil utilisé, et qu'ils se prêtent à l'exercice de proposer un alignement avec au moins un autre modèle logique (fig. 84).



83. Exemple de description et d'alignement de champs descripteurs d'EDARc, ODS et bdB (extrait de Reich *et al.* 2022)



84. Mise en relation du vocabulaire structuré en thésaurus de bdB avec les modèles logiques de EDArc et ODS (extrait de Reich et al. 2022)

Actuellement, les premiers prototypes ont été réalisés à partir de l'outil Opentheso¹⁰¹. Cette méthode représente néanmoins un des éléments centraux de SIAMOIS (*Système d'Information Archéologique Mutualisé et Ouvert reposant sur l'Intelligence Sémantique*), un projet très récent soutenu par le plan d'investissement France 2030 auquel Bibracte a associé plusieurs grands acteurs de l'archéologie française¹⁰². Avec ce projet, Bibracte et ses partenaires souhaitent s'appuyer sur les résultats de Bibracte Numérique pour mettre en place les solutions techniques pour codifier les données archéologiques, indépendamment des méthodes ou des outils utilisés. L'ambition à terme est de proposer un Système d'Information Archéologique personnalisable, évolutif, interopérable et libre, couvrant tout le cycle de vie de la donnée. En ce sens, Il devra être capable de s'adapter aux spécificités méthodologiques de chacun, tout en s'appuyant sur les possibilités offertes par la technologie numérique pour optimiser l'échange de données.

101 <https://opentheso.hypotheses.org/> ; consulté le 22/10/2023

102 <https://www.bibracte.fr/actualite/bibracte-laureat-du-programme-dinvestissement-france-2030> ; consulté le 22/10/2023

Pour revenir aux outils de saisie, dans les deux cas, l'expérimentation s'est faite sur une seule tablette. Il est possible d'utiliser ODS sur plusieurs terminaux simultanément, mais Bibracte ne possédait pas de solution réseau clé en main au moment des tests. Le support unique n'a pas significativement impacté l'enregistrement, mais il est apparu qu'il pouvait freiner la formation des étudiants aux techniques de documentation archéologique. Avec les fiches papiers, chaque étudiant est libre de remplir sa documentation de fouille avant de la faire contrôler et de la joindre à la documentation générale. Avec un seul dispositif numérique, le temps nécessaire à chaque fouilleur pour remplir sa documentation est souvent incompatible avec les besoins réels d'enregistrement sur les chantiers. Dans ces conditions, le risque est que l'usage de la tablette se limite à l'équipe encadrante, une pratique en contradiction avec les aspects pédagogiques défendus sur les chantiers universitaires qui interviennent sur le mont Beuvray. Ainsi, qu'il s'agisse de papier ou de numérique, il apparaît nécessaire de multiplier les dispositifs d'enregistrement sur le terrain à Bibracte.

Un autre point limitant, lié à la nature du support, a émergé lors d'une discussion sur les outils de saisie numériques avec une collègue archéologue responsable d'opération. Lors d'une fouille, les archéologues sont confrontés à de multiples informations, qu'il est nécessaire de systématiquement réintégrer dans un contexte plus général, afin de construire *in fine* des interprétations sur l'évolution du site. Les chercheurs doivent donc se prêter à l'exercice de synthétiser constamment leurs observations, et formuler presque simultanément des hypothèses afin d'orienter la stratégie de fouille. Dans ce processus, la documentation archéologique est utilisée pour établir une vue d'ensemble des données. Elle permet à l'observateur de poser les différents éléments pour construire un raisonnement archéologique. Ainsi, nombreux sont les archéologues qui ont appris à "penser avec leur stylo" sur le terrain. C'est par l'écriture que nous clarifions, que nous structurons, que nous vérifions la cohérence de nos observations et de nos premières déductions. La psychologie cognitive a déjà démontré que le passage du stylo au clavier peut avoir des répercussions sur la construction de textes au sens large et sur les processus d'apprentissage (Chang *et al.* 2015 ; Tricot 2020, p. 17-19 ; Bouriga 2020, p. 61-80). Il est difficile d'estimer dans notre cas l'effet qu'il peut avoir sur les mécanismes déductifs des archéologues, mais il y a fort à parier qu'il ne soit pas sans conséquence.

Pour palier à ce problème, une solution pourrait être d'utiliser des dispositifs d'enregistrement équipés de stylets numériques avec reconnaissance automatique de l'écriture manuscrite. Le stylet permet ainsi de la sensation de rédaction, à condition que la surface reste

assez proche du papier, et la reconnaissance automatique de l'écriture manuscrite permet d'encoder directement les caractères dans un format interopérable (UTF-8). Durant nos travaux de recherche, nous avons eu l'occasion de tester une tablette à papier électronique pour la prise de note (remarkable). La sensation est effectivement très proche du papier, mais le format de ce modèle de tablette était selon nous trop petit. Comme il n'est pas possible d'étaler côte à côte les "pages" de note, comme des feuilles de papier, nous avons trouvé ce facteur particulièrement limitant pour parcourir rapidement les informations enregistrées. Aussi, si la reconnaissance automatique marchait très bien pour caractères alphanumériques, le logiciel n'était pas capable de retranscrire correctement l'agencement du texte, et encore moins pour les croquis, parfois essentiels pour comprendre les notes.

Enfin, si les deux outils sont prometteurs, et plus particulièrement EDArc pour sa légèreté et son adaptabilité, aucun des deux n'est libre. Pourtant, la meilleure solution pour assurer le suivi et la pérennité d'un outil repose sur sa base d'utilisateurs. Plus un logiciel est utilisé, plus sa communauté en assure l'évolution logique et technologique.

3.6. Bilan sur les expérimentations numériques à Bibracte

En définitive, les expérimentations menées durant les campagnes 2017-2020 ont permis de livrer un protocole de relevé numérique opérationnel pour les chantiers de Bibracte. Ce dernier présente plusieurs avantages. Premièrement, l'ensemble de la démarche peut être réalisée directement sur le terrain, face aux vestiges à enregistrer. Aucune action n'est décentralisée au centre. Deuxièmement, en ce qui concerne sa mise en œuvre, il a été conçu pour fonctionner avec des dispositifs numériques déjà exploités dans les processus archéologiques (appareil photo numérique, ordinateur, tachéomètre). Il ne nécessite pas non plus l'achat de logiciel particulier, car tous les outils sont libres et supportés activement par des communautés de développement (en dehors des outils de saisie, qui ne sont pas nécessaires pour le relevé). De plus, les outils spécialisés (photogrammétrie, SIG) s'adressent en priorité à un public scientifique, et mettent en avant la précision des résultats, ainsi que la transparence et la reproductibilité des manipulations. Troisième et dernier point, ce protocole ne nécessite pas de niveau d'expertise particulier en informatique (programmation) ou en ingénierie (calculs complexes, fabrication d'appareils spécifiques).

Nous souhaitons aussi revenir sur une idée préconçue que nous avons maintes fois entendue : le passage au numérique n'entraîne pas de gain de temps sur la phase d'enregistrement

sur le terrain, du moins pas dans un premier temps. En effet, si l'on prend en compte toutes les étapes du protocole, depuis le levé topographique jusqu'à la centralisation des données dans le SIG, la démarche est rarement plus rapide que le dessin sur papier millimétré. Il en va de même pour les outils de saisie de terrain. La navigation dans les données est en général plus confortable, mais le temps de réflexion et de rédaction reste sensiblement le même.

Par contre, on observe dans la plupart des cas une amélioration de la qualité des enregistrements. Par exemple, dans le cas du protocole de relevé numérique, le support orthophotographique offre une précision centimétrique en tout point de l'espace représenté. Parfois même au point de modifier nos pratiques. Par exemple, avec le relevé manuel classique, les archéologues représentent leurs observations sur du papier millimétré standardisé, sans aucun lien direct avec la réalité. Ils ont la liberté de choisir les éléments à figurer, pour créer une représentation des vestiges qui correspond à leurs interprétations. À l'inverse, l'orthomosaïque est un support bien plus fidèle à la réalité. Elle offre un niveau de détail si poussé qu'il incite parfois à approfondir ses propres observations : "dois-je intégrer cette pierre à cette US ?", "Ce changement de couleur sur l'image est-il observable sur le terrain ?" En pratique, ce type de détail s'est souvent révélé excessivement minutieux, complexifiant le dessin sans pour autant apporter de nouvelles informations à l'étude du site. Cependant, dans quelques cas, la mise en cohérence du levé orthophotographique et du relevé technique a permis d'approfondir notre compréhension des qualités intrinsèques de certaines US, telles que les éléments caractérisant une bordure de voie ou les contours de surfaces d'épandage. Cette nouvelle précision, en tout point de l'espace représenté, change notre rapport au relevé. Elle nous pousse à remettre en question nos automatismes de dessin technique et à acquérir de nouveaux réflexes de travail.

Un dernier point important, le relevé numérique offre un gain de confort aux personnes qui ne sont pas à l'aise avec le dessin technique. L'objet à relever étant déjà représenté, il suffit d'en suivre les contours sur l'image pour l'identifier. Cette méthode est d'ailleurs si naturelle pour le dessinateur qu'elle peut conduire à l'une des erreurs les plus courantes observées avec le protocole. Certains utilisateurs ont en effet tendance à se détacher de leurs observations sur le terrain pour effectuer le relevé uniquement à partir de l'orthomosaïque, et ce même s'ils se trouvent face à la couche archéologique. Selon nous, cette situation est vraiment problématique, car l'image, aussi proche soit-elle de la réalité, n'en reste pas moins une représentation, avec ses caractéristiques propres (définition, luminosité, couleurs, aberrations). Rajouter un médium intermédiaire entre l'observateur et son objet d'étude a forcé-

ment des répercussions sur la qualité du résultat final. Ainsi, il est primordial d'utiliser le levé orthophotographique comme on utiliserait le papier millimétré, en reportant des éléments observés et compris sur le terrain.

Pour conclure, le protocole propose une méthode permettant de faciliter la critique des relevés archéologiques. En effet, comme nous l'avons déjà souligné, le relevé comporte une part d'interprétation. Avec l'association de l'orthomosaïque, bien plus proche de la réalité qu'une photographie, et du SIG, qui permet d'organiser les couches interprétatives, le protocole permet de justifier les données qui ont mené à cette interprétation. Afin d'augmenter le potentiel de réutilisation de la donnée produite, une dernière étape reste à réaliser : la documentation du corpus par un Plan de Gestion de Données.

3.7. Diffusion des données primaires d'un chantier archéologique : l'exemple de la *domus* PC2

3.7.1. Présentation de la donnée

Le corpus de données présenté ici est issu des 4 campagnes de fouille de la *domus* PC2 durant lesquelles se sont déroulées les expérimentations autour des technologies numériques, de 2017 à 2020. Le chantier a débuté en 2016 et il est toujours en cours au moment où nous écrivons ces lignes, mais ces 4 années correspondent à notre présence sur le terrain en tant que responsable de la fouille. Au final, les archives de fouilles 2017-2020 sont composées de :

- 4 référentiels analytiques qui compilent les descriptions d'Unités de Fouille, les minutes (68 au total) et les diagrammes de Harris de chaque campagne. Pour l'année 2020, il comprend 752 UF, qui correspondent dans le système d'enregistrement de Bibracte, aux US (UF positive, UF négative) et aux faits et secteurs (UF englobante) ;
- 405 photographies de terrain ;
- 219 orthomosaïques (construites à partir de 4846 photos) ;
- 1 projet QGIS qui renseigne 566 UF (les UF observées uniquement en coupe ne sont pas représentées sur le SIG).

Ces données sont actuellement diffusées de données sur l'entrepôt Nakala de Bibracte, accessible à l'adresse suivante : <https://nakala.fr/10.34847/nkl.8d81q3es>. Le référentiel analytique est pour l'instant le procédé employé pour la diffusion de données numériques archéologiques à Bibracte (fig. 85) (*cf. supra* 1.2.1. Prog de recherche européen). Cependant, le projet Bibracte Numérique et cette recherche doctorale ont permis de mettre en évidence que des solutions plus adaptées existent, comme les technologies du Web des Données (Bizer *et al.* 2009). C'est pourquoi cette URL est un lien temporaire, qui restera actif le temps que le projet SIAMOIS¹⁰³ (*Système d'Information Archéologique Mutualisé et Ouvert reposant sur l'Intelligence Sémantique*) prépare les processus pour le futur de la diffusion scientifique à Bibracte, avec des données documentées, libres et interopérables.

Pour le décrire et faciliter sa réutilisation, nous nous sommes tourné vers le modèle de PGD conçu par le consortium MASA pour les données archéologiques. Techniquement, le PGD est un document à même d'évoluer avec l'avancement d'un projet de recherche et comporter plusieurs versions (*cf. supra* 2.3. PGD). Toutefois, sachant que le chantier de la *domus* PC2 était déjà entamé avant notre arrivée dans l'équipe de fouille et que nous nous sommes formé à la démarche des PGD durant cette recherche doctorale, le document présenté dans la partie suivante ne présente qu'une version finale de PGD. Nous avons tout de fois eu l'occasion de construire un PGD au début d'un autre projet de recherche, celui du PCR *La confluence Saône-Doubs à l'âge du Fer (VIe s. av. J.-C. au Ier siècle de notre ère)* (Verriez 2020a).

3.7.2. Présentation du modèle de Plan de Gestion de Données MASA

Le modèle de PGD MASA a été conçu à partir d'un premier document développé au sein de l'Inrap en 2018. L'ensemble se matérialise par un formulaire à plusieurs onglets. Une première ébauche de PGD a été réalisée à partir du modèle "MASA – Modèle de PGD (français)" (version 3, publiée le 16 mai 2022), avant d'être mise à jour vers le modèle "MASA – Modèle de PGD (français) V2" (version 7, publiée le 5 mai 2023)¹⁰⁴. La nouvelle version présente 9 onglets. 5 onglets sont génériques à tous les modèles hébergés sur DMP OPIDoR (Informations générales, Contributeurs, Produits de recherche, Partager, Télécharger) et 4 onglets correspondent au modèle de MASA (*Contexte, Gestion des données, Description des données, À l'issue du projet*) (fig. 86).

103 <https://www.bibracte.fr/le-projet-siamois> ; consulté le 21/10/2023

104 https://dmp.opidor.fr/public_templates?page=1&search=masa ; consulté le 22/10/2023

RAPPORT ANNUEL 2020 DU PROGRAMME QUADRIENNAL DE RECHERCHE 2017-2020 SUR LE MONT BEUVRAY
 RÉFÉRENTIEL ANALYTIQUE
 LA DOMUS PC2 (INTERVENTION 953)

- B2018.41.262** (positive)
Interprétation : PA73 - Foyer simple
Description : charbons ; noir rougeatre ; argile rubefiée
Commentaires : foyer simple circulaire sur la couche cendreuse 228 au nord de la partie est du local x.
Datation : _
Relations observées : sert d'appui à 2018.41.171 ; s'appuie contre 2018.41.228
Représentée sur les minutes : _
Inventaire préliminaire du mobilier : _
- garder la chaleur par inertie.
Datation : _
Relations observées : sert d'appui à 2018.41.224 ; s'appuie contre 2018.41.228 ; s'appuie contre 2019.41.477
Représentée sur les minutes : _
Inventaire préliminaire du mobilier :
 - terre cuite architecture non conservée : 2360 g ; 24 reste(s)
- B2018.41.263** (positive)
Interprétation :
Description :
Commentaires : annulee
Datation : _
Relations observées : _
Représentée sur les minutes : _
Inventaire préliminaire du mobilier : _
- B2018.41.264** (positive)
Interprétation : PA4 - Remblai
Description : tca, pierres (5-10cm) ; jaune brun ; limon argileux
Commentaires : niveau sur lequel a stoppe la fouille de 2018 dans le local x, a l'est de la berme. couche épaisse découverte sous la couche charbonneuse 228. remblai de construction de la domus pc2 dans local x.
Datation : _
Relations observées : s'appuie contre 2018.41.124 ; s'appuie contre 2018.41.129 ; identique à 2018.41.135 ; est recoupé par 2018.41.151 ; équivalent de 2018.41.264 ; est recoupé par 2018.41.300 ; est recoupé par 2018.41.302 ; est recoupé par 2018.41.270 ; identique à 2019.41.327
Représentée sur les minutes : _
Inventaire préliminaire du mobilier : _
- B2018.41.265** (positive)
Interprétation : PA4 - Remblai
Description : clou, amphore, pierres ; brun ; limon heterogene compact
Commentaires : couche tres partiellement exploree dans local x, contre mur 123. d'abord pris pour comblement de tranchee bulliot 253. mais correspond a couche en dessous. equivalent de remblai 267 ?
Datation : _
Relations observées : est recoupé par 2018.41.253 ; équivalent de 2018.41.267 ; s'appuie contre 2018.41.123
Représentée sur les minutes : _
Inventaire préliminaire du mobilier :
 - bois non travaillé : 8 g ; 10 reste(s)
 - fer autre objet : fer plat ; 1 reste(s)
 - fer clous : 10 reste(s)
 - terre cuite amphore gardée : 31 reste(s)
 - terre cuite architecture non conservée : 520 g ; 9 reste(s)
 - terre cuite vaisselle : 7 reste(s)
- B2018.41.266** (positive)
Interprétation : PA71 - Sol construit
Description : cailloutis ; beige jaune ; mineral
Commentaires : niveau de sol construit du local x (mortier ou chaux). sol tres mal conserve, observe uniquement dans l'angle sud est de la partie a l'est de la berme du local x. dalle tres lisse et adree, sans decors.
Datation : _
Relations observées : sert d'appui à 2018.41.171 ; équivalent de 2018.41.227 ; s'appuie contre 2018.41.228 ; sert d'appui à 2018.41.237 ; est recoupé par 2018.41.253
Représentée sur les minutes : _
Inventaire préliminaire du mobilier : _
- B2018.41.267** (positive)
Interprétation : PA4 - Remblai
Description : tca, pierres (5-10cm) ; jaune brun ; limon argileux
Commentaires : niveau sur lequel a stoppe la fouille de 2018 dans le local x, a l'est de la berme. couche épaisse découverte sous la couche charbonneuse 228. remblai de construction de la domus pc2 dans local x.
Datation : _
Relations observées : s'appuie contre 2018.41.124 ; s'appuie contre 2018.41.129 ; identique à 2018.41.135 ; est recoupé par 2018.41.151 ; équivalent de 2018.41.264 ; est recoupé par 2018.41.300 ; est recoupé par 2018.41.302 ; est recoupé par 2018.41.270 ; identique à 2019.41.327
Représentée sur les minutes : _
Inventaire préliminaire du mobilier : _
- B2018.41.268** (positive)
Interprétation :
Description : noir ; limon argileux compact
Commentaires : couche noire dans local x observe uniquement dans la coupe ouest de la berme, au niveau du sondage realise a la pelle mecanique. equivalente aux couches cendreuses 228 a l'ouest du sondage et 258 a l'est de la berme. notee 228 dans la coupe dessinee du rapport 2018 (illu. 7).
Datation : _
Relations observées : équivalent de 2018.41.228 ; sert d'appui à 2018.41.171 ; équivalent de 2018.41.258 ; équivalent de 2019.41.396 ; remplit 2018.41.270
Représentée sur les minutes : _
Inventaire préliminaire du mobilier : _
- B2018.41.269** (positive)
Interprétation : PA4 - Remblai
Description : pierres (10-15cm), tca, gravier ; brun jaune ; limon heterogene
Commentaires : couche heterogene avec des nombreuses pierres dans la partie nord est du local v. contre le mur 120. a la fouille en plan, il a ete possible de diffencier cette couche de cinq autres (uf 246, 279, 290, 291, 296), mais impossible d'observer cette demarcation en coupe. ces niveaux tres pertubes semble etre des remblai pour combler l'espace v. englobe par 143.
Datation : _
Relations observées : s'appuie contre 2018.41.120 ; sert d'appui à 2018.41.246 ; est recoupé par 2018.41.248 ; sert d'appui à 2018.41.272 ; s'appuie contre 2018.41.279
Représentée sur les minutes : _

The screenshot shows the DMP OPIDoR interface for the project "Bibracte, Mont Beuvray. Campagnes de fouilles 2017-2020 de la domus PC2". The interface is in French and includes a navigation menu with options like "Informations générales", "Contributeurs", "Produits de recherche", "Contexte", "Gestion des données", "Description des données", and "À l'issue du projet". The main content area is divided into two columns. The left column, titled "Renseignements sur le projet", contains a checkbox for "projet de test, d'entraînement ou à des fins de formation", a text field for "Titre du projet" (filled with "Bibracte, Mont Beuvray. Campagnes de fouilles 2017-2020 de la domus PC2"), a text field for "Acronyme du projet" (filled with "PC2"), and a rich text editor for "Résumé du projet" containing a detailed description of the archaeological site and the project's goals. The right column, titled "Sélection des recommandations du plan", provides instructions and a list of organizations to select from, including CNRS, CY Cergy Paris Université, Digital Curation Centre, INRAE, and INRIA, each with a checkbox.

86. Présentation du modèle de PGD MASA

Un des objectifs principaux d'un PGD est de faciliter la réutilisation des données en exploitant les possibilités des outils numériques. Or les technologies numériques gèrent mal les imprécisions ou les manques. C'est pourquoi les entités nommées (personnes, institutions, données) ainsi que leurs qualificatifs doivent au maximum se rapporter à des référentiels, afin d'éviter les confusions par homonymie, les contresens, *etc.* Pour faciliter la rédaction du document, la plupart des questions présentent deux sources d'aide : des exemples, disponibles sous l'encart de réponse en cliquant sur *MASA exemple de réponse +* ; des *Commentaires & Recommandations*, à droite de l'écran, qui propose des conseils, des sources ou des référentiels sur lesquels s'appuyer pour construire sa réponse (fig. 87).

Responsabilités

Cette section doit être utilisée pour désigner les personnes en charge de la gestion des données pendant la durée du projet, contrairement au champ "Contact pour les données" de la phase "Renseignement sur le projet" qui désigne le responsable des données une fois le projet achevé.

[tout développer](#) | [tout réduire](#)

Responsable(s) de la gestion des données au cours du projet de recherche

B I U A **☰ ☱** **🔗** **📄**

Verriez, Quentin (q.verriez@bibracte.fr) / UMR 6298 ARTEHIS. <https://orcid.org/0000-0002-6823-4183>

Enregistrer

MASA exemple de réponse –

- Responsable des données CNRS : Hugo, Saturnin (saturnin.hugo@cnsr.fr) /USR 1000 du CNRS <http://orcid.org/0000-0003-4317-1800>
- Responsables des données l'INRAP : Edmond, Jeanne (jeanne.edmond@inrap.fr) / INRAP. <http://orcid.org/0000-0003-4317-1700>

Répondu il y a 3 mois par qverriez@gmail.com

Recommandations **Commentaires**

MASA

Le porteur du projet désignera un gestionnaire de données qualifié pour gérer les données collectées, traitées et analysées.

Le responsable de la gestion de données doit être impliqué dans le pilotage du projet (data stewardship), de sa création à son aboutissement. Il est responsable de la collecte, l'organisation et la conservation des données. Il peut être différent du responsable scientifique.

Dans le cas d'un projet associant des partenaires sur d'autres sites ou dans d'autres institutions, les responsabilités entre les partenaires doivent être précisées ici.

87. Les aides à la rédaction du modèle de PGD MASA

Le modèle n'étant qu'un moyen d'accompagner la rédaction d'un PGD, et non un document officiellement formalisé, il n'est pas obligatoire de répondre à toutes les questions. Certaines questions ne conviennent d'ailleurs qu'à des situations particulières, qui sont généralement identifiées par un astérisque ou dans l'énoncé de la question. Cependant, nous conseillons fortement aux auteurs de s'interroger sur les raisons qui les poussent à omettre une question, car le PGD permet de mettre en évidence des aspects souvent inconnus ou sous-estimés pour une bonne gestion du cycle de la vie de la donnée archéologique.

Les 3 premiers onglets, mis en place par l'Inist, l'institut à la base d'OPIDoR, amènent une série de questions pour identifier le PGD, son contenu et ses auteurs de la manière la plus normalisée possible. Cette première partie permet d'optimiser le traitement du PGD par des outils de moissonnage des données (*web scraping* ou *web harvesting* en anglais), qui consiste à extraire automatiquement les données d'un contenu, le plus souvent disponible en ligne,

pour l'exploiter dans un autre système (MESR 2022). ISIDORE¹⁰⁵ est un bon exemple d'outil de moissonnage. Il s'agit d'un moteur de chercheur pour les données en SHS développé par Huma-Num, qui moissonne un grand nombre de bases de données, telles que les portails HAL, theses.fr, Persée, Gallica ou encore OpenEdition Journals (Pouyllau *et al.* 2021).

Les 4 onglets suivants sont propres au modèle MASA. Le premier, *Contexte*, permet de caractériser en une douzaine de questions le cadre de création ou d'exploitation du jeu de données. On y retrouve quelques répétitions avec les onglets génériques de l'Inist, plus des informations spécifiques au domaine archéologique, comme la chronologie ou l'identification de l'opération archéologique le cas échéant. Le second onglet *Gestion des données* identifie en 4 questions les personnes responsables des données en cours et après le projet, ainsi que les moyens techniques et financiers alloués pour la création du corpus.

L'onglet suivant, *Description des données*, représente le cœur du modèle avec 33 questions. Une première partie s'applique à décrire les caractéristiques intrinsèques du corpus, soit l'origine, la structuration et le volume de données qu'il représente. Cette partie se termine avec la description des données, qui regroupe, selon nous, les questions requérant le plus d'investissement, mais aussi parmi les plus importantes pour maîtriser son corpus de données numériques : les méthodes d'acquisition, le format et le nommage des fichiers.

On trouve ensuite une partie sur les métadonnées (*cf. supra* 2.1.2. Métadonnées). Il s'agit d'un volet important, en 7 questions, qui commence par rappeler le but des métadonnées :

"décrire le document numérique pour assurer sa traçabilité et connaître :

- *les caractéristiques techniques du fichier ;*
- *son **identifiant**, et le **contenu intellectuel** auquel il correspond ;*
- *le **contexte** dans lequel il a été produit ou reçu ;*
- *sous quel régime de propriété se trouve le fichier.*
- *Elles sont la plupart du temps rassemblées dans un "**fichier de métadonnées**" associé au fichier qu'elles décrivent".*

Les métadonnées sont primordiales pour le moissonnage, donc la visibilité, l'accessibilité et la réutilisation des données. En tant qu'informations principalement à destination des machines, leur construction est de fait très normalisée. C'est pourquoi les questions du modèle insistent sur la description des standards de métadonnées sélectionnés et les méthodes et outils de production employés pour les produire.

La dernière partie de l'onglet *Description des données* regroupe un ensemble de questions sur les moyens techniques et juridiques mis en place pour assurer leur conservation et définir les conditions de diffusion : stockage, sécurité, accès, données sensibles et cadre juridique des données.

Le quatrième onglet du modèle MASA se nomme *À l'issue du projet*. Il développe en 9 questions tous les aspects à prendre en compte pour l'archivage et la dissémination des données, soit respectivement la conservation de l'intégrité des données sur le long terme et leur partage à large échelle selon des conditions prédéfinies. Contrairement à ce que le nom de l'onglet suggère, il peut être complété dès la construction d'un projet de recherche. Dans ce cas, le modèle MASA permet d'anticiper les différentes finalités des données, qui sont généralement déterminantes pour définir l'environnement numérique du projet.

Les deux derniers onglets, *Partager* et *Télécharger* sont des outils de gestion pour les auteurs du PGD. Le premier sert à inviter d'autres collaborateurs à éditer, commenter ou simplement lire le document, et définir la visibilité du PGD (privé, restreint à un groupe, public). Le second permet de formater et de télécharger une version du PGD construit à l'aide du modèle.

3.7.3. Le PGD de la fouille de la *domus* PC2 : retour d'expérience

Le PGD des campagnes de fouille 2017-2020 de la *domus* PC2 accompagne les archives de fouilles disponibles sur Nakala¹⁰⁶. Il présente un exemple de PGD abouti pour décrire un jeu de données issu créé dans le cadre d'une fouille en archéologie programmée. Après nous être prêté à l'exercice proposé par le modèle de PGD MASA, il apparaît que le formulaire est bien construit et plutôt explicite. Il faut cependant nuancer ce constat : nous sommes aujourd'hui à l'aise avec les nombreux concepts mentionner dans ce type de document, d'autres chercheurs pourraient livrer un retour d'expérience différent. La nouvelle version du modèle MASA a amélioré son organisation générale, rendant le déroulé des questions plus fluide pour le rédacteur.

106 <https://nakala.fr/10.34847/nkl.8d81q3es> ; consulté le 22/10/2023

Selon notre expérience, plusieurs points sont à prendre en compte pour gérer au mieux la rédaction d'un PGD en archéologie. Tout d'abord, la convention de nommage des fichiers apparaît comme un élément central de la bonne gestion des données numériques. Au début des expérimentations, nous avons concentré nos efforts sur l'arborescence de classement des fichiers, en partant du postulat que des dossiers bien organisés avec des noms explicites permettraient de faciliter le classement des données. Ce postulat s'est révélé en partie vrai. Dans le cas d'un projet accueillant de multiples producteurs de données, l'arborescence de classement permet de garder de la cohérence dans la structuration du corpus. Néanmoins, il reste très complexe d'évaluer avec précision le nombre et la pertinence des dossiers à créer en amont d'un projet. Dans la pratique, l'arborescence de classement a donc tendance à "s'étoffer" naturellement, avec pour conséquence paradoxale de complexifier la lisibilité et la navigation dans les fichiers. À l'inverse, une convention de nommage n'impose pas de structure des données. Si chaque fichier possède un nom unique et explicite, il est possible de rassembler et de classer automatiquement toutes les données dans un seul dossier. Les utilisateurs sont ensuite libres de les trier comme ils le souhaitent.

Un autre aspect important auquel il faut porter attention est le niveau de finesse des métadonnées. Pour les données de la *domus* PC2, nous avons fait le choix de rassembler tous les fichiers en une seule archive documentée par un fichier de métadonnées au format Dublin Core. En effet, à l'heure actuelle, il n'existe pas, en archéologie, d'outil ou de méthode pour produire efficacement des métadonnées pour chaque fichier. Produire une archive globale reste toutefois une démarche cohérente, car le jeu de données a été produit dans le cadre d'une problématique générale, et comme nous l'avons déjà évoqué, les données archéologiques n'ont généralement que peu de valeur si elles sont isolées les unes des autres.

Enfin, nous souhaitons insister sur l'importance de mettre en place un PGD au moment de la conception du projet de recherche. Durant cette phase, le PGD permet d'anticiper plusieurs aspects souvent sous-estimés pour la gestion du cycle de vie des données numériques en archéologie. On parle ici des règles d'échange pour assurer un partage de données fluide entre les chercheurs du projet, mais aussi de sensibiliser aux risques de perte d'information pour des raisons techniques ou humaines. Le PGD implique ainsi les chercheurs dans la construction de stratégies de conservation des fichiers. Il permet également de définir les règles de diffusion des fichiers – qu'est-ce qui va être diffusé, sous quelle forme et selon quelles conditions ? – avec toutes les personnes impliquées dans le projet.

3.8. Le développement des chantiers archéologiques numériques à Bibracte

Les expériences menées sur le chantier de la domus PC2 ont démontré qu'il est possible de créer de nouvelles données numériques adaptées aux problématiques scientifiques des archéologues sur le terrain, en adéquation avec les principes de la science ouverte, le tout en respectant le temps imparti à la campagne de fouille. Ce résultat a toutefois été rendu possible par un suivi constant du processus de création des données numériques sur le chantier. Il est certain que notre double position de responsable scientifique et de chargé des expérimentations numériques a facilité ce suivi. Nous avons pu adapter le déroulement de la fouille, afin de dégager le temps nécessaire à la formation des fouilleurs aux différents dispositifs et logiciels mis en œuvre. Nous avons également pu mesurer rapidement les problèmes engendrés par les technologies déployées et prendre les décisions nécessaires pour ajuster le protocole. Ainsi, pour arriver aux mêmes résultats dans un autre contexte, il nous semble essentiel qu'au moins une partie de l'équipe de fouilles, de préférence des membres avec un pouvoir de décision, porte un intérêt certain aux nouvelles méthodes d'enregistrement numérique et à la science ouverte. Ces personnes doivent posséder les connaissances suffisantes pour assurer la formation des autres archéologues, la maintenance des outils et le suivi de la qualité des données. Si de telles personnes ne sont pas présentes sur le terrain, il existe un risque pour que les méthodes d'enregistrements numériques ne soient pas appliquées. Par exemple, l'étape du relevé sur SIG a été progressivement abandonnée sur le chantier de la domus PC2 à partir de 2021, pour revenir à des techniques de dessin plus traditionnelles, bien que l'équipe encadrante ait été formée à cette technique.

Ce travail expérimental a également confirmé la difficulté pour bdB, le système de gestion et d'archivage de Bibracte, à gérer les types de données complexes produites par la photogrammétrie et les SIG. Les technologies numériques permettent aujourd'hui une gestion et un traitement avancé de ces informations spatiales numériques qui se présentent à la fois sous forme de texte et d'images matricielles et vectorielles. Or, comme nous l'avons mentionné, l'archéologie et l'espace sont deux concepts étroitement liés. Les usages de l'imagerie numérique et des SIG vont donc continuer à se développer dans les années à venir en archéologie, en même temps que les technologies vont se perfectionner. Il devient alors indispensable de faire évoluer le système bdB vers des technologies capables de gérer de l'information spatiale numérique. D'une part pour freiner la perte de données qui a déjà lieu à cause de l'absence

de catégories d'enregistrement, mais aussi pour intégrer les entités géographiques dans les éléments sujets à requête, car elles peuvent centraliser un grand nombre d'informations archéologiques sous des modes de représentation intuitifs pour les utilisateurs.

En ce qui concerne la diffusion des archives de fouille, en nous basant sur les principes FAIR (Findable, Accessible, Interoperable, Reusable), nos recherches ont permis de mettre en évidence que l'actuel processus à Bibracte est capable de produire des jeux de données trouvables et accessibles, mais limité en matière d'interopérabilité et de réutilisation. Le modèle du référentiel permet en effet de partager un grand nombre de données textuelles structurées et de documenter l'existence d'archives physiques, sous la forme d'un document partagé sur les entrepôts en ligne de Bibracte. Cependant, son format PDF et l'organisation des données sont optimisés pour la consultation par les utilisateurs, mais contraint très fortement l'exploitation automatique des données. La production de données primaires aux formats libres, interopérables et documentés sur les chantiers archéologiques de Bibracte devient alors une nécessité pour les prochaines évolutions de bdB. Ces formats permettent l'intégration des données directement dans le système, plutôt que de les représenter sous forme d'inventaires textuels. Il s'agit d'une étape fondamentale pour automatiser le processus de diffusion des archives de fouilles, en plus de donner tout son sens à la production de Plans de Gestion de Données (PGD).

Enfin, il est important de noter qu'avec les technologies numériques testées, leur apparente efficacité entraîne quasi systématiquement des dérives dans les pratiques d'enregistrement archéologique. Face aux dispositifs numériques, nous avons pu observer que les utilisateurs avaient une tendance naturelle à abandonner une partie de leurs démarches d'analyses. En effet, les levés effectués à l'aide de technologies numériques offrent des résultats si proches d'une représentation réaliste du monde que certains archéologues estiment ne pas avoir besoin de les compléter par des annotations interprétatives. Cette tendance interroge sur la représentation des technologies numériques en tant qu'outil dans nos pratiques archéologiques. Quelles sont les raisons qui nous poussent à substituer nos compétences scientifiques à des dispositifs numériques ? Ce type de questionnement nous a amenés à considérer les technologies numériques au-delà du champ de la technique, pour explorer les implications sur l'utilisateur et son environnement au sens large.

QUATRIÈME PARTIE

Repenser l'archéologie numérique : analyse et recommandations

4.1. L'humain e(s)t la machine : critique de l'usage des technologies numériques en archéologie

La conception du protocole de relevé numérique sur le chantier de la *domus* PC2 représente plusieurs années de réflexions, de tests plus ou moins fructueux et de mise en routine de nombreux outils numériques sur une fouille archéologique (fig. 88). Nous avons également pu recueillir les expériences quotidiennes des multiples fouilleurs de PC2 et d'autres chantiers sur lesquelles nous nous sommes investi, telles que la plateforme PC14 ou l'occupation protohistorique de l'éperon de Blis à Eymet (Dordogne, France). Enfin, ce sujet a suscité de nombreuses discussions avec des collègues archéologues, à Bibracte et ailleurs, afin de confronter nos observations avec le plus de réalités professionnelles possible. Ce temps de recherche a été pour nous l'occasion d'explorer les avantages, mais aussi les limites de ces technologies au-delà de la production de connaissances archéologiques.



4 campagnes de fouilles



8 mois de fouilles
16 mois de post-fouille



70 fouilleurs ont effectué les manipulations



219 acquisitions orthophotographies,
pour environ 150h de travail



566 Unités Stratigraphiques représentées sur SIG



2 outils de saisie numérique testés

88. Le protocole de relevé numérique établi sur PC2 en quelques chiffres

4.1.1. La place de la machine pour l'archéologue

Pour comprendre notre démarche, il est nécessaire de revenir à la période de conception du sujet. Au début de la thèse, en 2018, notre postulat concernant les technologies, et plus particulièrement les technologies numériques, était le suivant :

- **Omniscience** : le concepteur et l'utilisateur connaissent à l'avance toutes les conséquences de l'usage d'une technologie.
- **Contrôle** : ces conséquences sont sous le contrôle du concepteur et de l'utilisateur.
- **Neutralité** : les technologies sont des moyens neutres. La responsabilité morale repose sur l'usage qu'en fait l'utilisateur.

En d'autres termes, selon ce postulat, l'innovation technologique était synonyme de progrès. Pourtant, les critiques envers ce parti pris existaient déjà depuis plusieurs siècles. Pour le comprendre, il est nécessaire de revenir au sens étymologique de la technologie, qui signifie étude des techniques (du grec *tekhne*, "technique", et *logos*, "discours" ou "étude"). Par extension, ce terme désigne également "*l'ensemble des techniques propres à un domaine, à une science, à un métier*¹⁰⁷". Or, déjà au XIXe s., la philosophie marxiste défendait l'idée que le développement des techniques, ou modes de production, déterminait l'organisation des sociétés humaines : "*en acquérant de nouvelles forces productives, les hommes changent leur mode de production, et en changeant le mode de production, la manière de gagner leur vie, ils changent tous leurs rapports sociaux. Le moulin à bras vous donnera la société avec le suzerain ; le moulin à vapeur, la société avec le capitalisme industriel*" (Marx 1908, p. 155-156).

Depuis le milieu du XXe s., le langage binaire a remplacé la vapeur. L'apparition de l'informatique et de toutes les possibilités techniques qui en découlent a fait émerger une nouvelle génération de technologues (Jarrige 2016). Tous s'accordent pour démontrer qu'il existe des interactions fortes entre systèmes techniques et sociétés humaines, même si les avis divergent sur leurs répercussions. L. Winner, chercheur en science politique spécialisé sur les questions sociales et politiques qui entourent les changements technologiques modernes, offre un exemple pour illustrer ce propos dans son premier ouvrage consacré aux limites de la technologie (Winner 2002, p. 54-57). Durant le XXe s., l'Université de Californie a

107 <https://www.cnrtl.fr/definition/technologie> ; consulté le 22/10/2023

conduit plus de 50 années de recherche pour concevoir et améliorer une machine à ramasser les tomates, afin d'améliorer les conditions de travail des producteurs locaux. Au-delà du ramassage du fruit, la mise en place de la machine a eu plusieurs conséquences importantes sur le système de production et le territoire :

- les chercheurs ont dû élaborer une variété de tomates à la chair plus ferme, mais moins savoureuse pour correspondre aux contraintes techniques de la machine ;
- la machine a favorisé la production des très grandes exploitations, au détriment des producteurs locaux plus modestes, ce qui a mené à la perte de 32 000 emplois dans le secteur.

En réponse, un groupe d'agriculteurs a porté plainte contre l'Université de Californie, en avançant que les chercheurs utilisaient des fonds publics pour concevoir des solutions qui ne profitaient qu'à quelques grands acteurs privés. L'Université s'est défendue en avançant l'argument qu'on ne pouvait pas interdire toutes les recherches ayant des applications pratiques potentielles. À l'origine, les chercheurs avaient l'intention d'améliorer le quotidien des agriculteurs grâce à la technique, pas de créer un système pour détruire les petits producteurs de tomates. Pourtant, l'introduction de cette nouvelle technologie, conjuguée à des choix politiques et économiques, a mené à une redistribution des richesses et des rapports sociaux dans la communauté agricole de Californie.

Si la machine à ramasser les tomates représentait donc un progrès technique, les inégalités engendrées par ce dispositif ne permettaient pas d'en dire autant sur les plans sociaux et économiques. À travers leur plainte, les agriculteurs ont exprimé que la machine incarnait la transformation sociétale en cours, qui n'avaient pour eux rien de progressiste. L. Winner a donc avancé que les technologies représentaient bien plus que de simples objets techniques et que l'innovation technique en elle-même ne suffisait pas à définir le progrès : "*les opposants à des innovations comme la machine à ramasser des tomates se font taxer d'adversaires de la technologie ou du progrès. Car cette machine n'est pas seulement le symbole d'un ordre social qui récompense certains et en punit d'autres, elle est une véritable incarnation de cet ordre social*" (Winner 2002, p. 57).

L'auteur a employé le terme "incarnation", car, selon lui, les technologies représentent de véritables "*formes de vies*", dans le sens où l'intégration de nouvelles technologies engendre invariablement des changements dans les comportements et les institutions humaines

(Winner 2002, p. 32). Comme lors d'une greffe d'organe, lorsque l'intégration fonctionne, les transformations s'opèrent rapidement, jusqu'au point où "*on aurait beaucoup de mal à vivre sans elles*". Pour aller dans le sens de L. Winner, il n'y a qu'à s'imaginer réaliser une quelconque étude archéologique sans ordinateur ou connexion internet pour comprendre en quoi ces technologies sont devenues des éléments constitutifs de notre discipline.

À l'instar de l'écosystème d'un monde vivant, un système technique représente donc un ensemble d'outils, de moyens et de ressources en interactions constantes entre eux. Supprimer ou rajouter un élément implique des changements pour le système et ses utilisateurs. L'historien et philosophe J. Ellul, spécialiste des techniques, est allé encore plus loin en développant le concept de "*Système technicien*" (Ellul 2004). Selon lui, les sociétés humaines modernes, en constante recherche d'efficacité, ont progressivement mis en place un système technique si complexe qu'il en est devenu autonome et universel. Pour J. Ellul, la technique ne se résume pas seulement aux objets techniques, elle englobe également les méthodes d'organisation de la vie sociale (technique sportive), du travail (taylorisme) et de la politique (bureaucratie). "*Autrement dit, partout où il y a recherche et application de moyens nouveaux en fonction de critère d'efficacité on peut dire qu'il y a Technique*" (Ellul 2004, p. 38). Dans cet ouvrage, il défend l'idée que la Technique dans son ensemble forme un système globalisé, en expansion et qui tend à s'uniformiser, particulièrement depuis l'avènement de l'informatique et la mise en relation quasi automatique des technologies. Dans ce contexte, l'auteur porte un regard particulièrement pessimiste sur les interactions entre les êtres humains et la Technique. Pour lui, le *système technicien*, dans sa recherche rationnelle et implacable de logique et d'efficacité, s'est émancipé de l'ordre naturel. Il constitue un environnement à part entière, auquel l'être humain est contraint de s'adapter : "*l'homme apparaissant à la conscience trouve la technique comme un "déjà là". La technique constitue pour lui un milieu dans lequel il entre, où il s'insère. Il est parfaitement vain de dire que la technique n'est pas un vrai milieu : quoi que cet homme voie ou utilise, c'est un objet technique*" (Ellul 2004, p. 320).

Toujours sur le même sujet, nous pouvons également mentionner les travaux des sociologues B. Latour, M. Callon et M. Akrich autour de la "*sociologie de la traduction*" ou "*théorie de l'acteur-réseau*" (Akrich *et al.* dir. 2013). Ces trois chercheurs se sont interrogés, dans les années 1980, sur le processus de production des objets scientifiques et techniques ainsi que leurs finalités. Dans ce cadre, ils ont montré que l'objet technique ou le fait scientifique provient et évolue dans un réseau qui rassemble les intérêts d'acteurs humains (concepteur, utilisateur, prestataire, *etc.*) et non humains (objet, technologie, *etc.*). "*La théorie de l'ac-*

teur-réseau valorise la flexibilité interprétative, la controverse et surtout le rôle des réseaux et des groupes sociaux pour analyser l'émergence d'une technique. Celle-ci fait donc partie d'un tout complémentaire, agencé et n'obéissant à aucune hiérarchie : la technique et la société se définissent et se construisent en même temps et la distinction entre les deux est dissoute. D'où la construction simultanée du matériel et du social, et la cohabitation d'humains et de non-humains dans des réseaux complexes, cohérents et égalitaires" (Mahil, Tremblay 2017). Plus optimistes que J. Ellul, les auteurs ont donc mis en place une méthode pour analyser un système technique, ou selon leur théorie, un réseau, afin de le faire évoluer en tenant compte de tous les "actants" (humains et non-humains).

L'usage de technologies, en archéologie ou ailleurs, est donc au cœur d'une multitude d'interactions entre le naturel et la technique, qui concerne des implications autant pratiques que sociales, économiques et environnementales pour les usagers et leur communauté. Ces implications ne sont pas et ne peuvent pas être systématiquement anticipées, tant les effets indirects peuvent être larges et issus de facteurs variés.

Dans ses récents travaux de thèse, Ch. Tufféry a utilisé la sociologie de la traduction pour analyser la relation entre les dispositifs numériques et les utilisateurs sur les chantiers archéologiques (Tufféry 2022, p. 211-215). En prenant en compte "*un réseau entre archéologues, dispositifs numériques, mais aussi tous les autres types de dispositifs, les méthodes de travail, les observations, les supports de notations, les postures, les gestes, les discours officiels et officieux, les codes et règlements, les rites, etc.*", il considère que les pratiques numériques en archéologie relèvent d'une "hybridation". Ce terme, emprunté à B. Latour, désigne un "*processus durable de dialogue entre les humains et les dispositifs techniques, qui s'inscrit dans un réseau de relations où les affects s'expriment autant que la raison*".

Les recherches de Ch. Tufféry ont en effet démontré que les professionnels de l'archéologie sont dans l'attente d'un "*sens à leur action, le sentiment que leur engagement dans leur travail leur permet de trouver de l'épanouissement, du bien-être, du contentement, de la satisfaction, et même parfois du plaisir*" (Tufféry 2022, p. 211-213). D'après les observations de l'auteur, les outils d'enregistrement numériques qui laissent peu de place à la personnalisation sont, dans la plupart des cas, rejetés par les utilisateurs. Dans ces situations, les outils donnent l'impression aux usagers d'être de simples techniciens, dont la tâche consiste à remplir des cases préconçues, sans pouvoir exprimer "*des émotions et des affects dans les nouveaux savoirs archéologiques numériques*". Il résulte de l'utilisation de ces dispositifs une perte de sens général pour l'enregistrement archéologique, qui pousse au désintérêt de l'outil.

Dans d'autres situations, c'est la remise en question d'un savoir-faire par un outil numérique qui peut être vécu comme une contrainte par les archéologues. Il peut s'agir d'une réticence à intégrer une nouvelle technologie numérique dans les pratiques, comme de faire évoluer ces compétences numériques. Dans le premier cas, l'auteur cite les témoignages de plusieurs archéologues qui estiment que "*la photogrammétrie entraîne une perte de certains savoir-faire traditionnels*", comme le relevé pierre-à-pierre (Tufféry 2022, p. 213-214). Selon l'auteur, cette activité est encore une fois productrice de sens pour l'archéologue qui exerce autant ses talents manuels qu'intellectuels pour relever et interpréter la structure. Abandonner cette compétence au profit d'un nouvel outil mal maîtrisé peut revenir à perdre le sens de l'action, même si le résultat reste inchangé. On retrouve le même schéma dans le second cas, lorsqu'il s'agit de faire évoluer une compétence numérique acquise. Cette fois-ci, Ch. Tufféry décrit le refus de la quasi-totalité des topographes de l'Inrap pour remplacer leurs logiciels de travail par des solutions libres (Tufféry 2022, p. 157-158). Parmi les raisons invoquées pour justifier ce refus, les topographes mentionnent que "*la maîtrise de ces logiciels est le fruit de compétences accumulées depuis des années qui font partie de leur identité professionnelle*". Ainsi, introduire une nouvelle technologie, numérique ou non, peut être vécu comme une remise en cause de cette identité professionnelle.

4.1.2. Le poids environnemental des technologies numériques

En sortant du domaine archéologique, il nous semble également important d'évoquer le mythe moderne qui consiste à penser le système numérique comme une entité "dématérialisée", un "tout virtuel" qui ne consomme que peu d'espace, de ressources ou d'énergie (Marroone *et al.* 2023). Très récemment, plusieurs institutions ont tenté de mesurer l'incidence écologique des technologies numériques à l'échelle mondiale et nationale (Ferreboeuf 2018 ; Bordage 2019a ; Bordage *et al.* 2021 ; Ferreboeuf *et al.* 2021 ; Lees Perasso *et al.* 2023). Leur première tâche a été de définir les équipements qui constituent la sphère du numérique. Ils sont généralement découpés en trois catégories :

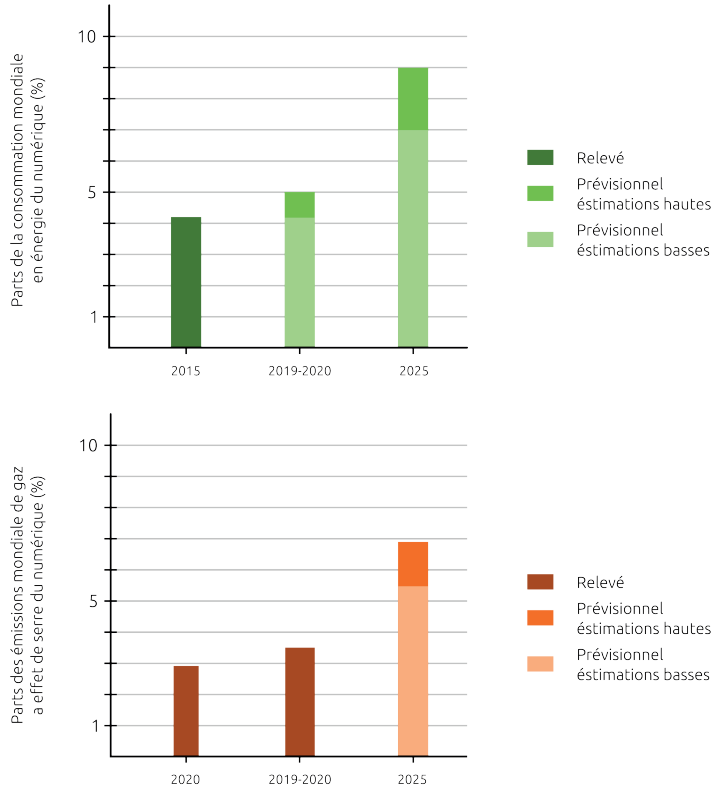
- Les **terminaux** : cette catégorie comprend tous les dispositifs possédant une interface (physique ou virtuel) qui permet à l'utilisateur d'interagir avec l'équipement : les ordinateurs, les smartphones, les tablettes, les écrans connectés, *etc.* Les objets connectés (ou IoT, *Internet of Things*) sont également souvent inclus dans cette catégorie. Il s'agit d'objets équipés de capteurs numériques reliés à un réseau, le plus souvent internet, afin d'enregistrer ou de réagir à des situations du monde physique : les outils domotiques (éclairage, thermostat, prises), les montres connectées, les porte-clés localisables, l'électroménager, *etc.* En 2019, le nombre de terminaux circulant sur la planète était estimé à environ 34 milliards (Bordage 2019a, p. 8).
- Les **réseaux** : chaque dispositif relié à internet ou à un réseau local exploite un équipement réseau. Les équipements réseau sont eux-mêmes déclinés en trois catégories, les réseaux filaires, mobiles et satellitaires.
 - . Le réseau filaire exploite des câbles de télécommunication pour relier des terminaux et objets connectés à travers le monde entier. À titre d'exemple, le site *Submarine Cable Map*¹⁰⁸ répertorie actuellement plus de 1,4 million de kilomètres de fibre optique déposée au fond des océans pour relier les continents. Ce réseau se ramifie encore plus sur terre, avec par exemple un câble pour chaque routeur internet classique (les *boxs*) qui équipe nos logements et lieux de travail.
 - . Le réseau mobile est distribué par un ensemble d'antennes relais, elles-mêmes alimentées par un réseau filaire. Elles sont réparties sur un territoire pour diffuser une connexion sans fil, accessible grâce à un terminal équipé d'une carte SIM. Un support est généralement équipé de plusieurs antennes, correspondant aux différents fournisseurs d'accès.
 - . Le réseau satellitaire est le plus récent. Cette technologie nécessite d'envoyer des satellites en orbite autour de la terre pour diffuser le réseau partout à la surface de la planète, tant qu'antenne réceptrice au sol peut offrir une vue directe et dégagée vers le ciel.
- Les **centres informatiques** : aussi nommés *data centers*, les centres informatiques accueillent de grandes quantités de serveurs connectés dans la très grande majorité des cas à internet. Ces serveurs hébergent toutes les données et services accessibles à distance (*cloud computing*), tels que les espaces de stockage, les services de vidéo en ligne ou encore les réseaux sociaux.

Le numérique, avec ces dizaines de milliards d'équipements, représente donc un ensemble matériel tangible, avec un impact environnemental mesurable. La figure suivante présente une synthèse des résultats des différentes études précédemment citées (fig. 89). Afin d'identifier la tendance actuelle, elle réunit les chiffres des modélisations fondées sur le développement non régulé des technologies numériques. Tous les résultats s'accordent sur le fait que l'empreinte écologique du numérique est déjà significative. Par exemple, les experts de *Green IT* estiment que pour l'année 2020, l'impact des équipements numériques représente environ 25 % du forfait d'émissions de gaz à effet de serre estimé soutenable pour un Français (Bordage *et al.* 2021, p. 9). Les analyses montrent également que cette tendance augmente graduellement, principalement en raison de la production d'équipements en constante progression. Rien que pour la France, l'étude de l'Ademe et l'Arcep prévoit que le nombre de terminaux passe de 795 millions en 2020 à plus de 4 milliards en 2050 (Lees Perasso *et al.* 2023, p. 51).

Le problème avec cette augmentation d'équipements est que l'industrie du numérique est particulièrement gourmande en métaux et terres rares, des ressources abiotiques non renouvelables qui se trouvent aujourd'hui au centre de situations géopolitiques et environnementales particulièrement complexes (Lepesant 2018). D'abord, face à la demande accrue, la consommation mondiale de ces matières premières s'est nettement intensifiée à partir des années 1960, à tel point que certains spécialistes parlent d'une "explosion" de la consommation (Ancre 2015, p. 9-10). Pourtant leur extraction nécessite souvent des produits polluants, comme c'est le cas pour le mercure, employé pour agglomérer les paillettes d'or, alors qu'un 1g de mercure suffit à polluer 1m³ de terre pendant 50 ans (Bordage 2019b, p. 120)

Les principaux pays extracteurs ne possèdent pas non plus des situations politiques toujours stables, ce qui entraîne des dérives importantes dans le commerce des minéraux. Le cas le plus connu est probablement celui du cobalt en République Démocratique du Congo (RDC). Le cobalt est un composant essentiel des batteries lithium-ion installées dans la très grande majorité des équipements numériques, depuis les smartphones jusqu'aux voitures électriques. La RDC possède 50 à 60 % des ressources mondiales connues de ce minerai. Dans une enquête menée en 2016 sur les mines de cobalt en RDC, Amnesty International et l'Observatoire africain des ressources naturelles ont révélé les dérives en matière d'atteinte aux droits humains de ce secteur d'activité (Amnesty International, Afrewatch 2016). Les mineurs creusent la plupart du temps avec des outils rudimentaires, sans protections particulières, dans des galeries étroites sur des profondeurs de plusieurs dizaines de mètres. Ces

Impact environnemental du numérique à l'échelle mondiale



Energie primaire
Energie disponible dans la nature avant transformation : les combustibles, l'énergie nucléaire, solaire, éolienne, etc.

Energie finale
Energie utilisée par les consommateurs

Ressources abiotiques
Ressources naturelles non-vivantes : minéraux, fossiles

Gaz à effet de serre
Gaz absorbant et réémettant une partie des rayons solaires, qui retiennent une partie de la chaleur émise par la planète

Impact environnemental du numérique à l'échelle de la France

Catégories d'impacts	2020	2030	2050
Consommation énergie finale (Terawatt/heure)	40 à 52 TWh 8,3 à 10,3% de la conso. nationale	54 TWh Jusqu'à + 35 %	93 TWh Jusqu'à + 70 %
Emissions GES (Mégatonne équivalent CO ₂)	17,2 à 24 Mt eq. CO ₂ 2,5 % à 3,2 % des émissions nationales	25 Mt eq. CO ₂ Jusqu'à + 45 %	49 Mt eq. CO ₂ Jusqu'à + 96 %
Consommation ressources abiotiques (Tonne équivalent antimoine)	833 à 952 t eq. Sb	1 080 t eq. Sb	1 510 t eq. Sb
Production de déchet (Tonne)	20 200 000 t	25 200 000 t	48 900 000 t

89. Synthèse sur l'impact environnemental du numérique (d'après Ferreboeuf 2018 ; Bordage 2019a ; Bordage et al. 2021 ; Ferreboeuf et al. 2021 ; Lees Perasso et al. 2023)

conditions de travail sommaires entraînent de nombreux accidents mortels dans les mines, ainsi que des problèmes sanitaires, notamment à cause de la poussière inhalée. On compte également des dizaines de milliers d'enfants parmi les mineurs, que le rapport classe parmi "*les pires formes de travail des enfants*" (Amnesty International, Afrewatch 2016, p. 18-19).

Enfin, le recyclage des équipements numériques reste encore très minoritaire, en partie à cause de la présence d'un grand nombre de matériaux assemblés dans des proportions infimes. À titre d'exemple, un smartphone emploie une quarantaine de métaux différents, dans des quantités allant de quelques milligrammes à quelques dizaines de grammes (Ferreboeuf 2018, p. 18-19). À ces densités-là, les opérations de tri et de retraitement des déchets numériques sont malheureusement plus complexes et coûteuses que d'exploiter de la matière première.

4.1.3. La technique comme solution miracle ?

Pour revenir à notre postulat de départ, à savoir que le développement et l'usage de nouvelles technologies numériques relevaient d'une pratique omnisciente, contrôlée et neutre, il semble au contraire qu'un regard plus réaliste sur celles-ci soit :

- **Systèmes complexes** : chaque nouvelle technologie s'inscrit dans un ensemble de dynamiques techniques, sociales, environnementales et économiques complexes à appréhender.
- **Systèmes difficiles à contrôler** : il est encore plus complexe de garder le contrôle sur ces dynamiques à l'échelle individuelle. Elles entraînent quasi systématiquement des conséquences imprévues.
- **Partialité des concepteurs et des utilisateurs** : pourtant, ne pas tenter de prendre en compte un maximum ces dynamiques avant de concevoir et de déployer une technologie revient à nier la réalité. Une technologie est créée dans la poursuite d'un système de valeur et son usage peut aboutir à la création de nouvelles valeurs.

L'innovation n'est donc pas forcément synonyme de progrès. L. Winner met en garde contre cette croyance qu'il nomme "*l'optimisme technologique*", qui consiste à envisager que l'arrivée d'une technologie va être en soit l'origine d'une transformation profonde et raisonnée de la société (Winner 2002, p. 89). Pour l'auteur, cette croyance est symptomatique de la compréhension du moderne comme toujours supérieur à l'ancien. En suivant ce raison-

nement, il n'est pas question d'adapter les technologies aux besoins des usagers, mais au contraire "de rénover les besoins de l'homme pour les adapter à ce que les scientifiques et les ingénieurs venaient de rendre disponible" (Winner 2002, p. 257). Le chercheur en histoire des sciences et journaliste E. Morozov pousse le concept encore plus loin, avec le "*solutionnisme technologique*", ou "*l'art d'apporter des solutions à des problèmes inexistantes*", une approche qu'il estime caractériser le domaine de l'innovation numérique (Morozov 2014 ; Dessus 2015, § 4). Si chaque technologie a des conséquences environnementales, sociétales et économiques, une solution technique apparente a de grandes chances d'apporter son lot de nouvelles contraintes, surtout si le problème identifié n'en était pas vraiment un. Il ne semble pas non plus judicieux d'espérer qu'une nouvelle solution technique émerge dans le futur pour résoudre ces contraintes, car cette position entraîne une interminable fuite en avant.

4.1.4. Le futur de l'archéologie numérique : vers une pratique raisonnée

À ce stade de la réflexion, il semble légitime de se poser la question : "Alors faut-il arrêter d'utiliser des technologies numériques en archéologie ?". En prenant en compte les conséquences écologiques, il s'agit probablement de la solution la plus simple et la plus efficace. Toutefois, l'archéologie ne représente qu'une goutte d'eau dans l'océan du numérique mondial. Ce type d'action, même soutenue par l'ensemble des archéologues de la planète, n'aurait probablement qu'une portée symbolique.

De plus, il est difficile d'imaginer aujourd'hui exercer notre métier sans l'aide des technologies numériques. Déjà en 1984, les participants de la rencontre sur la publication archéologique ont noté que la discipline a atteint une certaine forme de limite à travers les ouvrages de synthèse (Tchernia *et al.* 1986). L'article relate en effet les débats et les divergences d'opinions des participants autour de cette question. Tous se sont néanmoins accordés sur un point : il est nécessaire de faire évoluer le système de publication en archéologie, car les ouvrages sont trop volumineux et trop nombreux. Les arguments avancés sont divers :

- La quantité d'informations archéologiques produites requiert trop de temps de mise au propre pour être systématiquement intégrée dans un ouvrage de synthèse. Il n'est d'ailleurs pas pertinent d'intégrer toutes les données dans ce type d'ouvrage.
- Le rythme des publications atteint déjà un seuil critique. Un chercheur n'est plus en mesure de consulter l'ensemble des publications qui concerne ses problématiques de recherche.
- Les synthèses archéologiques nécessitent des ouvrages conséquents en nombre de pages et d'illustrations. Les moyens financiers risquent de manquer dans les années à venir pour continuer de produire ce type de publication.

Pour plusieurs des participants, cette évolution doit passer par l'utilisation de microfiches ou d'outils informatiques associés aux publications, afin de structurer et d'alléger les ouvrages. On relève d'ailleurs dans le débat des concepts qui prendront bien des années plus tard les noms d'*Open Data* ou de *Data Papers*.

Quarante ans plus tard, le volume d'informations n'a fait qu'augmenter, en partie grâce au développement des technologies numériques. Ce qui était vrai pour les publications archéologiques à l'époque l'est pour la gestion des données archéologiques aujourd'hui. Avec la multiplication des enregistrements, la professionnalisation de l'archéologie préventive et la création de nouveaux moyens d'échange, la quantité de données archéologiques disponibles a dépassé les capacités de mémorisation d'un être humain. Les cas sont rares où un chercheur est encore capable de retenir et de synthétiser toutes les informations en lien avec son sujet d'étude sans l'aide d'un ordinateur ou d'une connexion internet. Avec leur capacité de traitement de l'information quasiment sans limite à l'échelle des archéologues (les astrophysiciens traitent des lots de données sans commune mesure avec les nôtres), les outils numériques sont devenus indispensables pour continuer à produire une recherche archéologique de qualité.

De plus, l'objectif principal de l'archéologie ne concerne pas le profit de quelques-uns, mais bien une mission de sauvegarde de patrimoine au nom de tous. Par leur travail, les archéologues fournissent des éléments utiles pour construire les connaissances nécessaires au devoir de mémoire de chacun. Il s'agit d'une science fondamentale qui permet de comprendre le passé pour mieux appréhender le présent et l'avenir de l'humanité. En ce sens, il nous semble justifié de pouvoir employer les outils adaptés pour répondre à nos problématiques scientifiques et techniques.

Toutefois, il ne s'agit pas non plus d'une justification en soi pour utiliser des technologies numériques de manière débridée en archéologie. Il en va de notre responsabilité scientifique de prendre conscience de leur nature d'outils, sur lesquels exercer un regard critique qui dépasse largement le cadre technique. Nous devons dorénavant essayer de mesurer au maximum les implications pratiques, sociales et environnementales de ces outils, tant dans leur conception – les ressources et l'énergie consommées pour développer un prototype sont-elles justifiées par un vrai gain d'efficacité ultérieur ? Est-ce que cet outil est suffisamment évolutif pour s'adapter à la diversité des savoir-faire et objectifs du public ciblé ? – que dans leur diffusion – est-ce que la multiplication d'un tel dispositif ne risque pas de mener à une remise en question de certains utilisateurs ? Quels sont les risques sur l'organisation du travail et la production des connaissances ? Est-ce écologiquement et économiquement viable ? –. C'est seulement en entraînant ce regard critique que nous serons capables, en pleine connaissance de cause, de justifier l'utilisation d'outils numériques comme le moyen le plus adapté pour répondre à une problématique pertinente et clairement définie.

Des solutions commencent à émerger dans le monde de l'Enseignement Supérieur et de la Recherche pour mesurer le poids des technologies numériques sur l'environnement. On peut notamment citer ici les travaux de deux Groupements de Recherche du CNRS, *Labos 1point5*¹⁰⁹ et *EcoInfo*¹¹⁰. Tous les deux se sont donné pour objectif de réduire les impacts environnementaux et sociétaux des activités de recherche. *Labos 1point5* propose par exemple *GES 1point5*¹¹¹, un outil en ligne pour calculer le bilan de gaz à effet de serre d'un laboratoire à l'instant t, ainsi que *Scénario 1point5*¹¹², pour simuler la trajectoire d'évolution de l'empreinte carbone de ce même laboratoire jusqu'en 2030. Le groupement *EcoInfo* se concentre de son côté sur l'impact environnemental du parc informatique d'un établissement de l'Enseignement Supérieur et de la Recherche. Leur service *EcoDiag*¹¹³ permet de mesurer l'empreinte carbone des équipements informatiques d'un établissement sur une année en prenant en compte les émissions liées à la fabrication et à l'acheminement du matériel, ainsi que celles liées à la production d'énergie pour les faire fonctionner. Il serait particulièrement intéressant que l'EPCC Bibracte, qui représente de nombreux aspects de la chaîne opératoire archéologique, se prête aux exercices proposés par ces trois outils. Le résultat pourrait four-

109 <https://apps.labos1point5.org/> ; consulté le 22/10/2023

110 <https://ecoinfo.cnrs.fr/> ; consulté le 22/10/2023

111 <https://apps.labos1point5.org/ges-1point5> ; consulté le 22/10/2023

112 <https://apps.labos1point5.org/scenario-1point5> ; consulté le 22/10/2023

113 <https://ecoinfo.cnrs.fr/ecodiag-calcul/> ; consulté le 22/10/2023

nir un réel témoin du poids environnemental des activités de recherche et de valorisation archéologiques, et servir de point de départ pour mener une réflexion plus globale sur le sujet avec les nombreux partenaires de l'établissement (services de l'état, collectivités territoriales, opérateurs d'archéologie préventive).

Sur les aspects humains, la formation et l'accompagnement des usagers restent les meilleures solutions pour prévenir les risques de perte de sens ou de remise en cause de l'identité professionnelle. Il est important d'accompagner l'utilisateur afin qu'il comprenne par lui-même que l'outil numérique ne définit pas ce qu'il sait ni qui il est, mais qu'il s'agit uniquement d'une solution technique qui peut répondre à ses objectifs archéologiques. D'un autre côté, le formateur ne doit pas non plus tomber dans l'enthousiasme technologique et chercher à imposer à tout prix l'outil qu'il soutient : "*limiter la transformation des pratiques à l'apprentissage à tel ou tel logiciel ne suffit pas si les personnes formées ne sont pas capables en même temps d'exercer un regard critique sur ces nouvelles pratiques*" (Tufféry 2022, p. 233). C'est seulement en exerçant ce regard critique, qui prend en compte les dimensions pratiques, sociales, économiques et environnementales, qu'il pourra discerner chez la personne formée ce qui relève de la croyance et ce qui relève de l'argument scientifique appuyé sur l'expérience et la maîtrise des objectifs archéologiques.

Finalement, l'utilisation de technologies numériques, en plus de représenter un poids environnemental avéré, a de grandes chances de transformer les pratiques des archéologues, mais aussi le regard qu'ils portent sur l'exercice de leur profession. Pourtant, ces technologies sont devenues, depuis 40 ans, presque indispensables pour mener des recherches archéologiques. Il est donc désormais essentiel d'adopter une perspective critique à l'égard des technologies numériques déjà en usage en archéologie, ainsi que celles à venir, afin de saisir et d'anticiper leur impact sur l'évolution de nos pratiques professionnelles et sur l'environnement. Cet exercice nous incite à optimiser notre utilisation d'appareils numériques et à rationaliser notre production de données. L'approche de la science ouverte soutient cette démarche de rationalisation, car elle encourage l'utilisation de logiciels évolutifs et implique la sélection, la validation et la justification des données générées. Or ce travail nous a conduit à identifier un ensemble de points à mettre en place pour établir un processus de production numérique sur un chantier archéologique conforme aux principes de la science ouverte.

4.2. Recommandations pour des données archéologiques conformes à la science ouverte

Selon nous, cette synthèse ne pouvait pas se présenter sous la forme d'un guide, qui accompagne pas à pas le lecteur dans une démarche, car les situations possibles pour créer, traiter, partager et archiver des données numériques sont bien trop nombreuses sur les chantiers archéologiques. Nous livrons plutôt ici un ensemble de recommandations qui permettent aux archéologues d'acquérir les bons réflexes pour inscrire leurs démarches numériques dans une stratégie de rationalisation et d'ouverture. Ces recommandations résument notre expérience de 5 années de manipulations de données dans différents contextes de fouilles archéologiques. L'objectif est de produire les résultats les plus structurés, les plus libres et les plus transparents possibles, en proposant des solutions souples et légères qui ne transforment pas radicalement les savoir-faire des archéologues.

4.2.1. Matériels et logiciels

Le matériel

Comme nous l'avons expliqué précédemment, il semble compliqué de pouvoir se passer de dispositifs numériques pour pratiquer l'archéologie. Connaissant l'empreinte écologique actuelle et à venir de ce type d'équipements, il est donc important de limiter au maximum le nombre d'exemplaires en circulation parmi les archéologues, toutes utilisations confondues. Plusieurs leviers existent pour y parvenir.

Le premier levier est de toujours chercher à définir au mieux les besoins scientifiques, en fonction de la problématique archéologique établie, tout en prenant en compte les contraintes sociales, environnementales, techniques et économiques présentes dans le projet. Les dispositifs numériques n'étant finalement que des outils, tous les problèmes rencontrés dans la pratique des métiers de l'archéologie ne peuvent pas trouver de solutions techniques. Bien souvent, le dialogue, la formation et l'organisation des équipes de fouille peuvent s'avérer tout aussi efficaces que l'introduction d'une nouvelle technologie numérique pour accomplir une tâche. Lorsque le choix d'employer un dispositif numérique est justifié, il est aussi important de garder à l'esprit qu'une réponse technique se suffit rarement à elle-même pour résoudre un problème. La mise en place d'une technologie numérique

doit nécessairement s'accompagner de formations et d'accompagnement, afin que la majorité des usagers comprennent et valident les raisons de son introduction, en plus de maîtriser son fonctionnement.

Le deuxième levier consiste à explorer les solutions disponibles avec le matériel existant, quitte à en détourner le premier usage, pour répondre aux objectifs archéologiques. Cette approche nécessite parfois un temps de recherche et développement, mais permet également de mettre en place des solutions ingénieuses, conçues "sur mesure" pour répondre à des problématiques précises, tout en s'assurant le contrôle sur l'outil. Nous pensons par exemple à la démarche de J.-M. Friedt, professeur et ingénieur du département Temps-Fréquence du laboratoire FEMTO-ST (UMR 6174), qui a développé, à l'aide d'un smartphone et d'une antenne GNSS "*faible coût*", un système de géolocalisation centimétrique en temps réel, soit une précision équivalente au matériel de topographie haut de gamme (Friedt 2023). Ainsi, l'utilisation du matériel déjà existant présente donc deux avantages : limiter l'achat de nouveau matériel et aborder les usagers avec des équipements qu'ils maîtrisent déjà partiellement.

Si l'achat de nouveau matériel est incontournable pour la réussite d'un projet, un troisième levier consiste à mutualiser au maximum l'équipement en s'appuyant sur l'expérience de structures fédératrices comme les Maisons des Sciences de l'Homme. Plusieurs de ces structures mettent en effet à disposition des chercheurs issus des laboratoires membres des ressources matérielles pouvant être exploitées dans plusieurs disciplines. Le soutien au personnel et le renforcement des ressources financières de ce type d'organisme sont essentiels pour garantir une utilisation optimale et une maintenance de haute qualité des équipements de pointe indispensables à la recherche archéologique.

Le marché de l'occasion est généralement peu fourni pour les technologies numériques, car il s'agit la plupart du temps d'équipements fragiles présentant une faible durée de vie. Un quatrième levier consiste donc à allonger significativement cette durée de vie. La méthode la plus simple consiste à s'assurer d'offrir une seconde vie au matériel encore fonctionnel lors du renouvellement d'équipements. Il peut s'agir de don ou de revente, mais aussi, pour les grandes équipes, d'assurer un roulement dans le temps des machines des postes de travail les plus gourmands en ressources informatiques vers des postes qui nécessitent des moyens plus classiques. Par exemple, une personne qui traite de la 3D ou des images haute définition a besoin de plus de capacités de calcul qu'une personne qui utilise essentiellement du traitement de texte et des tableurs.

Pour garantir la durabilité des équipements et assurer ce roulement, il est recommandé en cinquième levier d'opter pour des technologies de haute qualité, réputées pour leur robustesse et leur longévité lors de l'achat de nouveaux dispositifs. Ces qualités sont souvent liées à trois facteurs essentiels : la simplicité, la facilité de maintenance et la réparabilité des équipements. Une veille technologique active est nécessaire pour identifier ces aspects, en particulier dans un domaine où les innovations sont fréquentes. Il peut donc être judicieux de créer des services chargés d'évaluer la pertinence et la faisabilité de l'adoption de nouvelles technologies en archéologie. Leurs équipes seraient composées de chercheurs maîtrisant à la fois les problématiques archéologiques et les aspects systémiques (techniques, sociaux, environnementaux, économiques) des technologies numériques. Ils assureraient des missions de conseil, consistant à évaluer l'impact et le degré de maturité d'un outil pour répondre à un questionnement défini par les archéologues : il est prêt et pertinent ; il est prometteur, mais pas encore pleinement développé ; il est inadapté ; il s'agit d'une nouvelle itération d'un système déjà existant.

Les logiciels

L'expérience réalisée à travers cette thèse de doctorat a démontré qu'il est possible de mener une recherche approfondie avec des logiciels libres, sans faire de compromis sur les capacités des outils. Nous recommandons donc de migrer, chaque fois que cela est possible, vers des logiciels libres soutenus par des communautés actives. À l'usage, le plus gros avantage des outils libres a été de pouvoir équiper toutes les machines avec lesquelles nous ou nos partenaires avons travaillé avec les mêmes outils, sans contraintes techniques ou économiques. Cette simplicité a permis d'assurer une continuité dans nos pratiques professionnelles et de simplifier la formation de nos partenaires.

En effet, il nous est apparu qu'en passant des logiciels propriétaires aux logiciels libres, que le cap le plus difficile à franchir est la recherche systématique de fonctionnalités strictement d'un type à l'autre. Une fois ce cap dépassé, nous nous sommes aperçu que les environnements logiciels libres conviennent bien mieux à un public scientifique qui recherche la qualité, la précision, la transparence, l'interopérabilité et le contrôle sur ses outils.

Nous proposons ici une liste d'outils informatiques qui conviennent à la grande majorité des usages numériques associés aux données archéologiques de terrain. Excepté les systèmes d'exploitation et le gestionnaire de logiciel, tous les logiciels cités sont compatibles avec Windows, Mac et Linux.

- **Système d'exploitation** : la distribution libre Ubuntu¹¹⁴ de Linux est la plus connue et fonctionne très bien pour les archéologues ayant l'habitude de travailler avec un environnement Windows. Durant cette thèse, nous avons majoritairement utilisé la version Ubuntu 20.04.6 LTS. Pour la suite, nous pensons évoluer vers Arch Linux¹¹⁵, une distribution qui vise un public plus avancé, en mettant en avant une expérience utilisateur transparente. Le principe est de réduire volontairement l'utilisation de l'interface graphique, qui a tendance à masquer le fonctionnement de certains outils. Cette opacité réduit la capacité de l'utilisateur à identifier et régler les problèmes qui peuvent survenir durant l'utilisation de l'ordinateur.
- **Gestionnaire de logiciel** : la méthode d'installation des logiciels en paquet et dépendances est un élément perturbant pour les nouveaux utilisateurs de distributions Linux habitués au fonctionnement de Windows. Le magasin d'application Flathub¹¹⁶ propose les versions mises à jour de nombreux logiciels libres au format Flatpak, un système d'installation "générique" permettant d'installer et d'assurer la stabilité des logiciels sur les distributions Linux les plus populaires. Tous les logiciels cités dans les paragraphes suivants sont disponibles gratuitement sur Flathub.
- **Suite bureautique** : la suite LibreOffice¹¹⁷ remplit parfaitement les rôles de traitement de texte et de gestion de tableurs, en plus d'être compatible avec des logiciels tiers comme Zotero¹¹⁸ pour la gestion de références bibliographiques. Un autre point notable, Calc gère plus facilement l'importation de fichiers CSV que son équivalent Microsoft Excel.
- **DAO** : le logiciel Inkscape¹¹⁹ permet de numériser les dessins techniques archéologiques avec la même précision que son concurrent Adobe Illustrator. Depuis la version 1.2, il propose même de configurer les raccourcis clavier sur le modèle d'Adobe Illustrator, pour permettre à l'utilisateur de retrouver rapidement ses habitudes de travail. Nous avons également noté que le module de vectorisation automatique des images matricielles est généralement plus performant sur Inkscape.

114 <https://ubuntu.com/> ; consulté le 22/10/2023

115 <https://archlinux.org/> ; consulté le 22/10/2023

116 <https://flathub.org/fr> ; consulté le 22/10/2023

117 <https://www.libreoffice.org/> ; consulté le 22/10/2023

118 <https://www.zotero.org/> ; consulté le 22/10/2023

119 <https://inkscape.org/> ; consulté le 22/10/2023

- **Traitement d'image** : GIMP¹²⁰ est le plus connu des logiciels libres de manipulation et de retouche d'image. Il remplit parfaitement les fonctions basiques parfois utiles pour la retouche de photographies archéologiques de terrain, comme la mise à l'échelle, le recadrage, la retouche des couleurs ou de la luminosité ou encore la modification des éléments du lettré. Pour le développement numérique d'images "brutes" (RAW), nous conseillons le logiciel Darktable¹²¹.
- **Base de données** : Dbeaver¹²² propose une alternative libre aux systèmes de gestion de base de données communément utilisés en archéologie. Il s'agit d'un outil beaucoup plus sobre, qui nécessite de maîtriser les bases du langage SQL, mais qui a l'avantage de pousser l'utilisateur à comprendre les principes fondamentaux d'une base de données relationnelle, contrairement à FileMaker. Il est donc plus complexe à utiliser au premier abord, mais cela se traduit par une interopérabilité de données considérablement accrue.
- **SIG** : QGIS¹²³ est aujourd'hui largement répandu dans la communauté archéologique. Il permet de gérer une très grande partie de la chaîne opératoire des données spatiales archéologiques, avec la production et la gestion de bases de données, des fonctions d'analyse spatiale avancées et un système d'édition cartographique très complet. La communauté d'utilisateurs, très active, met également à disposition une multitude d'extensions permettant de personnaliser au maximum l'expérience.
- **Orthophotographie et 3D** : Nous avons déjà présenté le logiciel de photogrammétrie MicMac¹²⁴ (*cf. supra* 3.3.1. Logiciel MicMac). Pour résumer ici, il s'agit du logiciel de photogrammétrie le plus précis et le plus transparent sur la qualité des résultats que nous avons eu l'occasion de tester en contexte archéologique. De plus, il est le seul à proposer une fonction automatique pour calculer des orthophotographies verticales, adaptées aux coupes ou aux élévations. Le logiciel Cloud Compare¹²⁵ permet de visualiser et d'éditer les nuages de point produit avec MicMac.

120 <https://www.gimp.org/> ; consulté le 22/10/2023

121 <https://www.darktable.org/> ; consulté le 22/10/2023

122 <https://dbeaver.io/> ; consulté le 22/10/2023

123 <https://www.qgis.org/fr/site/> ; consulté le 22/10/2023

124 <https://micmac.ensg.eu/index.php/Accueil> ; consulté le 22/10/2023

125 <https://www.cloudcompare.org/> ; consulté le 22/10/2023

- **PAO** : Scribus est souvent cité comme alternative libre pour les logiciels de publication par ordinateur. Toutefois, nous ne le recommandons pas, car il gère encore mal les listes, comme les titres ou les légendes de figure, généralement nombreux dans les manuscrits académiques. Il existe également le LaTeX¹²⁶, un langage de composition de document, mais nous n'avons jamais eu l'occasion de le pratiquer. Sinon, Writer, de la suite LO, offre des fonctions de mise en page à la manière de Microsoft Word.

4.2.2. Les pratiques

Convention de nommage et arborescence de classement

La première mission que nous nous sommes imposé lors de la reprise du chantier de PC2 en 2018 a été de définir une arborescence de classement des fichiers, que nous justifions alors par deux raisons :

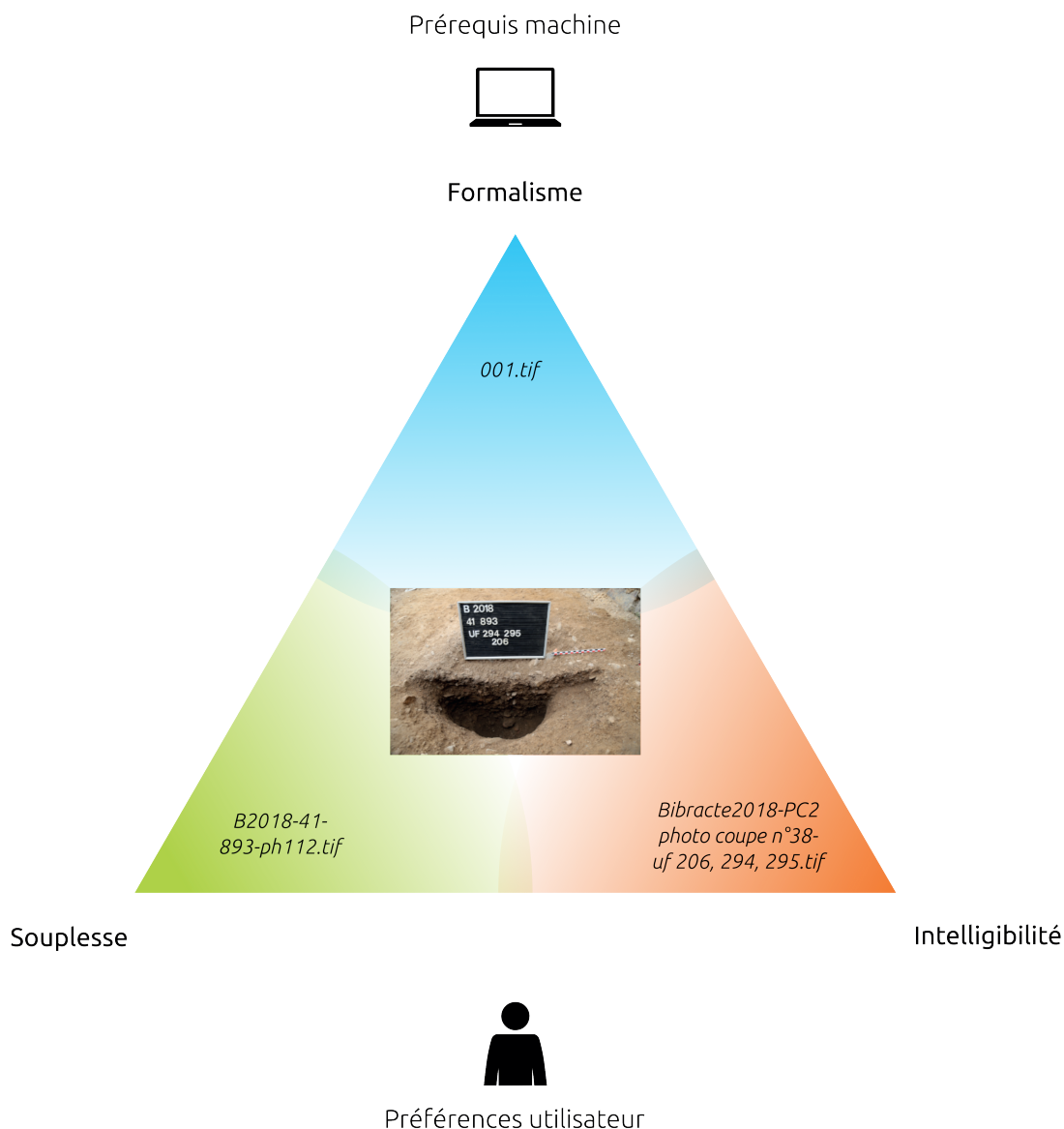
- si l'arborescence de classement est claire, les archéologues vont naturellement ranger les fichiers dans les bons dossiers ;
- nous maîtrisons les méthodologies et de notre discipline et notre objet d'étude, nous pouvons donc anticiper tous les dossiers nécessaires pour un classement optimal.

Pourtant, à l'usage, ces deux points de vue se sont rapidement révélés erronés. En réalité, une arborescence de classement reflète avant tout la logique de tri de son créateur, qui, en plus d'être personnelle, a tendance à évoluer avec le temps. De plus, l'archéologie rassemble de nombreuses spécialités, avec chacune leurs spécificités méthodologiques. Il semble donc compliqué dans ces conditions de définir une arborescence de classement intuitive et universelle. C'est en tout cas le constat que nous avons fait sur le chantier de PC2, où la structure de l'arborescence a évolué d'année en année, parfois pour rajouter des répertoires, parfois pour simplifier un ensemble devenu trop ramifié.

En parallèle, nous nous sommes progressivement rendu compte que le nom d'un fichier permettait de classer les documents de manière quasi automatique, sans imposer de structure de répertoires. En effet, une arborescence de classement est avant tout orientée pour un usage humain. Pour une machine, tous les fichiers peuvent se trouver à la racine du disque dur sans que ça n'affecte, en théorie, son bon fonctionnement. L'enjeu d'une bonne convention de

126 <https://www.latex-project.org/> ; consulté le 22/10/2023

nommage consiste donc à trouver un équilibre entre le formalisme, la souplesse et l'intelligibilité des noms de fichiers (fig. 90). Le formalisme permet d'exploiter les fonctions de tri automatique des machines, la souplesse offre la capacité d'adaptation aux nouvelles formes de données et l'intelligibilité sert l'utilisateur qui comprend le type et le contenu du fichier intuitivement.



90. Convention de nommage : équilibre entre formalisme, souplesse et intelligibilité

En se plaçant du point de vue de la machine, la seule contrainte à prendre en compte pour définir une convention de nommage est que tous les fichiers d'un même répertoire possèdent un nom unique. Ce nom n'a pas besoin d'être explicite sur le contenu du fichier, c'est pourquoi les noms générés automatiquement par les ordinateurs sont souvent des lignes de caractères

alphanumériques correspondant à la date et l'heure de création (par exemple "20230815-100542.jpg"), au matériel ou au logiciel utilisé pour la génération ("epson_xcr2f_001.pdf") ou même parfois sans logique apparente ("1000XCFB9.dbf"). À l'inverse, l'être humain cherche intuitivement à définir un nom qui explicite le contenu du fichier le plus précisément possible, afin de l'identifier le plus rapidement possible dans un jeu de données ("photo coupe 22 – Bibracte 2018 – PC2.jpg").

Cette deuxième solution, si elle semble naturelle, présente en réalité plusieurs problèmes pour la gestion et l'exploitation de jeux de données numériques. Le problème le plus évident est celui de la quantité de données à nommer. Il est possible de taper manuellement le nom de quelques fichiers, mais la tâche devient beaucoup plus laborieuse lorsqu'on dépasse la centaine de fichiers, ce qui est souvent le cas en chantier de fouille. Les fichiers sont donc souvent nommés au fil de leur création, sans respecter une logique stricte.

Cette disparité nominale mène au deuxième problème : si les noms d'un même jeu de données ne comportent pas un minimum de formalisation, il n'est plus possible d'utiliser les fonctions de tri automatique des machines. De plus, pour accéder aux données, les outils informatiques adressent des requêtes aux fichiers "conteneurs", en les identifiant par leur nom. Lorsque ces requêtes sont formulées par les usagers, plus le nom est long et informel, plus la requête est complexe à rédiger et plus les risques d'erreur sont nombreux. Il n'est pas non plus recommandé de modifier régulièrement les noms des fichiers, car de nombreux logiciels comme les systèmes de gestion de base de données, les SIG et les solutions de publication assistée par ordinateur l'utilisent pour créer le lien avec les données.

Pour résumer, une convention de nommage efficace est constituée des éléments suivants :

- Un nom **unique** pour chaque fichier du jeu de données.
- Un nom **le plus court possible**, d'abord pour faciliter la gestion des fichiers, mais aussi parce que certaines configurations du système d'exploitation Windows n'acceptent pas les chemins d'accès au fichier de plus de 260 caractères. Or, pour composer ce chemin, les systèmes d'exploitation renseignent la suite de répertoires à suivre pour atteindre le fichier depuis la racine du disque dur, par exemple "[c:/MesFichiers/Projets/Bibracte-2018/minute42.jpg](#)". Avec un nombre de répertoires importants et un nom de fichier trop long, la machine risque de ne pas pouvoir accéder au document. Dans la pratique, il est conseillé de ne pas dépasser 30 caractères pour un nom de fichier. Un nom qui ne comporte **pas de caractères spéciaux**, tels que " ! é / ^ * #". Cette res-

triction comprend donc les symboles, les lettres accentuées, les points et les espaces. Il s'agit d'une limitation technique imposée par les systèmes d'exploitation. Tous interdisent d'utiliser des caractères spéciaux dans les noms de fichier, mais pas toujours les mêmes entre chaque système. De plus, ces caractères spéciaux ne sont pas encodés de la même manière d'un système d'exploitation à l'autre, ou même parfois d'une version à l'autre d'un même système. Le plus simple est donc de proscrire complètement les caractères spéciaux, pour limiter les erreurs et améliorer l'interopérabilité des fichiers. Il est conseillé d'utiliser un tiret bas "_" (*underscore*) ou éventuellement un tiret "-" pour remplacer un espace. Un espace, même s'il prend un caractère, reste le moyen le plus efficace pour l'œil humain pour identifier les mots. Le cas du point "." est un peu particulier. La plupart des logiciels réservent son usage pour déclarer une extension de fichier, soit un suffixe pour identifier son format (.txt, .doc, .pdf, *etc.*). Rajouter un point dans le nom d'un fichier peut donc entraîner des erreurs lors de l'ouverture du fichier.

- Un nom **codifié au maximum**, particulièrement sur les premiers éléments qui le constituent. Le nom se divise généralement en 3 sections, elles-mêmes divisées en une ou plusieurs clés d'identification. La première section est généralement composée d'une clé répétée sur tous les fichiers, afin d'individualiser le lot de données. Il est intéressant d'utiliser une date comme première clé quand l'information est pertinente, car il est ensuite possible de classer les documents dans n'importe quel gestionnaire de fichiers. La deuxième section correspond la plupart du temps à un identifiant explicite renvoyant à l'objet d'étude, comme l'acronyme d'un site archéologique. La troisième section renseigne le type de fichier et éventuellement un indice sur son contenu. Cette dernière partie est celle qui varie le plus en fonction de la nature des fichiers.
- Un **nombre de caractères** identique pour les clés récurrentes entre plusieurs fichiers. Par exemple, dans le cadre d'une énumération, il est conseillé d'utiliser "001", "002", "003" plutôt que "1", "2", "3", ce qui nécessite d'anticiper le nombre d'itérations possibles. Ce détail facilite nettement les requêtes de fichiers et permet à l'œil humain d'effectuer rapidement le tri dans une liste de fichiers en se concentrant uniquement sur la troisième section des noms.
- Même s'il ne s'agit pas d'un élément déterminant, il est **déconseillé d'utiliser des majuscules**. En général, l'écriture en lettres capitales est considérée comme agressive et l'utilisation de majuscules ponctuelles dans un nom augmente les risques d'erreur lors de l'enregistrement ou de la requête du fichier.

Nous proposons ici un modèle de convention de nommage suivant ces recommandations et adapté aux fichiers issus d'opérations archéologiques : "*numéro d'arrêté de prescription d'opération*"_ "*acronyme du site*"_ "*(optionnel) données_complexes*"_ "*type de fichier*". "*format de l'extension*". Chaque élément est décrit comme suit :

- *Numéro d'arrêté de prescription d'opération* : Il s'agit du numéro de l'arrêté qui autorise l'opération archéologique, qu'il s'agisse d'une prospection ou d'une fouille. Ce numéro est lisible sur le document du même nom, fourni par les services de l'état à chaque responsable d'opération, habituellement sous la forme "2022/385". Nous conseillons d'utiliser un format à 8 chiffres, qui peut se traduire pour l'exemple précédent par "20220385". Il est intéressant d'utiliser une date comme première clé, quand cela est pertinent, car vous pouvez ensuite classer simplement les documents dans n'importe quel gestionnaire de fichiers. Dans le cas de Bibracte, qui regroupe plusieurs chantiers sous un même numéro d'opération, nous nous sommes appuyé sur le système d'enregistrement par intervention et chantier pour définir cette première section. Pour la campagne de la *domus* PC2 en 2019, soit l'intervention 921 du chantier 41, cela se traduit par "int921_exc41".
- *Acronyme du site* : Afin d'identifier rapidement le site, il est courant en archéologie d'utiliser un acronyme, sur les plaques photo par exemple. C'est cet identifiant que nous proposons de réutiliser ici, avec 4 lettres en minuscule, comme "vsld" pour Verdun-sur-le-Doubs.
- *(optionnel) Données_complexes* : Certaines données peuvent faire l'objet de plusieurs traitements, qui peuvent être regroupées grâce à cette clé optionnelle. C'est le cas par exemple des données photogrammétriques. Un site peut faire l'objet de plusieurs acquisitions et traitements photogrammétriques lors d'une opération, numérotés de 1 à *n*, sous la forme "sc001", "sc002", "sc003", *etc.*
- *Type de fichier* : Cette clé permet d'identifier le type de contenu du fichier. Il peut s'agir par exemple de "inv" pour les inventaires, "pho" pour les photos, "min" pour les minutes, *etc.* Si le jeu de données présente plusieurs fichiers du même type, comme les photos, il suffit de les numérotés de 1 à *n* : "pho001", "pho002", *etc.*

Le tableau suivant, non exhaustif, résume les différentes possibilités offertes par une convention de nommage qui respecte ce modèle (fig. 91). Finalement, la convention de nommage est un élément déterminant de la structuration des fichiers, car avec un ensemble de codes et de noms logiques, elle permet aux humains comme aux machines de s'affranchir du "rangement" pour se concentrer sur l'exploitation des fichiers. Même si cette pratique nécessite une certaine rigueur, elle offre en contrepartie un jeu de données clair au fort potentiel de réutilisation.

Catégorie de données	Type de données	Modèle conv. nommage	Nom de fichier
Données simples	Photo de terrain	"numéro d'arrêté de prescription d'opération"_"acronyme du site"_pho"numéro d'itération"."format de l'extension"	20220385_vsl_d_pho001.jpg
	Minute de terrain	"numéro d'arrêté de prescription d'opération"_"acronyme du site"_min"numéro d'itération"."format de l'extension"	20220385_vsl_d_min001.jpg
	Plan vectorisé	"numéro d'arrêté de prescription d'opération"_"acronyme du site"_map"numéro d'itération"."format de l'extension"	20220385_vsl_d_map001.svg
	Illustration	"numéro d'arrêté de prescription d'opération"_"acronyme du site"_ill"numéro d'itération"."format de l'extension"	20220385_vsl_d_ill001.svg
	Planche mobilier	"numéro d'arrêté de prescription d'opération"_"acronyme du site"_pmb"numéro d'itération"."format de l'extension"	20220385_vsl_d_pmb001.svg
	Documentation	"numéro d'arrêté de prescription d'opération"_"acronyme du site"_"type de documentation"_"version"."format de l'extension"	20220385_vsl_d_rapport_v1.svg
Données complexes	Inventaire	"numéro d'arrêté de prescription d'opération"_"acronyme du site"_inv_"type de fichier inventorié"."format de l'extension"	20220385_vsl_d_inv_photo.csv
			20220385_vsl_d_inv_mobilier.csv
			20220385_vsl_d_inv_minutes.csv
	Protocoles	"numéro d'arrêté de prescription d'opération"_"acronyme du site"_prt_"type de protocole décrit"."format de l'extension"	20220385_vsl_d_prt_conv_nom.odt
			20220385_vsl_d_prt_photogram.odt
			20220385_vsl_d_prt_acquis_photo.odt
	Données spatiales	"numéro d'arrêté de prescription d'opération"_"acronyme du site"_spa_"type de contenu vectorisé"."format de l'extension"	20220385_vsl_d_spa_us.shp
			20220385_vsl_d_spa_coupe.shp
			20220385_vsl_d_spa_topo.shp
	Photogrammétrie	"numéro d'arrêté de prescription d'opération"_"acronyme du site"_"numéro de scène"_"type de fichier"."format de l'extension"	20220385_vsl_d_sc01_ph01.jpg
20220385_vsl_d_sc01_ort.tif			
20220385_vsl_d_sc01_prc.jpg			

91. Exemple de nommage de fichiers

L'arborescence de classement n'est donc pas essentielle pour partager un jeu de données. Toutefois, elle peut s'avérer pratique pour fluidifier leur utilisation en cours de projet. Comme pour la convention de nommage, une arborescence de classement efficace respecte un équilibre, qui repose sur le nombre de niveaux de classement et de répertoires dans chaque niveau. Nous proposons trois règles à appliquer pour définir une arborescence de classement :

- des **noms clairs, courts et uniques** : Si la convention de nommage a été respectée, seul le répertoire général, à l'origine de l'arborescence, est obligatoire. Les sous-répertoires peuvent être supprimés et les fichiers rassemblés dans le répertoire général, sans que cela n'affecte techniquement la gestion des fichiers. C'est pourquoi le nom du répertoire général doit respecter la première section de la convention de nommage (*numéro d'arrêté de prescription d'opération_acronyme du site*). Le choix des noms des sous-répertoires est plus libre, mais doit tout de même respecter les autres règles de la convention de nommage : unique, le plus court possible, sans caractères spéciaux et sans majuscules ;
- une arborescence à **5 niveaux maximum** : il est déconseillé de dépasser 4 niveaux de sous-répertoire depuis le répertoire général du jeu de données. Au-delà de cette limite, l'arborescence devient généralement trop lourde à l'utilisation. De plus, elle risque de produire des chemins d'accès qui dépassent la limite des 260 caractères ;
- des **répertoires adaptés** : il est probable qu'un répertoire contenant moins de 10 fichiers ne soit pas essentiel pour le système de rangement. Ces fichiers peuvent certainement être clairement identifiés par la convention de nommage et déplacés dans un répertoire de niveau supérieur.

Pour terminer avec la convention de nommage, le cas de l'archéologie préventive en France est particulier, car certains Services Régionaux de l'Archéologie imposent déjà des conventions de nommage. Toutefois, tant que ces normes ne seront pas partagées et documentées à l'échelle nationale, elles ne contribuent pas à simplifier la lisibilité des données dans un contexte de partage à large échelle.

Les formats de fichier

En informatique, les données sont encodées dans des fichiers. Les formats de fichiers sont donc des structures conventionnelles de lignes de code qui mettent en forme les données (Boulétreau, Habert 2014). Les logiciels utilisent ces formats pour trouver les éléments dans les données afin d'afficher du texte, de l'image, de la vidéo ou de la 3D. Les formats sont donc à la base de l'exploitation des données, indispensables à prendre en compte pour partager de l'information numérique. En archéologie comme d'autres disciplines scientifiques, les données primaires archivées ou diffusées doivent pouvoir être réutilisées. Les formats libres favorisent cette réutilisation, car ils permettent de lire les données sans contraindre l'utilisateur dans son choix de logiciel. Ainsi, quel que soit le logiciel producteur, il est vivement recommandé d'enregistrer les données dans un format libre le plus universel possible.

L'un des formats libres les plus simples et les plus polyvalents qui existent pour partager des données archéologiques est probablement le format XML (*Extensible Markup Language*). Il s'agit en réalité d'un langage informatique générique permettant de structurer de l'information à partir de balises (fig. 92). Cette structure en balisage n'est pas la plus optimisée techniquement, mais elle a l'avantage d'être relativement explicite pour les utilisateurs, même débutants. Ainsi, il est conseillé d'utiliser le format XML chaque fois que cela est possible, afin d'assurer la pérennité et surtout l'interopérabilité du fichier. Cependant, l'exploitation de formats XML n'est pas toujours adaptée aux usages quotidiens des archéologues. Dans ce cas, nous recommandons les formats suivants :

```
1 <?xml version="1.0"?>
2 <unite_stratigraphique>
3   <numero>
4     0001
5   </numero>
6   <type>
7     positive
8   </type>
9   <!-- Note : penser à rajouter des éléments descripteurs -->
10 </unite_stratigraphique>
```

92. Exemple de fichier XML décrivant l'US 0001, de type positive

- **Documents textuels** : pour les documents éditables, il est recommandé d'utiliser l'Open Document Texte, avec son extension ".odt". Pour l'archivage, le Portable Document Format (.pdf) reste la référence.
- **Données tabulées** : pour un usage quotidien avec des fichiers à plusieurs feuilles de calcul, nous recommandons l'Open Document Spreadsheet (.ods). Pour la diffusion et l'archivage, il faut préférer le CSV (*Comma Separated Values*, .csv). Il s'agit en réalité d'un simple fichier Texte mis en forme, soit un format stable et très polyvalent.
- **Photographies et dessins techniques numérisés** : le format JPEG (.jpg) est pratique pour un usage quotidien, car il est optimisé pour un affichage rapide. Toutefois, il est fondé sur une méthode de compression avec perte, qui réduit la qualité des images d'origine. Le format TIFF (.tif) utilise une méthode de compression sans perte. Les images sont plus volumineuses, mais conservent leur qualité d'origine, ce qui en fait un format plus adapté à la photographie archéologique et aux numérisations de dessins techniques.
- **Dessins vectoriels** : le format libre adapté au DAO est le SVG (*Scalable Vector Graphics*, .svg), une version du XML adapté pour le dessin vectoriel.
- **Illustrations** : le PNG (*Portable Network Graphics*, .png) est un format libre d'image numérique sans perte. Il gère la transparence, ce qui en fait un format adapté pour les illustrations combinant des images matricielles et des images vectorielles.
- **Base de données** : la très grande majorité des bases de données archéologiques sont fondées sur le modèle logique relationnel. Un des outils les plus courants en informatique pour exploiter ce type de base de données est le langage SQL (*Structured Query Language*). Il permet à la fois de déclarer et de gérer des données organisées sous forme de tables affichant des relations hiérarchiques entre elles. Il est donc conseillé de se tourner vers ce type de langage et ses dérivés, tels que MySQL, SQLite et PostgreSQL pour la gestion de données archéologiques.
- **Données spatiales** : les données spatiales comportent plusieurs types de fichier. Pour les images matricielles, il est conseillé d'utiliser le GeoTIFF (.tif), un format dérivé du TIFF contenant des informations de géoréférencement et de géolocalisation. Pour les entités vectorielles, le format le plus répandu est le Shapefile (.shp), qui a été conçu et ouvert par l'entreprise ESRI, propriétaire du logiciel SIG ArcGIS. Toutefois, le

Shapefile est associé au format DBF (.dbf) pour stocker les données attributaires, qui n'est pas un format libre. C'est pourquoi il est recommandé d'évoluer vers le Geopackage (.gpkg), un format défini par l'Open Geospatial Consortium qui peut combiner des images vectorielles, matricielles et des données textuelles dans une base de données SQLite¹²⁷.

Toutes les ressources textuelles présentes dans ces fichiers doivent être encodées en UTF-8, l'un des systèmes de codage de caractère informatique le plus répandu au monde. Tous les formats recommandés ici, sauf le Shapefile et le Geopackage, sont compatibles avec les normes de la plateforme d'archivage PAC¹²⁸ du CINES, la référence française pour la conservation des données numériques de l'ESR. Aucun format équivalent au Shapefile ou au Geopackage n'est pour le moment proposé pour l'archivage au CINES. Probablement d'abord parce qu'il s'agit d'un type de données très spécifique, mais aussi, car elles peuvent être converties en format texte, comme le GeoJSON (.json) ou le CSV.

Les métadonnées

Comme nous l'avons déjà mentionné, les métadonnées, des données qui renseignent les données, sont principalement destinées aux machines (*cf. supra* 2.1.2. Métadonnées). Elles permettent aux programmes informatiques de rechercher, d'identifier et de gérer des données venant de sources diverses, afin de créer un réseau d'informations interopérables. Dans la pratique, les machines numériques ne sont pas conçues pour interpréter des nuances. Ainsi, ces métadonnées descriptives respectent des formats, ou normes, qui déterminent leur structure interne, comme les fichiers. En théorie, tous les fichiers partagés sont censés s'accompagner de métadonnées. La très grande majorité des plateformes de dépôt et d'échange de données archéologiques numériques, comme Nakala en France, utilisent la norme Dublin Core¹²⁹. Cette dernière a vu le jour à l'occasion d'un atelier à Dublin, en Ohio, en mars 1995, qui s'est ensuite organisé autour de l'institution Dublin Core Metadata Initiative (DCMI)¹³⁰. Depuis 2003, le Dublin Core fait l'objet de la norme internationale ISO 15836¹³¹. En France, la BnF propose également un document pour découvrir les spécificités des éléments Dublin Core, sur lequel nous nous sommes appuyé pour la présentation qui suit (BnF 2008).

127 <https://www.ogc.org/standard/geopackage/> ; consulté le 22/10/2023

128 <https://facile.cines.fr/> ; consulté le 22/10/2023

129 <https://documentation.huma-num.fr/nakala/> ; consulté le 22/10/2023

130 <https://www.dublincore.org/about/history/> ; consulté le 22/10/2023

131 <https://www.iso.org/fr/standard/71339.html> ; consulté le 22/10/2023

Le Dublin Core prévoit 15 éléments tous facultatifs et tous répétables, qui portent sur la description :

- du contenu : *Title, Subject, Description, Source, Language, Relation, Coverage* ;
- de la propriété intellectuelle : *Creator, Contributor, Publisher, Rights* ;
- de l'instanciation : *Date, Type, Format, Identifier*.

Ils sont définis par les *DCMI Metadata terms*¹³². Les éléments sont facultatifs dans le sens où la présence des 15 éléments n'est pas requise pour respecter la norme. Toutefois, en archéologie, les données présentent toujours, au minimum, les éléments suivants : *Title, Creator, Subject, Description, Publisher, Date, Type, Format, Identifier, Coverage, Rights*. Ils sont aussi répétables, ce qui signifie que lorsqu'un élément présente plusieurs valeurs, il peut être répété autant de fois que nécessaire. Par exemple, si deux auteurs sont à l'origine de la donnée, il suffit de répéter deux fois l'élément *Creator*. Il existe également la norme *Qualified Dublin Core*, qui ajoute 3 éléments supplémentaires, ainsi que des qualificatifs pour préciser les éléments. Cependant, son usage n'est pas systématique, donc nous ne la présentons pas ici.

Parmi ces 15 éléments, seule la description est réellement prévue pour accueillir du texte libre. Tous les autres font appel à des valeurs contrôlées et normalisées, afin de répondre à l'objectif principal des métadonnées : créer un réseau de données structurellement hétérogènes à partir d'un langage de description commun (fig. 93). Il est donc important qu'une information corresponde à une valeur unique. Pourtant, dans le langage courant, un objet ou un concept peut être désigné par plusieurs mots ou expressions ou, à l'inverse, un mot ou une expression peut désigner plusieurs objets ou concepts à la fois. Le terme "photographie" est un bon exemple. Il peut signifier à la fois la technique ou l'objet qui en résulte, et s'exprimer en plusieurs langues : *photographie, photography, fotografia, etc.* Or, dans la pratique, les machines gèrent mal les synonymes, les homonymes ou les changements de langue d'expression. C'est pourquoi il existe plusieurs référentiels en ligne qui permettent d'individualiser un certain nombre de concepts. Il s'agit de bases de données nationales ou internationales, qui fournissent la plupart du temps un lien pérenne (*Uniform Resource Identifier*, ou URI) vers une description contrôlée et normalisée de la ressource mentionnée. Lorsque la base référentielle est organisée en thésaurus, la description propose également des concepts liés à la ressource, ainsi que des traductions dans d'autres langues.

132 <https://www.dublincore.org/specifications/dublin-core/dcmi-terms/#http://purl.org/dc/terms> ; consulté le 22/10/2023

Élément Dublin Core	Définition	Normalisation
Title	Nom donné à la ressource	En fonction de la convention de nommage définie
Creator	Responsable de la création du contenu de la ressource (personne, organisation ou service)	Nom et URI de l'entité
Subject	Thème du contenu de la ressource	Mots-clés issus d'une liste de vocabulaire contrôlé, avec l'URI de chaque terme
Description	Présentation du contenu de la ressource (résumé, référence à une représentation graphique du contenu, exposé du contenu)	Texte libre
Publisher	Entité responsable de la mise à disposition ou diffusion de la ressource (personne, organisation ou service)	Nom et URI de l'entité
Contributor	Entité responsable de contributions au contenu de la ressource (personne, organisation ou service)	Nom et URI de l'entité
Date	Date d'un événement dans le cycle de vie de la ressource. Il s'agit généralement de la date de création ou de mise à disposition de la ressource	Format AAAA-MM-JJ, AAAA-MM ou AAAA
Type	Nature ou genre du contenu de la ressource	Au minimum un élément conforme aux types DCMI (collection, dataset, event, image, interactive resource, moving image, physical object, service, software, sound, still image, text)
Format	Matérialisation de la ressource (type de medium)	Selon la liste de vocabulaire contrôlé des types de médias internet, ou types MIME
Identifier	Référence univoque et pérenne à la ressource	URI de la ressource
Source	Référence à une ressource dont la ressource décrite est dérivée	URI de la ressource source
Language	Langue du contenu intellectuel de la ressource	Selon la norme ISO 639-3 (code à trois lettres)
Relation	Signifie la relation entre la ressource décrite et une autre ressource, qui est liée à la ressource décrite d'une manière ou d'une autre. Cette relation peut être exprimée de manière réciproque, mais cela n'est pas obligatoire et dépend du type de relation.	URI de la ressource en relation
Coverage	Couverture spatio-temporelle de la ressource (lieu d'invention, culture archéologique)	Mots-clés issus d'une liste de vocabulaire contrôlé, avec l'URI de chaque terme
Rights	Informations sur les droits associés à la ressource	URI de la licence associée à la ressource

93. Définitions et normalisation des éléments Dublin Core

Nous proposons ici une liste de référentiels à prendre en compte pour compléter des métadonnées de documents archéologiques, ainsi qu'une aide à la création des éléments *Type*, *Format* et *Language* en annexe (Annexe 5) :

- **Personne** : ORCID¹³³ est un service international qui fournit sur demande un identifiant numérique persistant (ORCID iD) à toutes les personnes en lien avec le domaine de la recherche. Cet identifiant se présente sous la forme suivante : Quentin Verriez, <https://orcid.org/0000-0002-6823-4183>. Le même type de service existe également pour les auteurs de l'enseignement supérieur et la recherche en France, qui se nomme IdRef¹³⁴ (par exemple : Quentin Verriez, <http://www.idref.fr/192400355/id>).
- **Lieu** : GeoNames¹³⁵ est une base de données géographique collaborative qui permet d'identifier des millions de lieux à travers le monde. Par exemple, pour le mont Beuvray, <https://www.geonames.org/3032913/mont-beuvray.html>
- **Organisation** : La référence mondiale pour les institutions de recherche est le Research Organization Register¹³⁶. Par exemple, pour l'Université Bourgogne Franche-Comté, <https://ror.org/02dn7x778>. En France, il existe également le répertoire national des structures de recherche¹³⁷, mais il ne fournit pas d'URI, seulement un code associé à la structure. Par exemple, pour Bibracte EPCC : 200824243W.
- **Archéologie** : le thésaurus Pactols¹³⁸ est une liste de vocabulaire contrôlée fondée à l'origine par la Fédération et Ressources sur l'Antiquité (Frantiq). Le service est aujourd'hui librement accessible par l'interface d'Opentheso, grâce au soutien du consortium MASA. On y retrouve les concepts fondamentaux de notre discipline (matériels, culturels, temporels et spatiaux), ainsi que des éléments participant à la méthodologie de l'archéologie. Par exemple, l'unité stratigraphique, <https://ark.frantiq.fr/ark:/26678/pcrttppkhp72xq>.

133 <https://orcid.org/> ; consulté le 22/10/2023

134 <https://www.idref.fr/> ; consulté le 22/10/2023

135 <https://geonames.org> ; consulté le 22/10/2023

136 <https://ror.org/> ; consulté le 22/10/2023

137 <https://appliweb.dgri.education.fr/rnsr/> ; consulté le 22/10/2023

138 <https://pactols.frantiq.fr/opentheso/> ; consulté le 22/10/2023

- **Licence** : Les *Creative Commons*¹³⁹ sont les licences les plus utilisées pour le partage de ressources numériques. Nous recommandons l'usage des licences libres CC BY 4.0¹⁴⁰ et CC BY-SA 4.0¹⁴¹. La première permet de partager et modifier une œuvre à condition d'en créditer la source, et la seconde oblige en plus à réutiliser la même licence que l'œuvre originale.
- **Langage** : La DCMI recommande l'usage de la norme ISO 639-3¹⁴², un code à trois lettres pour spécifier la langue d'expression. Nous proposons en annexe un document qui rassemble les principales langues utilisées dans le monde occidental (Annexe 5).
- **Type** : La nature du contenu de la ressource est normalisée par la norme Dublin Core, qui propose 12 types (Annexe 5). En archéologie, la très grande majorité des fichiers sont de type *Dataset* (tableurs, bases de données, etc.), *Still Image* (dessins, planches, cartes, etc.), *Text* (rapports, articles, publications, etc.) ou *Collection*, une agrégation de plusieurs types (par exemple une archive ZIP).
- **Format de fichier** : Le type *Multipurpose Internet Mail Extensions* (**type MIME**) est un standard permettant d'indiquer la nature et le format d'un document. L'*Internet Assigned Numbers Authority* (IANA) est l'organisme officiel responsable du suivi de l'ensemble des types MIME¹⁴³ officiels existants. Nous proposons également en annexe une liste des types MIME des formats recommandés précédemment (Annexe 5).

Pour résumer, le tableau suivant présente l'exemple de métadonnées au format Dublin Core accompagnant le rapport de Bibracte de la campagne 2020 (fig. 94).

139 <https://creativecommons.org/share-your-work/cclicenses/> ; consulté le 22/10/2023

140 <https://creativecommons.org/licenses/by/4.0/deed.fr> ; consulté le 22/10/2023

141 <https://creativecommons.org/licenses/by-sa/4.0/deed.fr> ; consulté le 22/10/2023

142 https://iso639-3.sil.org/code_tables/639/data ; consulté le 22/10/2023

143 <https://www.iana.org/assignments/media-types/media-types.xhtml> ; consulté le 22/10/2023

Élément Dublin core	Exemple
Title	Rapport annuel 2020 du programme quadriennal de recherche 2017-2020 sur le Mont-Beuvray
Creator	Vincent Guichard (https://orcid.org/0000-0003-3269-269X)
Subjet	Archéologie (https://ark.frantiq.fr/ark:/26678/pcrty05M9SVnLu)
Subjet	Oppidum (https://ark.frantiq.fr/ark:/26678/pcrtrolpDW7j36)
Subjet	Bibracte (https://ark.frantiq.fr/ark:/26678/pcrtTNpasdzlib)
Description	Ce volume, qui résume les acquis de la campagne de recherche de 2020 sur le Mont-Beuvray en suivant l'articulation du programme quadriennal 2017-2020, est accompagné d'un Référentiel, qui compile la documentation primaire de la campagne et qui est livré sur support électronique
Publisher	Bibracte EPCC (200824243W)
Date	2021
Type	Text
Format	PDF (application/pdf)
Identifiant	https://hal.science/BIBRACTE/halshs-03261895v1
Language	fra
Coverage	Mont Beuvray (https://www.geonames.org/3032913/mont-beuvray.html)
Coverage	Nièvre (https://www.geonames.org/2990371/nievre.html)
Coverage	Saône-et-Loire (https://www.geonames.org/2976082/saone-et-loire.html)
Coverage	France (https://www.geonames.org/3017382/republic-of-france.html)
Coverage	Second âge du Fer (https://ark.frantiq.fr/ark:/26678/pcrtxVmyWBbIQq)
Coverage	1er s. av. J.-C. (https://ark.frantiq.fr/ark:/26678/pcrtzSW6S4S5D)
Coverage	Civilisation de La Tène (https://ark.frantiq.fr/ark:/26678/pcrtRJxa5OT7xR)
Rights	CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/deed.fr)

94. Les métadonnées du rapport de synthèse 2020 du programme de recherche du mont Beuvray

Nous avons d'abord choisi le format tableur pour plus de clarté, mais en réalité les métadonnées Dublin Core doivent se présenter sous la forme d'un fichier XML, comme dans l'exemple suivant (fig. 95). Toutes les informations pour construire le balisage XML au format Dublin Core sont documentées dans le guide d'utilisation mis à disposition par la DCMI¹⁴⁴. Un outil libre, *Dublin Core Generator*, est également disponible en ligne pour construire le code XML à partir d'un formulaire¹⁴⁵.

```
<?xml version="1.0" encoding="UTF-8"?>

<dc:title>Rapport annuel 2020 du programme quadriennal de recherche 2017-2020 sur le Mont-Beuvray</dc:title>
<dc:creator>Vincent Guichard</dc:creator>
<dc:creator>https://orcid.org/0000-0003-3269-269X</dc:creator>
<dc:subject>Archéologie</dc:subject>
<dc:subject>https://ark.frantiq.fr/ark:/26678/pcrty05M9SVnLu</dc:subject>
<dc:subject>Oppidum</dc:subject>
<dc:subject>https://ark.frantiq.fr/ark:/26678/pctrlpDW7j36</dc:subject>
<dc:subject>Bibracte</dc:subject>
<dc:subject>https://ark.frantiq.fr/ark:/26678/pcrtTNpasdzlib</dc:subject>
<dc:description>Ce volume, qui résume les acquis de la campagne de recherche de 2020 sur le Mont-Beuvray en suivant l'articulation du programme quadriennal 2017-2020, est accompagné d'un Référentiel, qui compile la documentation primaire de la campagne et qui est livré sur support électronique</dc:description>
<dc:publisher>Bibracte EPCC (200824243W)</dc:publisher>
<dc:date xsi:type="dcterms:W3CDTF">2021</dc:date>
<dc:type xsi:type="dcterms:DCMIType">Text</dc:type>
<dc:format xsi:type="dcterms:IMT">application/pdf</dc:format>
<dc:identifiant xsi:type="dcterms:URI">https://hal.science/BIBRACTE/halshs-03261895v1</dc:identifiant>
<dc:language xsi:type="dcterms:ISO639-3">fra</dc:language>
<dc:coverage>Mont Beuvray</dc:coverage>
<dc:coverage>https://www.geonames.org/3032913/mont-beuvray.html</dc:coverage>
<dc:coverage>Nièvre</dc:coverage>
<dc:coverage>https://www.geonames.org/2990371/nievre.html</dc:coverage>
<dc:coverage>Saône-et-Loire</dc:coverage>
<dc:coverage>https://www.geonames.org/2976082/saone-et-loire.html</dc:coverage>
<dc:coverage>France</dc:coverage>
<dc:coverage>https://www.geonames.org/3017382/republic-of-france.html</dc:coverage>
<dc:coverage>Second âge du Fer</dc:coverage>
<dc:coverage>https://ark.frantiq.fr/ark:/26678/pcrtxVmyWBblQq</dc:coverage>
<dc:coverage>1er s. av. J.-C.</dc:coverage>
<dc:coverage>https://ark.frantiq.fr/ark:/26678/pcrtzSW6S4S5D</dc:coverage>
<dc:coverage>Civilisation de La Tène</dc:coverage>
<dc:coverage>https://ark.frantiq.fr/ark:/26678/pcrtRJxa5OT7xR</dc:coverage>
<dc:rights>CC BY-NC-SA 4.0</dc:rights>
<dc:rights xsi:type="dcterms:URI">https://creativecommons.org/licenses/by-nc-sa/4.0/deed.fr</dc:rights>
```

95. Les métadonnées du rapport de synthèse 2020 du programme de recherche du mont Beuvray au format XML

144 <https://www.dublincore.org/resources/userguide/> ; consulté le 22/10/2023

145 https://nsteffel.github.io/dublin_core_generator/index.html ; consulté le 22/10/2023

En théorie, un fichier XML au format Dublin Core est censé accompagner chaque fichier archivé et diffusé. Ainsi, quelle que soit la nature, la méthode de production ou la finalité du fichier, il peut normalement s'intégrer dans un vaste réseau de ressources numériques hétérogènes décrites par un langage commun. Ce langage commun, le Dublin Core, est suffisamment clair pour être compris par l'utilisateur, tout en étant structuré, contrôlé et normalisé pour permettre une exploitation par les machines à grande échelle. Cependant, d'après notre expérience, les archéologues ne possèdent actuellement ni les compétences informatiques ni les solutions techniques pour produire efficacement des métadonnées pour tous les fichiers qu'ils souhaitent diffuser. En attendant des solutions plus appropriées, nous considérons que l'échelle la plus pertinente pour générer des métadonnées est celle du jeu de données résultant d'un projet archéologique cohérent, tel qu'une campagne de fouilles. En effet, un seul fichier de métadonnées complet pour l'ensemble du jeu de données permet aux outils numériques de le moissonner. L'archéologue peut ensuite naviguer dans le jeu de données pour identifier les différentes ressources disponibles. Cette navigation est généralement facilitée par les nombreux inventaires que les archéologues réalisent, tels que des inventaires de photos, de minutes, de points topographiques, ou encore des inventaires préliminaires du mobilier. Pour l'instant, ces inventaires jouent un rôle de métadonnées au sein d'un jeu de données cohérent et sont essentiels dans nos pratiques.

Entrepôt pérenne

Il existe plusieurs solutions pour rendre des données archéologiques accessibles en ligne. La priorité est de différencier les services construits comme des entrepôts de données pérennes et ceux orientés vers la logique de fonctionnement des réseaux sociaux. Les premiers proposent généralement un espace de stockage avec des processus de dépôt rigoureux, nécessitant une production de métadonnées. Une fois les données déposées, il est possible de les diffuser sous une adresse unique et stable. Cela signifie que même si les données sont mises à jour ou déplacées, par l'auteur ou par le service pour des besoins de maintenance, le lien reste assuré de manière permanente. Ce type de relation est possible grâce aux technologies comme le *Digital Object Identifier*¹⁴⁶ (DOI), l'*Handle*¹⁴⁷ ou encore l'*Archival Resource Key*¹⁴⁸ (ARK). Les entrepôts pérennes mettent donc en avant le libre accès à des données interopérables, dans le respect des principes de la science ouverte.

146 <https://www.doi.org/> ; consulté le 22/10/2023

147 <https://www.handle.net/> ; consulté le 22/10/2023

148 <https://arks.org/> ; consulté le 22/10/2023

À l'inverse, pour les services conçus sur le modèle des réseaux sociaux, tels que *Academia*¹⁴⁹ ou *ResearchGate*¹⁵⁰, l'objectif est d'accumuler le plus d'utilisateurs. Ils se montrent donc moins exigeants sur la qualité des données déposées et la production de métadonnées associées. Les données sont de plus en plus associées à des identifiants uniques et permanents, comme les DOI sur *ResearchGate*, mais les données ne sont pas interopérables. Enfin, il s'agit d'organisations commerciales, qui financent actuellement leur modèle à travers la publicité ou les formules d'abonnement premium en échange de divers services. L'accès aux données est le moment majoritairement gratuit, mais il existe toujours le risque que le modèle économique évolue vers un système plus restrictif dans les années à venir. Il convient également de ne pas confondre les entrepôts numériques pérennes avec les gestionnaires de fichiers en ligne, ou *Cloud*, à l'image de *DropBox*¹⁵¹ ou *Google Drive*¹⁵². Ce type d'outil agit comme un espace de travail décentralisé, accessible depuis plusieurs machines grâce à internet. Ils s'adressent à une utilisation individuelle, ou pour un groupe restreint. L'organisation des fichiers est trop souple pour permettre la mise en place de liens uniques et permanents.

Ainsi, pour les données archéologiques en France, nous recommandons l'utilisation du service Nakala de l'IR* Huma-Num, un entrepôt de données de recherche pour les Sciences Humaines et Sociales¹⁵³. Le service est développé pour assurer à la fois la conservation, la diffusion et la réutilisation des données. Il est ouvert à toutes les personnes justifiant d'un statut en lien avec la recherche en SHS (chercheurs, enseignants, doctorants) et d'un projet scientifique producteur de données numériques. La documentation disponible pour l'utilisation de Nakala se trouve à l'adresse suivante : <https://documentation.huma-num.fr/nakala/>. Les dépôts et les métadonnées peuvent être créés, gérés et administrés soit par l'interface web dédiée, soit par une API (Application Programming Interface) pour les utilisateurs expérimentés¹⁵⁴. Nakala est organisé en données, qui peuvent être rassemblées en collections. Le plus important à prendre en compte est qu'une fois qu'une donnée est déposée et rendue publique, elle ne peut plus être supprimée par l'utilisateur. En effet, comme les ressources partagées sont associées à un DOI, Huma-Num bloque cette fonction afin d'éviter les "liens morts", qui ne pointent vers aucune ressource.

149 <https://www.academia.edu/> ; consulté le 22/10/2023

150 <https://www.researchgate.net/> ; consulté le 22/10/2023

151 <https://www.dropbox.com/> ; consulté le 22/10/2023

152 <https://www.google.com/intl/fr/drive/> ; consulté le 22/10/2023

153 <https://nakala.fr/> ; consulté le 22/10/2023

154 <https://api.nakala.fr/doc> ; consulté le 22/10/2023

Pour la diffusion, il suffit ensuite de partager le DOI d'une donnée si l'on souhaite la partager de manière isolée, ou le DOI d'une collection pour partager l'ensemble des données. Comme nous l'avons déjà mentionné dans nos recommandations sur les métadonnées, le dépôt et la création de métadonnées fichier par fichier nous semblent inadaptés à nos pratiques. La plupart des archéologues ne possèdent ni le temps ni les compétences pour effectuer ce type de démarche par l'API, et le nombre de fichiers produits dans le cadre d'une fouille est souvent trop important pour la réaliser un à un par l'interface web de Nakala. Il paraît au contraire bien plus réaliste de rassembler les données dans une archive informatique, par exemple au format ZIP (.zip), et de déposer le jeu en une fois avec des métadonnées complètes et cohérentes. Cette manière de faire limite grandement l'interopérabilité des données elles-mêmes, mais elle permet néanmoins de diffuser des jeux de données archéologiques contrôlés, identifiés et documentés, en accord avec les principes de la science ouverte. Malheureusement, cette pratique limite également les possibilités d'archivage numérique sur le long terme. Les données hébergées sur Nakala peuvent en effet faire l'objet, sur demande, d'un dépôt sur la plateforme d'archivage du CINES (PAC). Cependant, à ce jour, PAC n'accepte aucun format d'archivage informatique. Nakala s'engage toutefois sur la préservation des données et nous n'avons jamais constaté de problème à ce niveau depuis le début de cette recherche doctorale.

En France, d'autres solutions de dépôt adaptées à certains types de données archéologiques existent, mais nous avons fait le choix de les évoquer dans une partie spécifique, qui dresse l'état actuel des Humanités Numériques en archéologie (*cf. infra* partie 4.3 Humanités Numériques).

Plan de gestion de données

Deux parties sont déjà consacrées aux Plans de Gestions de Données dans ce manuscrit (*cf. supra* 2.3. PGD et 3.7. Diffusion données primaires). Aussi nous résumons ainsi la fonction du PGD : le PGD est à l'utilisateur ce que les métadonnées sont aux machines. Il permet aux producteurs et aux utilisateurs, pendant et après un projet de recherche, de constituer, d'identifier, de partager et d'exploiter un jeu de données. C'est pourquoi la prise en compte de toutes les recommandations précédentes facilite la rédaction d'un PGD, car elles reprennent dans leurs grandes majorités des éléments constitutifs de ce type de document. C'est le cas pour le modèle du consortium MASA¹⁵⁵, le plus abouti actuellement pour documenter un corpus de données issu de fouilles archéologiques.

155 https://dmp.opidor.fr/public_templates?search=MASA ; consulté le 22/10/2023

La contrainte la plus importante pour la rédaction d'un PGD est pour le moment le rôle qu'on lui attribue. Il est souvent vécu par les chercheurs comme un nouveau document administratif imposé pour justifier une source de financement, mais sans apport réel pour leur recherche. Pourtant, ce type de document propose aussi une méthode au service des archéologues. Le bon déroulement d'une fouille archéologique repose sur le respect de règles et le partage de responsabilités plus ou moins tacites entre les différents acteurs. De la même manière, il est indispensable de définir des règles et des standards à respecter pour faciliter la création, la gestion et l'échange de données numériques durant le projet comme après la diffusion des résultats. Cependant, contrairement au terrain, la bonne gestion des données dans une démarche de science ouverte n'est pas intégrée dans les pratiques des archéologues. Un PGD, en posant les questions adaptées, peut donc réellement les accompagner dans la mise en œuvre des "bonnes pratiques" en ce qui concerne les ressources numériques. Il intervient à toutes les étapes du projet :

Au début du projet :

- Le PGD permet d'anticiper les règles qui vont déterminer le cycle de vie des données numériques (collecte et création, gestion, diffusion et archivage) au sein d'un projet, tel qu'une opération archéologique.
- Il établit les responsabilités et tâches des différents acteurs du projet.

En cours de projet :

- Il propose un manuel pour faciliter le partage de données entre les acteurs du projet.
- Il permet de prendre du recul et d'adopter un regard critique sur la qualité et la quantité des données produites.
- Il permet d'adopter les bons réflexes pour la diffusion et le partage des données.

En fin de projet :

- Il authentifie le jeu de données renseigné, avec toutes les informations pour identifier le ou les auteurs et les conditions de diffusion et de réutilisation.
- Il justifie les motifs de création et de diffusion des données, afin d'encourager la critique et la validation par les pairs.
- Il documente les processus de création et d'administration des données, afin d'en faciliter la réutilisation.

Pour que le PGD soit pertinent pour les chercheurs, il doit donc être anticipé dès la phase de conception d'un projet, soit dès les études documentaires préalables à une opération de terrain en archéologie. Cette prise en compte dès l'émergence d'un projet permet d'anticiper des questions essentielles sur la gestion des données, et ainsi de limiter la perte d'informations durant le projet et la rétrodocumentation au moment de la publication des données. De plus, faire évoluer les PGD avec le projet permet de répartir la charge de travail nécessaire à sa rédaction sur l'ensemble du projet.

Il convient d'intervenir sur le PGD à chaque événement ayant une répercussion notable sur le cycle de vie des données. Toutefois, la rédaction d'un PGD n'est pour le moment pas intégré dans la chaîne opératoire archéologique. Nous vous recommandons donc de définir en amont les critères qui déclenchent l'actualisation des données, telle que :

- l'apport de nouvelles données ;
- un changement dans les équipes de producteurs ou de gestionnaires ;
- la modification d'une norme de production ;
- une évolution dans le système de traitement et de gestion des données ;
- un changement dans les conditions de sélection des données destinées à l'archivage.

En plus, il est important de définir une période régulière de révision complète du PGD, par exemple :

- une première version dans les six mois suivant le début du projet de recherche ;
- une version à chaque évaluation périodique, comme le rendu annuel d'un rapport d'opération ;
- une version finale accompagne le rapport final du projet de recherche.

4.2.3. La science ouverte en archéologie : un horizon inaccessible ?

La transition vers la science ouverte représente un tournant significatif pour la communauté archéologique. L'adoption des recommandations que nous venons de décrire pourrait avoir un impact significatif sur la manière dont nous partageons les résultats de nos recherches et accédons aux données archéologiques. Respecter ces recommandations permet d'imaginer un scénario où chaque rapport de fouille, au lieu d'être un document isolé, serait systématiquement accompagné d'un jeu de données en formats libres. Ces données seraient organisées selon une convention de nommage claire, garantissant une cohérence dans l'ensemble des informations archéologiques.

Ce jeu de données serait complété par des métadonnées détaillées, facilitant l'exploitation de son contenu et de son contexte. Un Plan de Gestion de Données (PGD) accompagnerait également ce jeu de données, décrivant en détail la manière dont les données sont collectées, stockées, traitées et partagées. Le PGD servirait de guide pour garantir la qualité et l'intégrité des données aux utilisateurs et protéger les intérêts des producteurs.

Le rapport, le jeu de données et le PGD seraient hébergés sur des entrepôts pérennes, conçus pour garantir la disponibilité à long terme des informations. Ces entrepôts seraient ouverts et accessibles à tous. Ils permettraient de mettre fin aux barrières qui limitent actuellement l'accès aux données archéologiques.

L'impact d'une telle transition vers la science ouverte serait considérable. Elle permettrait de multiplier les opportunités de partage de connaissances et d'applications pratiques. Les chercheurs, les étudiants, les institutions archéologiques et patrimoniales pourraient explorer ces données, les analyser, et les utiliser pour des projets de recherche, d'éducation ou de préservation du patrimoine. La communauté archéologique franchirait ainsi une étape majeure vers une démarche scientifique plus ouverte, transparente et collaborative, tout en contribuant à la préservation et à la diffusion du patrimoine archéologique.

Cependant, ce type de scénario se confronte encore à la réalité des Humanités Numériques françaises en archéologie, qui présentent encore des limites pour l'hébergement et le partage pérenne d'archives de fouilles numériques.

4.3. L'évolution des Humanités Numériques en archéologie française

L'environnement numérique pour l'archéologie française a considérablement évolué durant cette thèse, signe qu'une partie de la communauté s'est approprié les questions relatives à la science ouverte. Nous présentons ici les principales plateformes de partage adaptées aux données archéologiques, tout en évaluant leurs capacités à héberger des archives de fouilles numériques comme celles produites dans le cadre des expérimentations sur le chantier de la *domus* PC2.

4.3.1. HAL science ouverte

HAL¹⁵⁶ est aujourd'hui l'une des principales plateformes de partage de ressources textuelles et graphiques en libre accès destinées aux données de la recherche en langue française. Au 31 décembre 2023, elle recensait 1 132 483 documents consultables¹⁵⁷. Alors que la plateforme était principalement reconnue dans les années 2010 pour le partage d'articles prépublication (*preprints*), elle a connu de nombreuses évolutions récemment. Une partie de ces évolutions est attribuable au soutien apporté par le Plan National pour la Science Ouverte du Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation, qui a désigné HAL comme le service national d'archives numériques ouvertes pour la communauté de recherche (MESRI 2018, p. 4-5 ; MESRI 2021b, p. 9-10).

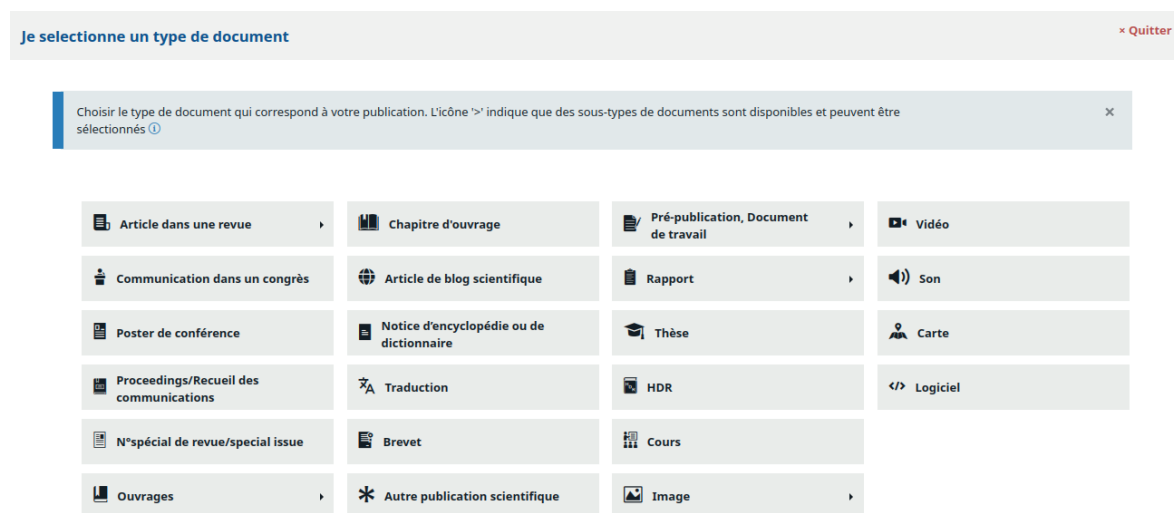
Dans cette optique, le second plan vise à atteindre 100 % de publications en accès ouvert d'ici 2030. HAL a donc subi une refonte significative pour améliorer l'ergonomie et les procédures de dépôt, dans le but d'encourager chercheurs et institutions à partager des documents conformes aux normes de la science ouverte, dans toutes les disciplines. Ces améliorations font aujourd'hui de HAL la plateforme à privilégier pour le dépôt de littérature scientifique, telle que les rapports, les articles ou les travaux universitaires.

Cependant, son orientation demeure principalement axée sur la présentation de résultats plutôt que sur les données de la recherche. Il est probable que ces limitations évoluent rapidement, étant donné les développements récents de HAL, mais lors du dépôt, les types de document ne proposent pas de sections pour les données primaires (fig. 96). Il est toute-

156 <https://hal.science/> ; consulté le 22/10/2023

157 <https://www.ccsd.cnrs.fr/2023/03/ccsd-publication-du-rapport-d-activite-2022/> ; consulté le 22/10/2023

fois possible de déclarer dans les métadonnées du document déposé un lien vers un jeu de données associé. Pour les archives de fouille numériques, HAL est donc à prioriser pour le partage du rapport d'opération, tout en renseignant le ou les liens vers d'autres entrepôts pérennes qui hébergent les jeux de données associés.



96. Types de dépôts possibles sur HAL

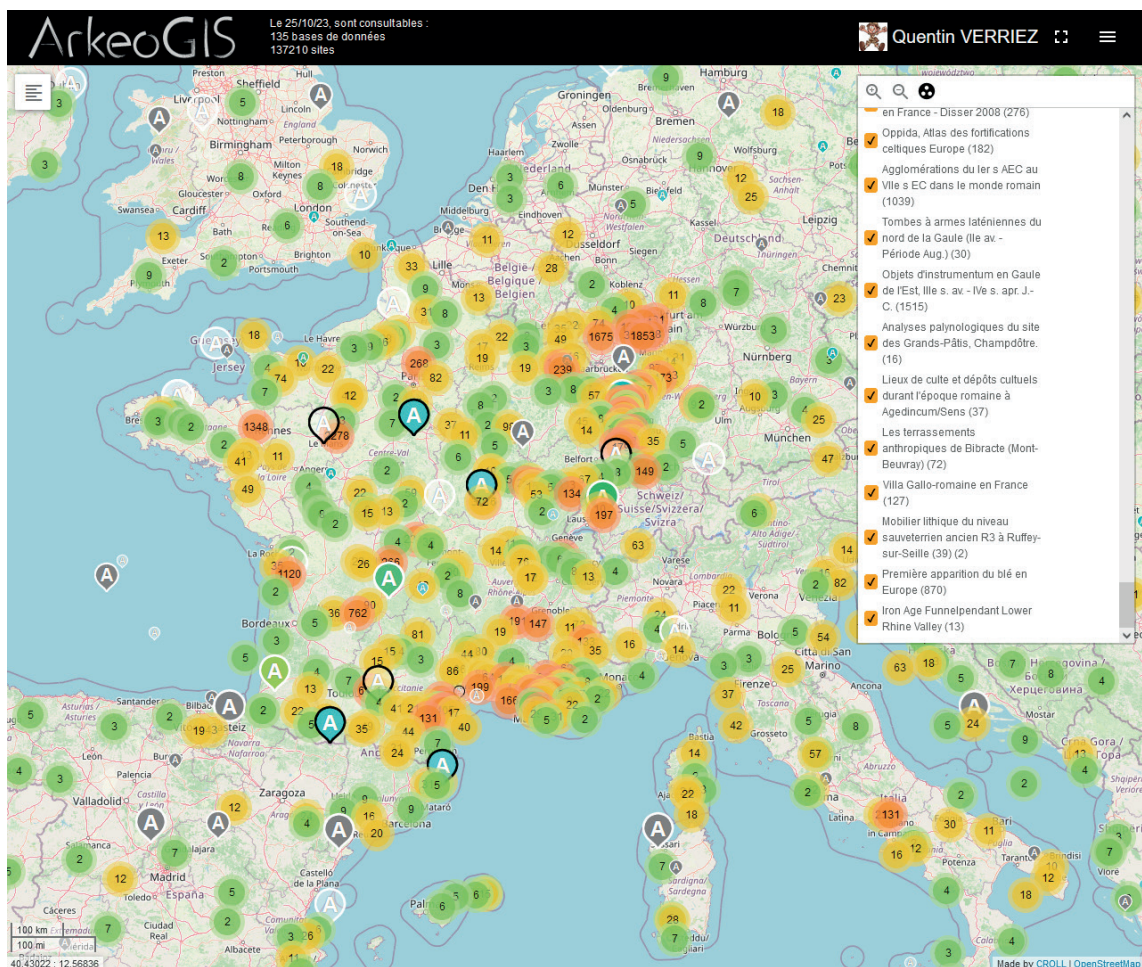
4.3.2. ArkeoGIS

ArkeoGIS¹⁵⁸ est une solution en ligne pour partager et interroger des données spatialisées "sur le passé". La plateforme affiche une volonté transdisciplinaire, puisqu'elle accueille les données issues de plusieurs disciplines étudiant des événements anthropiques anciens, tels que l'archéologie, l'histoire, la géographie ou le paléoenvironnement. L'origine de ces données est aussi variée, car il peut s'agir de travaux de chercheurs, d'étudiants en master et doctorat, d'opérateurs d'archéologie préventive ou encore de services de l'État.

Lors de sa création en 2006, la plateforme était gérée conjointement par la Maison Internuniversitaire des Sciences de l'Homme (MISHA) et l'Université de Strasbourg. Elle a depuis été transférée sur les serveurs d'Huma-Num, afin d'assurer la pérennité de l'outil et des données hébergées. Elle a également intégré la suite d'outils soutenue par le consortium MASA, confirmant la position d'ArkeoGIS en tant que plateforme nationale pour l'ouverture des données archéologiques.

158 <https://arkeogis.org/> ; consulté le 22/10/2023

ArkeoGIS fonctionne sur le principe du WebSIG. Une fois connecté par un accès sécurisé, l'utilisateur peut naviguer sur une carte interactive qui permet de visualiser les données disponibles (fig. 97). La plateforme souhaite ainsi offrir aux utilisateurs des éléments pour contextualiser leurs propres travaux de recherche et renforcer la mise en commun des résultats. Pour cela, les concepteurs ont fait le choix de ne présenter que des données spatiales simples. Les données sont localisées sur des fonds de carte génériques sous forme d'entités ponctuelles et d'emprises au sol, afin de signaler la présence d'un site, d'une structure, d'un objet étudié ou d'une étude spécialisée (fig. 98). À ce jour, la plateforme héberge 135 jeux de données regroupant 137 210 sites répartis principalement en Europe et au Proche-Orient (données relevées le 23/10/2023). Depuis 2022, les ressources partagées sous licences libres sur ArkeoGIS sont également accessibles sans accès sécurisé par l'interface ArkeOpen¹⁵⁹. Cette dernière présente l'avantage de fournir l'accès direct aux données sur Nakala, afin de pouvoir les télécharger librement (fig. 99).



97. Exemple de données signalées pour l'âge du Fer sur ArkeoGIS

**BIBRACTE** (863)

Bdd: Interventions archéologiques à Bibracte
(2010-2018) - Bibracte Numérique
Auteur(s): Matthieu THIVET

PÉRIODE DE DÉBUT
-120

PÉRIODE DE FIN
-20



Latitude: 46.92405847351997 Longitude: 4.034677219303598

Nom de la commune: Glux-en-Glenne

Précision des coordonnées: Fine

Occupation: Non renseigné

Date de création de la fiche: 17/9/19 23:28

Période de début: -120 / -120 jusqu'à la période de fin: -20 / -20

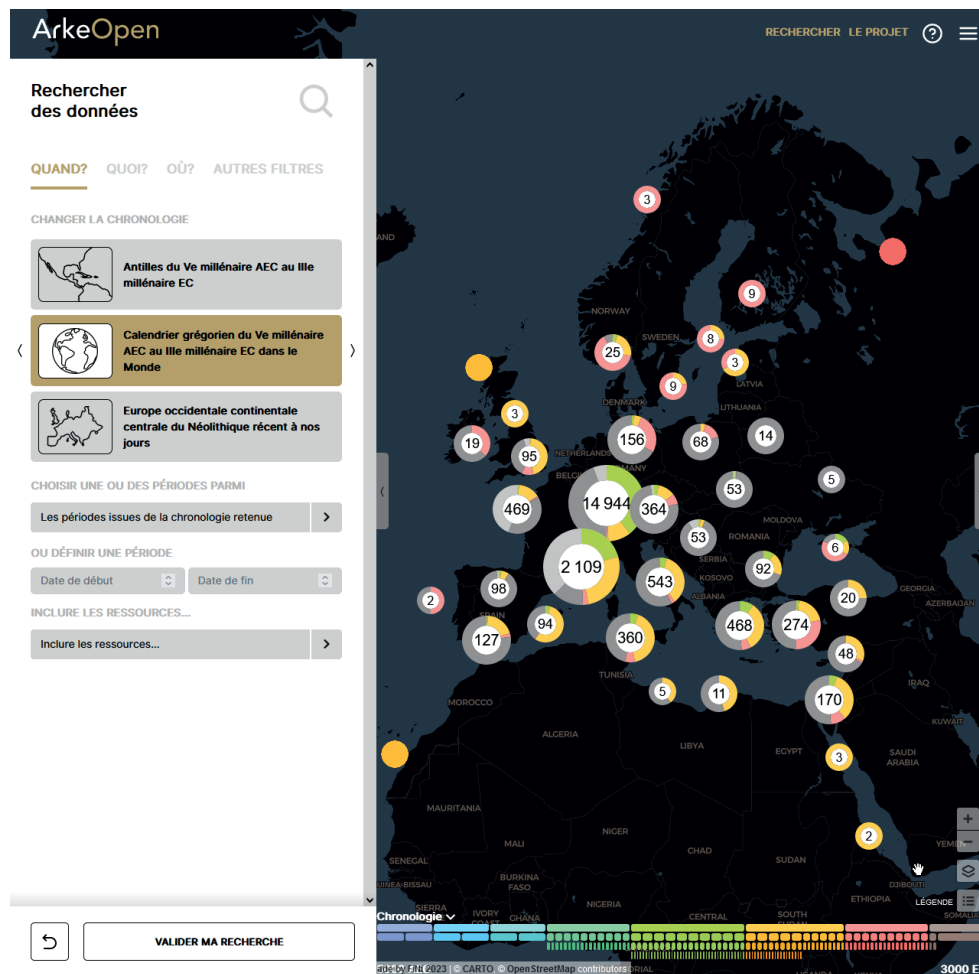
Immobilier**Immobilier / Habitat / Groupé / Enclos**

État des connaissances: Fouillé

Commentaire: année d'intervention : 2017 / responsable d'intervention : Ch. Martini , Q. Verriez / nombre d'UF : non renseigné

Bibliographie: <https://halshs.archives-ouvertes.fr/halshs-01671478>

98. Exemple de données signalées sur ArkeoGIS : la campagne de fouille 2017 de la domus PC2



99. Page d'accueil de ArkeOpen

Les données sont déposées par les utilisateurs. Pour être importés, les jeux de données doivent respecter un formatage précis défini par la plateforme. L'ensemble se présente sous la forme d'un tableur CSV¹⁶⁰ de 22 colonnes, dont 10 sont obligatoirement à remplir (Bernard, Jullien 2022). Tous les termes à employer pour caractériser une entité sont organisés en respectant la norme des thésaurus. Les thésaurus ArkeGIS sont alignés sur les thésaurus PACTOLS et Getty, en plus d'être compatibles avec le modèle sémantique de référence CIDOC CRM¹⁶¹. De plus, les dépositaires doivent renseigner les métadonnées des jeux de données lors de l'importation, en respectant le format Dublin Core.

160 <https://arkeogis.org/manuel/importation/> ; consulté le 22/10/2023

161 <https://nakala.fr/collection/10.34847/nkl.f62682m1> ; consulté le 22/10/2023

Le formatage des données, s'il peut paraître lourd pour certains contributeurs, assure le bon fonctionnement des deux finalités principales d'ArkeoGIS. Premièrement, la mise en forme assure l'interopérabilité des jeux de données déposés sur la plateforme. Ils peuvent ainsi tous faire l'objet de requêtes par catégorie, par zone géographique et/ou par fourchette chronologique. Il est ensuite possible d'exporter le résultat d'une requête dans un seul nouveau fichier CSV formalisé. Le deuxième aspect concerne la visibilité et l'accessibilité des jeux de données déposés. Comme ArkeoGIS repose sur l'infrastructure d'Huma-Num, les jeux de données peuvent être automatiquement moissonnés par des portails de recherche externes à la plateforme, tels qu'Isidore.

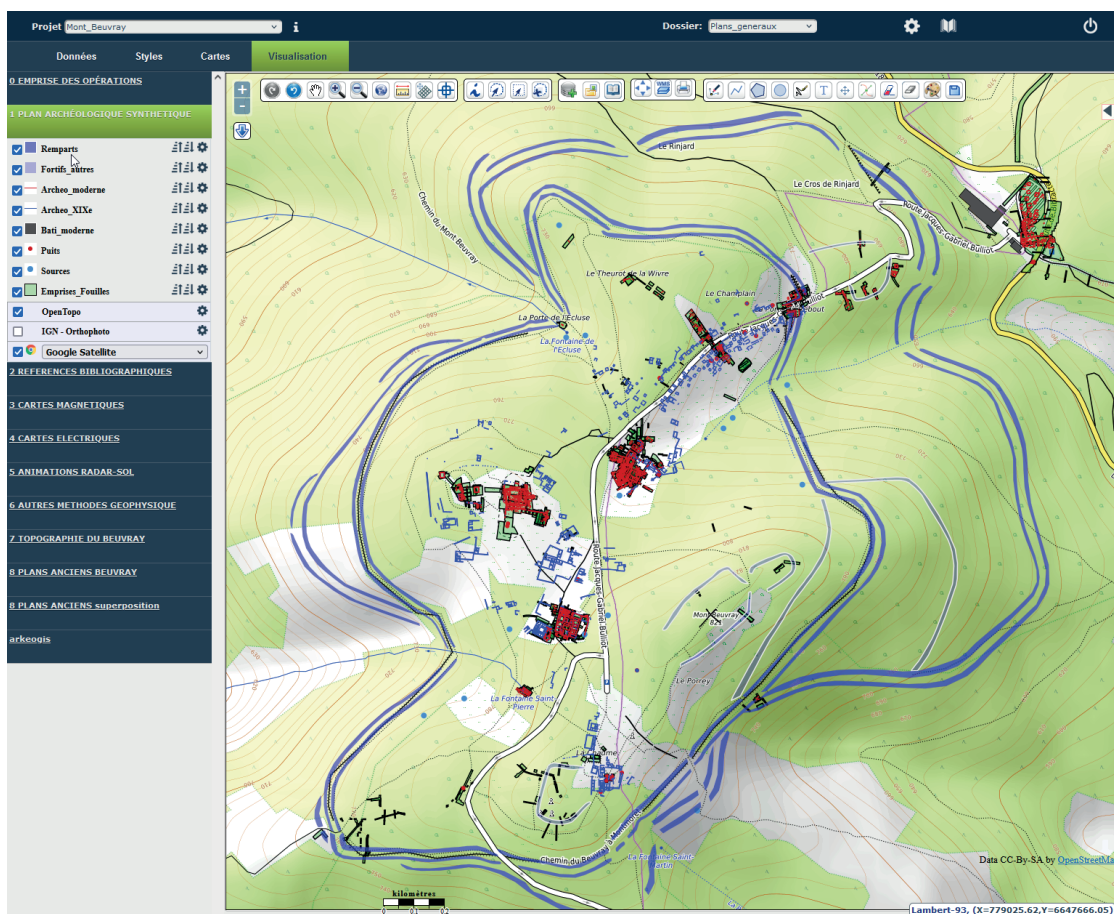
Par sa nature, ArkeoGIS a pour vocation de signaler l'existence et de mettre en relation des jeux de données, plutôt que de permettre de les exploiter en ligne. La grille de lecture de la plateforme se place volontairement à l'échelle du site archéologique ou de l'objet étudié. ArkeoGIS est donc un bon outil de contextualisation, capable de rassembler dans une base unique un grand nombre de données sur une aire géographique définie. C'est pourquoi la plateforme n'est pas conçue pour interroger des données au niveau de l'US, qui seraient beaucoup plus complexes à formater pour permettre l'interopérabilité entre les jeux de données. Pour ce type de données, il est possible de renseigner un lien vers d'autres jeux de données une fois le site renseigné sur ArkeoGIS. La domus PC2, comme l'intégralité des interventions archéologiques ayant eu lieu entre 2010 et 2018, est donc renseignée sur le portail ArkeoGIS. Pour une solution web-SIG adaptée à l'échelle du site archéologique, il est possible de s'orienter vers Chronocarto.

4.3.3. Chronocarto

La plateforme Chronocarto¹⁶² est fondée au départ sur la solution GCserver de la société Geocarta, afin de proposer en 2006 un outil de web-SIG au sein du laboratoire AOrOc (UMR 8546). Depuis, l'outil s'est largement développé et prend aujourd'hui l'aspect d'un portail en ligne qui accueille plusieurs applications collaboratives pour la gestion et l'exploitation de données numériques scientifiques. Par les origines de la plateforme, les données disponibles sont issues des domaines de l'archéologie et de la géologie, mais Chronocarto souhaite aujourd'hui ouvrir ses outils à l'ensemble de la communauté scientifique. Comme pour ArkeoGIS, Chronocarto est soutenu par Huma-Num, ce qui signifie que l'intégralité des outils et des données est hébergée sur les serveurs de l'IR*.

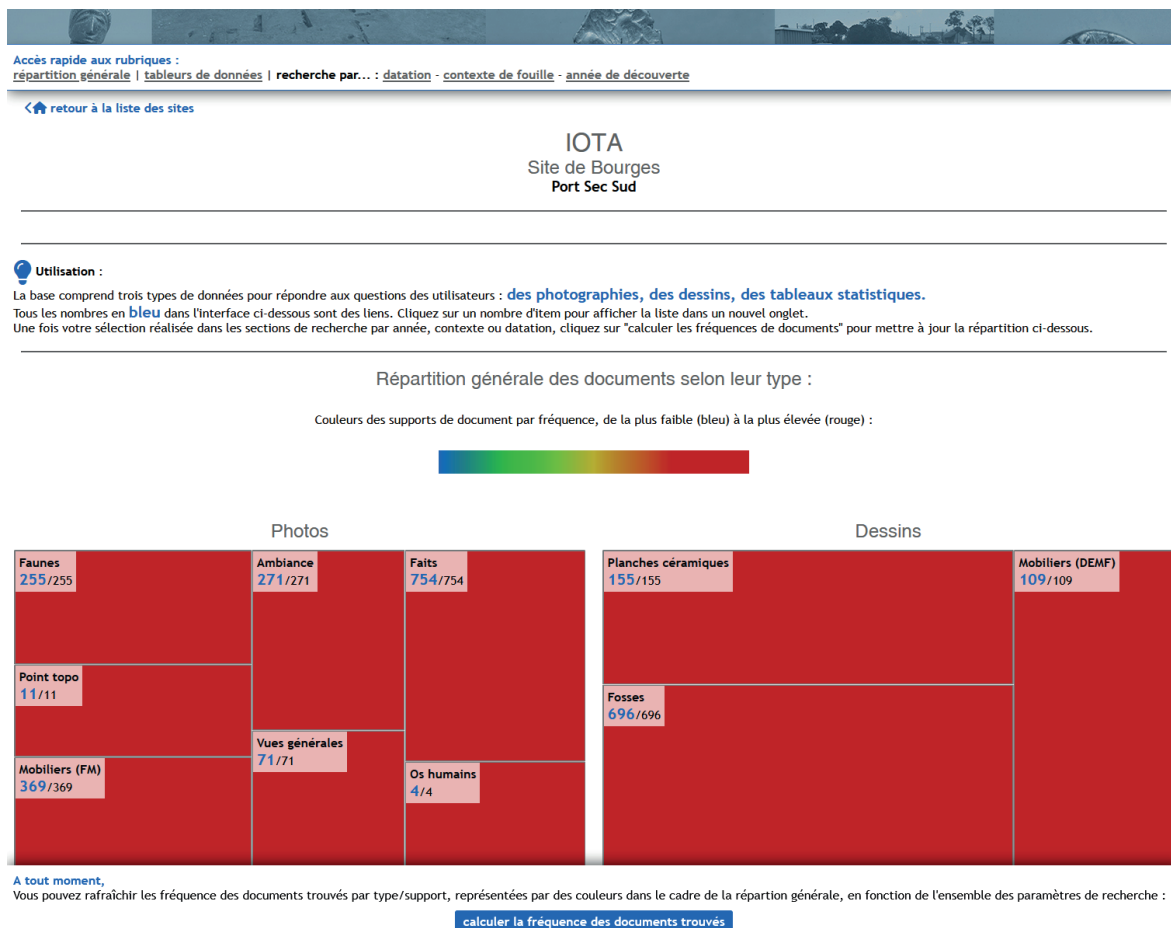
162 <https://www.chronocarto.eu/> ; consulté le 22/10/2023

Le web-SIG reste pour le moment l'application centrale de Chronocarto. Il s'agit d'une version allégée des logiciels de SIG, comme QGIS, qui fonctionne intégralement en ligne (fig. 100). Du fait du public visé, l'application priorise les fonctionnalités de cartographie les plus utilisées pour l'étude de données archéologiques. La demande d'accès au service se fait sur demande, après avis d'un comité éditorial. Ainsi, chaque groupe de contributeurs peut accéder à son projet afin de visualiser, d'importer ou de modifier des données à référence spatiale sur un espace commun. Cet espace permet de construire collaborativement et de partager les cartes d'un projet de recherche en cours. Les projets peuvent être accessibles au public ou sécurisés par un mot de passe selon la volonté des propriétaires de projet. Les différents projets affichés sur le site sont organisés par dossiers thématiques, eux-mêmes classés selon les échelles internationales, régionales ou locales.



100. Exemple du projet web-SIG du programme de recherche du mont Beuvray sur Chronocarto (données M. Dabas, Ch. Sanchez)

Chronocarto réunit également plusieurs bases de données d'envergure nationale et internationale, telles que la *Base Fer*¹⁶³ (inventaire des sites de l'âge du Fer) ou *AcoR*¹⁶⁴ (Atlas des techniques de construction romaine). Ces bases ont fait l'objet d'un travail avancé pour permettre leur mise à disposition en ligne. Chacune possède sa propre interface graphique, ainsi que sa logique de navigation et d'exploitation. Dans ce cas, Chronocarto agit avant tout comme un portail qui fédère les accès vers ce type de bases de données. Cependant, Chronocarto propose aussi un service d'ouverture de données primaires grâce à son moteur de recherche graphique et textuel IOTA (Imagerie de l'Objet au Territoire Archéologique)¹⁶⁵. Cet outil permet en effet de créer une interface générique pour explorer des jeux de données archéologiques déposés sur un entrepôt Nakala (fig. 101).



101. Le moteur de recherche IOTA, accessible depuis Chronocarto

163 <https://www.chronocarto.eu/spip.php?article8&lang=fr> ; consulté le 22/10/2023

164 <https://www.chronocarto.eu/spip.php?article3&lang=fr> ; consulté le 22/10/2023

165 <http://s802828443.onlinehome.fr/iota3/> ; consulté le 22/10/2023

Enfin, depuis peu, Chronocarto propose une plateforme d'édition en ligne, avec la volonté de produire des publications qui exploitent pleinement le potentiel du numérique. À titre d'exemple, le premier article publié, qui présente les résultats de la fouille de "Port Sec" à Bourges, affiche une série de 12 plans éditorialisés et interactifs, directement issus du web-SIG du projet disponible sur Chronocarto (Buchenschutz, Batardy 2022).

Finalement, Chronocarto a adopté une politique différente d'ArkeoGIS pour le dépôt des données. Les contributeurs de Chronocarto ont accès à plusieurs catégories de dépôt (cartographie, jeu de données, publication) et la structuration interne des données n'est pas imposée. Cette souplesse répond à la volonté de créer des espaces de travail collaboratifs, mais limite l'interopérabilité des données entre les projets déposés. Chronocarto est donc une des solutions les plus abouties pour visualiser et partager des données à référence spatiale détaillées à l'heure actuelle. Toutefois, l'aspect collaboratif nous semble encore assez limité. En effet, les fonctionnalités restent trop peu nombreuses pour créer et gérer un SIG complet en cours de projet, surtout à plusieurs mains. Pour effectuer des modifications, la meilleure solution reste encore de télécharger le projet, le traiter sur une machine locale puis de le téléverser à nouveau sur Chronocarto. Enfin, dans l'état actuel, le jeu de données de la *domus* PC2 ne peut pas être intégralement déposé sur le projet réservé à Bibracte, car il comporte beaucoup d'orthophotographies volumineuses. Ces dernières doivent être traitées et allégées avant un éventuel dépôt, sous peine de saturer les capacités d'affichage de l'outil sur le web.

4.3.4. Conservatoire National des données 3D

Le Conservatoire National des données 3D¹⁶⁶ (CND3D) est présenté par le Consortium 3D SHS comme un espace pour assurer la conservation et le partage des données 3D produites dans le cadre de projets de recherche en SHS (fig. 102). En tant que consortium soutenu par Huma-Num, la pérennité des outils et des données hébergés sur le CND3D est encore une fois assurée par l'infrastructure de l'IR*. Pour sa conception, le consortium s'est fortement inspiré du projet similaire ArcheoGRID, créé en 2007 par la plateforme Archeovision¹⁶⁷. C'est pourquoi cette structure, qui réunit des spécialistes de l'acquisition et de la restitution de données 3D dans le domaine du patrimoine historique et archéologique, est également en charge du CND3D.

166 <https://3d.humanities.science/> ; consulté le 22/10/2023

167 <https://www.archeogrid.fr/> <https://www.chronocarto.eu/>

Carnyx Tintignac

Nature de la ressource	dépôt
Date du dépôt	2008-01
Créateur du dépôt	Sarah Tournon-Valiente 
Entité dépositante	Archeovision
Nb objets virtuels 3D	1
Nombre de fichiers	0
Nom du site (texte libre)	TINTIGNAC
Date projet	2008
Date archéologique min.	-600
Date archéologique max.	499
Commanditaire	Service Régional d'Archéologie (SRA)
Nature du dépôt	ACQUISITION_RESTITUTION
PeriodO	Antiquité



Localisation



Citation: **Pascal Mora,Loïc Espinasse,François Daniel, Christophe Maniquet (INRAP),Martine Fabioux (DRAC Limousin) (2008).** Carnyx Tintignac , *Archeovision*.  10.34969/CND3D/216007.d.2008

Contenu du dépôt (1 Objet Virtuel 3D)

Objet Virtuel 3D  Carnyx Tintignac



Titre	Carnyx Tintignac
Créateur(s)	Pascal Mora Loïc Espinasse François Daniel
Date 3D	2008
Date archéologique	-600 - 499
Version	V2
Contributeur(s)	Christophe Maniquet (INRAP) Martine Fabioux (DRAC Limousin)
Date creation 3D min	2008-01-01
Date creation 3D max	2008-12-31
Categorie	Monument

102. Exemple de ressource hébergée sur le CND3D : modèle 3D du Carnyx de Tintignac (données Mora et al. 2008)

Dans le cadre de la science ouverte, le premier objectif du CND3D est d'offrir aux données 3D déposées une existence en tant que production scientifique, au même titre que les publications. Les dépôts se présentent ainsi sous la forme de pages d'inventaire liées à un identifiant unique (DOI), qu'il est possible de citer en tant que source bibliographique ou sitographique. Cependant, l'accès aux fichiers sources n'est que très rarement disponible. Pourtant, la solution technique existe, car quelques modèles affichent un lien de téléchargement, comme les numérisations d'ossements gravettiens de la collection GRAVETT'OS¹⁶⁸ du Muséum national d'Histoire naturelle (Thibeault, Villotte 2021). Il s'agit donc probablement d'une volonté des auteurs de dépôt, qui dans leur grande majorité ne souhaitent pas partager directement les fichiers sources.

Dans le détail, un dépôt regroupe trois types d'information :

- un ou plusieurs fichiers 3D ;
- les fichiers jugés utiles pour la compréhension des données 3D, soit les fichiers sources ayant mené à la production du ou des fichiers 3D ;
- les métadonnées nécessaires pour assurer la bonne gestion des fichiers dans le cadre du CND3D et de la science ouverte.

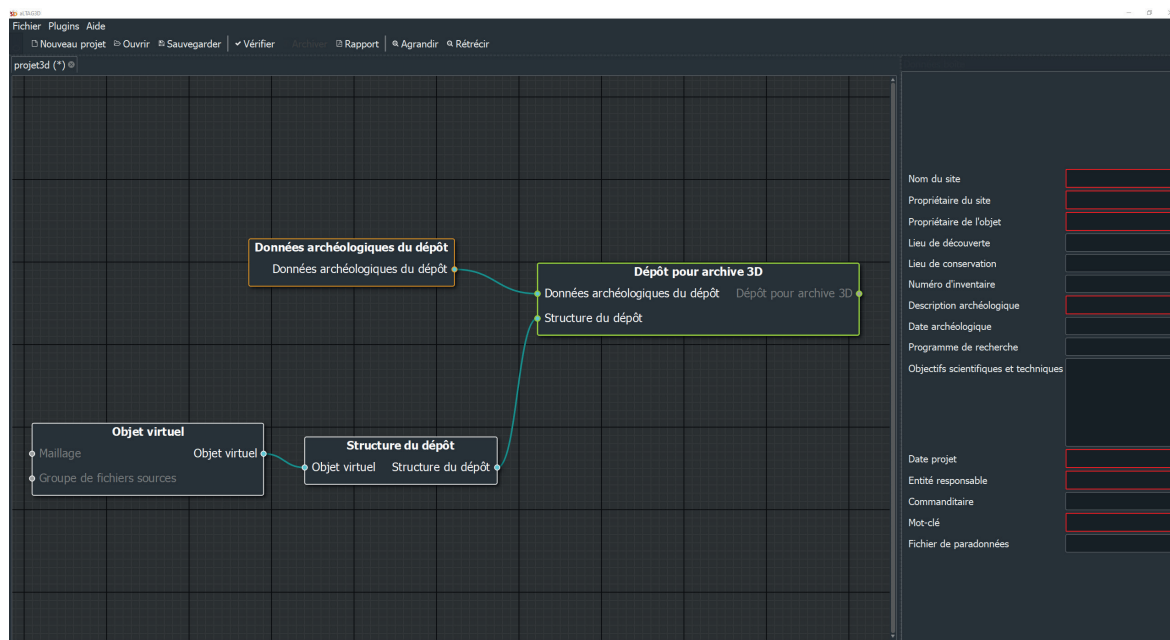
Le CND3D n'impose pas de formats de conservation pour les fichiers 3D et n'effectue pas de contrôle sur la qualité des données. Il revient au déposant, et éventuellement à la personne qui souhaite citer la ressource, d'assumer cette responsabilité. Pour cela, les contributeurs ont besoin d'un compte, qui nécessite de justifier d'une appartenance à un organisme de la recherche en SHS.

Le dépôt peut être effectué de deux façons, soit en utilisant un formulaire en ligne, soit en utilisant le logiciel dédié aLTAG3D¹⁶⁹. Cet outil permet de renseigner sous la forme d'un schéma UML (*Unified Modeling Language*) le processus de production d'un objet 3D et de le traduire automatiquement en un fichier XML qui respecte les métadonnées définies par le Consortium 3D SHS (fig. 103). Ces métadonnées sont alignées sur les critères d'archivage de l'outil PAC du CINES, ce qui permet d'automatiser la constitution d'une archive numérique à long terme. Sur demande, il est également possible d'intégrer une interface de manipulation en ligne de l'objet 3D renseigné.

Dans sa forme actuelle, il semble donc que le CND3D ne réponde que partiellement aux objectifs fixés par le Consortium 3D SHS. S'il s'agit effectivement d'un lieu de conservation pérenne des modèles 3D, la dimension de partage reste plus discutable. Certes, les données sont visibles, mais elles sont, pour la plupart, inaccessibles. Sachant qu'aucun contrôle qualité n'est effectué sur les lots déposés, il est ardu de critiquer la qualité des modèles 3D dans la plupart des cas. Sans ce retour critique possible par des pairs, il apparaît alors problématique de citer "objet 3D" comme source dans un texte scientifique. Sur les aspects pratiques en revanche, le CND3D a su trouver, selon nous, le juste équilibre entre la rigidité technique des machines et le besoin de souplesse des utilisateurs. Le logiciel aLTAG3D en est une bonne illustration. Il s'agit d'une solution ingénieuse pour produire intuitivement, à l'aide d'une in-

169 <https://altag3d.huma-num.fr/> ; consulté le 22/10/2023

terface graphique, une archive numérique normalisée prenant en compte tout le cycle de vie de données 3D. Ce type de logiciel devrait être ouvert à toutes les données manipulées en SHS, car, à l'heure actuelle, la production de métadonnées exploitables reste complexe, alors qu'il s'agit d'un point essentiel pour le développement de pratiques numériques responsables. Toutefois, il est évident que le travail mené par le Consortium 3D SHS pour définir un schéma de métadonnées validé par l'ensemble de la communauté d'utilisateurs est indispensable pour arriver au résultat d'aLTAG3D. La même démarche doit être appliquée à tous les autres types de données avant d'envisager le développement d'un logiciel capable de produire des métadonnées normalisées de manière semi-automatique.



103. Logiciel aLTAG3D, pour la documentation du processus de production et d'archivage de modèles 3D

En ce qui concerne un éventuel dépôt des données des fouilles de Bibracte sur le CND3D, comme nous l'avons déjà évoqué, nous n'avons pas produit de modèles 3D dans le cadre du protocole d'enregistrement numérique (*cf. supra* 3.3.3 Photogrammétrie). En effet, selon notre expérience, à l'heure actuelle, les acquisitions et les traitements 3D sont chronophages et nécessitent trop de ressources informatiques et d'espace de stockage, pour un résultat similaire aux techniques de représentation archéologique en 2D.

4.3.5. Bilan sur les plateformes de dépôt

En premier lieu, il est intéressant de noter que des plateformes d'envergure nationale comme Chronocarto et le CND3D, qui se veulent transdisciplinaires, ont à l'origine été fondés par des archéologues. Cet état de fait peut signifier deux choses : soit, les archéologues considèrent que toutes les données scientifiques numériques possèdent une structure unique ; soit, qu'être capable de traiter des données archéologiques revient à être capable de traiter un grand nombre de données. En effet, l'archéologie emprunte ses méthodes à de nombreuses autres disciplines, pour explorer des dimensions à la fois matérielles, sociales, temporelles, géographiques, sémantiques, *etc.* De plus, par la nature des vestiges étudiés, les données sont pour la plupart très fragmentaires. Ainsi, les outils informatiques capables de gérer la complexité, le volume et l'hétérogénéité des données archéologiques, tout en respectant le besoin de précision souvent élevé lors du traitement des résultats, peuvent s'avérer fonctionnels pour d'autres applications scientifiques.

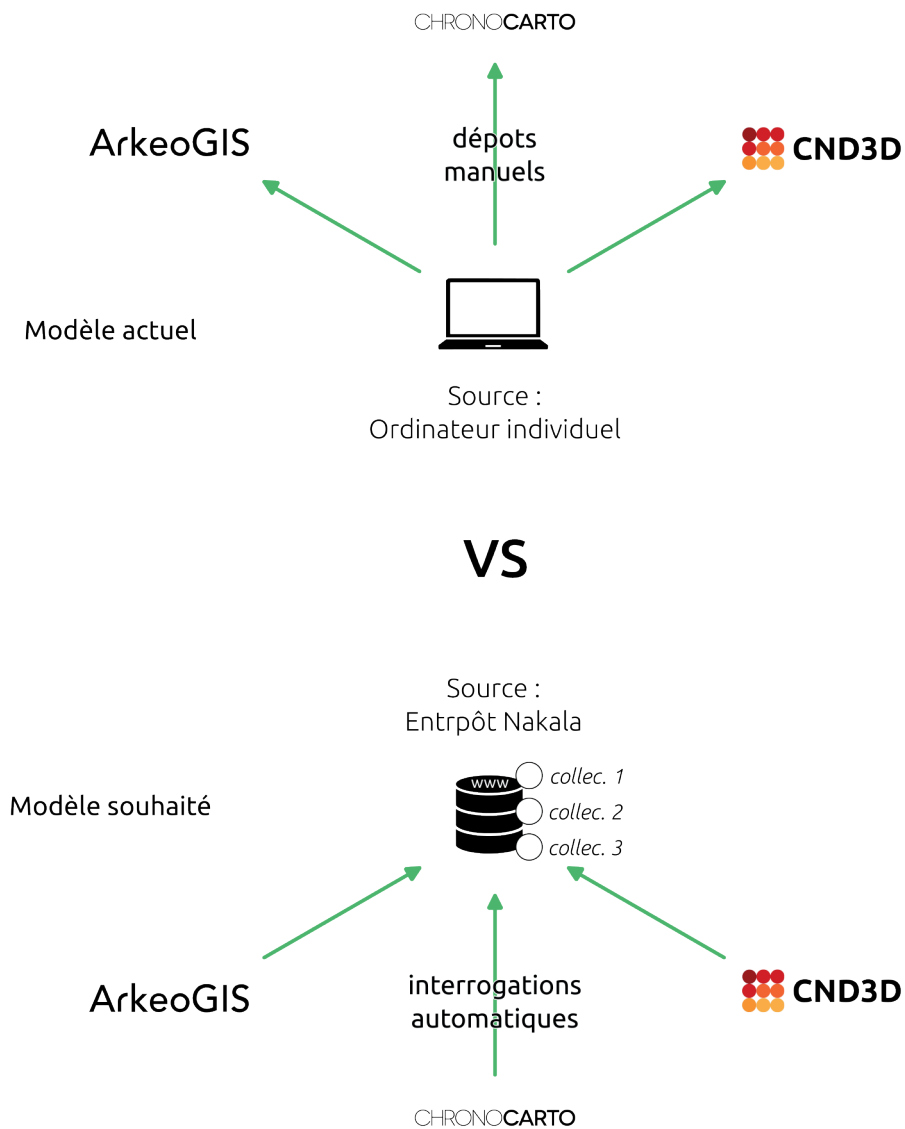
Le fait de devoir gérer un dépôt de données par plateforme est aussi un aspect contraignant pour les gestionnaires de données. Même si nous reconnaissons qu'il est techniquement et économiquement difficile de faire autrement pour construire une plateforme, cette situation complexifie grandement la tâche aux archéologues déposants. Ils doivent en effet :

- multiplier les dépôts en découpant ou dédoublant un ensemble de données cohérent ;
- répéter la tâche de mise aux normes en respectant les règles de formatage qui diffèrent d'une plateforme à l'autre ;
- assurer les mises à jour en cas de modification du jeu de données sources.

Dans ces conditions, il est difficile pour un archéologue, dont la conservation et la diffusion des données ne sont pas les seules missions, de maintenir la qualité des lots déposés. Au lieu de multiplier les dépôts, il serait donc plus cohérent de centraliser le jeu de données sources sur un entrepôt unique en ligne, dont la gestion serait assurée par l'archéologue. Il pourrait ensuite choisir de lier son entrepôt aux différentes plateformes, pour rendre visibles et partager ses données ou collaborer avec d'autres chercheurs. Sachant que les plateformes sont déjà hébergées par Huma-Num, toute l'infrastructure existe pour mettre en place cette solution. En effet, les données proposées à ArkeoGIS, Chronocarto et le CND3D sont conservées sur les espaces Nakala respectifs de ces trois structures. Avec le modèle proposé, l'objectif est

que les chercheurs ou les institutions possèdent leur propre espace Nakala, et de mettre en place les solutions techniques pour que les plateformes viennent les interroger (fig. 104). Il s'agit, selon nous, d'une solution "gagnant-gagnant", car elle permet :

- au déposant de garder le contrôle sur les données
- aux plateformes de se libérer de la gestion des données pour se concentrer sur le développement d'outils et de fonctionnalités favorisant la visibilité et l'interopérabilité des données.



104. Modèles de gestion des dépôts de données sur les plateformes nationales

Évidemment, pour fonctionner, ce modèle requiert des développements techniques, mais également de définir les bonnes pratiques et les règles sémantiques pour échanger des informations numériques au sein de la communauté archéologique. C'est pourquoi cette problématique est au cœur du nouveau projet ANR SIAMOIS (*Système d'Information Archéologique Mutualisé et Ouvert reposant sur l'Intelligence Sémantique*) coordonnée par Bibracte depuis 2023.

Enfin, ce tour d'horizon des différentes plateformes met en évidence l'importance qu'a pris l'IR* Huma-Num dans le paysage numérique français en SHS ces dernières années. L'équipe d'Huma-Num a su mesurer les enjeux et mettre en place les solutions techniques adaptées, comme le démontre l'intérêt de la communauté SHS pour leurs services. Il est évident que le fait d'avoir accès à des outils collaboratifs, des solutions d'hébergements et des espaces de stockage en ligne de manière quasi illimitée représente un réel atout pour nos domaines scientifiques. D'autant plus quand les fournisseurs de ces services laissent aux chercheurs la possibilité de s'organiser en consortiums pour définir eux-mêmes leurs pratiques plutôt que de les imposer.

Toutefois, comme toujours avec le numérique, si tout accessible en ligne, rien n'est virtuel et tout à un coût. Pour exister, les services d'Huma-Num nécessitent un personnel qualifié, capable d'assurer le développement et la maintenance d'une infrastructure matérielle et logicielle à l'échelle nationale. Classiquement, ce genre de service est assuré par les GAFAM américains (Google, Amazon, Facebook, Apple, Microsoft), ou plus récemment leur équivalent chinois, les BATX (Baidu, Alibaba, Tencent et Xiaomi). L'IR* rivalise donc avec ces "géants" privés en conservant un personnel et une infrastructure basée en France et financée par des initiatives publiques.

En résumé, le modèle proposé par Huma-Num est précieux, car il assure la pérennité et l'indépendance de la recherche française en SHS sur les aspects liés au cycle de vie des données numériques. Ainsi, il est naturel que l'IR* devienne le centre névralgique numérique de cette communauté. Cependant, si la concentration des services facilite l'évolution technique et l'harmonisation progressive des pratiques numériques, elle présente aussi plusieurs risques :

En comparaison du nombre de services proposés, Huma-Num est actuellement constitué d'une petite équipe. Si les membres de la communauté scientifique continuent d'appuyer leurs recherches sur l'infrastructure de l'IR*, mais sans raisonner leur utilisation, il existe

un risque que le personnel soit dépassé et ne puisse plus assurer une telle qualité de service.

Le choix d'Huma-Num de soutenir des consortiums est pour le moment très favorable et particulièrement adapté aux besoins de la communauté scientifique. Mais ce choix est rendu possible, entre autres, grâce à des politiques nationales et européennes pour le moment très en faveur de la science ouverte, une situation qui peut cependant évoluer très rapidement.

Pour assurer la gratuité des services, le financement d'Huma-Num repose entièrement sur des fonds publics qui peuvent varier à tout moment. Dans le pire des scénarios, un arrêt des financements signifie un arrêt total de l'infrastructure, se traduisant par des outils et des données inaccessibles.

Il est donc indispensable que les utilisateurs mesurent la chance d'avoir accès à des services indépendants, souverains et à l'écoute des besoins des acteurs de la recherche en SHS, tout en prenant conscience de la fragilité du modèle pour être prêt à le défendre en cas de besoin.

CONCLUSION ET PERSPECTIVES

L'intensification des pratiques numériques en archéologie depuis les années 2000 est le résultat d'une longue évolution des méthodes de travail dans le domaine. Dès la création des premiers ordinateurs, dans les années 1960, des archéologues ont exploité ces outils pour collecter, organiser et traiter des informations sous forme numérisées. Ils voyaient dans ces machines des moyens adaptés pour traiter leurs données par définition volumineuses, complexes et fragmentaires. À peine une vingtaine d'années plus tard, des spécialistes, comme J.-C. Gardin, ont estimé que la quantité et le rythme de production d'informations archéologiques avaient déjà dépassé les capacités de traitement des êtres humains (Tchernia *et al.* 1986). La miniaturisation et la baisse de prix des équipements numériques, au début du XXI^e s., ont simplement permis de répondre à ce besoin de traitement en démocratisant l'utilisation des ordinateurs et autres dispositifs au sein de la discipline. L'analyse des évolutions numériques de l'*Archaeological Data Service* au Royaume-Uni et de l'*Inrap* en France révèle que cette transformation s'est réalisée de manière progressive, davantage en réaction aux besoins qu'en anticipant véritablement l'intégration des technologies numériques dans les pratiques professionnelles des archéologues. Le nombre de données archéologiques numériques en circulation a donc augmenté au fur et à mesure des progrès technologiques, tels que l'augmentation exponentielle des capacités de calcul des microprocesseurs et l'essor d'internet, et des progrès ergonomiques, avec des interfaces logiciels simplifiant de plus en plus les usages.

En revanche, l'utilisation des technologies numériques en archéologie se caractérise par le fait que la plupart des archéologues n'ont jamais pleinement intégré ces outils dans leur pratique professionnelle. Les contraintes de temps, de ressources humaines et financières ne leur ont laissé que peu de marge pour former adéquatement leurs compétences informatiques. De plus, le volume et le niveau de complexité inhérents aux données archéologiques ont probablement été à la fois la cause et la conséquence de leur incapacité à déléguer le développement d'outils numériques adaptés à leurs besoins à des experts en informatique. Ils ont ainsi élaboré leurs propres outils informatiques et ajusté leurs méthodes de travail en fonction de leurs compétences préalables, même si cela impliquait parfois d'utiliser ces technologies de manière détournée. Ces méthodes, souvent employées de manière individuelle, ont conduit à la création d'outils et de données diversifiés, qui s'avèrent souvent difficiles à exploiter ou à comprendre en dehors de la communauté de chercheurs qui les a créés et qui les utilise directement.

Dans le même temps, le volume de données numérisées a continué d'augmenter. L'apparition de la photogrammétrie aéroportée sur les chantiers de fouilles de Bibracte en est un exemple significatif. Entre 2010 et 2015, les volumes de données produits par certaines équipes ont été multipliés par 10, passant de 5 à 50 Go. Nous nous retrouvons ainsi dans une situation paradoxale : les technologies numériques, issues des sciences de l'information et de la communication, ouvrent la voie à un partage de connaissances à une échelle sans précédent en archéologie. Pourtant, les archéologues continuent à générer un volume croissant de données numériques qui peinent à interagir entre elles. En d'autres termes, si l'on doit considérer une rupture, celle-ci réside moins dans une révolution de nos méthodes de travail provoquée par les technologies numériques que dans la menace grandissante de ne plus pouvoir gérer, dans un avenir proche, la masse croissante de données archéologiques numériques produites au quotidien.

La science ouverte apparaît comme la solution idéale pour répondre aux défis d'administration, de conservation et de partage des données archéologiques numériques. Nous proposons de définir ce mouvement comme essentiel pour l'archéologie et pour la science en général, car il a pour objectif d'assurer la pérennité et le partage universel de la connaissance, sans barrières cognitives, techniques ou financières, et ce afin de maintenir une recherche de qualité et de promouvoir l'innovation. En trois piliers, qui sont les logiciels libres, les données ouvertes et le libre accès aux publications scientifiques, la science ouverte définit un ensemble de pratiques pour mettre en œuvre ces principes.

Nous avons donc entrepris de numériser, à partir de 2018, le protocole d'enregistrement du chantier de la *domus* PC2, sur le mont Beuvray, afin de vérifier s'il était possible d'intégrer ces principes dans nos pratiques de fouilles. L'expérience a été réalisée sous la forme d'un audit : nous avons volontairement maximisé l'utilisation de technologies numériques afin d'en faire ressortir les avantages et les limites dans un contexte de fouille archéologique. Cinq technologies ont ainsi été mises en œuvre sur les campagnes de 2017 à 2020 : le drone, la photographie numérique et la photogrammétrie, les SIG et les outils de saisie numérique. Pour la partie enregistrement graphique, la mise en routine du protocole a permis d'effectuer le relevé de chaque couche archéologique directement sur SIG, en s'appuyant sur la précision d'orthophotographies. Pour la partie textuelle, les tests ont été effectués avec deux outils de saisie numérique, EDArc et ODS, en parallèle de la production de la documentation papier. Toutes les actions, de l'acquisition au traitement des données, ont été effectuées sur le terrain, au plus près des vestiges archéologiques. Le protocole est opérationnel et documenté, la partie sur la photogrammétrie ayant même fait l'objet d'une publication, sous la forme d'un guide méthodologique accessible librement en ligne (Verriez *et al.* 2023).

Trois points essentiels ressortent de cette expérience :

Il est possible de réaliser une fouille archéologique en respectant les principes de la science ouverte. Tous les logiciels utilisés dans le cadre du protocole sont libres, les données ont été produites dans des formats ouverts et les archives de fouilles sont disponibles et accessibles sur l'entrepôt Nakala de Bibracte¹⁷⁰. Cette expérience nous a également fait réaliser qu'à l'échelle de l'archéologue de terrain, la science ouverte ne repose pas sur la maîtrise d'outils informatiques complexes ou des dernières innovations technologiques. Tous les outils numériques que nous avons utilisés sont déjà intégrés dans les pratiques de nombreux archéologues. Nous nous sommes contentés de mettre en place des habitudes de travail pour les faire communiquer dans le cadre d'un protocole d'enregistrement numérique complet sur le terrain. Notre démarche en direction de la science ouverte a donc essentiellement consisté à acquérir les bons réflexes en matière de création, de traitement, de gestion, de conservation et de diffusion des données archéologiques numériques, que nous avons synthétisé sous la forme de recommandations à la fin de ce manuscrit.

170 <https://nakala.fr/10.34847/nkl.8d81q3es> ; consulté le 22/10/2023

La préparation des données pour leur diffusion est l'étape la plus complexe à appréhender. Dans la plupart des cas, les équipes d'archéologues fonctionnent de manière relativement isolée durant le temps de la fouille. La production et l'échange des données se concentrent essentiellement au sein de l'équipe. En ce sens, intégrer les principes de la science ouverte revient essentiellement à transposer les logiciels ou les formats de fichiers couramment employés par les membres de l'équipe du modèle propriétaire à celui du libre. Cependant, le premier intérêt d'utiliser des logiciels et des formats libres est de pouvoir diffuser largement des résultats clairs et réutilisables. Pour réussir cette étape, il est nécessaire d'acquérir des compétences qui ne sont pas celles des archéologues. Il s'agit par exemple de maîtriser les normes des métadonnées Dublin Core, d'identifier les principaux référentiels numériques persistants, comme ORCID, GeoNames ou les thésaurus Pactols, ou de connaître les éléments attendus dans un Plan de Gestion de Données. De plus, même si la diffusion intervient à la fin de la chaîne opératoire, la plupart de ces éléments prennent forme dès la conception d'un projet de fouille. La gestion rétrospective des données, en vue de leur diffusion à la fin du projet, s'avère généralement lourde et complexe. Il est donc essentiel pour les archéologues d'acquérir ces compétences s'ils souhaitent créer un réseau de partage qui assure la pérennité et le libre-échange des connaissances archéologiques.

L'expérience sur les chantiers de Bibracte démontre également que développer la science ouverte nécessite souvent de faire évoluer nos outils de gestion et de conservation des archives de fouilles numériques. En effet, les données produites dans le cadre de cette thèse sont techniquement libres, car elles sont construites autour de formats ouverts, documentées par un PGD et hébergées en ligne sur un entrepôt pérenne et accessible. Toutefois, ce travail de recherche a mis en évidence que l'actuel processus de diffusion de l'établissement est capable de produire des données lisibles, mais qui limite encore fortement les possibilités de réutilisation, du fait des formats employés, comme le PDF. En d'autres termes, le processus de diffusion manque d'interopérabilité. Ce résultat a offert une base de réflexion pour le nouveau projet numérique de Bibracte et de ses partenaires, qui ambitionne de construire un *Système d'Information Archéologique Mutualisé et Ouvert reposant sur l'Intelligence Sémantique* (SIAMOIS). SIAMOIS met en avant la documentation des méthodes plutôt que la normalisation des outils, afin d'améliorer la capacité des technologies numériques à fonctionner en dehors d'interactions prédéfinies. Il ne s'agit pas non plus d'imposer un langage de documentation commun, mais au contraire d'utiliser les technologies du Web sémantique pour faire communiquer les différentes documentations des outils et méthodes exploitées en archéologie. Ce type de démarche est notamment rendu possible par l'utilisa-

tion de thésaurus polyhiérarchiques, capables de renseigner et de catégoriser les relations entre un grand nombre de concepts. En ce sens, les démarches comme celles proposées par SIAMOIS représentent, selon nous, le futur des pratiques numériques en archéologie. Le Web sémantique et les thésaurus polyhiérarchiques permettent à chacun d'utiliser les méthodes de travail, les outils et les structures de données qu'il considère adaptés à ses besoins, tout en permettant d'intégrer son système logique dans un réseau pouvant être traité automatiquement par les machines.

L'intégration des outils numériques dans les méthodes archéologiques ne se traduit pas automatiquement par une amélioration globale de l'efficacité au travail et de la qualité des données. Notre étude des différentes technologies a mis en évidence que, sans un suivi actif, les outils entraînent certaines dérives méthodologiques et imposent de nouvelles contraintes liées à la nature numérique des données. Par exemple, les orthomosaïques offrent un mode de représentation d'une rare précision pour les archéologues. Nous avons ainsi observé que certains fouilleurs les jugent suffisantes en tant qu'enregistrement graphique. Pourtant, les orthomosaïques ne participent pas d'un processus interprétatif et ne peuvent donc pas remplacer les relevés archéologiques. De plus, en tant qu'abstractions du monde réel, elles présentent des limites techniques intrinsèques (luminosité, colorimétrie, déformations, *etc.*), qui ne devraient pas permettre de les supplanter aux observations directes des vestiges. Enfin, il s'agit de données particulièrement volumineuses, et leur diffusion et leur conservation posent des questions à la fois logistiques, pour l'organisation des espaces de stockage, et environnementales, par l'énergie et les ressources consommées pour produire, alimenter et connecter au réseau ces espaces de stockages.

Ce constat nous a poussé à considérer les technologies numériques au-delà de leur dimension technique. Le travail de plusieurs spécialistes démontre que les technologies en général, et particulièrement les technologies numériques, s'organisent en systèmes complexes, presque indépendamment de la volonté des utilisateurs. Une première conséquence de ce constat est l'augmentation exponentielle du volume de données en circulation. Les systèmes techniques multiplient les échanges et s'agrègent, faisant naître le risque de perdre le contrôle sur les données qu'ils exploitent. De plus, ces systèmes techniques sont impliqués dans un ensemble de réseaux ayant des implications souvent insoupçonnées, à la fois sociales, économiques et environnementales. Sur les aspects humains, cela peut se traduire chez les archéologues par une perte de sens dans la pratique professionnelle lorsque les outils sont trop directifs, ou par la remise en question de leur identité professionnelle lors de l'introduction de nouvelles

technologies dans les méthodes de travail (Tufféry 2022). Du point de vue environnemental, plusieurs études d'envergure internationale ont récemment souligné le poids du numérique à l'échelle mondiale. Elles estiment que le cycle de vie des technologies numériques, de la production au recyclage, représentait déjà en 2020 plus de 4 % des énergies primaires consommées sur la planète, et 3,5 % des émissions de gaz à effet de serre. Plus préoccupant encore, en continuant sur les modèles de consommation actuels, ces valeurs risquent de doubler en à peine cinq ans, à l'horizon 2025. Il apparaît donc nécessaire de rationaliser l'usage de ces technologies, en considérant les outils et les données numériques dans leurs dimensions techniques, mais aussi sociales, économiques et environnementales. C'est seulement dans ce cadre que les chercheurs pourront inscrire durablement l'archéologie numérique dans une approche scientifique éthique et de qualité.

La science ouverte offre aux archéologues des moyens de maximiser l'impact positif des technologies numériques. Nos recommandations se focalisent sur des aspects concrets, en harmonie avec les réalités scientifiques et logistiques auxquelles les archéologues sur le terrain sont confrontés, tout en prenant en compte les implications des technologies numériques. Cependant, étant donné la nature en constante évolution de ces technologies, une partie substantielle de cette thèse risque rapidement de perdre de sa pertinence. Par conséquent, la leçon essentielle à retenir est qu'il est impératif de maintenir un contrôle sur les outils et les données numériques, tout en adoptant une perspective critique à leur égard. La démarche de la science ouverte permet d'exercer ce regard critique. Ainsi, nous souhaitons profiter de cette conclusion pour discuter de son implémentation dans le domaine de l'archéologie à diverses échelles, en nous fondant sur cinq années à défendre ces principes auprès de notre communauté.

Au niveau individuel, il est essentiel d'adopter une approche progressive et patiente dans l'adoption des principes de la science ouverte. En ce qui concerne les outils, nous préconisons de les intégrer un par un, en commençant par des alternatives proches des outils déjà maîtrisés. L'aspect esthétique ne doit pas être un facteur déterminant ni la recherche d'un seul outil tout-en-un. Il est au contraire important de se concentrer sur la définition précise des fonctionnalités recherchées, pour identifier les outils les plus adaptés à ces besoins. Dans notre cas, la transparence caractéristique des outils libres, permettant de comprendre les logiques de fonctionnement derrière chaque fonction, est rapidement devenu un critère de choix plus important que l'ergonomie dans le choix des logiciels pour mener nos recherches. De la même manière, lors de la gestion des métadonnées et de la rédaction de Plans de Ges-

tion de Données, il est recommandé de ne pas s'obliger à renseigner chaque fichier, mais plutôt de justifier le choix du niveau de détail donné à la documentation en fonction des ses compétences ou de la destination du corpus numérique.

Au niveau du groupe de travail ou de l'institution, l'objectif consiste à inciter les membres à partager des outils et des normes respectant les principes de la science ouverte tout au long du cycle de vie des données (collecte et création, traitement et analyse, archivage, partage, réutilisation). Les éléments clés à prendre en compte ne sont donc pas les technologies ou les données, mais leurs utilisateurs. Cette approche se construit à partir d'échanges, de formations et de suivis. Les échanges interviennent dans un premier temps, afin de comprendre les méthodes de travail des membres, leur place dans la chaîne opératoire, la définition qu'ils donnent de leurs objectifs dans cette chaîne opératoire, et enfin, comment est-ce qu'ils se positionnent par rapport à une évolution de leurs méthodes de travail ou du processus de création de connaissances archéologiques vers la science ouverte. Les réponses à ces questions permettent de définir les technologies numériques les plus adaptées, mais également de limiter fortement le risque de rejet "de principe", quel que soit le degré de pertinence scientifique des outils ou méthodes proposés. La formation vient ensuite, pour permettre aux usagers de s'appropriier les enjeux de la science ouverte, en intégrant les avantages qu'ils peuvent en tirer à leur niveau, sur le court, le moyen et le long terme. Par exemple, pour les archives de fouille numériques, à court terme, toute l'équipe peut avoir accès à des données claires et exploitables ; à moyen terme, les spécialistes et les gestionnaires ont accès à la même base d'informations ; à long terme, tous les rapports d'opérations et publications peuvent être accompagnés de leurs données primaires. Enfin, il est essentiel d'opérer un suivi auprès des membres une fois les technologies ou les méthodes déployées dans le groupe ou l'institution. En effet, les archéologues peuvent être convaincus du bien-fondé d'une démarche de science ouverte, mais pour autant abandonner rapidement les "bonnes pratiques". L'archéologie est une science qui nécessite déjà de maîtriser beaucoup d'outils et de méthodes, souvent empruntés à d'autres disciplines. Si les archéologues sont confrontés de manière trop récurrente à des problèmes de fonctionnement ou des incompréhensions qui restent sans réponse, ils retourneront systématiquement aux solutions qu'ils maîtrisent, libres ou non.

Au niveau national et international, la science ouverte en archéologie est déjà, paradoxalement, relativement bien organisée. Elle se structure autour d'établissements, comme l'*Achaological Data Service*, ou de groupements de spécialistes, tels que les consortiums Huma-Num. Ces deux formes d'organisation sont essentielles. D'abord, les consortiums per-

mettent de définir des règles communes de production, de gestion, de conservation et de diffusion des données archéologiques numériques. C'est pourquoi ils se doivent de regrouper tous les acteurs de l'archéologie : les chercheurs et universitaires, les représentants et les agents de l'État et des communautés territoriales, les opérateurs préventifs. Ils peuvent ainsi penser des règles qui prennent en compte les réalités de toutes ces communautés, afin de construire un langage et des méthodes d'échange stables et largement partagés au sein de notre communauté. Toute la difficulté réside dans le fait de coconstruire des règles suffisamment évolutives pour respecter les spécificités de tous les acteurs, tout en fondant les bases indispensables d'un réseau de partage et de pérennisation du savoir archéologique. Selon nous, une telle démarche ne peut être que collégiale, pour déboucher sur des consensus plutôt que sur des normes prenant en compte les intérêts d'une partie seulement de la communauté archéologique. Notre recherche doctorale a également démontré l'intérêt de développer des expérimentations en parallèle de ces discussions. La science ouverte et les technologies numériques ne sont pas qu'affaire de technique, et certaines limites n'apparaissent clairement qu'avec une mise en pratique en condition réelle. Toutes ces initiatives sont difficiles, longues et nécessairement coûteuses à mettre en place. Il est donc indispensable de continuer à soutenir et développer les consortiums pour, en reprenant le nom du projet européen éponyme, sauver l'archéologie européenne de l'âge de l'obscurité numérique¹⁷¹ (*SEADDA, Saving European Archaeology from the Digital Dark Age*).

Les établissements à l'image de l'*Archaeological Data Service* sont ensuite une manière d'appliquer et de pérenniser les règles éditées par les consortiums. En France, Huma-Num offre des fondations solides, mais il est essentiel de faire émerger des structures avec des équipes pleinement formées à la fois aux problématiques archéologiques et aux défis techniques de la science ouverte. Actuellement, l'existence des plateformes comme ArkeoGIS, Chronocarto ou le CND3D, est essentiellement assurée par des chercheurs investis, qui explorent des solutions technologiques pour exploiter le potentiel des données numériques, mais surtout pour garder le contrôle sur celles-ci. Or ces chercheurs n'ont pas toujours les compétences nécessaires pour maîtriser les solutions technologiques, car elles s'éloignent très largement de leur formation initiale. Pourtant, face au défi de la science ouverte et de l'augmentation continue de la masse des données, il n'est plus possible de continuer avec ce type d'approche expérimentale ou informelle. Les archéologues doivent maintenant s'associer durablement à des spécialistes des sciences de l'information et de la communication pour

171 <https://www.seadda.eu/> ; consulté le 22/10/2023

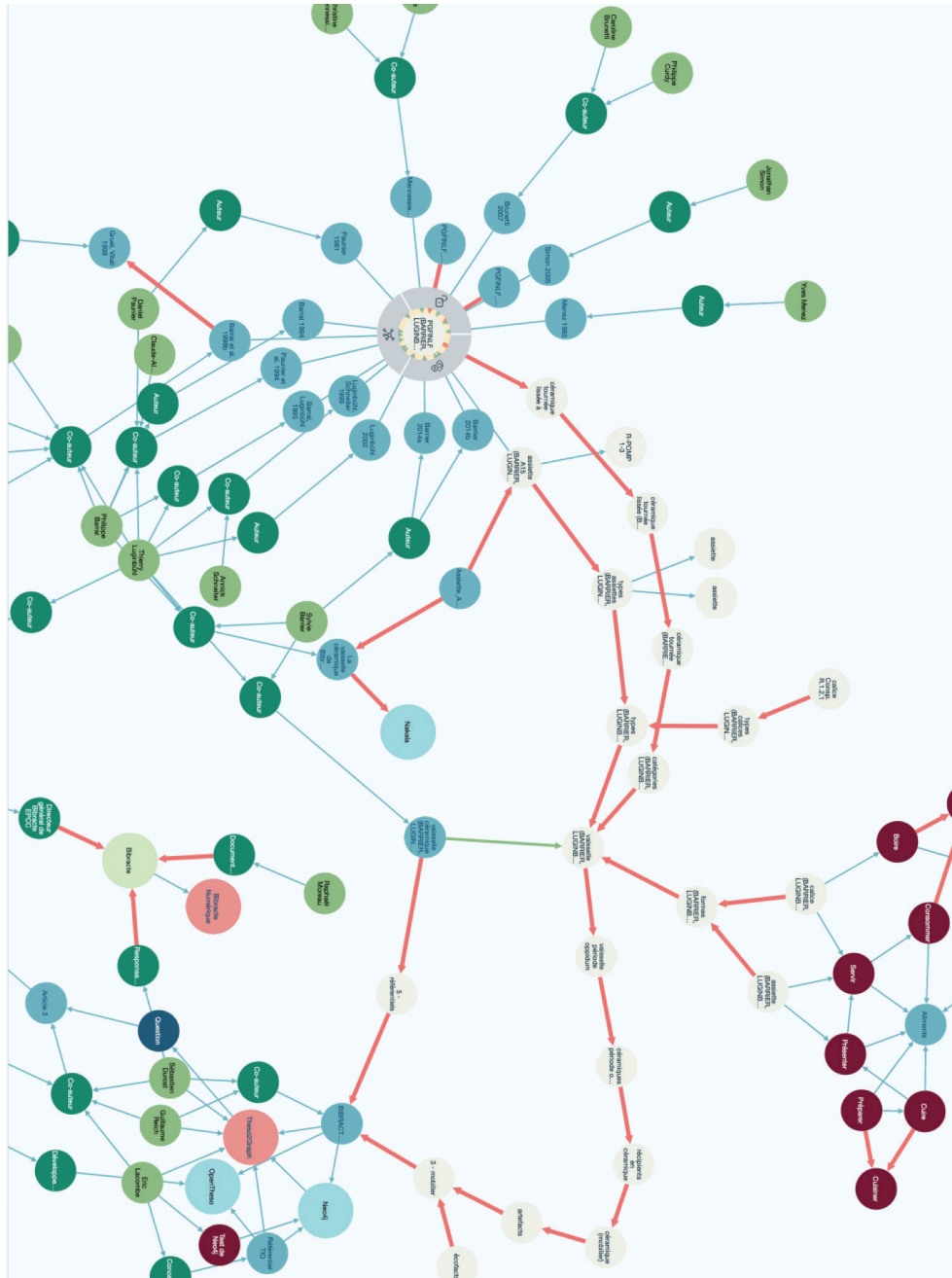
construire les solutions adaptées à l'exploitation, à la conservation et au partage des données archéologiques. C'est pourquoi nous appelons à la création d'un établissement qui regroupe les savoir-faire techniques, les connaissances scientifiques et l'infrastructure matérielle pour assurer la maîtrise et le partage libre et éthique des savoirs archéologiques.

Plus concrètement, la science ouverte et les technologies numériques en archéologie offrent de multiples perspectives de développement. Nous présentons ici ceux que nous souhaitons prioriser, en commençant par l'étape de la création et de la collecte de données numériques archéologiques. Notre expérience a une fois de plus démontré l'intérêt des SIG sur les chantiers archéologiques, autant comme outil de production que comme support d'analyse et de prise de décision (Rodier *et al.* 2011). Cependant, notre utilisation reste très limitée par rapport aux véritables capacités de l'outil. Par exemple, la mise en relation des données attributaires par des bases de données PostgreSQL/PostGIS permet une bien meilleure gestion des informations (fig. 105)(fig. 106). Actuellement, nous nous contentons, comme beaucoup d'archéologues, de fusionner les données utiles dans les tables attributaires des entités géographiques. Cette pratique a le mérite d'être intuitive et facile à mettre en œuvre, mais elle peut aussi rapidement mener à des erreurs ou des difficultés d'exploitation. L'import manuel (une "jointure") de chaque nouvelle donnée que nous souhaitons interroger dans la table attributaire s'avère fastidieux, et entraîne finalement des tables lourdes, parfois à la limite des capacités d'exploitation de QGIS. Ainsi, structurer les données textuelles dans des tables relationnelles apparaît bien plus logique. Le modèle logique relationnel est d'ailleurs le plus répandu dans la communauté archéologique pour la structuration de données. Le langage PostgreSQL/PostGIS reprend ce modèle, en intégrant nativement la possibilité de gérer des entités géographiques. Assimiler cette technologie est donc une évolution "naturelle" de nos méthodes, en ajoutant à notre panoplie l'outil adapté à une pratique que nous avons pour habitude de "bricoler". Un nouveau guide, à l'image de celui paru récemment sur la photogrammétrie en contexte archéologique, est déjà prévu sur le sujet (Verriez *et al.* 2023). Il traitera notamment de l'utilisation du SIG en chantier archéologique, en prenant en compte le système de gestion de base de données PostgreSQL/PostGIS.

En revanche, pour marquer une rupture plus nette avec les méthodes des archéologues, nous souhaitons également explorer les technologies de gestion des données basées sur la théorie des graphes, tels que Neo4j. Neo4j est un système de gestion de base de données libre orientée graphe (NoSQL), qui se spécialise dans le stockage et la récupération de données sous forme de graphes. Le modèle logique derrière ce type d'outil change du modèle relationnel. L'information textuelle n'est plus structurée en tables, mais en sommets (ou nœuds) et en arrêtes (ou liens) (fig. 107). Les graphes sont une structure de données puissante et non rigide pour modéliser et représenter des relations complexes entre différents types d'entités. Neo4j est couramment utilisé dans des domaines tels que la gestion des connaissances ou la gestion de réseaux sociaux.

Ainsi, Neo4j offre une perspective de développement intéressante pour l'archéologie en raison de la nature complexe et intriquée des données de la discipline. Les données archéologiques impliquent souvent une multitude d'entités (comme des artefacts, des sites, des couches stratigraphiques et des relations entre eux) et des connexions qui reflètent des contextes historiques et spatiaux. En utilisant un modèle de données de graphe, Neo4j permet de représenter ces entités, relations et structures de manière potentiellement plus naturelle et efficace que les bases de données relationnelles traditionnelles. L'archéologie implique également l'analyse de relations complexes entre des éléments, pour déterminer des concepts tels que les influences culturelles ou les échanges commerciaux. Neo4j est capable d'exploiter simplement les relations voisines, en effectuant des recherches de type "partir d'un nœud et parcourir le graphe". Il pourrait ainsi permettre aux archéologues de modéliser ces relations de manière explicite et de naviguer facilement dans les données pour faire apparaître des liens discrets.

Une première expérience a eu lieu à Bibracte avec cet outil, qui visait à documenter la mise en relation des raisonnements de plusieurs chercheurs conduisant à l'interprétation de vestiges archéologiques (Lacombe *et al.* 2022 ; Reich *et al.* 2022). L'objectif était de construire une méthode capable d'appuyer les démonstrations archéologiques sur un système plus structuré que des exposés en langage vernaculaire, afin de les codifier directement dans un système de gestion de la connaissance archéologique. Nous souhaitons à notre tour explorer les possibilités d'une base de données orientée graphe pour documenter les productions archéologiques au plus proche de la source, directement sur le chantier de fouilles. Le Neo4j présente en effet un réel potentiel pour fluidifier l'enregistrement, mais surtout la navigation dans les données, à travers des représentations graphiques qui présentent naturellement les



107. Exemple de base de données orientée graphe Neo4j
(données E. Lacombe, S. Durost, J.-P. Girard, G. Reich)

liens entre les données. Nous envisageons également le développement de fonctionnalités comme la prise en compte d'interprétations concurrentes, ou la vérification de la cohérence stratigraphique à la demande.

Sur la documentation des données, il est également important de concevoir des solutions pour simplifier la création des métadonnées et de description des méthodes de production. Une première étape pourrait être d'ouvrir aLTAG3D, l'outil de génération d'archive à long terme de projets 3D du consortium 3D SHS, à d'autres types de données. Il propose en effet une interface intuitive, là encore sous forme de graphes, pour documenter tout le processus de production d'un objet 3D. On peut alors envisager une interface similaire pour renseigner les données archéologiques. Toutefois, ce type de logiciel nécessite en amont la définition de standards d'échange et d'archivage, un travail qu'a dû effectuer le consortium 3D SHS en même temps que le développement de l'outil (Granier *et al.* 2019). C'est pourquoi l'orientation du projet SIAMOIS, qui souhaite mettre en place les solutions techniques pour formaliser les processus de production de la connaissance archéologique est particulièrement pertinente. Elle permettrait de limiter les standards aux données, sans homogénéiser les méthodes de production. Enfin, il apparaît maintenant opportun de s'intéresser aux possibilités offertes par l'intelligence artificielle (IA) pour gérer et documenter des données archéologiques. L'IA est souvent annoncée comme la nouvelle "rupture technologique", un point de basculement qui transformera notre manière de travailler. Il est donc primordial de s'intéresser dès maintenant à ce champ d'études, avec un regard critique et sans perdre de vue nos objectifs archéologiques, pour ne pas se le faire "imposer" de fait et de manière désorganisée dans quelques années. D'ailleurs, si le projet SIAMOIS parvient à formaliser les modèles logiques qui régissent les outils et des données, il devient alors possible et intéressant d'entraîner une IA capable de naviguer et d'extraire des informations dans des jeux de données complexes comme ceux des archéologues.

LISTE DES FIGURES

1. Les supports d'informations : analogique ou numérique.....	16
2. Page d'accueil du portail de l'Archaeology Data Service.....	21
3. Évolution du nombre de collections par années.....	23
4. Page d'accueil du service d'accès en ligne à l'index des fouilles archéologiques du Royaume-Uni (OASIS)	24
5. Page d'accueil du service de dépôt autonome ADS-easy	25
6. Exemples de prix pour le dépôt de fichiers grâce au service ADS-easy.....	26
7. Archsearch exploite plus de 1,3 million de ressources	30
8. Chaîne opératoire analytique pour le déploiement des SIG à l'Inrap (d'après Rodier et al. 2011)	37
9. Exemple de recherche du sujet "bibracte" sur le catalogue Dolia	39
10. Localisation et topographie générale du mont Beuvray (Saône-et-Loire, Nièvre, France)	44
11. Plan de l'oppidum de Bibracte d'après les relevés topographiques de H. d'Aboville (d'après Guillaumet 1996)	46
12. Exemple de planche d'objets photographiés issue de la publication de synthèse des travaux archéologiques de J.-G. Bulliot sur le mont Beuvray (d'après Bulliot et al. 1899)	48
13. Schéma du système Arkéoplan (d'après Buchsenchutz et Gruel 1994)	50
14. Levé LiDAR du mont Beuvray, réalisé en 2007 (Bibracte EPCC)	51
15. Exemple de synthèse cartographique des prospections magnétiques menées par l'université de Brno de 2011 à 2019 sur le mont Beuvray (d'après Sanchez et al. 2019).....	52
16. Comparaison entre l'orthophotographie de fin de chantier et les relevés numérisés sur le chantier PC14/PC15 à Bibracte en 2013 (cliché B.-N. Chagny, Kap-archéo)	54
17. La nécessité d'abriter les chantiers archéologiques sur le mont Beuvray (cliché M. Thivet)	55
18. Résultats non concluant des premiers tests pour un protocole d'acquisition photogrammétrique sous abri	55
19. Orthophotographie aéroportée de la terrasse PC15 (cliché M. Thivet, Q. Verriez)	56
20. Page d'accueil de bdB Consultation	58

21. Structure de bdB en 2001 (d'après Chaillou 2003)	60
22. Structure de bdB en 2021.....	60
23. Structuration du rapport annuel du programme de recherche du mont Beuvray depuis 2012 (d'après Desachy et al. 2012).....	64
24. Les projets associés à Bibracte Numérique	68
25. Plan de la domus PC2	69
26. Plan de localisation de la domus PC2 (DAO A. Meunier, Q. Verriez)	70
27. Plan d'ensemble du Parc-aux-Chevaux (d'après Guillaumet 1996)	71
28. Début de la fouille de la domus PC2 en 2017, face à l'angle sud-ouest de la domus PC1	72
29. La terrasse de la domus PC2, vue depuis le sud-ouest.....	72
30. Schéma de la domus PC2 réalisé par J.-G. Bulliot dans son carnet de fouille 1882	73
31. Plan de la domus PC2 après les fouilles de 1882 (d'après Bulliot 1899)	74
32. Plan de la domus PC2 actualisé	75
33. Organisation de la connaissance selon le modèle DIKW (d'après Rowley 2007).....	83
34. Différence entre les images matricielles et les images vectorielles	87
35. Les types de géométrie vectorielle.....	88
36. Le cycle de vie des données.....	91
37. Les éléments Dublin Core (d'après BnF 2008)	93
38. Les licences Creative Commons	104
39. Nombre de ressources déposées sur HAL contenant le mot clé "archéologie"	106
40. L'écosystème numérique MASA pour les données archéologiques (extrait de Marlet et al. 2022)	109
41. L'écosystème numérique de l'IR* Huma-Num (image d'après huma-num.fr).....	111
42. Le portail institutionnel de Bibracte sur HAL	112
43. Le drone OnyxStar FOX-C8HD de la MSHE Ledoux (cliché A. Maillier).....	126
44. La télécommande et le système de retour vidéo de l'OnyxStar FOX-C8HD (cliché A. Maillier)	126
45. Évolution chronologique du bâtiment monumental de la terrasse PC15. Plans dessinés sur l'orthomosaïque de fin de chantier de 2015 (extrait de Hantrais 2017)	128
46. Le drone Dji Phantom 4 Pro de la MSHE Ledoux	129
47. Conditionnement des batteries LiPo dans sacs ignifugés	132

48. Conditions pour le survol de site archéologiques par drone	133
49. Types d'appareils photo numérique.....	143
50. Exemple d'objectif interchangeable polyvalent : le Canon EF 24-70mm f/4L IS USM	144
51. La différence entre la définition, la résolution et la taille d'impression d'une photographie	149
52. Une photo archéologique à deux définitions différentes. Verdun-sur-le-Doubs, le Petit Chauvort, Saône-et-Loire, 2021 (cliché E. Hamon)	155
53. Les déformations optiques courantes. Bibracte, domus PC2, Nièvre, 2017	156
54. Photos redressées de la pièce XX de domus PC2, à Bibracte. Bibracte, la domus PC2, Nièvre, 2017	157
55. Orthophotographie de la pièce XX de domus PC2, à Bibracte. Bibracte, la domus PC2, Nièvre, 2017	159
56. Processus théorique de production d'une orthomosaïque.	161
57. Points homologues entre deux images voisines. Bibracte, plateforme PC14, Nièvre, 2021	162
58. Projection des points homologues dans un espace 3D. Bibracte, plateforme PC14, Nièvre, 2021.....	163
59. Schéma de représentation des principales surfaces de capteur photographique disponible sur le marché	167
60. Schéma de relation entre l'angle de champ d'une photographie et la longueur focale de l'objectif utilisé (au format 24 x 36 mm). (d'après Pamart et al. 2019).....	168
61. Schéma de la relation entre profondeur de champ et nombre d'ouverture	173
62. Principe d'un cheminement en “S”, recommandé pour les acquisitions aéroportées (DAO. M. Thivet)	176
63. Calcul de la résolution au sol.....	178
64. Placement optimal des points de contrôle au sol. Couan, Saint-Aubin-des-Chaumes, Nièvre, 2021	180
65. L'interface du logiciel MicMac, dans le terminal de Ubuntu,et sa première commande, pour lancer l'outil Tapioca.....	182
66. Les fenêtres d'analyse pour l'étude spatiale des sites archéologiques (extrait de Clarke 1977)	191
67. SIG de PC15 réalisé par J. Hantrais.....	193
68. Principe de la méthode du Single-Layer Plan (extrait de Harris 1989)	195
69. La stratigraphie d'un fossé de l'âge du fer recomposée en plan grâce aux SPL (New Road	

Site, Winchester 1975, extrait de Harris 1989).....	197
70. Principe de continuité spatiale : même s'il est quasi certain qu'il s'agit des mêmes couches de part et d'autre de la berme, ces dernières sont enregistrées dans des US différentes....	198
71. Processus de production des fichiers de forme selon le protocole de relevé numérique mis en place sur le chantier de la domus PC2	200
72. Les 566 US relevées sur le chantier de la domus PC2 durant les campagnes 2017 à 2020	201
73. Les tablettes Xplore Xslate B10 (à gauche) et Dell Latitude 7212 (à droite) de la MSHE Ledoux	201
74. La répartition des parts de marché entre les principaux systèmes d'exploitation (extrait de statcounter : https://gs.statcounter.com/os-market-share/desktop/worldwide/2022)	203
75. Une fouilleuse qui utilise une tablette comme support de relevé en 2019	205
76. Test de relevé pierre à pierre sur le mur effondré 655 en 2020.....	206
77. Un problème rencontré avec le protocole de relevé numérique pour la représentation de structures hautes ou profondes.....	210
78. Formulaire principal de la version EDArc pour Bibracte : l'enregistrement des UF ...	213
79. Architecture générale d'EDArc (extrait de Tufféry 2022)	214
80. ODS fonctionne sur de multiples supports (image 9heuresprecises.com).....	216
81. Exemple de gestion de l'avancement de l'enregistrement dans l'onglet "Reste à faire" de la version d'EDArc pour Bibracte.....	218
82. Exemple de gestion de l'avancement de l'enregistrement en mode tablette et en mode ordinateur sur ODS.....	219
83. Exemple de description et d'alignement de champs descripteurs d'EDArc, ODS et bdB (extrait de Reich et al. 2022)	220
84. Mise en relation du vocabulaire structuré en thésaurus de bdB avec les modèles logiques de EDArc et ODS (extrait de Reich et al. 2022)	221
85. Extrait du référentiel analytique de la campagne de fouilles 2020 de la domus PC2 ..	227
86. Présentation du modèle de PGD MASA.....	228
87. Les aides à la rédaction du modèle de PGD MASA	229
88. Le protocole de relevé numérique établi sur PC2 en quelques chiffres	236
89. Synthèse sur l'impact environnemental du numérique (d'après Ferreboeuf 2018 ; Bordage 2019a ; Bordage et al. 2021 ; Ferreboeuf et al. 2021 ; Lees Perasso et al. 2023).....	244
90. Convention de nommage : équilibre entre formalisme, souplesse et intelligibilité	256
91. Exemple de nommage de fichiers	260

92. Exemple de fichier XML décrivant l'US 0001, de type positive.....	262
93. Définitions et normalisation des éléments Dublin Core	266
94. Les métadonnées du rapport de synthèse 2020 du programme de recherche du mont Beuvray	269
95. Les métadonnées du rapport de synthèse 2020 du programme de recherche du mont Beuvray au format XML.....	270
96. Types de dépôts possibles sur HAL	278
97. Exemple de données signalées pour l'âge du Fer sur ArkeoGIS.....	279
98. Exemple de données signalées sur ArkeoGIS : la campagne de fouille 2017 de la domus PC2.....	280
99. Page d'accueil de ArkeOpen.....	281
100. Exemple du projet web-SIG du programme de recherche du mont Beuvray sur Chronocarto (données M. Dabas, Ch. Sanchez)	283
101. Le moteur de recherche IOTA, accessible depuis Chronocarto	284
102. Exemple de ressource hébergée sur le CND3D : modèle 3D du Carnyx de Tintignac (données Mora et al. 2008)	286
103. Logiciel aLTAG3D, pour la documentation du processus de production et d'archivage de modèles 3D.....	288
104. Modèles de gestion des dépôts de données sur les plateformes nationales.....	290
105. Exemple de SIG... ..	302
106. ... et sa base de données PostgreSQL/PostGIS associée (données A. Tomasinelli, Eveha International).....	303
107. Exemple de base de données orientée graphe Neo4j (données E. Lacombe, S. Durost, J.-P. Girard, G. Reich)	305

LISTE DES ABRÉVIATIONS

ADS : Archaeology Data Service

Afan : Association pour les fouilles archéologiques nationales

AFUL : Utilisateurs de Logiciels Libres

AGT : Archaeological Geophysics Toolbox

AHDS : Art and Humanities Data Service

ANR : Agence Nationale de la Recherche

AOrOc : Archéologie & Philologie d'Orient et d'Occident

APN : Appareils photo numériques

ARIADNE. ARIADNE+ : Archaeological Research Infrastructure for Archaeological Data Networking in Europe

ARK : l'Archival Resource Key

ArTeHiS : Archéologie, Terre, Histoire et Sociétés

BATX : Baidu, Alibaba, Tencent et Xiaomi

bdB : base de donnée Bibracte

CAVIAR : : Catalogue de visualisation de l'information archéologique

CC : Creative Commons

CCE : Centre de conservation et d'étude

CERN : organisation européenne pour la recherche nucléaire

CITERES-LAT : Équipe Laboratoire Archéologie et Territoires

Cifre : Convention industrielle de formation par la recherche

COST SEADDA : Saving European Archaeology from the Digital Dark Age

CNRS : Centre national de la recherche scientifique

DAO : dessin assisté par ordinateur

DEA : Diplôme d'études approfondies

DGAC : Direction générale de l'aviation civile

DOI : Digital Object Identifier

DPI : Dots Per Inch

DSI : Direction des Systèmes d'Information (Inrap)

DST : Direction Scientifique et Technique (Inrap)

EFA : École Française d'Athènes

ENSTP : École Nationale Supérieure des Travaux Publics

EPCC : Établissement Public de Collaboration Culturel

ERC : European Research Council

FAIR : Findable, Accessible, Interoperable, Reusable, ou Facile à trouver, Accessible, Interopérable et Réutilisable

FOSTER : Fostering the practical implementation of Open Science in Horizon 2020 and beyond

GAFAM : Google, Amazon, Facebook, Apple, Microsoft

GCP : Ground Control Point

GPS : Global Positioning System

GSD : Ground Sample Distance

HEIRNET : Heritage Information Resource

HTML5 : HyperText Markup Language 5

IA : Intelligence artificielle

Inrap : Institut national de recherches archéologiques préventives

LiDAR : Light Detection and Ranging

LO : LibreOffice

LO/OL : Licence Ouverte/Open Licence

MASA : Mémoire des Archéologues et des Sciences de l'Archéologie

MESRI : Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation

MNE : Modèle Numérique d'Élévation

MNT : Modèle Numérique de Terrain

MOOC : Massive Open Online Courses

MSHE : Maison des Sciences de l'Homme et de l'Environnement

NERC : National Environment Research Council

OASIS : Online Access to the Index of Archaeological Investigations

OdbL : Open Database Licence

OECD : Organisation de coopération et de développement économique

PAO : Publication assistée par ordinateur

PATOIS : Publications and Archives in Teaching : Online Information Sources

PC : Parc-aux-Chevaux

PDF : Portable Document Format

PGD : Plans de Gestion de Données

PNG : Portable Network Graphics

PnSO : Plan national pour la Science Ouverte

PPP : Pixels Par Pouce

REFIT : Resituating Europe's First Towns

RGI : Référentiel Général d'Interopérabilité

SCR : Système de Coordonnées de Référence

SGBD : Systèmes de Gestion de Base de Données

SHS : Sciences Humaines et Sociales

SIAMOIS : Système d'Information Archéologique Mutualisé et Ouvert reposant sur l'Intelligence Sémantique

SIG : Système d'Information Géographique

SLP : Single Context Plan

SRA : Services Régionaux de l'Archéologie

STAR : Semantic Technologies for Archaeological Ressources

STELLAR : Semantic Technologies Enhancing Links and Linked Data for Archaeological Resources

SVG : Scalable Vector Graphics

TAG : Transatlantic Archaeology Gateway

tDAR : the Digital Archaeological Record

THATcamp : The Humanities and Technology Camp

UAR : Unité d'Appui à la Recherche

UF : Unité de fouille

UML : Unified Modeling Language

UMR : unité Mixte de Recherche

UMS : Unité Mixte de Services

URI : Uniform Resource Identifier

US : Unité stratigraphique

XML : eXtensible Markup Language

XSL : eXtensible Stylesheet Language

ANNEXES

Annexe 1

Traduction de la définition du savoir libre selon l'*Open Knowledge Foundation*¹

Définition du savoir libre 2.1

La définition du savoir libre précise la signification du terme "ouvert" concernant la connaissance, en promouvant un patrimoine commun bien défini auquel tout le monde peut participer et où l'interopérabilité est maximisée.

Résumé : Les connaissances sont ouvertes si tout le monde est libre d'y accéder, de les utiliser, de les modifier et de les partager - sous réserve, tout au plus, de mesures qui préservent la provenance et l'ouverture.

Cette signification essentielle correspond à celle du terme "ouvert" pour les logiciels, comme dans la définition de l'Open Source², et est synonyme de "libre" ou "gratuit" comme dans la définition du logiciel libre³ et la définition des œuvres culturelles libres⁴ (Collaboratif 2019).

Le terme "œuvre" sera utilisé pour désigner l'objet ou l'élément de connaissance transféré.

Le terme "licence" désigne les conditions juridiques dans lesquelles l'œuvre est fournie.

Le terme "domaine public" désigne l'absence de droits d'auteur et de restrictions similaires, que ce soit par défaut ou par renonciation à toutes ces conditions.

Les mots clés "doit", "ne doit pas", "devrait" et "peut" dans ce document doivent être interprétés comme décrit dans RFC2119⁵.

1 <http://opendefinition.org/od/2.1/en/>, consulté le 22/10/2023

2 <https://opensource.org/osd> ; consulté le 22/10/2023

3 https://en.wikipedia.org/wiki/The_Free_Software_Definition ; consulté le 22/10/2023

4 https://en.wikipedia.org/wiki/Definition_of_Free_Cultural_Works ; consulté le 22/10/2023

5 <https://datatracker.ietf.org/doc/html/rfc2119> ; consulté le 22/10/2023

1. Œuvres libres

Une œuvre libre doit satisfaire aux exigences suivantes dans sa distribution :

1.1 Licence ou statut ouvert

L'œuvre doit être dans le domaine public ou être fournie sous une licence ouverte (telle que définie dans la section 2). Toute condition supplémentaire accompagnant l'œuvre (telle que des conditions d'utilisation ou des brevets détenus par le donneur de licence) ne doit pas contredire le statut de domaine public de l'œuvre ou les conditions de la licence.

1.2 Accès

L'œuvre doit être fournie dans son intégralité et à un prix ne dépassant pas un coût raisonnable de reproduction et devrait pouvoir être téléchargée gratuitement sur internet, sans frais additionnels. Toute information supplémentaire nécessaire à la conformité de la licence (telle que les noms des contributeurs requis pour la conformité aux exigences d'attribution) doit également accompagner l'œuvre.

1.3 Lisibilité par des machines

L'œuvre doit être fournie sous une forme facilement traitable par un ordinateur et où les éléments individuels de l'œuvre peuvent être facilement accessibles et modifiés.

1.4 Format ouvert

L'œuvre doit être fournie dans un format ouvert. Un format ouvert est un format qui n'impose aucune restriction, monétaire ou autres, à son utilisation et qui peut être entièrement traité avec au moins un outil libre.

2. Licences ouvertes

Une licence devrait être compatible avec d'autres licences ouvertes.

Une licence est ouverte si ses termes satisfont aux conditions suivantes :

2.1 Autorisations requises

La licence doit permettre (ou autoriser) de manière irrévocable les éléments suivants :

2.1.1 Utilisation

La licence doit permettre la libre utilisation de l'œuvre sous licence.

2.1.2 Redistribution

La licence doit permettre la redistribution de l'œuvre sous licence, y compris sa vente, qu'elle soit seule ou qu'elle fasse partie d'une collection constituée d'œuvres provenant de sources différentes.

2.1.3 Modification

La licence doit permettre la création de dérivés de l'œuvre sous licence et autoriser la distribution de ces dérivés aux mêmes conditions que l'œuvre originale sous licence.

2.1.4 Séparation

La licence doit permettre à toute partie de l'œuvre d'être librement utilisée, distribuée ou modifiée séparément de toute autre partie de l'œuvre ou de toute collection d'œuvres dans laquelle elle a été initialement distribuée. Toutes les parties qui reçoivent une reproduction d'une partie quelconque d'une œuvre dans le cadre des conditions de la licence originale devraient avoir les mêmes droits que ceux qui sont accordés en liaison avec l'œuvre originale.

2.1.5 Compilation

La licence doit permettre à l'œuvre sous licence d'être distribuée avec d'autres œuvres distinctes sans imposer de restrictions sur ces autres œuvres.

2.1.6 Non-discrimination

La licence ne doit pas discriminer une personne ou un groupe.

2.1.7 Propagation

Les droits attachés à l'œuvre doivent s'appliquer à tous ceux à qui elle est redistribuée sans qu'il soit nécessaire d'accepter des conditions juridiques supplémentaires.

2.1.8 Application à toute fin

La licence doit permettre l'utilisation, la redistribution, la modification et la compilation à toutes fins. La licence ne doit pas empêcher quiconque d'utiliser l'œuvre dans un domaine d'activité spécifique.

2.1.9 Gratuité

La licence ne doit pas imposer d'arrangement tarifaire, de redevance ou autre compensation ou rémunération monétaire dans le cadre de ses conditions.

2.2 Conditions autorisées

La licence ne doit pas limiter, rendre incertaines ou diminuer de quelque manière que ce soit les autorisations requises à la section 2.1, sauf dans les conditions autorisées suivantes :

2.2.1 Attribution

La licence peut exiger que les reproductions de l'œuvre incluent l'attribution des contributeurs, des détenteurs de droits, des sponsors et des créateurs, tant que de telles prescriptions ne sont pas onéreuses.

2.2.2 Intégrité

La licence peut exiger que les versions modifiées d'une œuvre sous licence portent un nom ou un numéro de version différent de celui de l'œuvre originale ou indiquent d'une autre manière les changements apportés.

2.2.3 Partage selon les mêmes conditions

La licence peut exiger que les reproductions de l'œuvre restent sous la même licence ou une licence similaire.

2.2.4 Mention

La licence peut exiger la conservation des mentions de droits d'auteur et l'identification de la licence.

2.2.5 Source

La licence peut exiger que toute personne distribuant l'œuvre fournisse aux destinataires un accès à la forme préférée pour effectuer des modifications.

2.2.6 Interdiction des restrictions techniques

La licence peut exiger que les reproductions de l'œuvre soient exemptes de toute mesure technique qui restreindrait l'exercice de droits autrement autorisés.

2.2.7 Non-agression

La licence peut exiger des modificateurs qu'ils accordent au public des autorisations supplémentaires (par exemple, des licences de brevet) nécessaires à l'exercice des droits autorisés par la licence. La licence peut également conditionner les autorisations à la non-agression des partis concernés dans le cadre de l'exercice de tout droit autorisé (là encore, par exemple, les litiges relatifs aux brevets)".

Annexe 2

Traduction de la définition des données ouvertes⁶

1. Complète

Toutes les données publiques sont mises à disposition. Les données publiques sont des données qui ne sont pas soumises à des restrictions valables en matière de confidentialité, de sécurité ou de privilèges.

2. Primaire

Les données sont telles qu'elles ont été collectées à la source, avec le plus haut niveau de granularité possible, et non sous forme agrégée ou modifiée.

3. En temps utile

Les données sont mises à disposition aussi rapidement que nécessaire pour préserver leur valeur.

4. Accessible

Les données sont disponibles pour le plus grand nombre d'utilisateurs et à des fins très diverses.

5. Traitable par machine

Les données sont raisonnablement structurées pour permettre un traitement automatisé.

6. Non discriminatoire

Les données sont accessibles à tous, sans obligation d'enregistrement.

7. Non-propriétaire

Les données sont disponibles dans un format sur lequel aucune entité n'a de contrôle exclusif.

⁶ https://public.resource.org/8_principles.html/ ; consulté le 23/10/2023

8. Libre de droits

Les données ne sont soumises à aucune réglementation en matière de droits d'auteur, de brevets, de marques ou de secrets commerciaux. Des restrictions raisonnables en matière de confidentialité, de sécurité et de privilèges peuvent être autorisées".

L'un des participants, J. Tauberer, a réuni les autres concepts existants sur le web autour de la donnée ouverte pour ajouter 7 principes⁷:

9. En ligne et gratuite

L'information n'est pas véritablement publique si elle n'est pas disponible sur internet gratuitement, ou du moins pas plus que le coût marginal de reproduction. Elle doit également pouvoir être retrouvée.

10. Permanente

Les données doivent être disponibles en ligne indéfiniment sur un support stable et dans un format de données stable aussi longtemps que possible.

11. Fiable

Le contenu publié doit être signé numériquement ou comporter une attestation de la date de publication/création, de l'authenticité et de l'intégrité.

12. Une présomption d'ouverture

La présomption d'ouverture repose sur des lois telles que la loi sur la liberté d'information, des procédures telles que la gestion des documents et des outils tels que les catalogues de données.

13. Documentée

La documentation sur le format et la signification des données contribue grandement à rendre ces dernières utiles.

7 <https://opengovdata.org/> ; consulté le 22/10/2023

14. Ouverture sécurisée

Les organismes publics qui publient des données en ligne doivent toujours chercher à les publier en utilisant des formats de données qui n'incluent pas de contenu exécutable. Le contenu exécutable des documents présente un risque pour la sécurité des utilisateurs des données, car il peut s'agir de logiciels malveillants (virus, vers, etc.).

15. Conçu avec la participation du public

Le public est le mieux placé pour déterminer quelles technologies de l'information seront les mieux adaptées aux applications qu'il entend créer pour lui-même. La contribution du public est donc essentielle pour diffuser l'information de manière à ce qu'elle ait de la valeur".

Annexe 3

Traduction de la définition du logiciel libre (*Open Source*) selon l'*Open Source Initiative*⁸

1. Redistribution gratuite

La licence n'empêche pas une partie de vendre ou de donner le logiciel en tant que composant d'une distribution logicielle globale contenant des programmes provenant de plusieurs sources différentes. La licence n'exige pas de redevance ou d'autres frais pour une telle vente.

2. Code source

Le programme doit inclure le code source, et doit permettre la distribution sous forme de code source aussi bien que sous forme compilée. Lorsqu'une certaine forme de produit n'est pas distribuée avec le code source, il doit exister un moyen bien publié d'obtenir le code source pour un coût de reproduction raisonnable, de préférence par téléchargement gratuit via internet. Le code source doit être la forme préférée sous laquelle un programmeur modifierait le programme. Le code source délibérément obscurci n'est pas autorisé. Les formes intermédiaires telles que la sortie d'un préprocesseur ou d'un traducteur ne sont pas autorisées.

3. Travaux dérivés

La licence doit autoriser les modifications et les travaux dérivés, et doit permettre qu'ils soient distribués selon les mêmes termes que la licence du logiciel original.

4. Intégrité du code source de l'auteur

La licence peut restreindre la distribution du code source sous une forme modifiée uniquement si elle autorise la distribution de "fichiers de correction" avec le code source dans le but de modifier le programme au moment de la compilation. La licence doit autoriser explicitement la distribution de logiciels construits à partir d'un code source modifié. La licence peut exiger que les travaux dérivés portent un nom ou un numéro de version différent de celui du logiciel original.

8 <https://opensource.org/osd> ; consulté le 23/10/2023

5. Pas de discrimination à l'encontre de personnes ou de groupes

La licence ne doit pas être discriminatoire envers une personne ou un groupe de personnes.

6. Non-discrimination à l'égard des domaines d'activité

La licence ne doit pas empêcher quiconque d'utiliser le programme dans un domaine d'activité spécifique. Par exemple, elle ne doit pas empêcher l'utilisation du programme dans une entreprise ou pour la recherche génétique.

7. Distribution de la licence

Les droits attachés au programme doivent s'appliquer à tous ceux à qui le programme est redistribué sans qu'il soit nécessaire que ces parties se soumettent à une licence supplémentaire.

8. La licence ne doit pas être spécifique à un produit

Les droits attachés au programme ne doivent pas dépendre du fait que le programme fasse partie d'une distribution logicielle particulière. Si le programme est extrait de cette distribution et utilisé ou distribué selon les termes de la licence du programme, toutes les parties auxquelles le programme est redistribué doivent avoir les mêmes droits que ceux qui sont accordés dans le cadre de la distribution originale du logiciel.

9. La licence ne doit pas restreindre d'autres logiciels

La licence ne doit pas imposer de restrictions aux autres logiciels distribués en même temps que le logiciel sous licence. Par exemple, la licence ne doit pas insister sur le fait que tous les autres programmes distribués sur le même support doivent être des logiciels libres.

10. La licence doit être neutre sur le plan technologique.

Aucune disposition de la licence ne doit être fondée sur une technologie ou un style d'interface particulier.

Annexe 4

Protocole de relevé numérique avec le logiciel QGIS⁹

Depuis 2017, un protocole de relevé numérique est développé sur le chantier de la *domus* PC2 à Bibracte (mont Beuvray, Nièvre, France). L'étape du relevé est maintenant pleinement opérationnelle et démontre plusieurs avantages significatifs :

- **Rapidité d'apprentissage** : malgré des séjours courts sur le chantier-école (15 jours), tous les fouilleurs sont formés et exploitent la méthode de relevé en routine.
- **Intégration dans la chaîne opératoire archéologique** : le relevé est réalisé en temps réel, sans ralentir la fouille.
- **Qualité de l'information** : en associant les observations directes de l'archéologue et l'orthophotographie comme base du relevé, les dessins sont d'une grande précision.
- **Cohérence** : avec les relevés nativement géoréférencés, le plan général de la fouille est mis à jour en temps réel.
- **Contrôle qualité** : toute l'information est centralisée dès l'étape de la fouille, permettant de contrôler chaque Unité de Fouille et produire un véritable SIG intra-site en fin de campagne.

L'objectif du manuel qui suit est de décrire le plus clairement possible les étapes de dessin sur le logiciel SIG QGIS. Ce manuel (version 1) a été écrit avec la version 3.16.0 *Hannover* de QGIS.

9 <https://www.qgis.org/fr/site/> ; consulté le 22/10/2023

Préparer le stockage des données

QGIS n'intègre pas les données interrogées dans le fichier du projet. Ces dernières restent indépendantes du projet. Il n'enregistre que les liens qui pointent vers les dossiers où sont stockés vos fichiers sur l'ordinateur.

Cela implique plusieurs choses :

- Pour déplacer un projet QGIS sur votre ordinateur, ou d'une machine à une autre, il ne suffit pas de déplacer juste le fichier projet (monprojet.qgz), mais tous les fichiers interrogés par ce projet.
- Si vous déplacez un fichier interrogé dans un projet QGIS, une fenêtre apparaît à l'ouverture de ce dernier pour signaler que le logiciel ne retrouve pas le fichier en question au chemin indiqué (fig. 1). Il est toujours possible de retrouver un fichier ou un dossier en cliquant sur **Parcourir**, mais la démarche peut être laborieuse en fonction du nombre de fichiers manquants.

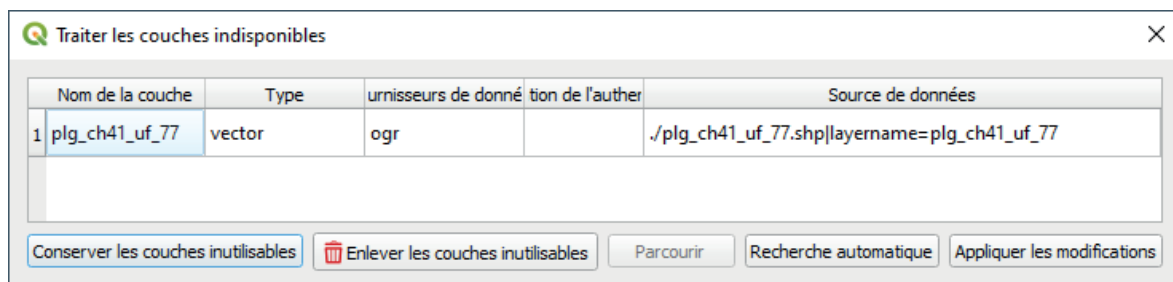


Figure 1

- Un fichier interrogé peut être modifié dans un autre projet QGIS ou un autre logiciel. Tant qu'il garde le même nom, il sera automatiquement mis à jour dans le projet en cours.
- L'utilisation courante d'un SIG produit beaucoup de fichiers. L'utilisateur peut rapidement s'y perdre s'il ne prend pas la peine de les organiser.

Voici deux astuces pour faciliter la gestion de votre projet :

- Créez un dossier qui contiendra toutes vos données, ainsi que le fichier du projet QGIS (monprojet.qgz).

- Définissez une arborescence de sous-dossiers pour gérer vos différents types de données. Sans être obligatoire pour le bon fonctionnement du logiciel, cela facilite grandement la gestion des données par les utilisateurs. À titre d'exemple, voici celle utilisée sur le chantier de la *domus* PC2 à Bibracte (fig. 2).



Figure 2

Installer QGIS

- Suivez le lien suivant : <https://www.qgis.org/fr/site/forusers/download.html>
- Choisissez votre système d'exploitation et votre version et suivez les étapes d'installation.
 - À savoir : deux versions sont disponibles (ce protocole fonctionne avec les deux) : la dernière version ou *unstable* et la version long terme ou *long term release* (LTR). La priorité de la dernière version est de proposer de nouvelles fonctionnalités, ce qui implique parfois quelques bugs. La LTR est une version vérifiée et mise à jour pour éliminer tous les bugs, mais qui ne propose pas les dernières fonctionnalités.

Télécharger le dossier « Bibracte GIS Tools »

Afin de faciliter l'utilisation de QGIS sur le terrain, nous avons préalablement paramétré plusieurs fichiers dont vous aurez besoin pour la suite du protocole.

- Ils sont rassemblés dans un dossier compressé (.zip) disponible en suivant le lien : <https://nakala.fr/10.34847/nkl.8d81q3es>
- Décompressez le dossier.

Modifier l'interface de QGIS

- Toutes les fonctionnalités de QGIS ne sont pas utiles dans notre cas. Une interface simplifiée a été mise en place pour faciliter le travail sur le terrain.
- Le dossier de modification « *archeo_draw* » se trouve dans le dossier « Bibracte_GIS_Tools » (*cf. supra* « Télécharger le dossier « Bibracte_GIS_Tools » »)
- Ouvrez QGIS. Dans l'onglet **Préférences**, sélectionnez **Profils Utilisateurs** puis :
- Sous Windows 10:
 - Ouvrez l'Explorateur de fichier, sélectionnez l'onglet **Affichage** , puis **cochez la case** « Éléments masqués ».
 - Ajoutez le dossier télécharger dans le répertoire suivant : C:\Users\qverr\AppData\Roaming\QGIS\QGIS3\profiles

- Ouvrez **QGIS Desktop 3.XX.XX with Grass 7.XX.XX**. Dans l'onglet **Préférence**, sélectionnez **Profil Utilisateur** puis « archeo_draw »
- Une nouvelle fenêtre QGIS s'ouvre avec l'interface simplifiée (fig. 3).

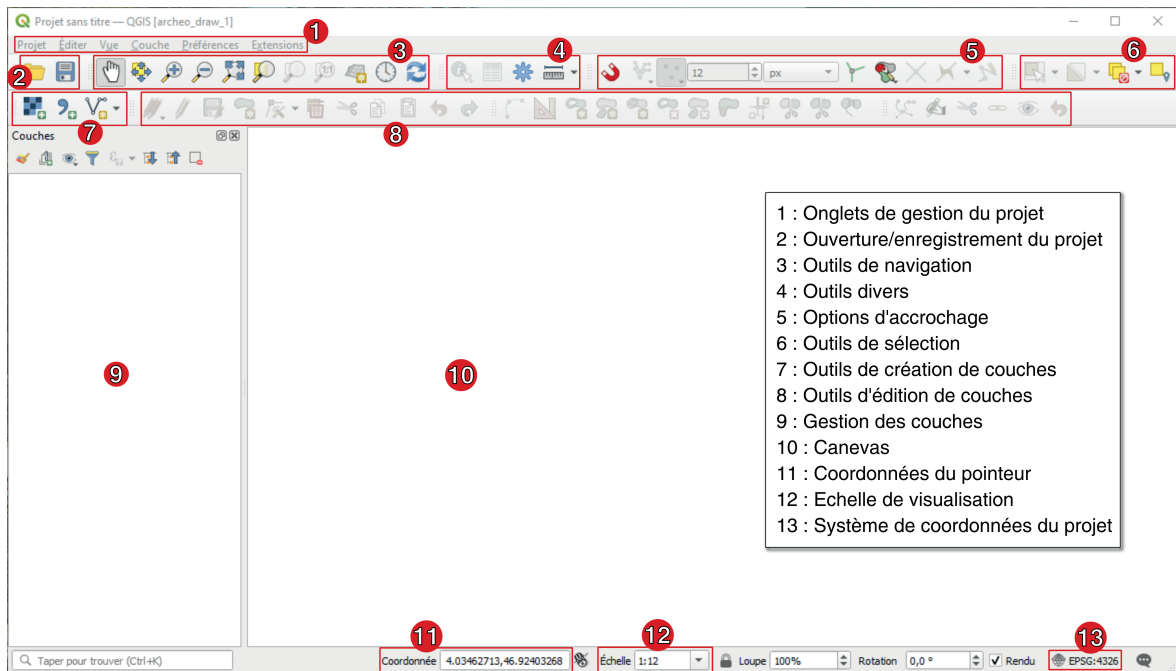


Figure 3

- Sous Mac :
 - Ouvrez le Finder et ajoutez le dossier télécharger dans le répertoire suivant : ~/Library/Application/Support/QGIS/QGIS3/profiles
 - Ouvrez QGIS. Dans l'onglet **Préférence**, sélectionnez **Profil Utilisateur** puis « archeo_draw »
 - Une nouvelle fenêtre QGIS s'ouvre avec l'interface simplifiée.
- Sous Linux Ubuntu :
 - Ouvrez l'application, sélectionnez l'onglet **Paramètres**, puis **cochez la case** « Afficher les éléments cachés ».
 - Ajoutez le dossier télécharger dans le répertoire suivant : Dossier personnel/.local/share/QGIS/QGIS3/profiles
 - Ouvrez QGIS. Dans l'onglet **Préférence**, sélectionnez **Profil Utilisateur** puis « archeo_draw »
 - Une nouvelle fenêtre QGIS s'ouvre avec l'interface simplifiée.

Créer et paramétrer un nouveau projet

- Ouvrez **QGIS Desktop 3.XX.XX with Grass 7.XX.XX**
- Dans l'onglet **Projet**, cliquez sur **Nouveau**.
- La fenêtre de QGIS se nomme maintenant « *Projet sans titre ». Le symbole * indique que votre projet n'est pas enregistré. Dans l'onglet **Projet**, cliquez sur **Enregistrer sous...** puis enregistrez le projet au format .qgz dans le dossier défini préalablement (cf. *supra* « Préparer le stockage des données »).
- Votre nouveau projet est créé, il faut maintenant le paramétrer. Dans l'onglet **Projet**, cliquez sur **Propriétés...**
- Général :
 - Dans la nouvelle fenêtre « Propriétés du projet », cliquez sur l'onglet latéral **Général**. (fig. 4).

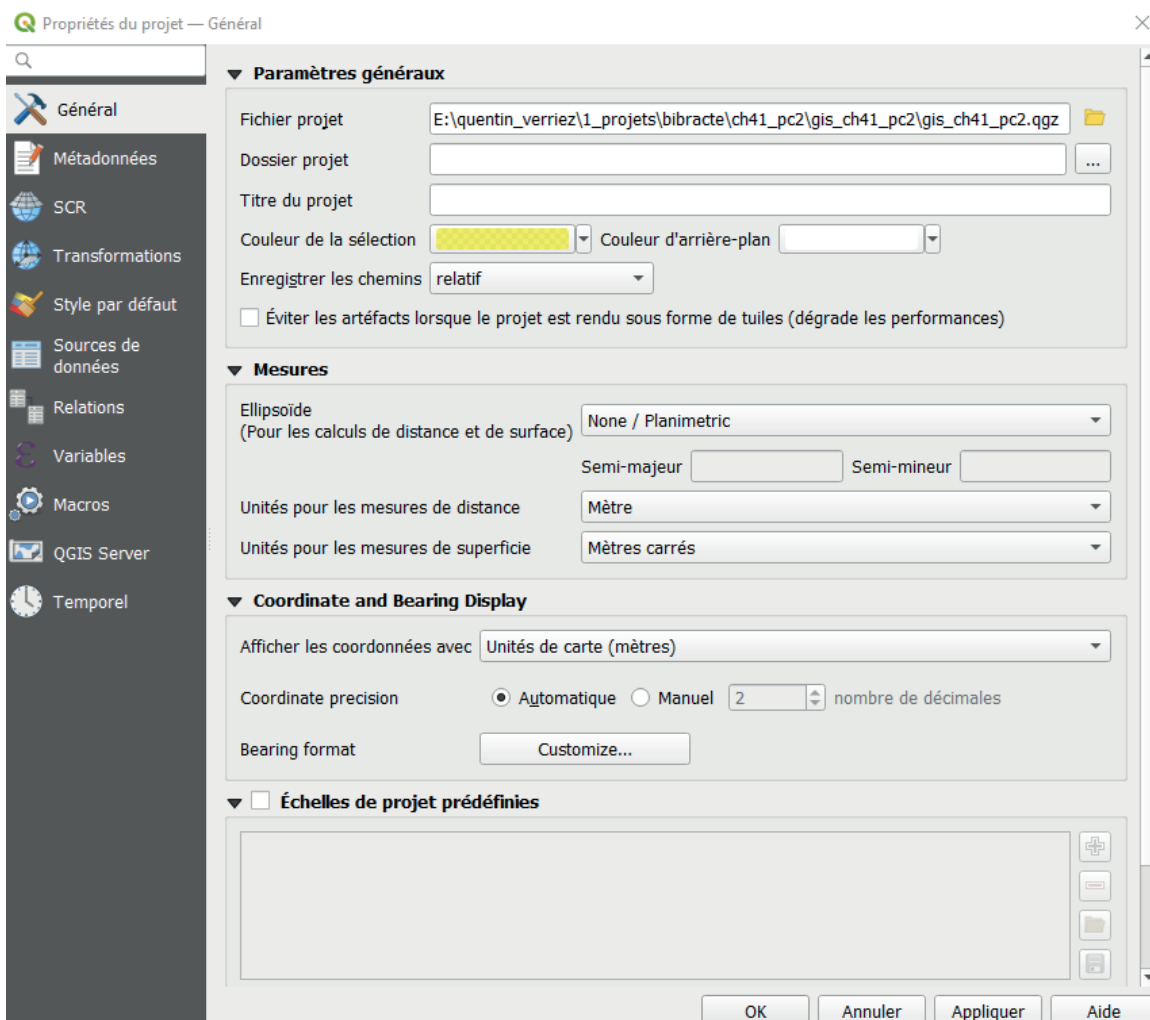


Figure 4

- **Fichier projet** indique le chemin vers votre fichier de projet .qgz.
- Définissez votre **Titre du projet** (évitez les caractères spéciaux, accents et espaces).
- Dans **Enregistrer les chemins**, choisissez **relatif**. Cela permet de déplacer le dossier de votre projet sur votre ordinateur ou une autre machine sans supprimer les liens avec les données interrogées.
- Ne modifiez pas les autres paramètres.
- Métadonnées :
 - Cette étape est essentielle pour documenter vos données.
 - Cliquez sur l'onglet latéral **Métadonnées** (fig. 5).

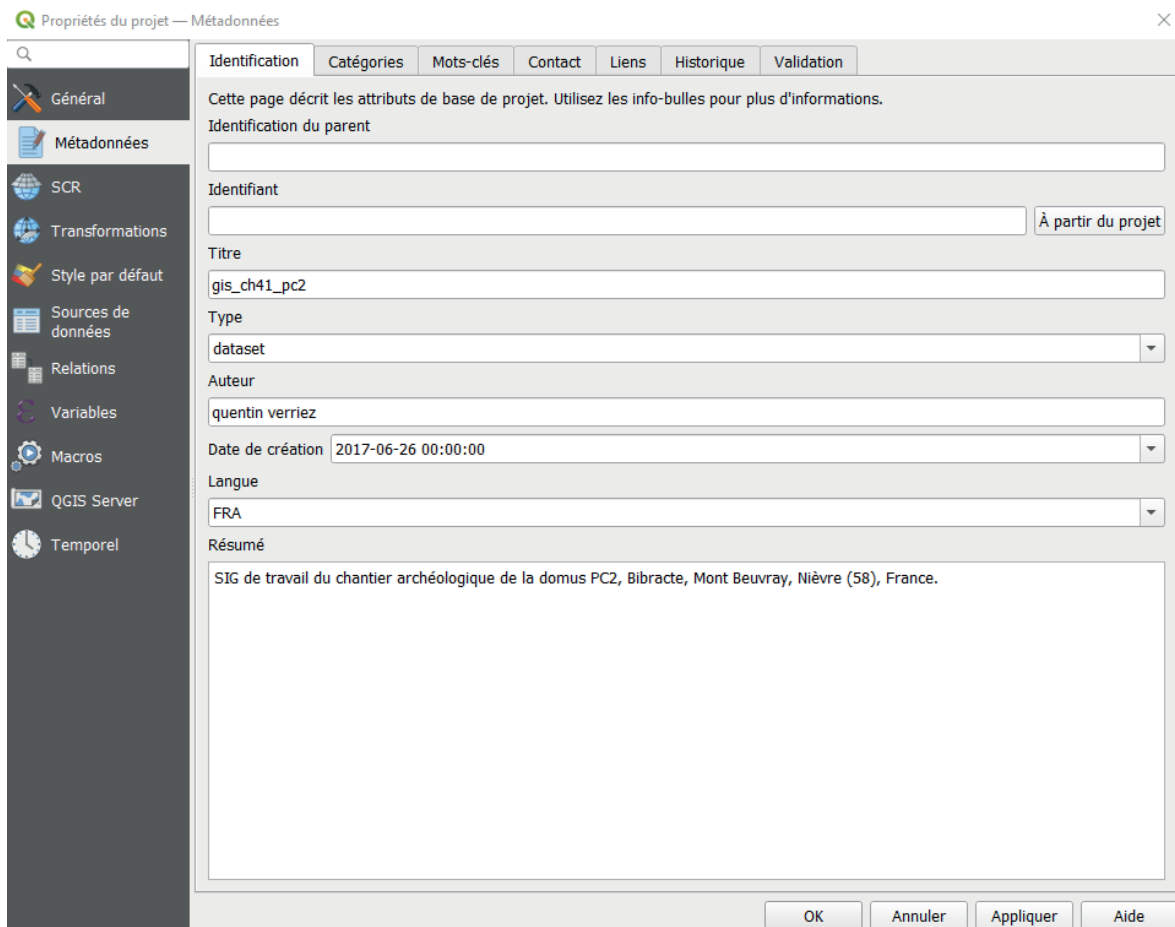


Figure 5

- Il est indispensable de remplir, à minima, les cases suivantes :
 - **Identification** : Titre, Type (choisissez « dataset »), Auteur, Date de création, Langue (« FRA » pour français, « ENG » pour anglais), Résumé.
 - **Contact** : Nom, Rôle (choisissez « owner »), Organisation, Position (votre statut), Courriel (fig. 6).

Propriétés du projet — Métadonnées

Identification Catégories Mots-clés **Contact** Liens Historique Validation

Contacts en lien avec projet.

Nom: Quentin Verriez

Rôle: owner

Organisation: Bibracte EPCC

Position: Archéologue

Courriel: q.verriez@bibracte.fr

Téléphone:

Fax:

Adresse

Type	Adresse	Code postal	Ville	one administratif	Pays

OK Annuler Appliquer Aide

Figure 6

- SCR :

Le Système de Coordonnées de Référence ou SCR permet d'associer les coordonnées numériques des fichiers interrogés à une position sur la surface de la Terre. QGIS est capable d'en gérer plus de 7000, définis par le European Petroleum Search Group (EPSG) et l'Institut National Géographique (IGN).

Il est indispensable de connaître le SCR dans lequel s'inscrivent vos données, sans quoi elles seront complètement déformées, à une mauvaise position, voire complètement illisibles.

- Pour définir le SCR de votre projet, cliquez sur l'onglet latéral **SCR**. (fig. 7), puis recherchez dans **Filtre** votre SCR.

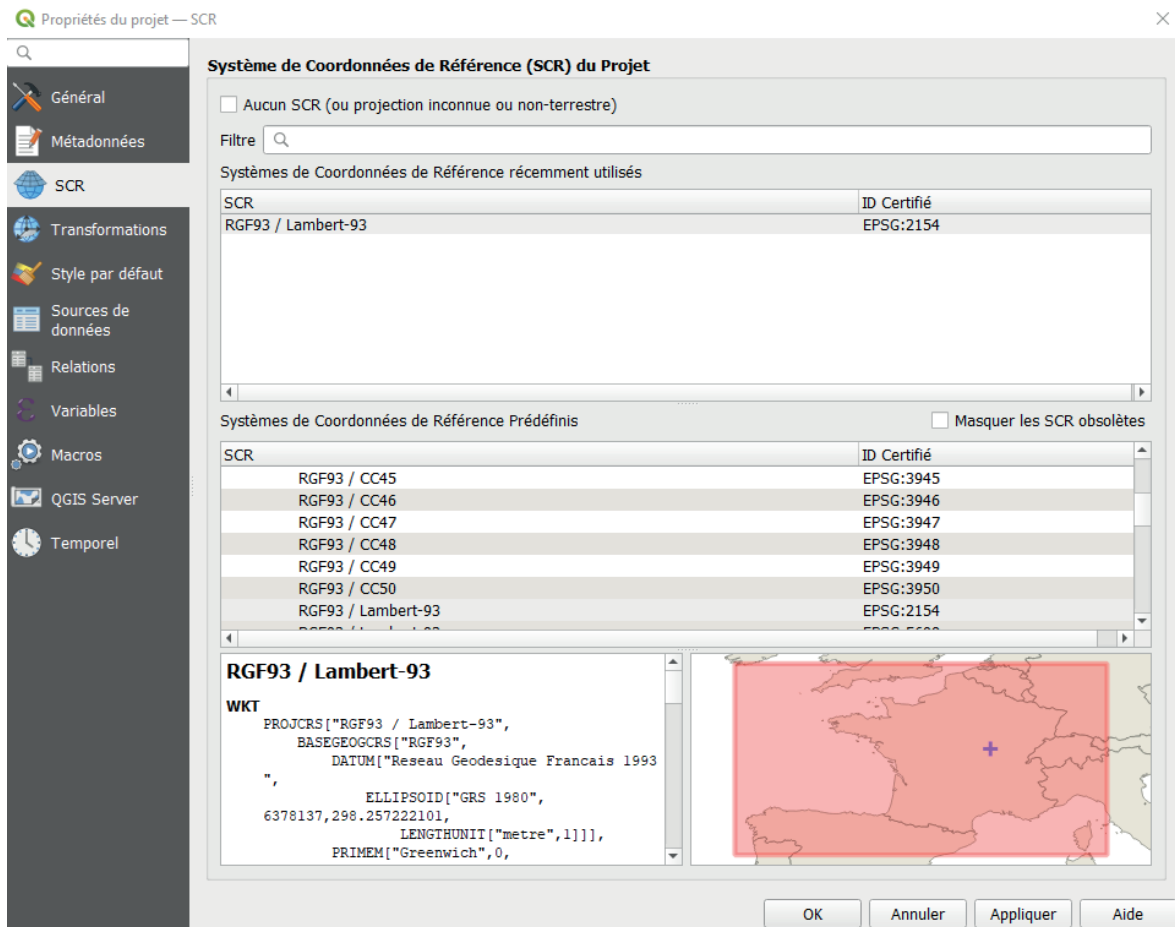


Figure 7

- Cliquez ensuite sur votre SCR dans la liste « Systèmes de Coordonnées de Référence Prédéfinis ». Un descriptif et une carte s'affichent pour vous montrer l'emprise générale associée à ce SCR.
- Cliquez sur **Appliquer**.
- Vous pouvez contrôler le SCR de votre projet en bas à droite de l'interface générale de QGIS (fig. 8).



Figure 8

Le SCR officiel le plus couramment utilisé en archéologie sur le territoire français est le RGF93 / Lambert-93 (EPSG : 2154). Le SCR de base utilisé par QGIS, comme par de nombreux logiciels ou machines générant de l'information spatialisée (GPS, tachéomètres, drones, etc.), est le WGS 84 (pour World Geodetic System 1984) (EPSG : 4326).

- Style par défaut
 - Cliquez sur l'onglet latéral **Style par défaut**.
 - Dans Options, réglez l'**Opacité** à 70 %. Cela facilite les relevés.
- Cliquez sur **OK** pour fermer la fenêtre « Propriétés du projet ».



Importer des données (raster)

Cette étape vous permet d'importer par exemple des fonds de carte ou des orthophotographies. Notez que seuls les rasters déjà géoréférencés peuvent être importés selon les méthodes suivantes.

Les formats les plus couramment exploités sont le GeoTIFF et le GeoJPEG. Attention, le préfixe « Geo » n'apparaîtra pas systématiquement devant l'intitulé du format de votre fichier. Des fichiers .tiff ou .jpeg peuvent être des rasters géoréférencés.

- Sélectionnez le raster à importer, maintenez **clic gauche** et faites **glisser** le fichier dans le canevas de QGIS.
- Si les SCR de votre projet et du raster correspondent, QGIS positionne automatiquement le raster à l'échelle.

OU

- Utilisez le raccourci  ou dans l'onglet **Couche**, sélectionnez **Ajoutez une couche**, puis **Ajoutez une couche raster...**
- Le Gestionnaire des sources de données s'ouvre dans une fenêtre. Sélectionnez **Fichier** dans « Type de source ».
- Dans « Sources », indiquez le raster à importer en cliquant sur  .
- Cliquez sur **Ajouter** pour importer le raster. Si le SCR de votre projet et du raster correspondent, QGIS positionne automatiquement le raster à l'échelle.

Importer des données (vecteur)

Cette étape vous permet d'importer des vecteurs déjà géoréférencés, donc produits grâce à un SIG. Les dessins vectoriels produits par DAO (grâce à Inkscape ou Illustrator) ne peuvent pas être importés avec cette méthode.

Le format le plus couramment utilisé est le Shapefile, ou ESRI Shapefile, du nom de l'entreprise créatrice du format. Un fichier Shapefile s'accompagne toujours de 4 autres fichiers portant le même nom (.dbf, .prj, .qpj, .shx) et d'un fichier de style optionnel (.qml). Vous devez conserver tous les fichiers, mais vous n'avez besoin d'utiliser que le fichier .shp.

- Sélectionnez le raster à importer, maintenez le **clic** et faites **glisser** le fichier dans le canevas de QGIS.
- Si les SCR de votre projet et du raster correspondent, QGIS positionne automatiquement le raster à l'échelle.

OU

- Dans l'onglet **Couche**, sélectionnez **Ajoutez une couche**, puis **Ajoutez une couche vecteur...**
- Le « Gestionnaire des sources de données » s'ouvre dans une fenêtre. Sélectionnez **Fichier** dans « Type de source » et laissez l'**Encodage** en automatique.
- Dans « Sources », indiquez le vecteur à importer en cliquant sur .
- Ouvrez le fichier vecteur et cliquez sur **Ajoutez**.


Importer des données (texte délimité)

Cette étape vous permet d'importer des fichiers CSV et Texte (des tableaux enregistrés en format texte) pour les transformer en points sur une carte. Pour visualiser des cibles de photogrammétrie ou des objets individualisés par exemple. Les fichiers Texte exploitables par QGIS présentent à minima les informations suivantes : le nom du point et ses coordonnées X (Est) et Y (Nord) (fig. 9).

Nom_point	X	Y	Z
Target 14	778701.040	6647641.856	786.829
Target 15	778699.872	6647640.818	785.787
Target 17	778700.306	6647642.758	785.903
Target 19	778693.659	6647643.671	785.122
Target 20	778687.842	6647641.806	784.482
Target 21	778680.799	6647644.319	783.897
Traget 22	778675.252	6647641.537	783.413

← Entête ou Header (optionnel)

Figure 9

- Utilisez le raccourci  ou, dans l'onglet **Couche**, sélectionnez **Ajoutez une couche**, puis **Ajoutez une couche de texte délimité...**
- Le « Gestionnaire des sources de données » s'ouvre dans une fenêtre (fig. 9).

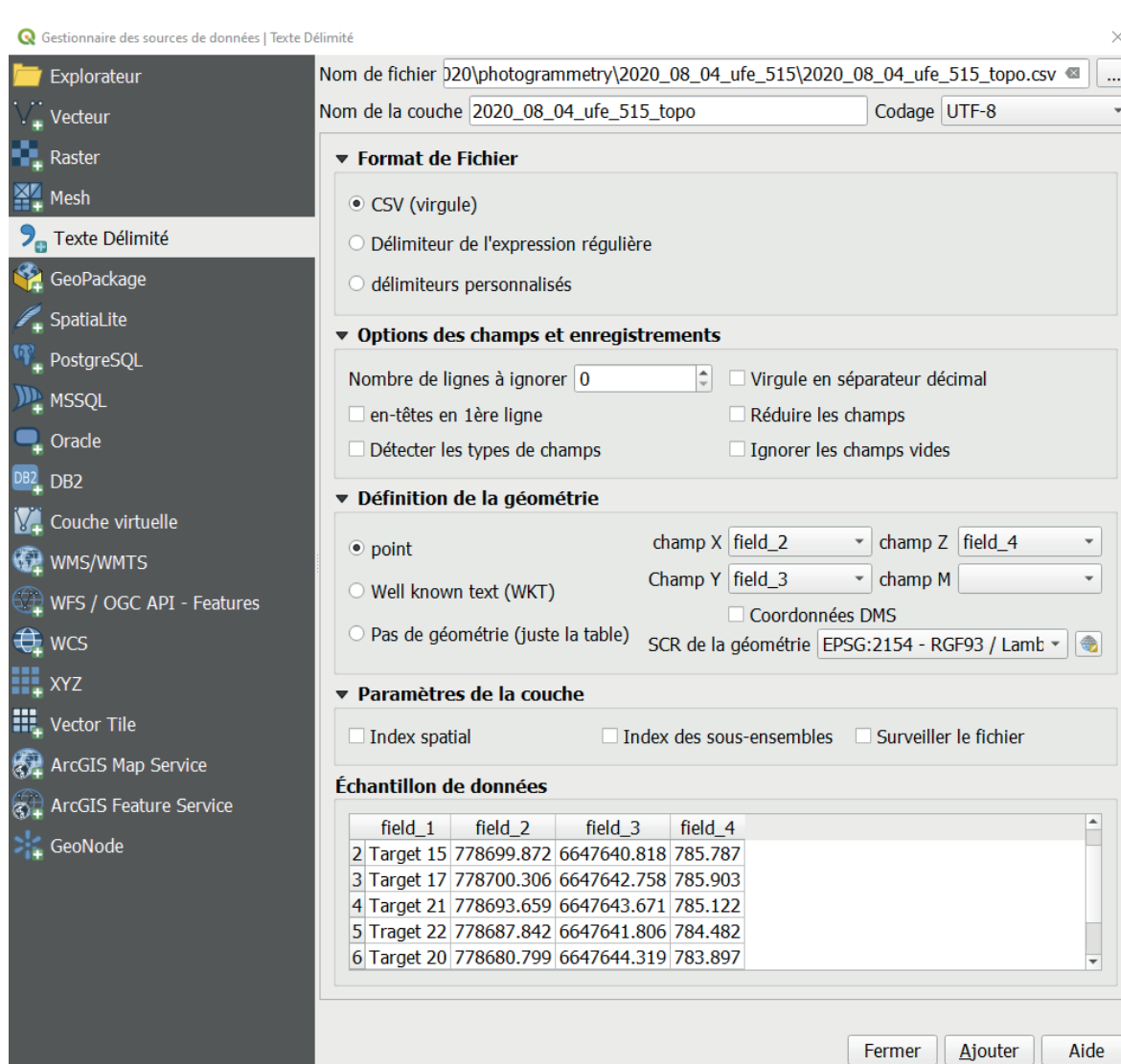
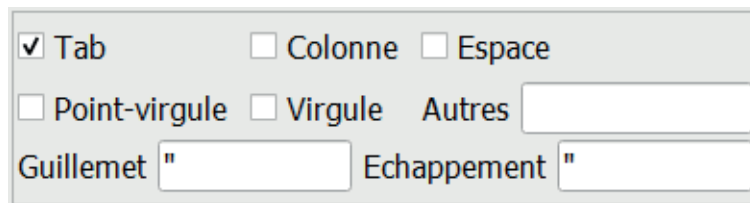


Figure 9

- Indiquez le fichier texte à importer en cliquant sur .
- Indiquez un **Nom de couche** si vous souhaitez qu'un nom différent de celui du fichier apparaisse dans QGIS.
- Sélectionnez le **Codage** (ou encodage) de votre fichier. Les encodages les plus courants en Europe sont l'UTF-8, Windows-1252 et latin-1.
- À tout moment, le tableur « Échantillon de données » en bas de page permet de vérifier comment QGIS interprète le fichier texte sélectionné en fonction des paramètres choisis.
- Dans « Format de Fichier » :
 - Cochez **CSV (virgule)** s'il s'agit d'un fichier CSV. QGIS devrait l'interpréter automatiquement.
 - Cochez **Délimiteur de l'expression régulière** pour un fichier texte dont les colonnes sont séparées par un ensemble de caractères (une expression) qui se répète systématiquement. Renseignez ensuite l'**Expression** en question dans le nouveau champ qui apparaît.
 - Cochez **délimiteurs personnalisés** pour un fichier Texte dont les colonnes sont séparées par un ou plusieurs caractères spéciaux (fig. 10).



The image shows a dialog box titled 'Format de Fichier' (File Format) with several options for delimiters. The 'Tab' option is checked, while 'Colonne' and 'Espace' are unchecked. Below these, there are three more options: 'Point-virgule' (unchecked), 'Virgule' (unchecked), and 'Autres' (checked), which is followed by a text input field. At the bottom, there are two more input fields: 'Guillemet' (with a double quote character) and 'Echappement' (with a double quote character).

Figure 10

- « Options des champs et enregistrements » :
 - **Nombre de lignes à ignorer** indique le nombre de lignes, à partir de la première, que QGIS ne doit pas prendre en compte lors de l'importation (pour les tableurs à en-têtes complexes par exemple).
 - En cochant **en-têtes en 1ère ligne**, vous indiquez à QGIS que la première ligne contient les intitulés des colonnes et non de la donnée directement. Il n'est pas obligatoire d'avoir des en-têtes de colonne. Si cette case n'est pas cochée, QGIS nommera automatiquement les colonnes « field_1 », « field_2 », etc.

- **Détecter les types de champs** permet à QGIS de détecter automatiquement le type d'information contenue dans un champ. Par exemple une suite de caractères, une date, une heure, etc.
- Si les champs contiennent des chiffres décimaux séparés par des virgules (et non par des points), cochez **Virgule en séparateur décimal**.
- Ne cochez pas **Réduire les champs**.
- Cochez **Ignorer les champs vides** pour ne pas prendre en compte les champs vides lors de l'importation.
- « Définition de la géométrie »
 - Cochez **point**.
 - Indiquez à quelle colonne correspond le **champ X** (Est), le **champ Y** (Nord) et le **champ Z** (Élévation). Laissez le champ M vide.
 - Sélectionnez-le **SCR de la géométrie**, soit le SCR dans lequel s'exprime les coordonnées de votre fichier texte. Il doit s'agir du même SCR que celui de votre projet.
- Cliquez sur **Ajouter** en bas à droite de la fenêtre pour importer la couche et voir apparaître les points sur le canevas (fig. 11).

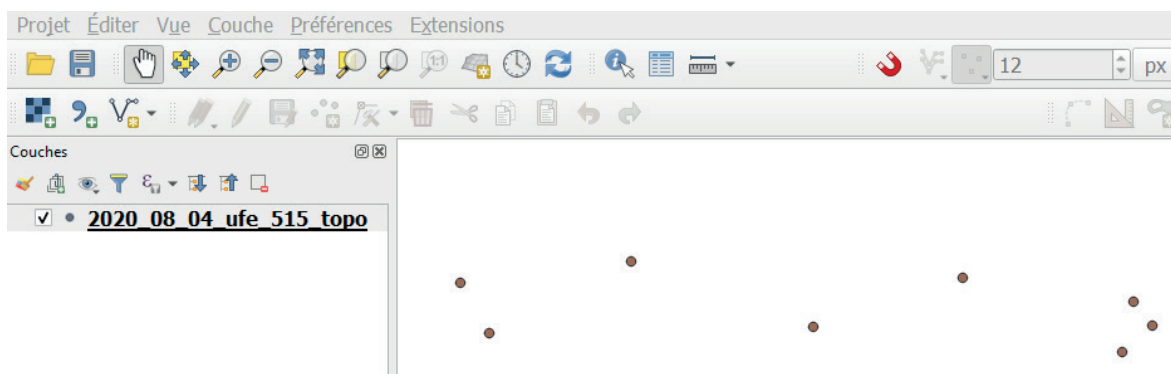


Figure 11

Préparer le dessin

Techniquement, QGIS pourrait gérer toutes les couches archéologiques de votre chantier dans un seul Shapefile. Toutefois, pour pouvoir dessiner sur plusieurs tablettes en même temps, il faudrait équiper le chantier avec un serveur et un réseau local ou une connexion internet. Pour s'affranchir de ces équipements, il suffit de s'en tenir à une règle simple : une unité d'enregistrement (US, UF, Fait, etc.) par Shapefile. Nous verrons par la suite comment fusionner simplement des Shapefile (*cf. infra* « Fusionner des couches »).

- Vous avez besoin du Shapefile « plg_chxx_ufxx » disponible dans le dossier « Bibracte_GIS_Tools » (*cf. supra* « Télécharger le dossier « Bibracte_GIS_Tools » »).
- Remplacez « chxx » dans le nom des 5 fichiers pour qu'ils correspondent à votre numéro de chantier (« plg_ch[mon numéro de chantier]_ufxx »).

Dans notre exemple, elle se nomme « plg_ch41_uf_xx ».

- Reprenez l'étape « Importer des données (vecteur) » pour importer le Shapefile dans votre projet QGIS.
- Importez une orthophotographie qui vous servira de support pour votre relevé. Pour cet exemple, nous utiliserons une orthophotographie du four de bronzier de la pièce XX de la *domus* PC2 (fig. 12).

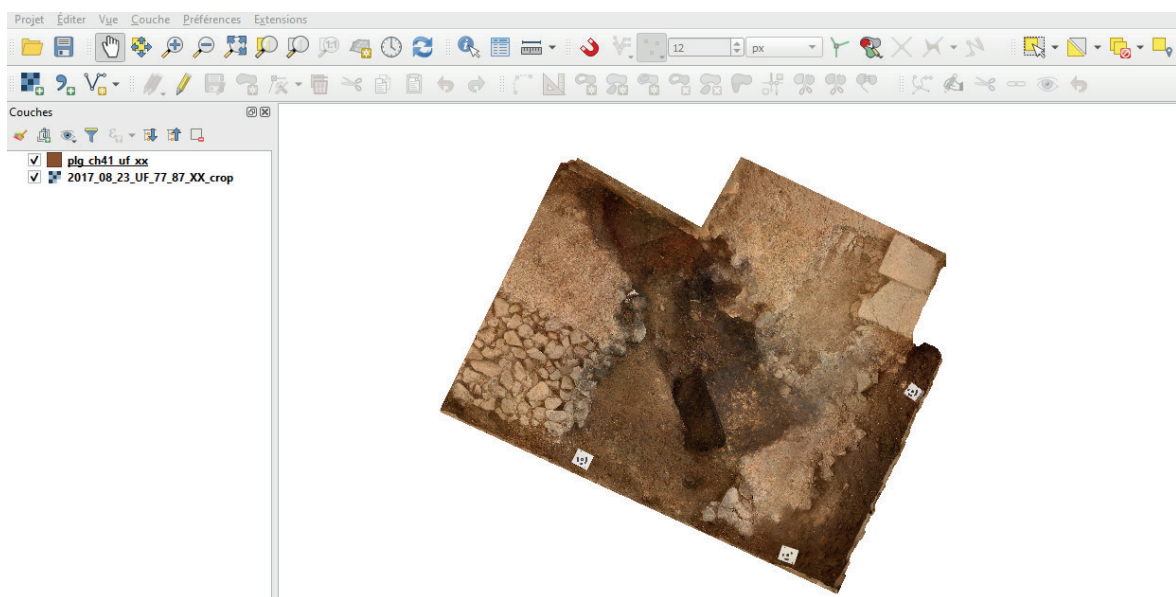





Figure 12

Dessiner une couche archéologique

Nous allons dessiner le creusement du four au centre de la pièce XX.

- Cliquez sur la couche « plg_chxx_ufxx » que vous venez d'importer dans le panneau **Couches** (de base à gauche de l'écran).
- Cliquez sur . Vous venez de passer la couche en mode édition.
 - À savoir : dans QGIS, aucune modification ne peut être faite sur une couche vecteur tant qu'elle n'est pas en mode édition.
- Dans QGIS, les formes créées s'appellent des entités. Pour créer une nouvelle entité, cliquez sur . Le pointeur de la souris devient .
- Faites un **clic gauche** pour commencer à dessiner l'entité.
- Un nouveau **clic gauche** permet de tracer une ligne droite entre les deux points cliqués. Répétez l'action pour tracer les limites de l'UF (fig. 13).

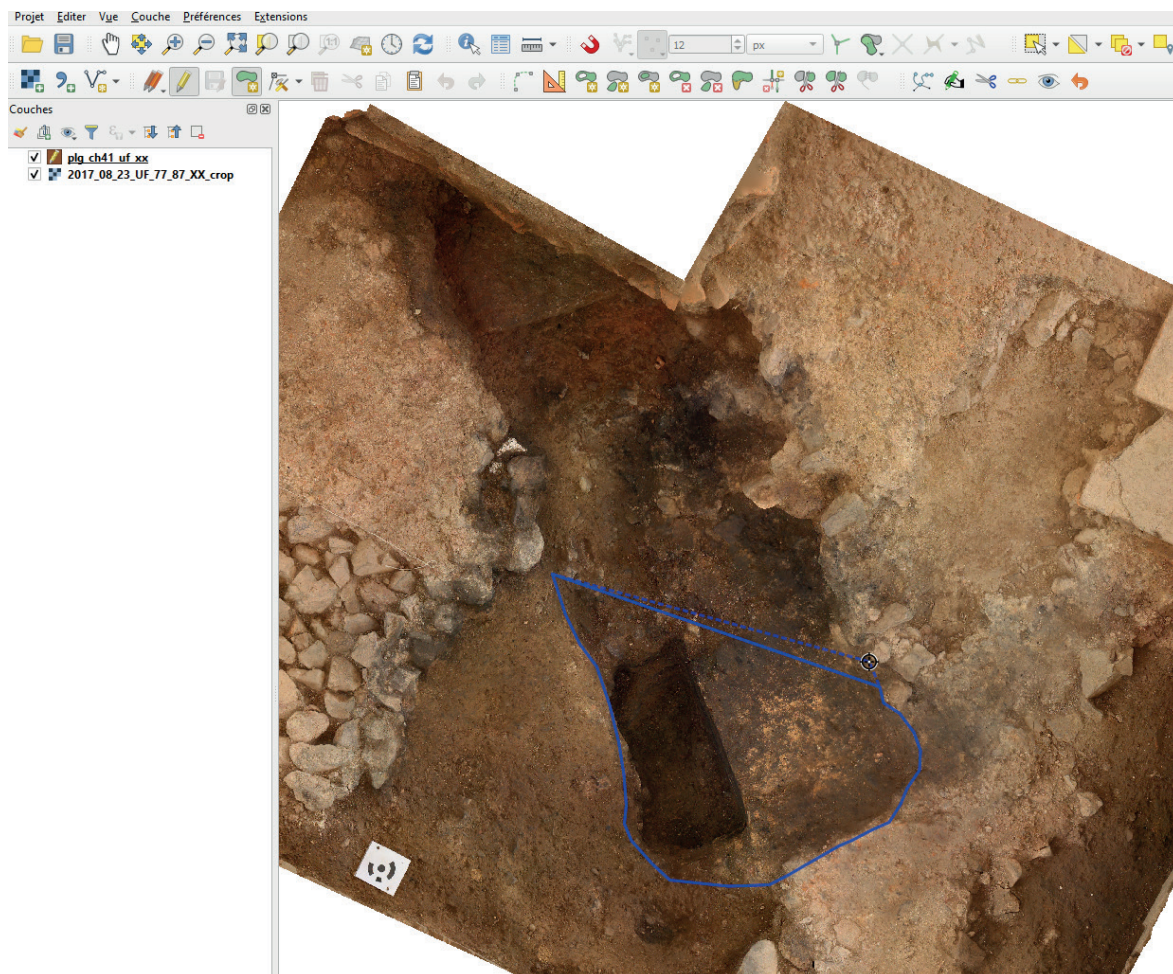


Figure 13

- La **molette** de la souris permet de zoomer. Pour vous déplacer, utilisez **clik molette** maintenu.
- La touche **supprimer** efface le dernier point cliqué. Plusieurs appuis suppriment les clics gauches précédents.
- Pour fermer et terminer la forme, faites **clik droit**. Peu importe l'emplacement de votre clic, l'emplacement n'est pas pris en compte.
- Une fenêtre « Attributs d'entités » s'ouvre automatiquement. Cette dernière présente 5 attributs (fig.14-15) :

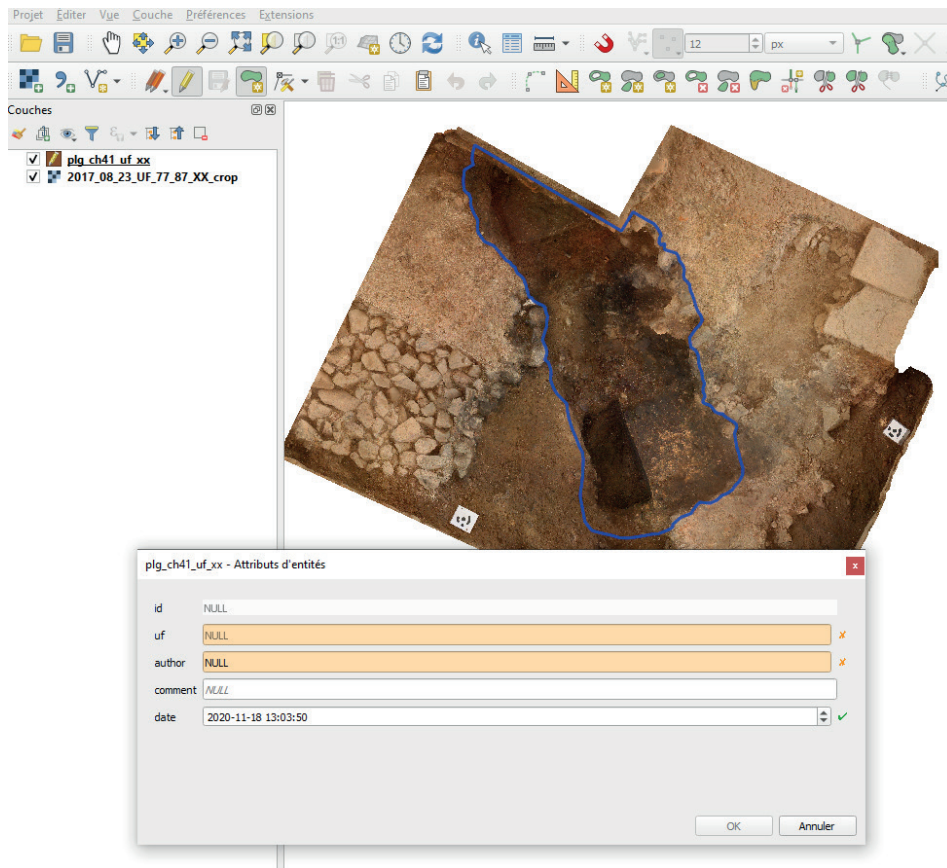


Figure 14

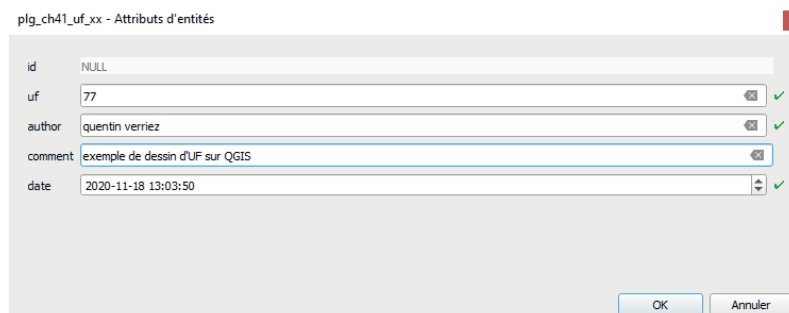


Figure 15

- **id** : cet attribut permettra de générer automatiquement un identifiant unique pour l'UF.
- **uf** : attribut obligatoire. Renseignez le numéro de l'UF dessinée. Cet attribut n'accepte que des chiffres.
- **author** : attribut obligatoire. Renseignez l'auteur du relevé.
- **comment** : attribut optionnel. Si vous souhaitez faire un commentaire sur le relevé.
- **date** : attribut généré automatiquement. Il enregistre la date et l'heure de création du Shapefile.
- Cliquez sur **OK**. Vous venez de créer votre première entité. Elle apparaît avec un fond de couleur uni.
- La couche/entité se trouve sur le Shapefile « plg_chxx_ufxx ». Il faut l'enregistrer sur un autre Shapefile.
 - Faites **clic droit** sur la couche « plg_chxx_ufxx », puis **Exporter** et **Sauvegarder les entités sous...** (fig. 16)

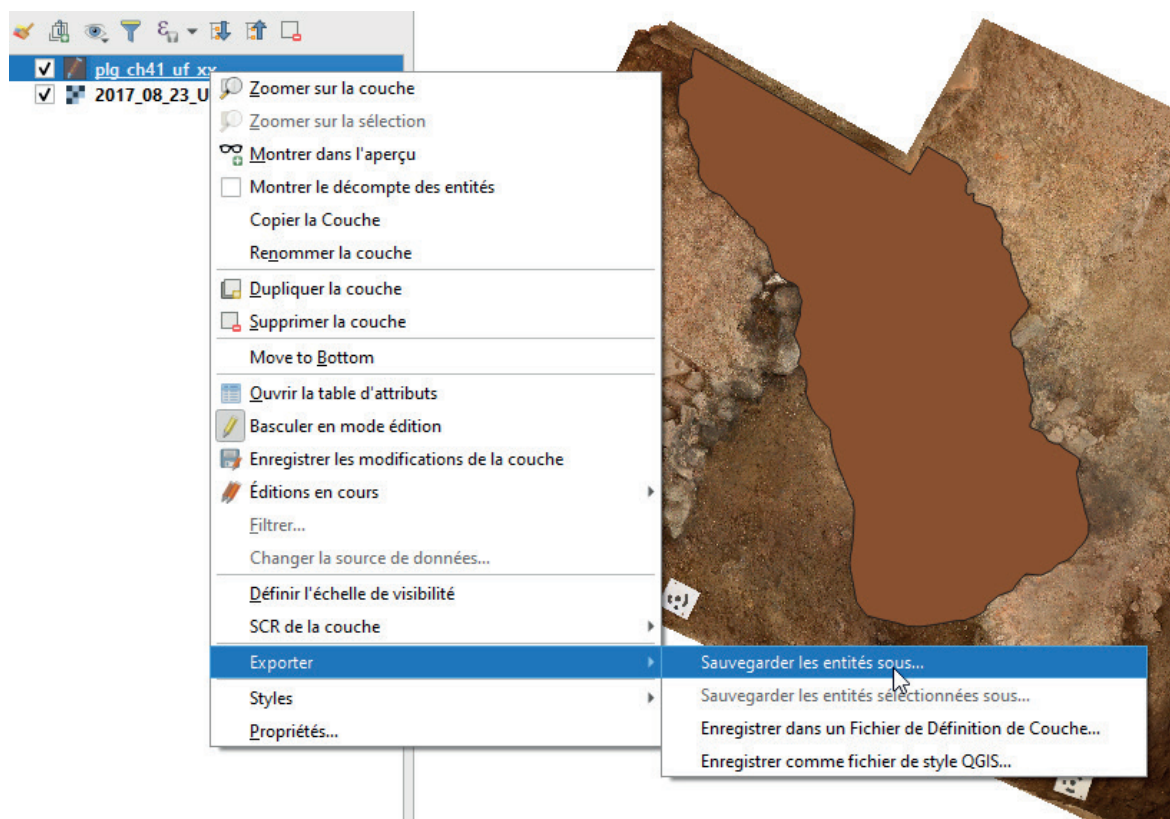


Figure 16

- Une fenêtre « Enregistrer la couche vectorielle sous... » s'ouvre (fig. 17).
- Vérifier les éléments suivants :
 - Format = ESRI Shapefile
 - SCR = EPSG 2154 – RGF93 / Lambert 93
 - Codage = UTF-8

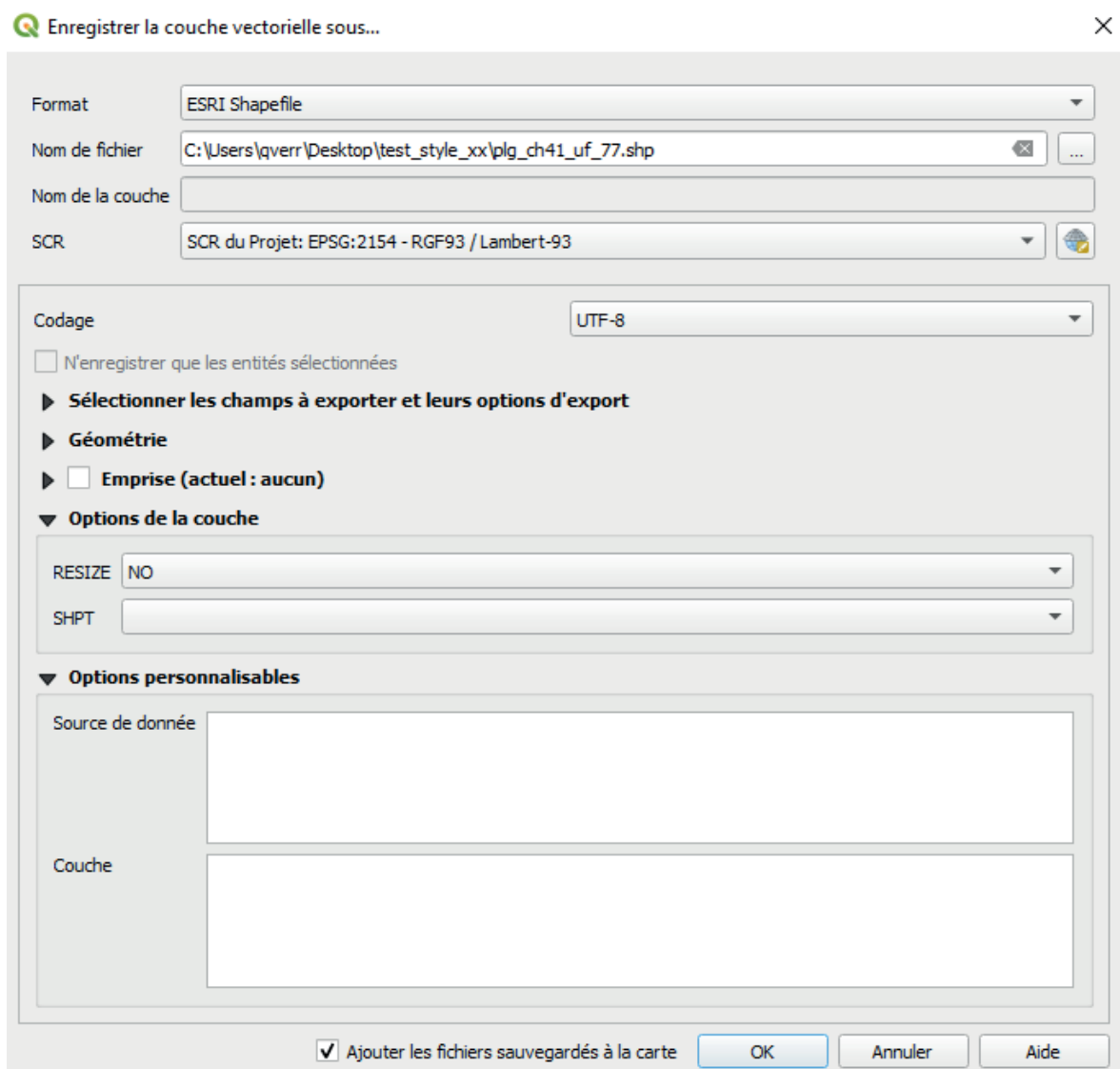


Figure 17

- Cliquez sur [...] pour sélectionner l'emplacement où vous souhaitez enregistrer le Shapefile.
- Ne modifiez pas les autres paramètres. Cliquez sur **OK**.
- Attention, n'utilisez pas **Dupliquer la couche** qui apparaît avec un **clic droit** sur une couche (fig. 18). Si cette fonction crée une nouvelle couche dans QGIS, elle ne crée pas de nouveau Shapefile. Le Shapefile reste le même pour la couche d'origine et le duplicata. Si vous enregistrez une modification sur la couche dupliquée, elle sera aussi effective sur la couche d'origine.

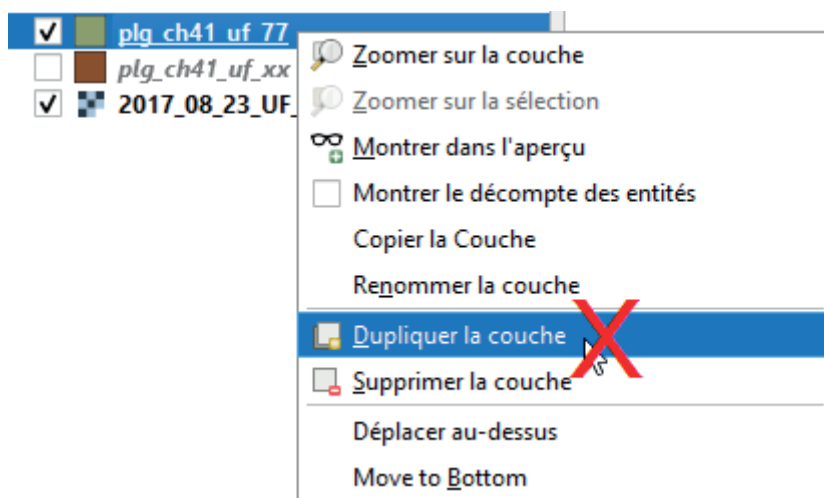


Figure 18

- Votre nouvelle couche apparaît dans le canevas et le panneau « Couches » de QGIS.
- Sélectionnez la couche « plg_chxx_ufxx » dans le panneau « Couches » et cliquez sur [Pencil]. Une fenêtre « Arrêter l'édition » s'ouvre. Cliquez sur **Ne pas tenir compte** afin de garder la couche « plg_chxx_ufxx » vierge pour vos prochains dessins (fig. 19).

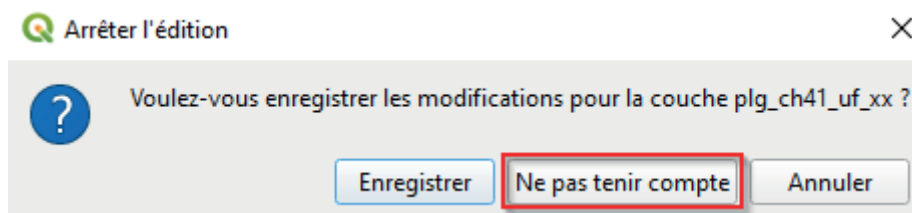


Figure 19

Dessiner avec des courbes

Takayuki Mizutani est l'auteur de l'extension officielle « *BezierEditing* » permettant de dessiner des entités avec des courbes de Bézier (fig. 20).

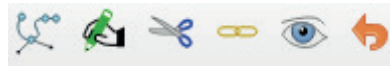





Figure 20

- Lors de la création d'une entité (cf. *supra* « Dessiner une couche archéologique »), cliquez sur  au lieu de .
- Le pointeur de la souris devient .
- Vous pouvez alors créer une entité en utilisant les courbes de Bézier. Le manuel complet de l'extension se trouve à l'adresse suivante : [https://github.com/tmizu23/BezierEditing/wiki/Document-\(English\)](https://github.com/tmizu23/BezierEditing/wiki/Document-(English))

Éditer une couche

Si vous avez besoin de modifier l'apparence d'une entité (couleur, symbole ou transparence), notez que QGIS n'enregistre pas ces modifications lors de la fermeture du projet. Pour conserver une apparence personnalisée, il faudra enregistrer un fichier de style (.qml) portant le même nom que votre Shapefile et ranger dans le même dossier.

- Faites un **double-clic** sur une couche pour ouvrir les paramètres (1) (fig. 21).
- Dans les onglets latéraux à gauche, sélectionnez **Symbologie** (2).
- Laissez **Symbole Unique** dans la première liste déroulante (3).
- Cliquez sur **Remplissage simple** pour ouvrir toutes les options de modification de base de votre entité (4).
- Par exemple, la transparence du fond se règle dans **Couleur de remplissage** (5).

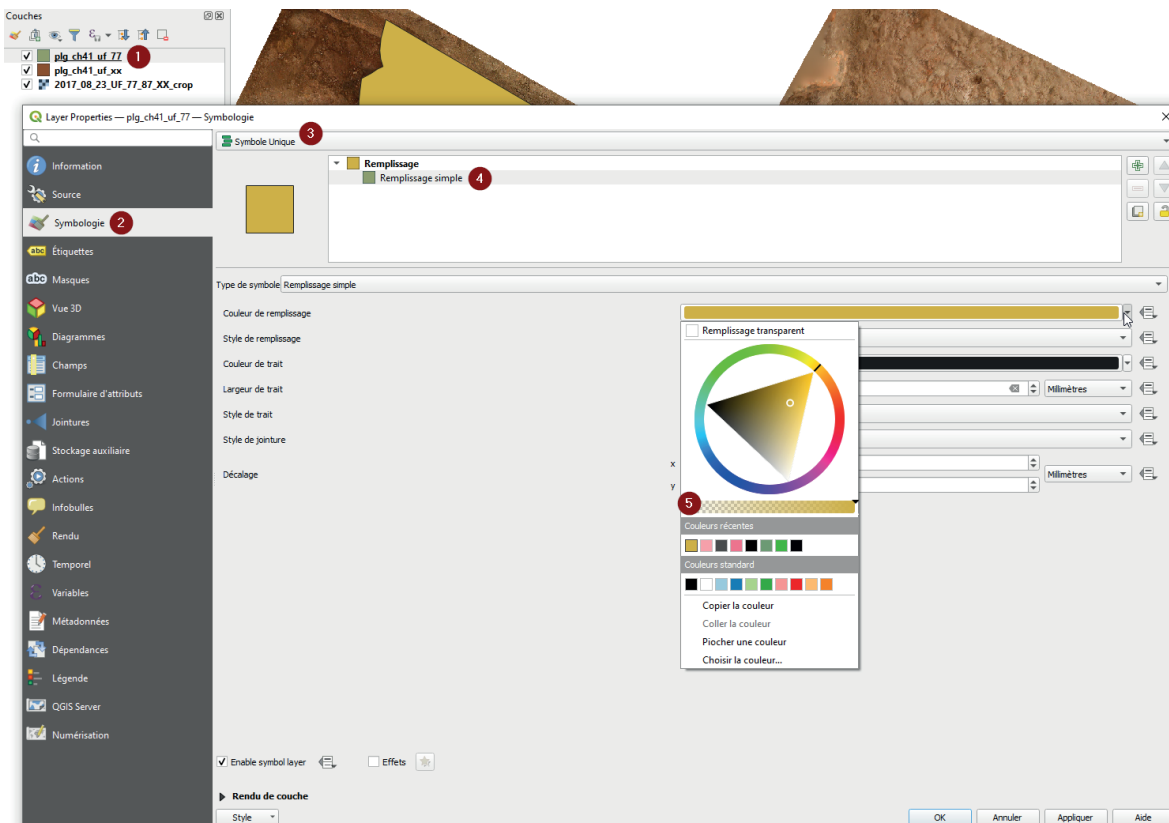


Figure 21

Deux outils permettent de modifier simplement la forme d'une entité dans QGIS :

- Outil de nœud  :

Lors de la création de votre entité, chaque **clik gauche** correspond à un nœud (ou sommet).

- Pour déplacer un nœud, sélectionnez-le avec **clik gauche** et faites un nouveau **clik gauche** sur le nouvel emplacement (fig. 22).

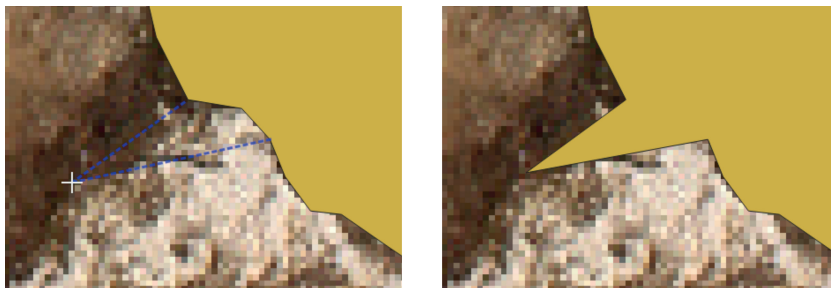


Figure 22

- Pour supprimer un nœud, sélectionnez-le avec **clik gauche** et utilisez la touche **supprimer**. Un nouvel appui sur la touche **supprimer** efface le nœud précédent.

- Pour ajouter un nœud, cliquez sur la **croix** au centre du segment et faites un nouveau **clik gauche** sur le nouvel emplacement (fig. 23).



Figure 23

- Pour déplacer un segment, sélectionnez-le avec **clik gauche** et faites un nouveau **clik gauche** sur le nouvel emplacement (fig. 24).

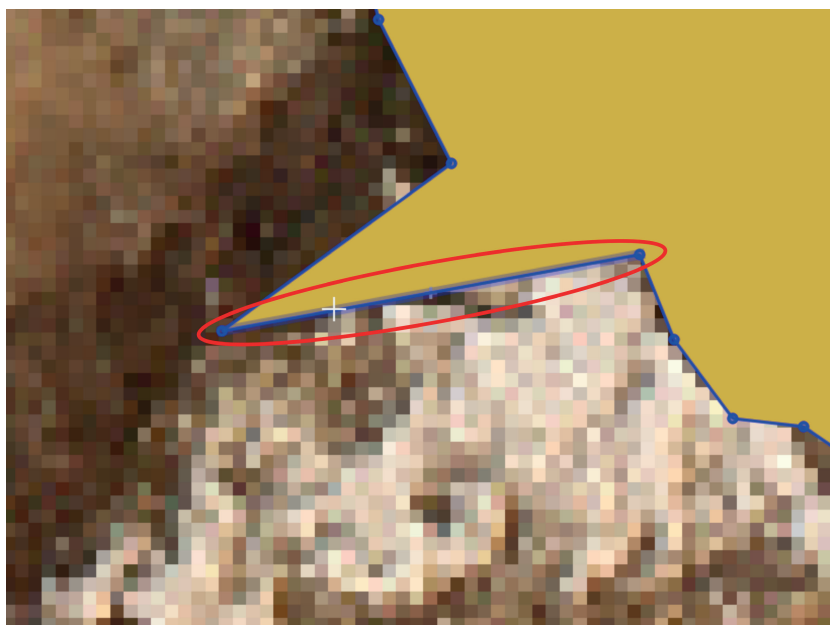


Figure 24

- Utilisez  ou les touches **ctrl + z** ou **cmd + z** pour annuler la dernière action.

- Une croix jaune vous indique qu'il y a une erreur sur le tracé de l'entité (fig. 25). Il peut s'agir soit de deux nœuds superposés, soit d'une boucle dans le tracé. Supprimez les nœuds en cause pour régler le problème.



Figure 25

- Remodeler les entités  :

Cette fonction permet d'étendre ou réduire facilement la superficie d'une entité.


- Pour étendre une entité, faites un premier **clik gauche** à l'intérieur de celle-ci.
- Dessinez l'extension avec des **cliks gauches** à l'extérieur de l'entité.
- Le dernier **clik gauche** doit revenir à l'intérieur de l'entité. Faites un **clik droit** pour créer l'extension (fig. 26).



Figure 26

- Pour réduire une entité, le processus est inversé. Faites un premier **clik gauche** à l'extérieur de l'entité.
- Dessinez l'extension avec des **cliks gauches** à l'intérieur de l'entité.
- Le dernier **clik gauche** doit être à l'extérieur de l'entité. Faites un **clik droit** pour supprimer une partie de l'entité.

Options d'accrochage

Une couche archéologique est forcément bordée par d'autres entités, qu'il s'agisse d'autres couches ou des limites de fouille. Lors de la création d'une couche/entité, il est nécessaire de pouvoir se baser sur les limites des entités voisines. En pratique, il est impossible de superposer parfaitement deux sommets ou segments « à la main ». Chaque point est associé à des coordonnées très précises, et cette précision est impossible à obtenir avec le pointeur de la souris, quel que soit le niveau de zoom. Pour « aimanter » un sommet sur une entité déjà existante, vous devez utiliser les options d'accrochage (fig. 27). Ainsi, dès que vous approchez la souris d'une entité « accrochable » avec le pointeur pour dessiner , il va automatiquement s'aimanter à cette entité (fig. 28).

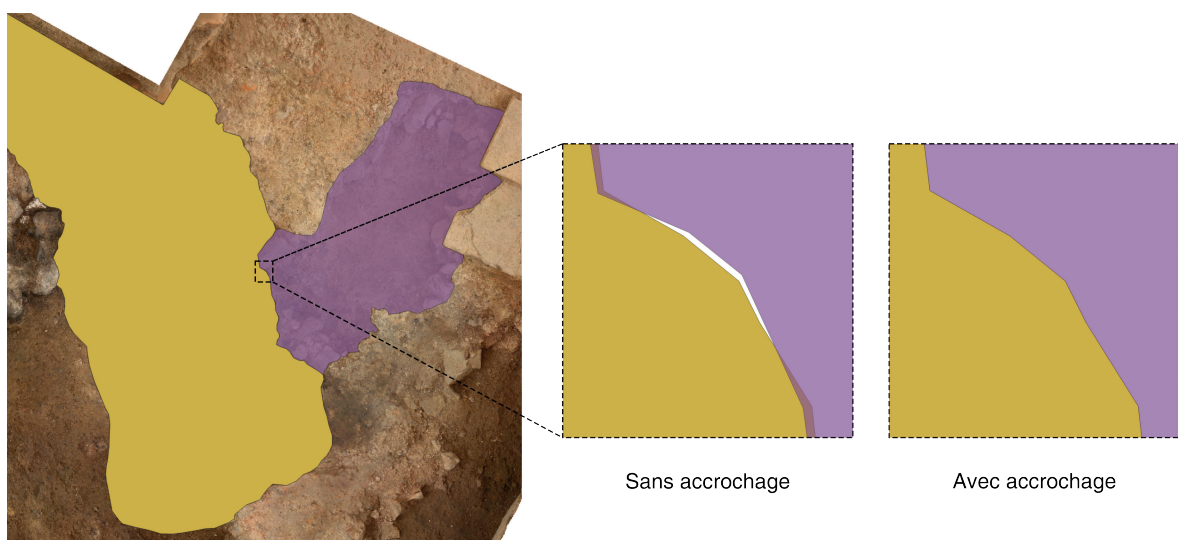


Figure 27

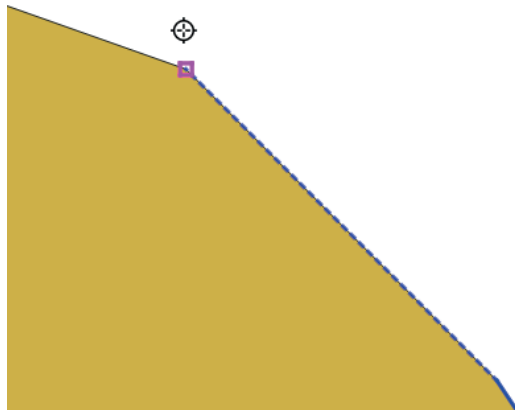





Figure 28

- Pour activer l'accrochage, cliquez sur .
- Le deuxième raccourci  est une liste déroulante :
 - **Toutes les couches** : l'accrochage est activé sur toutes les couches visibles sur le canevas
 - **Couche active** : lorsque vous créez plusieurs entités sur une seule couche, seules les entités de la couche en mode édition proposent l'accrochage. Cette option est inutile dans notre cas puisque nous ne dessinons qu'une entité par couche.
 - **Configuration avancée** : les options sont paramétrables couche par couche. C'est le mode à sélectionner dans notre cas.
 - **Ouvrir les options d'accrochage** : ouvre une fenêtre flottante pour gérer la configuration avancée.  ouvre la même fenêtre non flottante.
- Sélectionnez **Ouvrir les options d'accrochage** pour ouvrir la fenêtre de configuration avancée (fig. 28).

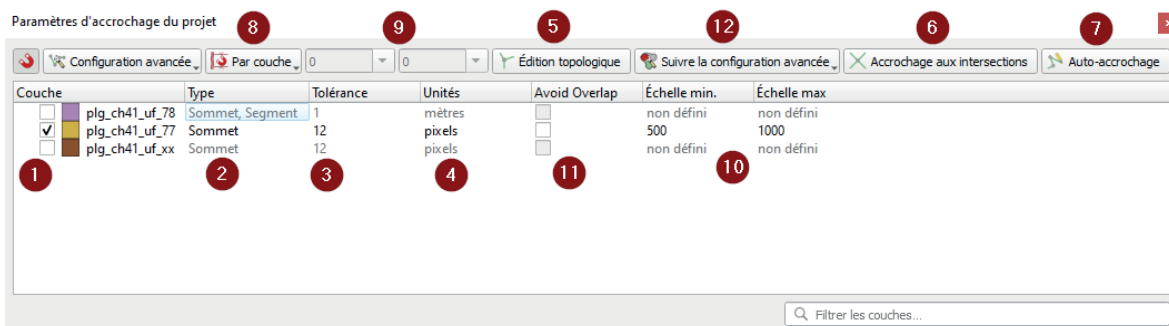


Figure 28

- La **case à cocher** (1) active les options d'accrochages pour la couche correspondante.

Dans notre exemple, la couche « plg_ch41_uf_77 » est activée.

- Le **Type** d'accrochage (2) ouvre une liste déroulante pour choisir quels éléments de la couche peuvent accrocher : les sommets, les segments, la surface, le centroïde ou le milieu des segments. Vous pouvez sélectionner plusieurs éléments.
- La **Tolérance** (3) et les **Unités** (4) fonctionnent ensemble. La **Tolérance** définit la taille du rayon dans lequel l'accrochage va s'activer autour d'un élément. Le rayon s'exprime en mètres ou en pixels (les **Unités**). Si vous n'avez pas de contrainte spécifique, il est conseillé de laisser ces paramètres à 12 pixels.

« plg_ch41_uf_78 » présente un rayon d'accrochage à un mètre autour des sommets et des segments de ses entités.

- L'**Édition topologique** (5), quand elle est activée, permet de déplacer conjointement des sommets et segments de plusieurs entités qui se superposent. Cela fonctionne aussi pour les entités de différentes couches, à condition que toutes les couches soient en mode édition.
- L'**Accrochage aux intersections** (6), comme son nom l'indique, permet de s'accrocher aux intersections de couches activées, même s'il n'y a pas de sommet à l'intersection.
- Il est possible de s'accrocher sur l'entité que vous êtes en train de dessiner si **Auto-accrochage** (7) est activé.
- Cela n'est pas très utile dans notre cas, mais il est possible de définir une échelle minimum et une échelle maximum entre lesquelles les options d'accrochages sont activées. Avec la **liste déroulante** (8), sélectionnez si vous souhaitez appliquer ces critères à toutes les couches actives (**Global**), ou les définir couche par couche (**Par couche**). Si vous choisissez l'option **Global**, indiquez les échelles min. et max. dans les **cadres** en (9). Si vous préférez le faire **Par couche**, utilisez les **cadres** en (10). Il est possible de taper manuellement l'échelle souhaitée.

« Éviter les chevauchements » (**Avoid Overlap**) (11) est une fonction très pratique pour dessiner une entité partageant ses limites avec d'autres. En l'activant, l'entité d'une couche ne peut plus être chevauchée. Donc si nous dessinons une nouvelle entité avec une partie qui se superpose à la première, QGIS découpera automatiquement cette partie (ill. 29).

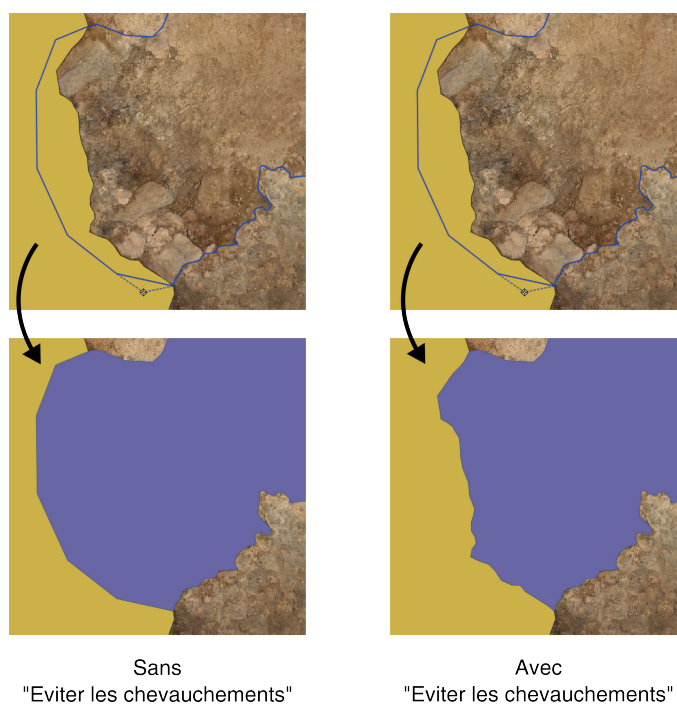


Figure 29


- Pour activer la fonction, cliquez sur la **case à cocher** d'**Avoid Overlap** (11) de la couche correspondante.
- La **liste déroulante** (12) permet de désactiver « Éviter les chevauchements », d'activer la fonction pour toutes les couches actives ou de suivre les règles de la configuration avancée couche par couche. Il est conseillé de choisir **Suivre la configuration avancée**.
- Attention, désactiver les options d'accrochage d'une couche ne désactive pas « éviter les chevauchements », même si la **case à cocher** apparaît en grisée (fig. 30). Pour désactiver la fonction, vous devez cliquer sur la **case à cocher** pour activer la couche, décocher **Avoid Overlap** et désactivez à nouveau la couche.

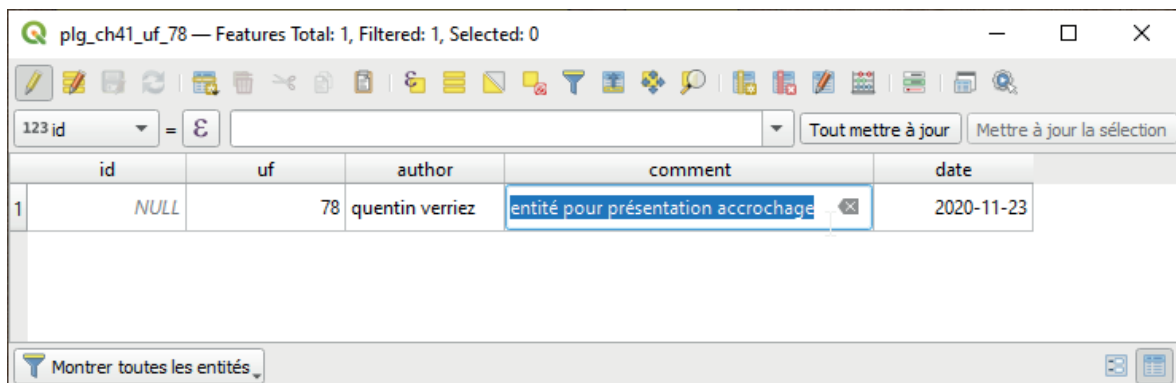
Couche	Type	Tolérance	Unités	Avoid Overlap
<input type="checkbox"/> plg_ch41_uf_78	Sommet	12	pixels	<input type="checkbox"/>
<input type="checkbox"/> plg_ch41_uf_77	Sommet	12	pixels	<input checked="" type="checkbox"/>
<input type="checkbox"/> plg_ch41_uf_xx	Sommet	12	pixels	<input type="checkbox"/>

Figure 30

Modifier la table attributaire



Il est possible de modifier les attributs que vous avez complétés lors de la création d’une entité (cf. supra « Dessiner une couche archéologique »).

- Sélectionnez la couche à modifier dans le **Panneau de gestion des couches**.
- Cliquez sur  pour ouvrir la **Table des attributs** (fig. 31).





id	uf	author	comment	date
1	NULL	quentin verriez	entité pour présentation accrochage	2020-11-23

Figure 31

- Cliquez sur  pour passer en **mode édition**.
- Faites un double-clic dans un champ pour le modifier. Veillez à respecter les informations attendues dans le champ à modifier (cf. supra « Dessiner une couche archéologique »).
- Cliquez sur  pour enregistrer les modifications et fermez la fenêtre.

Prendre des mesures

Il est possible de mesurer manuellement des longueurs, des aires ou des angles sur le canevas. Pour cela, cliquez sur la flèche à gauche de l’icône  et sélectionnez la fonction souhaitée :

-  pour **mesurer une longueur**. Deux **clics gauches** consécutifs créés un segment, dont la longueur est indiquée dans la fenêtre « Mesure » (fig. 32). Vous pouvez créer plusieurs segments à la suite. Un **clik droit** arrête la mesure en cours.

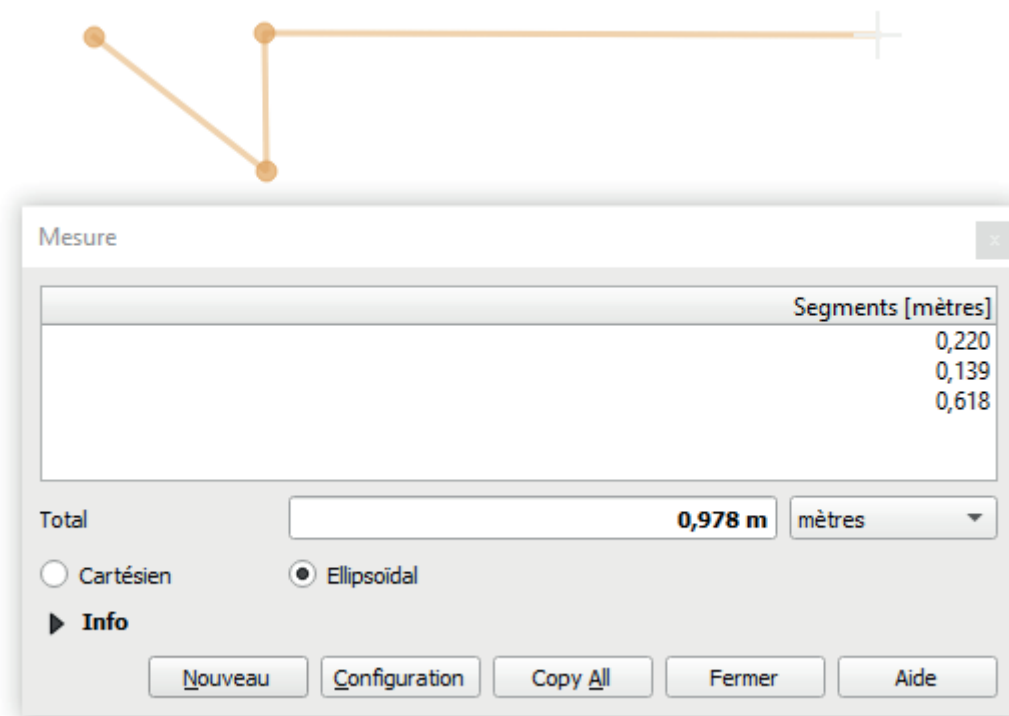


Figure 32



-  pour **mesurer une aire**. Comme lors de la création d'un polygone, utilisez le **clik gauche** pour dessiner l'emprise de l'aire à mesurer, puis faites un **clik droit** pour clore la forme (fig. 33).



Figure 33

-  pour **mesurer un angle**. Cet outil mesure l'angle entre deux segments, toujours dans le sens horaire (fig. 34). Notez que pour un angle de plus de 180 degrés, la mesure s'exprime en négatif, avec des valeurs comprises entre -0,1 et -179,999 degrés.

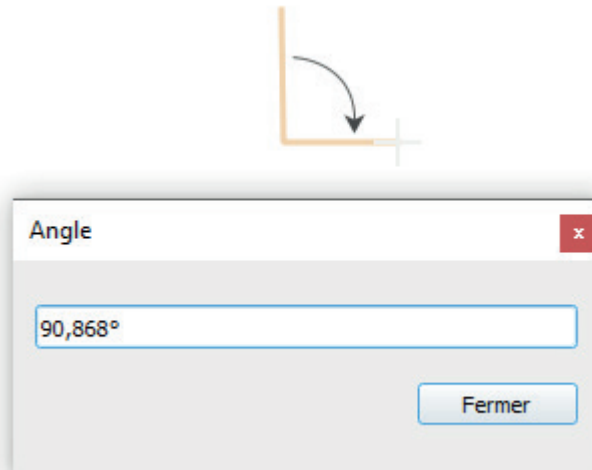


Figure 34

Organiser les couches

Lorsque le projet comporte beaucoup de couches, il devient utile de les organiser pour faciliter la visualisation de l'ensemble sur le canevas.

- Dans le panneau **Couches**, un **clik glissé** permet de trier une couche manuellement (fig. 35).

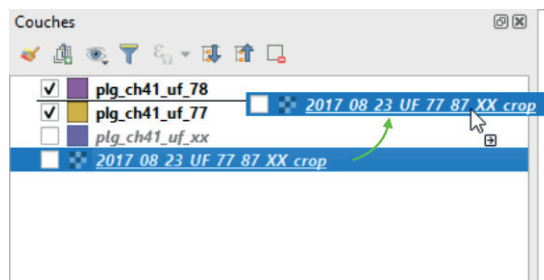


Figure 35

- Notez que la position des couches influe sur l'affichage dans le canevas. À la manière de feuilles de calques superposées, les couches supérieures apparaissent sur les couches inférieures (fig. 36).

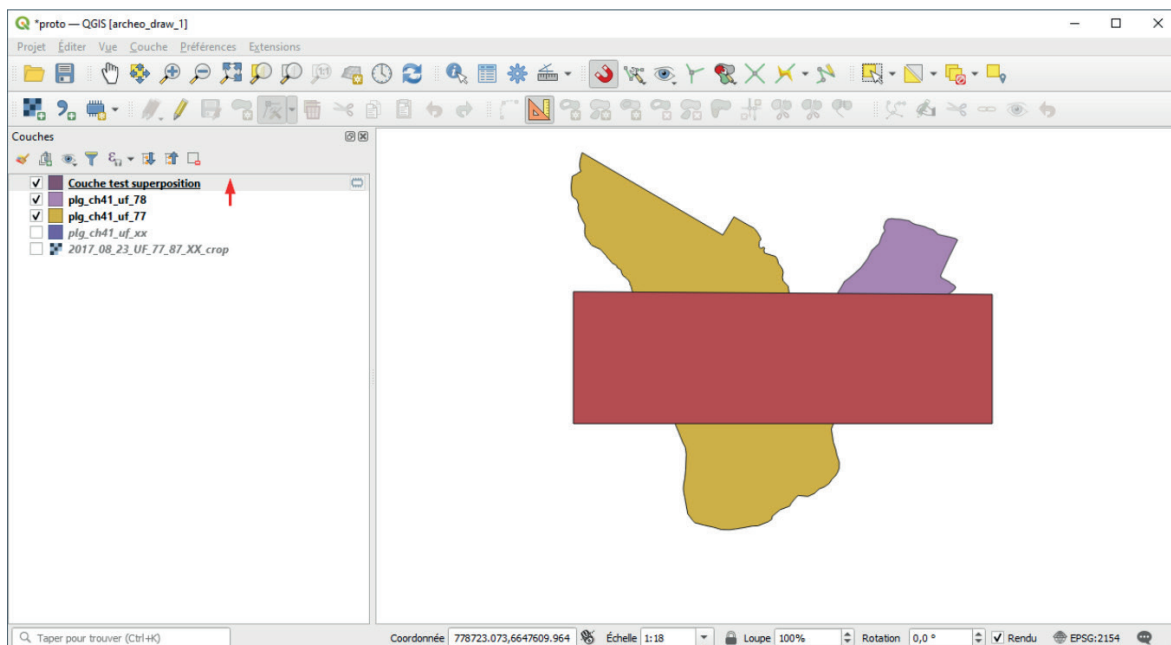
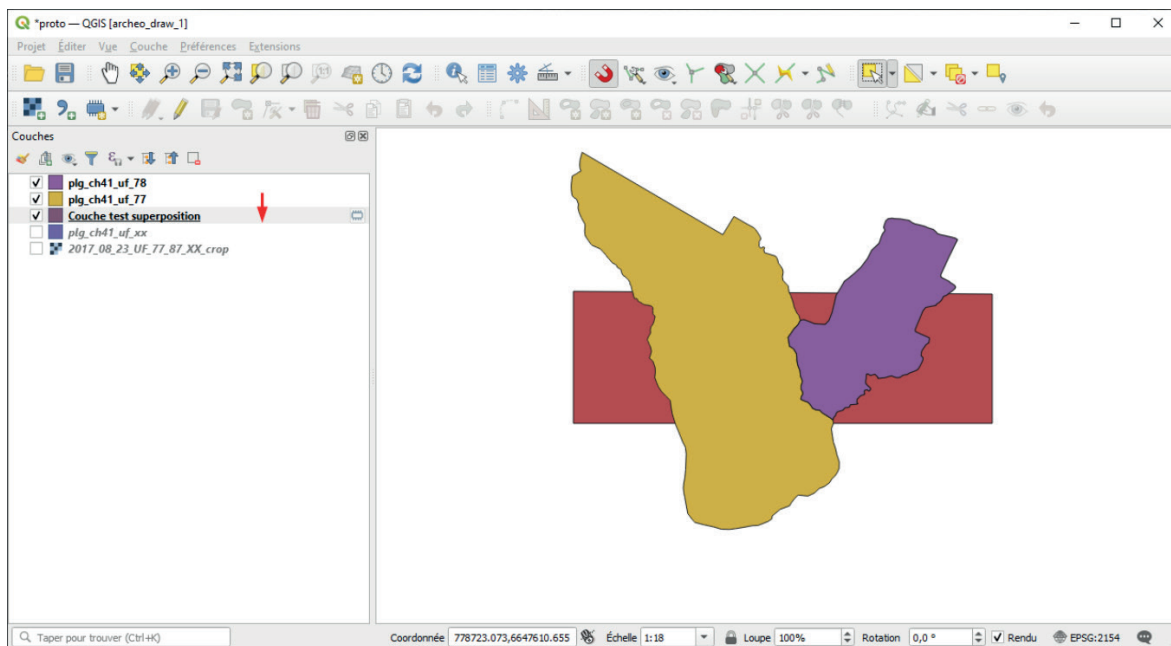




Figure 36

- Vous pouvez créer un groupe de couches en cliquant sur .
 - Le groupe créé, nommé par défaut « group1 », se place en bas de la liste.
 - Vous pouvez le renommer en faisant un **clic droit**, puis **Renommer le groupe**.
 - **Sélectionnez** ensuite les couches que vous souhaitez intégrer au groupe et faites un **clic glissé** sur sa couche du groupe.

- Les « thèmes de carte » enregistrent les couches visibles sur le canevas et vous permettent d’y revenir à tout moment en un clic. Cela permet de naviguer rapidement entre différents aperçus ou cartes de travail. Dans l’exemple suivant, nous allons créer un thème qui affichera les deux UF.
 - **Cochez** les couches à visualiser, et uniquement celles-ci. Dans notre cas, il s’agit des deux UF (fig. 37).



Figure 37

- Cliquez sur  dans la barre de raccourcis supérieure du panneau **Couches**, puis sur **Ajouter un thème...** (fig. 38).

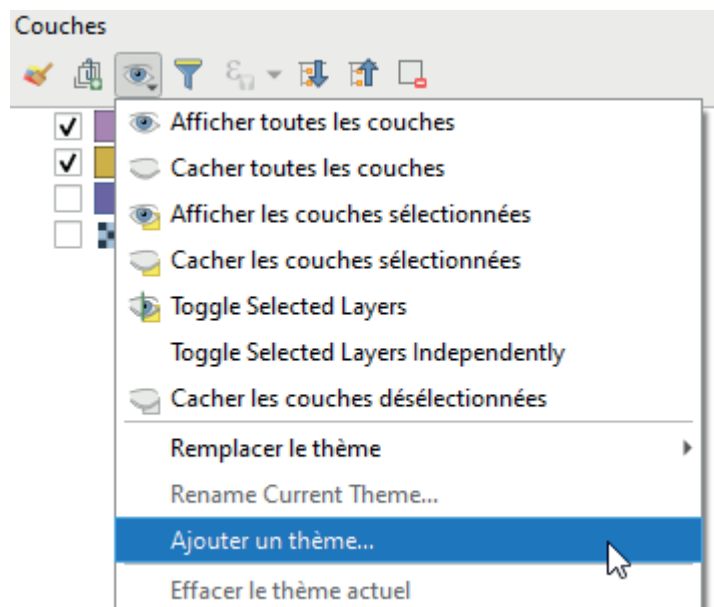


Figure 38

- Une fenêtre « Thèmes de la carte » s'ouvre. Indiquez le **Nom du nouveau thème** et cliquez sur **OK** (fig. 39).

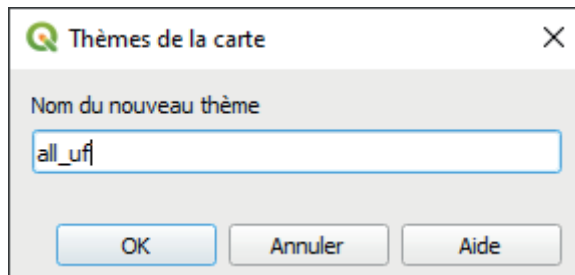



Figure 38

- Si vous cliquez de nouveau sur , le thème est maintenant disponible sous forme de **case à cocher** (fig. 39).

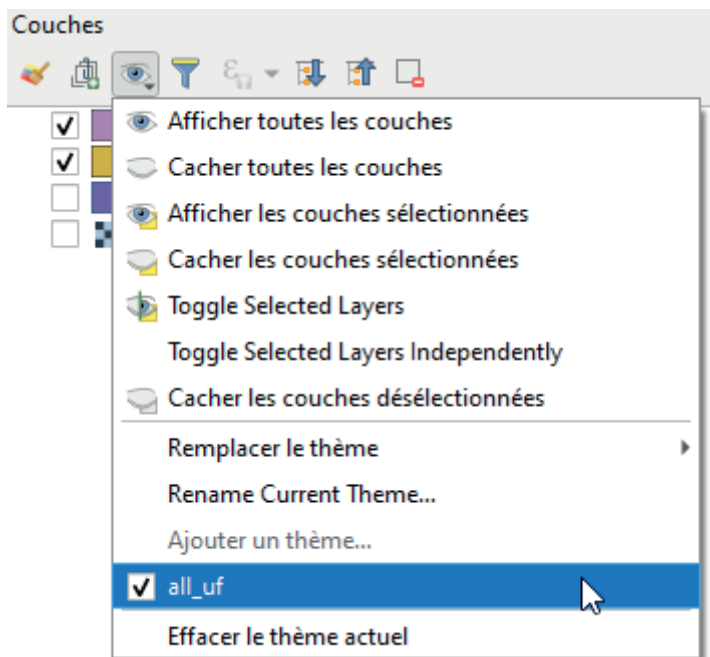


Figure 39

- Pour modifier un thème, reprenez depuis la première étape, sélectionnez **Remplacer le thème** à la place d'**Ajouter un thème...** puis choisissez le thème à actualiser dans la liste (fig. 40).

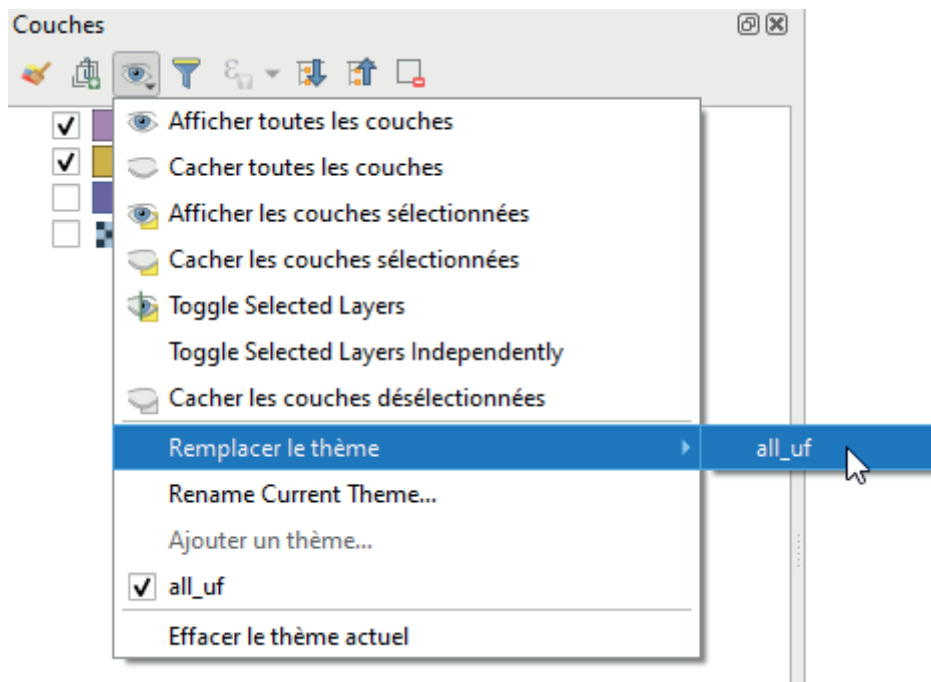





Figure 40

Identifier une couche

L'outil  identifie une couche sur le canevas, une fonction pratique lorsqu'on ne se rappelle plus à quoi correspond une entité visible.

- Cliquez sur . Le curseur de la souris devient .
- Faites un **clic droit** dans le canevas, sur la couche à identifier.
- Une liste s'affiche. Si plusieurs éléments se superposent, la liste comporte plusieurs éléments. En passant sur les différents éléments, les entités correspondantes s'allument en rouge sur le canevas. Sélectionnez l'entité à identifier (fig. 41).

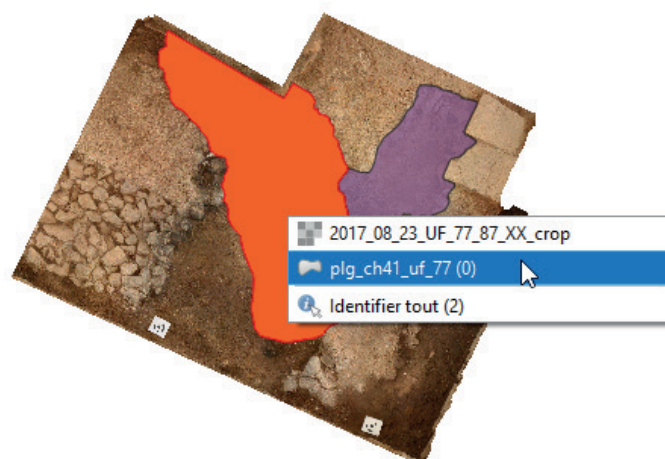


Figure 41

- Un panneau « Résultat de l'identification » s'ouvre, vous donnant le nom de l'entité ainsi que le contenu de sa table attributaire (fig. 42).

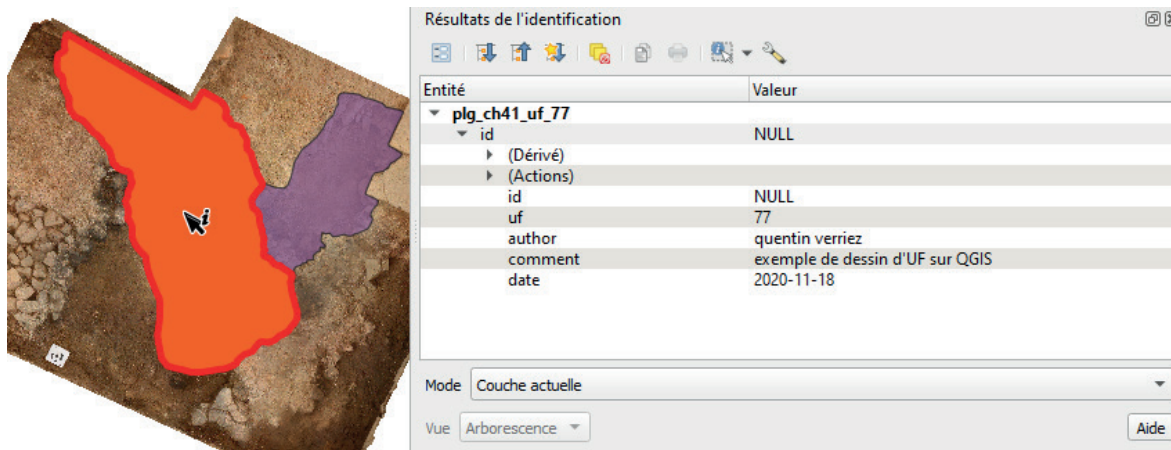


Figure 42

Déplacer le projet

Vous pouvez déplacer votre projet QGIS pour dessiner simultanément sur plusieurs machines (ordinateur ou tablette).

- Pour cela, **copiez** le dossier général de votre projet (*cf. supra* « Préparer le stockage des données ») sur un périphérique externe (clé USB, disque dur externe ou carte SD).
- **Collez-le** sur la nouvelle machine. L'emplacement du dossier général n'est pas important, mais vous ne devez pas changer les noms ou des sous-dossiers et des fichiers.
- Si QGIS n'est pas installé sur la nouvelle machine, reprenez l'étape « Installer QGIS » (*cf. supra*).
- Installez le fichier de l'interface simplifiée en reprenant l'étape « Modifier l'interface de QGIS » (*cf. supra*).
- Faites un **double-clic** sur le fichier du projet (monprojet.qgz) pour l'ouvrir.

Si vous utilisez plusieurs machines pour le même projet, définissez un ordinateur (ou une tablette) « dépôt » qui centralisera toutes les données produites.

Rassembler des données

Sans réseau ni serveur, les machines que vous utilisez pour travailler sur votre projet QGIS ne peuvent pas communiquer. Il faut donc rassembler manuellement les données produites sur différents postes. Il suffit pour cela de déplacer à nouveau votre dossier général vers l'ordinateur qui centralise les données.

- **Copiez** le dossier général de votre projet (*cf. supra* « Préparer le stockage des données ») depuis la machine productrice sur un périphérique externe (clé USB, disque dur externe ou carte SD).
- **Collez-le** vers l'ordinateur « dépôt », au même emplacement que le dossier général déjà existant.
- Une fenêtre « Remplacer ou ignorer les fichiers » s'ouvre (fig. 43). Cliquez sur **Remplacer les fichiers dans la destination**. L'ordinateur va automatiquement mettre à jour les dossiers et fichiers déjà présents et rajouter ceux nouvellement produits.

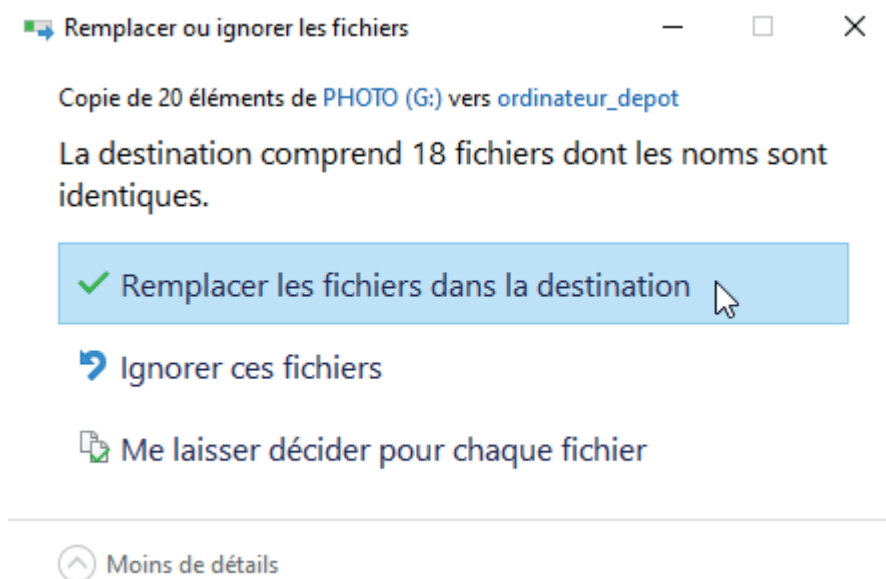


Figure 43

- Ouvrez votre projet QGIS avec un **double-clic** sur le fichier projet (monprojet.qgz). Vérifiez que les nouvelles couches que vous venez d'importer sont présentes. Sinon reprenez l'étape « importer des données » (*cf. supra*) pour les données manquantes, qui seront rangées au même endroit dans votre arborescence.

Fusionner des couches

- Cliquez sur  pour ouvrir la « Boîte à outils de traitements » (ill. 44).

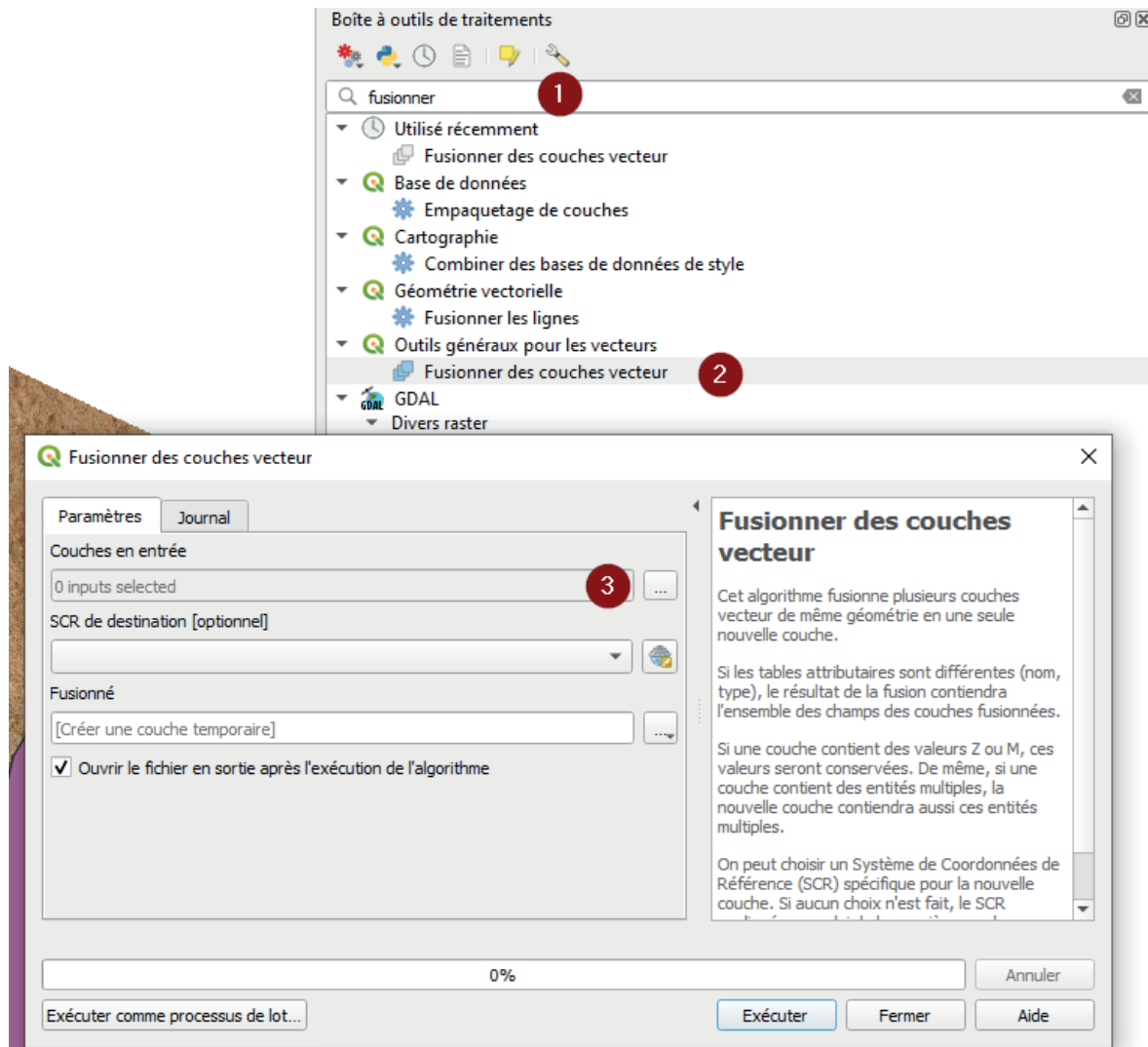



Figure 44

- **Tapez** « fusionner » dans la barre de recherche (1).
- Faites un **double-clic** sur « fusionner des couches vecteur », dans « Outils généraux pour les vecteurs » (2).
- La fenêtre « fusionner des couches vecteur » s'ouvre. Dans **Couches en entrée**, cliquez sur  (3) pour voir la liste des couches à fusionner (fig. 45).
- Sélectionnez les couches que vous souhaitez fusionner. Vous pouvez utiliser les **cases à cocher** et/ou les options **Sélectionner tout**, **Annuler la Sélection**, **Inverser la sélection**.

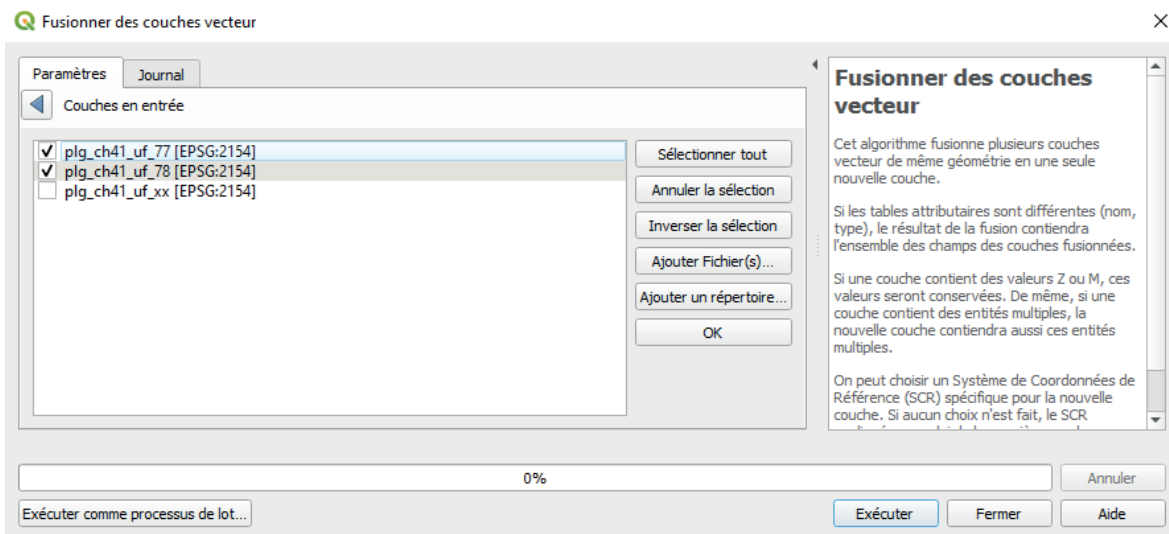


Figure 45

- **Ajouter Fichier(s)...** permet d'ajouter des Shapefile qui ne sont pas utilisés dans le projet.
- **Ajouter un répertoire...** permet d'ajouter tous les Shapefile d'un dossier.
- Cliquez sur **OK**. Cela vous renvoie vers la première fenêtre (fig. 46).

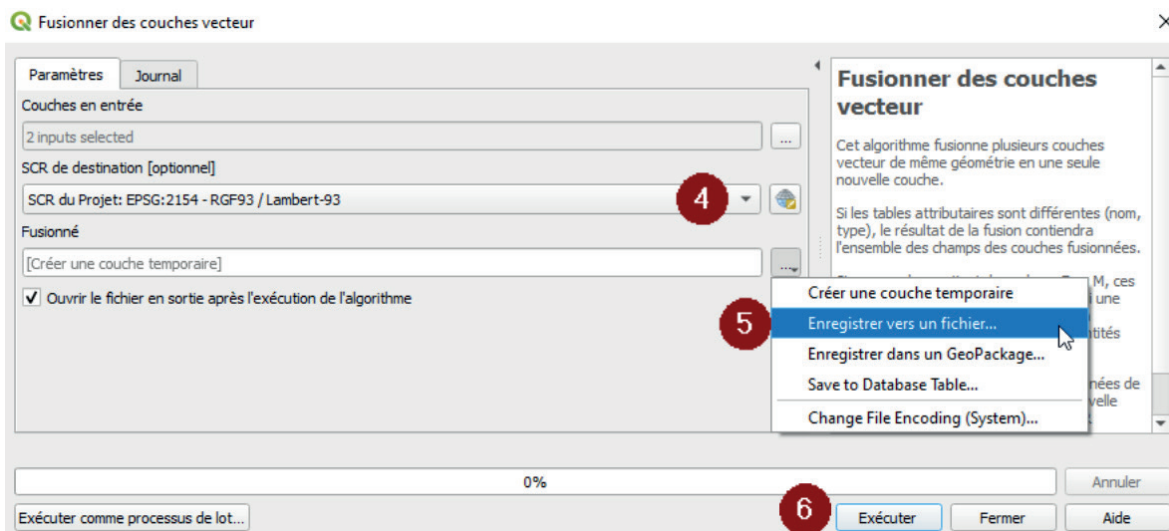


Figure 46

- Indiquez « EPSG 2154 – RGF93 / Lambert 93 » dans **SCR de destination [optionnel]** (4).

- Cliquez sur dans **Fusionné** pour indiquer le nom et l'emplacement d'enregistrement du nouveau fichier de couches fusionnées (5). Si cette case n'est pas remplie, QGIS créera une couche temporaire qui ne s'enregistrera pas à la fermeture du projet.
- Cliquez sur **Exécuter** pour lancer le processus (6).
- La nouvelle couche apparaît dans le canevas du projet.

Éditer une carte

QGIS possède un module pour éditer des cartes à partir du canevas. Cette fonction est pratique pour imprimer des plans à l'échelle.

- Créer une nouvelle carte :
 - Dans l'onglet **Projet**, cliquez sur **Nouvelle mise en page...**
 - Une fenêtre « Créer mise en page » s'ouvre. Nommez la carte et cliquez sur **OK** (fig. 47).

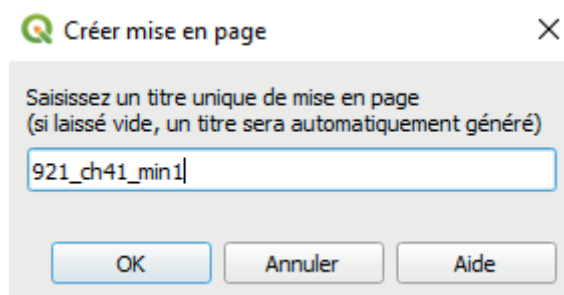


Figure 47

- Le module de mise en page s'ouvre dans une nouvelle fenêtre (fig. 48).

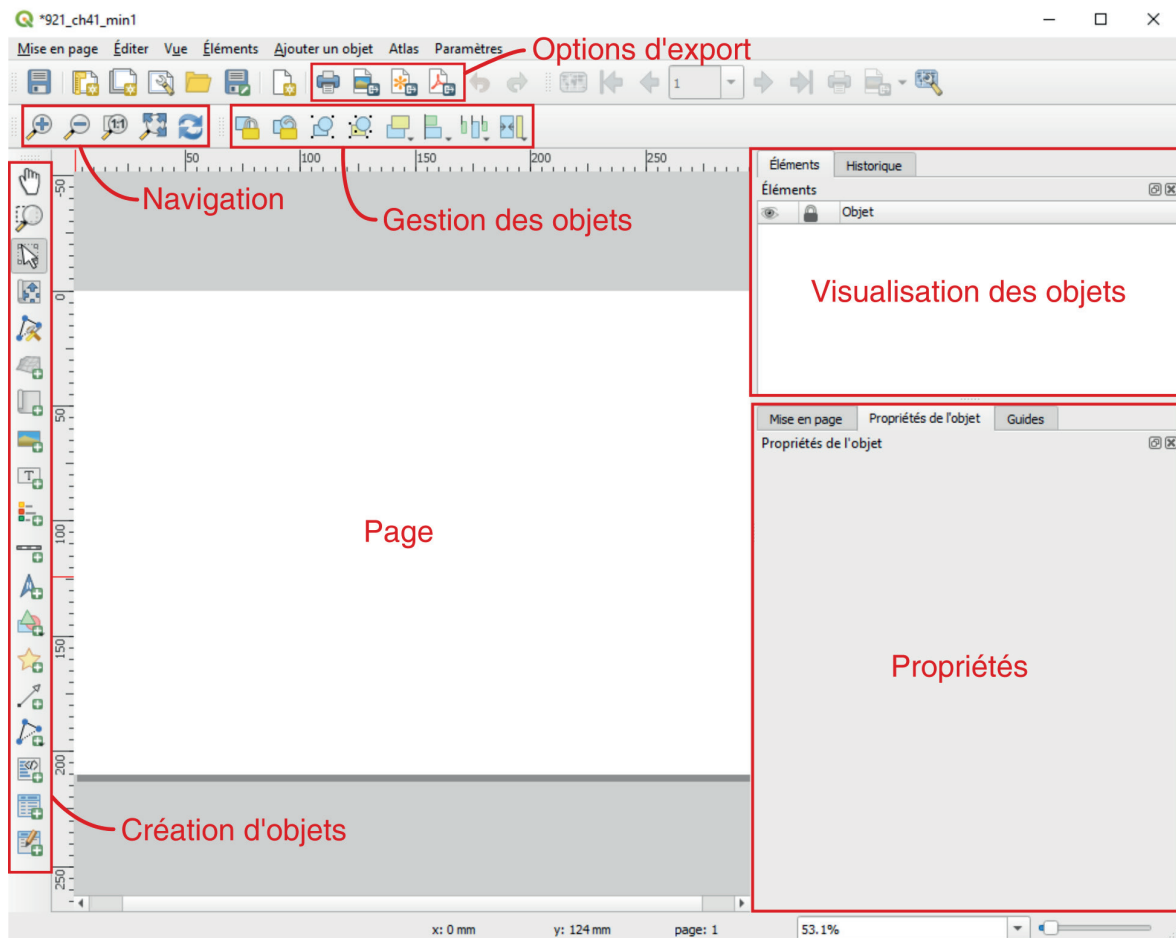


Figure 48

- Modifier la taille et la résolution de l'illustration :
 - Faites un **clic droit** sur la page et sélectionnez **Propriétés de la page....**
 - Les propriétés s'ouvrent dans le panneau **Propriétés** à droite, dans l'onglet **Propriétés de l'objet** (fig. 49).

Vous pouvez utiliser soit une taille prédéfinie (A4, A3, etc.) en portrait ou paysage, soit définir une taille personnalisée. Nous utiliserons ici les normes de la charte graphique de Bibracte : 172x220mm.

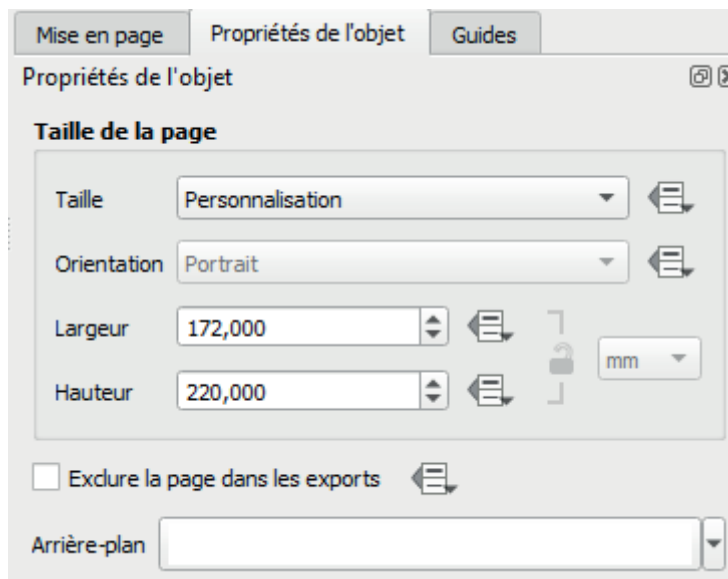







Figure 49

- Ajouter une carte :
 - Cliquez sur  et faites un **clik glissé** sur la page pour ajouter une nouvelle carte.
 - La carte, qui s'appelle par défaut « Carte 1 », est créée automatiquement. Dans QGIS, tous les éléments ajoutés sur la page s'appellent des objets.
 - Dans le panneau **Visualisation des objets**, deux **cases à cocher** sont présentes à gauche de l'objet « Carte 1 ».
 - L'œil rend visible ou invisible l'objet.
 - Le cadenas verrouille la position et la taille de l'objet.
 - Pour déplacer la carte (ou l'objet), utilisez . Pour déplacer uniquement le contenu de la carte, sans bouger le cadre, utilisez .
- Modifier une carte :
 - Les couches visibles sont identiques à celles présentes dans le canevas de QGIS. Vous pouvez retourner sur le canevas et modifier la visibilité ou la symbologie des couches. En revenant sur le module de mise en page, actualisez la carte avec .
 - Pour enrichir une carte, passez par le panneau **Propriétés**, onglet **Propriétés de l'objet**.

QGIS propose de nombreuses possibilités d'édition. Voici un exemple qui respecte les normes d'édition des illustrations de Bibracte :

- **Propriétés principales :**
 - **Échelle :** modifie l'échelle de la carte. Par exemple, indiquer 20,000 passera la carte au 1/20^{ème}.
- **Couches :**
 - **Verrouiller les couches :** coupe le lien de visibilité entre les couches du canevas et la mise en page. Si vous décochez une couche sur le canevas, cette couche restera quand même visible sur le module de mise en page (même après actualisation).
 - **Verrouiller les styles des couches :** coupe le lien de style entre les couches du canevas et la mise en page. Si vous modifiez la symbologie d'une couche sur le canevas, elle ne changera pas sur le module de mise en page (même après actualisation).
- **Grilles :**

Une grille vous permet de représenter les coordonnées X et Y sur l'emprise de votre carte.

 - Cliquez sur  pour ajouter une grille, nommée par défaut « Grille 1 ».
 - **Modifier la grille...** ouvre les propriétés de la grille (ill. 49).

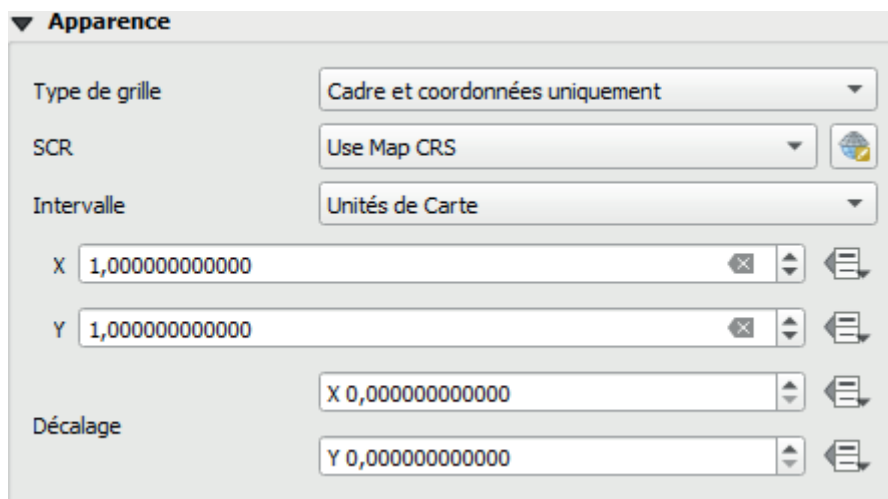
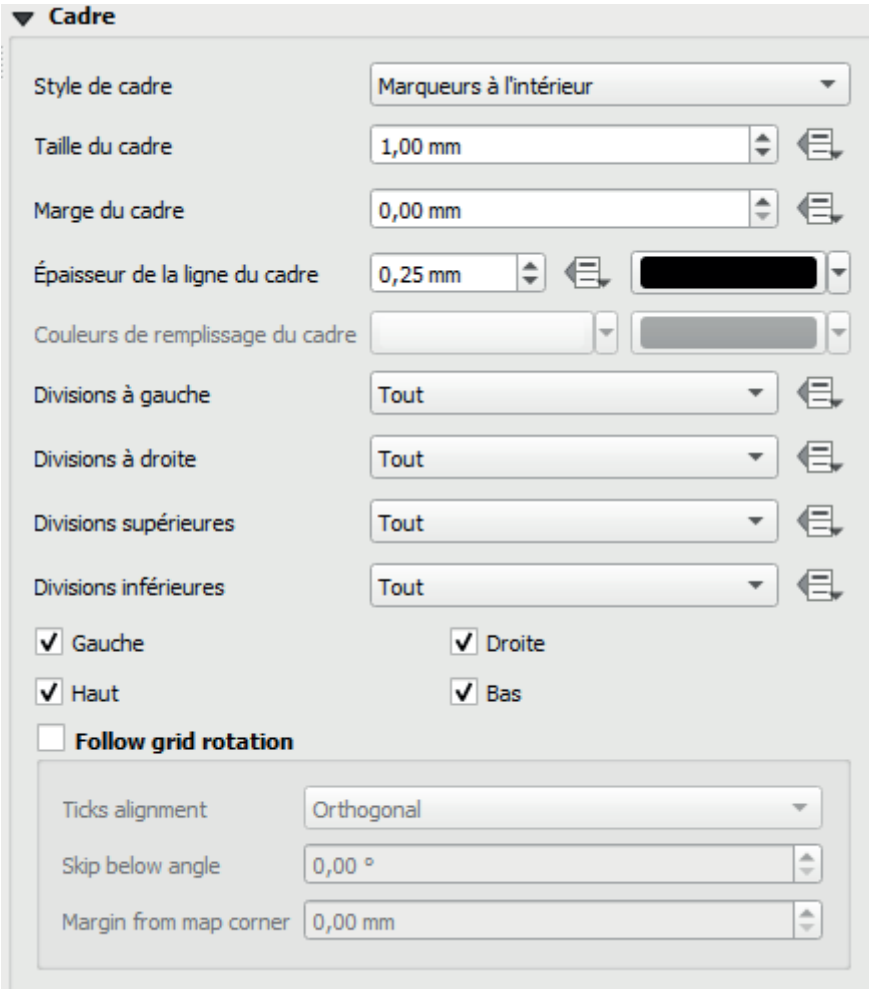


Figure 49

- **Type de grille** : type de repères pour représenter les coordonnées (lignes continues, croix aux intersections, symboles aux intersections ou marqueurs uniquement sur le cadre).

Sélectionnez **Cadre et coordonnées uniquement**.

- **X et Y** : mètres d'intervalle entre les repères. À définir en fonction de l'échelle de votre carte.
- **Décalage** : permet de décaler les repères.
- **Style de cadre** : type de représentation des repères sur le cadre de la carte. Sélectionnez **Marqueurs à l'intérieur** (fig. 50).



▼ Cadre

Style de cadre: Marqueurs à l'intérieur

Taille du cadre: 1,00 mm

Marge du cadre: 0,00 mm

Épaisseur de la ligne du cadre: 0,25 mm

Couleurs de remplissage du cadre: [] []

Divisions à gauche: Tout

Divisions à droite: Tout

Divisions supérieures: Tout

Divisions inférieures: Tout

Gauche Droite

Haut Bas

Follow grid rotation

Ticks alignment: Orthogonal

Skip below angle: 0,00 °

Margin from map corner: 0,00 mm

Figure 50

- **Épaisseur de la ligne du cadre** : modifie l'épaisseur des marqueurs. Indiquez 0,25mm.

- Cochez **Afficher les coordonnées** et suivez la configuration suivante pour les paramétrer (ill. 51) :

Afficher les coordonnées

Format: Décimal

Gauche: Tout afficher, Dans le cadre, Horizontal

Droite: Tout afficher, Dans le cadre, Horizontal

Au dessus: Tout afficher, Dans le cadre, Vertical Descending

En dessous: Tout afficher, Dans le cadre, Vertical Ascending

Police: Police

Espaceur du cadre: 1,00 mm

Précision des coordonnées: 0



Follow grid rotation

Annotations alignment: Orthogonal







Skip below angle: 0,00 °

Margin from map corner: 0,00 mm

Figure 51

- Dans **Police**, indiquez **Police Helvetica**, **Style Oblique** et **Taille 5 points**.
- Cliquez sur , puis de nouveau sur  pour retourner aux propriétés générales de l'objet.

- **Cadre :**
 - Cochez **Cadre**.
 - Modifiez l'**Épaisseur** à 0,25 point.
- Autres objets :

QGIS permet d'ajouter d'autres objets comme une échelle , une flèche nord  ou une légende . Toutefois, les éléments intégrés nativement dans le logiciel ne correspondent pas aux normes d'édition de Bibracte. Il est préférable d'exporter la carte et de rajouter les autres éléments à l'aide d'un logiciel de DAO (Inkscape ou Illustrator).
- Exporter la carte :
 - Dans le panneau **Propriété**, sélectionnez l'onglet **Mise en page**.
 - Dans **Paramètres d'export**, vous pouvez modifier la **Résolution d'exportation** de l'illustration, de base à 300dpi. Si votre illustration comporte un raster, il est conseillé de passer à 600dpi.
 - Pour conserver les propriétés d'édition des vecteurs lors de l'export, vous avez deux possibilités : exporter au format SVG (recommandé pour Inkscape) ou au format PDF (recommandé pour Illustrator).
 - Pour le SVG, cliquez sur , indiquez où vous souhaitez enregistrer le fichier et suivez les options suivantes (fig. 52) :
 - Pour le PDF, cliquez sur , indiquez où vous souhaitez enregistrer le fichier et suivez les options suivantes (ill. 53) :
 - Pour exporter un raster (.tiff, .png, .jpeg, etc.), cliquez sur , indiquez où vous souhaitez enregistrer le fichier et paramétrez les options d'export en fonction du format que vous avez sélectionné.

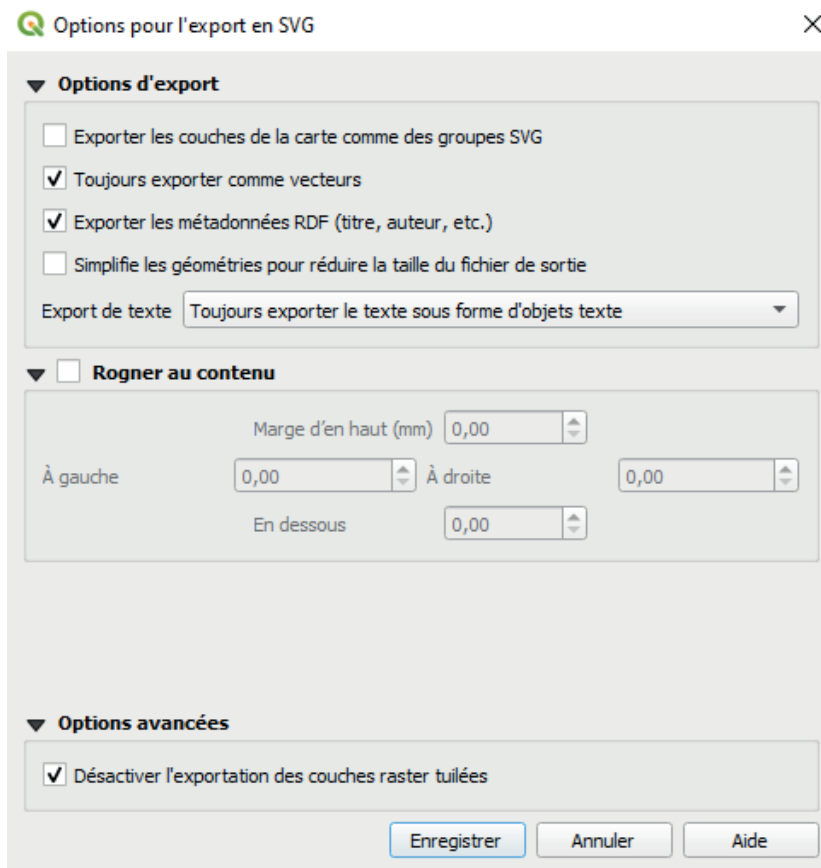


Figure 52

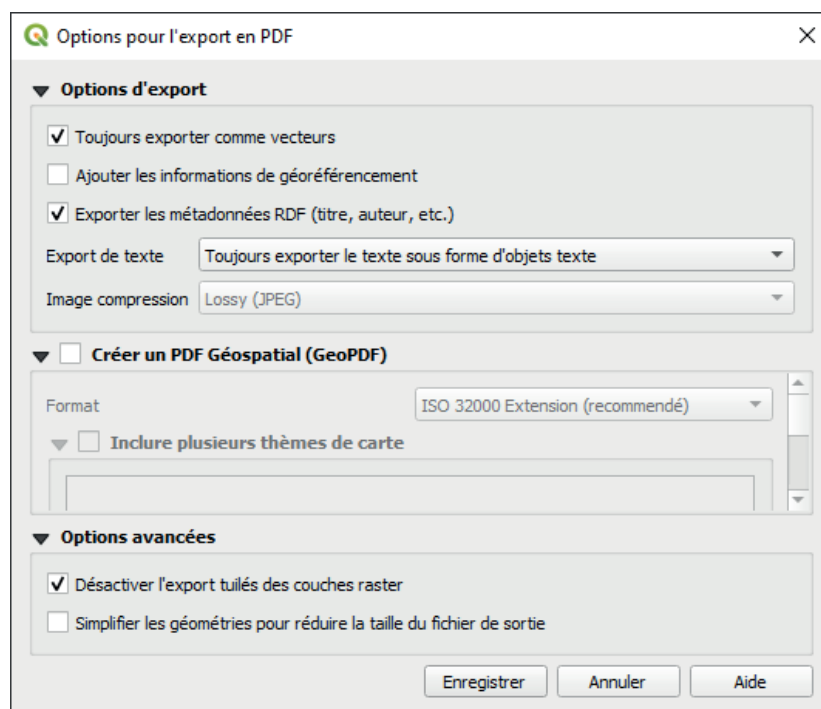


Figure 53

Annexe 5

Éléments de normalisation des métadonnées Dublin Core

Langue	ISO 639-3
Albanais	sqi
Allemand	deu
Anglais	eng
Biélorusse	bel
Bosnien	bos
Bulgare	bul
Croate	hrv
Danois	dan
Espagnol	spa
Estonien	est
Finois	fin
Français	fra
Grec moderne	ell
Hongrois	hun
Irlandais	gle
Islandais	isl
Italien	ita
Letton	lav
Lituanien	lit
Néerlandais	nld
Norvégien	nor
Norvégien	nor
Polonais	pol
Portugais	por
Roumain	ron
Russe	rus
Serbe	srp
Slovaque	slk
Slovène	slv
Suédois	swe
Suédois	swe
Tchèque	ces
Ukrainien	ukr

Aide à la création de l'élément Dublin Core *Language* : principaux langues de monde occidental et leur codification ISO 639-3.

Type	Définition	Remarque
<i>Collection</i>	Une agrégation de ressources	Une collection est décrite comme un groupe ; ses parties peuvent également être décrites séparément
<i>Dataset</i>	Données codées selon une structure définie	Les listes, les tableaux et les bases de données en sont des exemples
<i>Event</i>	Un événement non persistant dans le temps	Les exemples incluent une conférence, un atelier, une journée portes ouvertes, une bataille historique, etc.
<i>Image</i>	Une représentation visuelle autre qu'un texte	Les exemples incluent les images et les photographies d'objets physiques, les peintures, les gravures, les dessins, d'autres images et graphiques, les animations et les images animées, les films, les diagrammes, les cartes. Ce type peut comprendre des représentations électroniques et physiques.
<i>Still Image</i>	Une représentation visuelle statique	Les exemples incluent les peintures, les dessins, les conceptions graphiques, les plans et les cartes. Il est recommandé d'attribuer le type "Text" aux images de documents textuels. Il s'agit d'une sous-classe du type "Image"
<i>Moving Image</i>	Série de représentations visuelles donnant une impression de mouvement lorsqu'elles sont montrées successivement	Les exemples incluent les animations, les films, les programmes de télévision, les vidéos. Il s'agit d'une sous-classe du type "Image"
<i>Interactive Resource</i>	Une ressource nécessitant une interaction de la part de l'utilisateur pour être comprise, exécutée ou expérimentée	Il peut s'agir par exemple de formulaires sur des pages Web, d'objets d'apprentissage multimédia ou d'environnements de réalité virtuelle
<i>Physical Object</i>	Un objet ou une substance inanimé(e) et tridimensionnel(le)	Notez que les représentations numériques de ces objets, ou les substituts de ces objets, doivent utiliser "Image", "Text" ou l'un des autres types
<i>Service</i>	Un système qui assure une ou plusieurs fonctions	Il peut s'agir par exemple d'un service de photocopie, d'un service d'authentification, de prêts entre bibliothèques ou d'un serveur Web
<i>Software</i>	Un programme informatique sous sa forme source ou compilée	Il peut s'agir par exemple d'un fichier source en langage C, d'un exécutable MS-Windows ".exe" ou d'un script Python
<i>Sound</i>	Une ressource principalement destinée à être entendue	Les exemples incluent un format de fichier de lecture musicale, un disque compact audio et des paroles ou des sons enregistrés
<i>Text</i>	Une ressource composée principalement de mots à lire	Les exemples incluent les livres, les revues, les articles, les notes. Il est à noter que les fac-similés ou les images de textes font toujours partie du genre Texte

Aide à la création de l'élément Dublin Core *Type* : définition des types de données selon la norme Dublin Core (<https://www.dublincore.org/specifications/dublin-core/resource-typelist/>).

Type de document	Extension	Type MIME
Adobe Portable Document Format (PDF)	.pdf	application/pdf
Archive ZIP	.zip	application/zip
Binaire	.bin	application/octet-stream
Comma-separated values (CSV)	.csv	text/csv
GeoPackage	.gpkg	application/geopackage+sqlite3
JPEG	.jpeg .jpg	image/jpeg
JSON	.json	application/json
OpenDocument document texte	.odt	application/vnd.oasis.opendocument.text
OpenDocument feuille de calcul	.ods	application/vnd.oasis.opendocument.spreadsheet
OpenDocument présentation	.odp	application/vnd.oasis.opendocument.presentation
Polygon Model	.ply	text/plain
Portable Network Graphics	.png	image/png
Scalable Vector Graphics (SVG)	.svg	image/svg+xml
Shape File	.shp	application/vnd.shp, application/vnd.shx, application/vnd.dbf
Tagged Image File Format (TIFF)	.tif .tiff	image/tiff
Text File	.txt	text/plain
Wavefront Object File	.obj	text/plain
XML	.xml	application/xml

Pour les types MIME, voir <https://www.iana.org/assignments/media-types/media-types.xhtml>

Aide à la création de l'élément Dublin Core Format : Type MIME (Multipurpose Internet Mail Extensions) des formats de fichiers recommandés en archéologie.

LISTE DES ANNEXES

Traduction de la définition du savoir libre selon l'<i>Open Knowledge Foundation</i>.....	319
Traduction de la définition des données ouvertes.....	325
Traduction de la définition du logiciel libre (Open Source) selon l'<i>Open Source Initiative</i>	329
Protocole de relevé numérique avec le logiciel QGIS.....	333
Éléments de normalisation des métadonnées Dublin Core.....	381

BIBLIOGRAPHIE

Ackoff R. L. 1989 : From data to wisdom, *Journal of Applied Systems Analysis*, 16, p. 3-9.

ADS 1998 : *Archaeology Data Service Annual Report 1997-1998*, Archaeology Data Service, 12 p. [URL : <https://archaeologydataservice.ac.uk/library/browse/issue.xhtml?recordId=1196838>, consulté le 16 août 2022].

ADS 1999 : *Archaeology Data Service Annual Report 1998-1999*, Archaeology Data Service, 14 p. [URL : <https://www.archaeologydataservice.ac.uk/library/browse/issue.xhtml?recordId=1196840&recordType=MonographSeries>, consulté le 16 août 2022].

ADS 2000 : *Archaeology Data Service Annual Report 1999-2000*, Archaeology Data Service, 17 p. [URL : <https://archaeologydataservice.ac.uk/library/browse/issue.xhtml?recordId=1196841>, consulté le 16 août 2022].

ADS 2001 : *ADS Annual Report 2000-2001*, Archaeology Data Service, 14 p. [URL : <https://archaeologydataservice.ac.uk/library/browse/issue.xhtml?recordId=1196842>, consulté le 16 août 2022].

ADS 2002 : *ADS Annual Report 2001-2002*, Archaeology Data Service, 16 p. [URL : <https://archaeologydataservice.ac.uk/library/browse/issue.xhtml?recordId=1196843>, consulté le 16 août 2022].

ADS 2003 : *ADS Annual Report 2002-2003*, Archaeology Data Service, 14 p. [URL : <https://archaeologydataservice.ac.uk/library/browse/issue.xhtml?recordId=1196844>, consulté le 16 août 2022].

ADS 2004 : *ADS Annual Report 2003-2004*, Archaeology Data Service, 15 p. [URL : <https://archaeologydataservice.ac.uk/library/browse/issue.xhtml?recordId=1196845>, consulté le 16 août 2022].

ADS 2005 : *ADS/AHDS Annual Report 2004-2005*, Archaeology Data Service, 15 p. [URL : <https://archaeologydataservice.ac.uk/library/browse/issue.xhtml?recordId=1196846>, consulté le 16 août 2022].

ADS 2006 : *ADS/AHDS Annual Report 2005-2006*, Archaeology Data Service, 16 p. [URL : <https://archaeologydataservice.ac.uk/library/browse/issue.xhtml?recordId=1196847>, consulté le 16 août 2022].

ADS 2007 : *ADS/AHDS Annual Report 2006-2007*, Archaeology Data Service, 22 p. [URL : <https://archaeologydataservice.ac.uk/library/browse/issue.xhtml?recordId=1196848>, consulté le 16 août 2022].

ADS 2008 : *ADS Annual Report 2007-2008*, Archaeology Data Service, 21 p. [URL : <https://archaeologydataservice.ac.uk/library/browse/issue.xhtml?recordId=1196849>, consulté le 16 août 2022].

ADS 2009 : *ADS Annual Report 2008-2009*, Archaeology Data Service, 33 p. [URL : <https://archaeologydataservice.ac.uk/library/browse/issue.xhtml?recordId=1196850>, consulté le 16 août 2022].

ADS 2010 : *ADS Annual Report 2009-2010*, Archaeology Data Service, 38 p. [URL : <https://archaeologydataservice.ac.uk/library/browse/issue.xhtml?recordId=1196851>, consulté le 16 août 2022].

ADS 2011 : *ADS Annual Report 2010-2011*, Archaeology Data Service, 39 p. [URL : <https://archaeologydataservice.ac.uk/library/browse/issue.xhtml?recordId=1196851>, consulté le 16 août 2022].

ADS 2012 : *ADS Annual Report 2011-2012*, Archaeology Data Service, 38 p. [URL : <https://archaeologydataservice.ac.uk/library/browse/issue.xhtml?recordId=1196853>, consulté le 16 août 2022].

ADS 2013 : *ADS/Internet Archaeology Annual Report 2012-2013*, Archaeology Data Service, 52 p. [URL : <https://archaeologydataservice.ac.uk/library/browse/issue.xhtml?recordId=1196854>, consulté le 16 août 2022].

ADS 2014 : *ADS/Internet Archaeology Annual Report 2013-2014*, Archaeology Data Service, 52 p. [URL : <https://archaeologydataservice.ac.uk/library/browse/issue.xhtml?recordId=1196855>, consulté le 16 août 2022].

ADS 2015 : *ADS/Internet Archaeology Annual Report 2014-2015*, Archaeology Data Service, 52 p. [URL : <https://archaeologydataservice.ac.uk/library/browse/issue.xhtml?recordId=1196856>, consulté le 16 août 2022].

ADS 2016 : *ADS/Internet Archaeology Annual Report 2015-2016*, Archaeology Data Service, 55 p. [URL : <https://archaeologydataservice.ac.uk/library/browse/issue.xhtml?recordId=1196857>, consulté le 16 août 2022].

ADS 2017 : *ADS Terms of Use and Access Summary*, *Archaeology Data Service* [URL : <https://archaeologydataservice.ac.uk/advice/termsOfUseAndAccess.xhtml>, consulté le 16 juin 2022].

ADS 2018 : *ADS/Internet Archaeology Annual Report 2017-2018*, Archaeology Data Service, 40 p. [URL : <https://archaeologydataservice.ac.uk/library/browse/issue.xhtml?recordId=1196859>, consulté le 16 août 2022].

ADS 2019 : *ADS/Internet Archaeology Annual Report 2018-2019*, Archaeology Data Service, 23 p. [URL : <https://archaeologydataservice.ac.uk/library/browse/issue.xhtml?recordId=1196860>, consulté le 16 août 2022].

ADS 2020 : *ADS/Internet Archaeology Annual Report 2019-2020*, Archaeology Data Service, 24 p. [URL : <https://archaeologydataservice.ac.uk/library/browse/issue.xhtml?recordId=1196861>, consulté le 16 août 2022].

ADS 2021 : *ADS/Internet Archaeology Annual Report 2020-2021*, Archaeology Data Service, 42 p. [URL : <https://archaeologydataservice.ac.uk/library/browse/issue.xhtml?recordId=1196862>, consulté le 16 août 2022].

Akrich M., Callon M., Latour B. (dir.) 2013 : *Sociologie de la traduction : Textes fondateurs*, Paris, Presses des Mines (coll. Sciences sociales), 401 p. [URL : <http://books.openedition.org/pressesmines/1181>, consulté le 10 août 2023].

Alix S. 2000 : *Synthèse sur les prospections géophysiques réalisées au Mont Beuvray*, mémoire de DESS, Université de Bourgogne.

Alperin J. P., Barbour V., Chan L., Eve M., Gray E., Hagemann M., Joseph H., Kuchma I., McKiernan E., Prosser D., Shearer K., Shockey N., Suber P., Velterop J., Vidal I., Zhang X. 2017 : BOAI15, *Budapest Open Access Initiative* [URL : <https://www.budapestopenaccessinitiative.org/boai15-1>, consulté le 4 mai 2020].

Amnesty International, Afrewatch 2016 : « Voilà pourquoi on meurt ». *Les atteintes aux droits humains en république démocratique du congo alimentent le commerce mondial du cobalt*, Amnesty International, Observatoire africain des ressources naturelles, 73 p. [URL : <https://www.amnesty.org/fr/documents/afr62/3183/2016/fr/>, consulté le 11 septembre 2023].

Ancre 2015 : *Ressources minérales et énergie*, Alliance Nationale de Coordination de la Recherche pour l'Énergie, 65 p. [URL : https://www.allianceenergie.fr/wp-content/uploads/2021/06/2_2021_AG_RessourcesMinerales_B.pdf, consulté le 11 septembre 2023].

Anichini F., Banterle F., Buxeda i Garrigós J., Callieri M., Dershowitz N., Dubbini N., Lucendo Diaz D., Evans T., Gattiglia G., Green K., Letizia Gualandi M., Angel Herivas M., Itkin B., Madrid i Fernandez M., Miguel Gascón E., Remmy M., Richards J., Scopigno R., Vila L., Wolf L., Wright H., Zallocco M. 2020 : Developing the ArchAIDE application: A digital workflow for identifying, organising and sharing archaeological pottery using automated image recognition, *Internet Archaeology*, 52. [URL : <https://intarch.ac.uk/journal/issue52/7/index.html>, consulté le 10 août 2022].

ANR 2019a : *ANR - Modèle de PGD (français)*, Agence nationale de la recherche, 3 p. [URL : https://dmp.opidor.fr/template_export/1306665387.pdf, consulté le 22 novembre 2017].

ANR 2019b : *Règlement financier relatif aux modalités d'attribution des aides de l'ANR et valant conditions générales de ces aides*, Agence nationale de la recherche, 20 p. [URL : <https://anr.fr/fileadmin/aap/2019/ANR-RF-2019-1.pdf>, consulté le 22 octobre 2022].

Augry S. 2018 : Un espace urbain singulier : les abords de la cathédrale du Mans. Enjeux et méthodes, in Eusèbe S., Nicolas T., Gouranton V., Gaugne R. (dir.), *3ème séminaire scientifique et technique de l'Inrap*, Rennes, p. 7. [URL : <https://hal-inrap.archives-ouvertes.fr/hal-02126949>, consulté le 25 juillet 2022].

Ayache L., Chaillou A., Durost S., Guichard V., Moreau R., Verriez Q. 2017 : BibNum : un nouveau projet de développement numérique pour Bibracte, in Guichard V., *Rapport intermédiaire 2017 du programme quadriennal de recherche 2017-2020 sur le Mont Beuvray*, Glux-en-Glenne, Bibracte EPCC, p. 23-27. [URL : <https://shs.hal.science/halshs-01671478>, consulté le 3 octobre 2023].

Ayache L., Chaillou A., Durost S., Guichard V., Moreau R., Verriez Q. 2018 : Bibracte Numérique : projet de développement numérique, in Guichard V., *Rapport intermédiaire 2018 du programme quadriennal de recherche 2017-2020 sur le Mont Beuvray*, Glux-en-Glenne, Bibracte EPCC, p. 31-35. [URL : <https://hal.science/BIBRACTE/halshs-02001397v1>, consulté le 3 octobre 2023].

Barge O., Rodier X., Davtian G., Saligny L. 2004 : L'utilisation des systèmes d'information géographique appliquée à l'archéologie française, *ArchéoSciences, revue d'Archéométrie*, 28-1, p. 15-24.

Barker P. 1993 : *Techniques of archaeological excavation*, London, Routledge, 211 p.

Barral P., Beck P., Bernal J., Boyer F., Buchsenschutz O., Flouest J.-L., Laszlovszky J., Luginbühl T., Paratte C.-A., Paunier D., Quinn D., Ralston I. B. M., Szabó M., Vitali D., Wiethold J. 1995 : Les fouilles du Mont-Beuvray, Nièvre, Saône-et-Loire. Rapport biennal 1992-1993, *Revue Archéologique de l'Est et du Centre-Est*, 2-46, p. 217-293.

Bartoš K., Pukanská K., Sabová J., Weiss G. 2014 : The Application of Open-Source and Free Photogrammetric Software for the Purposes of Cultural Heritage Documentation, *Geo-Science Engineering*, 60-2, p. 19-26.

Bel V., Rochette M., Thomas B., Vacheret A. 2021 : *L'enregistrement des données sur tablettes avec le logiciel Isyslight : l'expérience de la fouille des Berges de la Robine à Narbonne*, Maison de l'Asie, Paris [URL : <https://www.resefe.fr/fr/node/4570>, consulté le 17 janvier 2023].

Belarbi M., Raymond P., Saulière N., Touquet R. 2012 : L'acquisition 3D par photogrammétrie en archéologie, *Archéopages : archéologie & société*, 34, p. 90-91.

Bellan G. 2010 : 25 ans de rapports archéologiques : auto-édition et publication, in *Le rapport de fouille archéologique : réglementation, conservation, diffusion*, Paris, De Boccard/Maison Archéologie & Ethnologie René-Ginouvès (coll. Travaux, 11), p. 129-135.

Benabou V.-L., Farchy J. 2007 : *La mise à disposition ouverte des oeuvres de l'esprit*, Conseil supérieur de la propriété littéraire et artistique, 47 p. [URL : <https://www.culture.gouv.fr/content/download/102481/file/Rapport%20CSPLA%20mise%20%C3%A0%20dispo%20ouverte%20des%20oeuvres.pdf?inLanguage=fr-FR>].

Bernard L., Jullien P. 2022 : *Fichier d'importation standard et fichier descriptif des champs pour ArkeoGIS 4.7* [URL : <https://nakala.fr/10.34847/nkl.ede43qsi>, consulté le 21 août 2023].

Berners-Lee T. 1997 : Axioms of Web Architecture: Metadata, W3C [URL : <https://www.w3.org/DesignIssues/Metadata.html>, consulté le 9 mai 2020].

Bernier S., Duthoit S., Ladet S., Baudet D. 2014 : Les concepts de base des Systèmes d'Information Géographique (SIG) : les données et les fonctions générales, *Cahier des Techniques de l'INRA*, p. 19-27.

Bizer C., Heath T., Berners-Lee T. 2009 : Linked Data - The Story So Far, *International Journal on Semantic Web and Information Systems (IJSWIS)*, 5-3, p. 1-22.

BnF 2008 : *Guide d'utilisation du Dublin Core (DC) à la BnF : Dublin Core simple et Dublin Core qualifié, avec indications pour utiliser le profil d'application de TEL*, Bibliothèque nationale de France, 16 p. [URL : <https://www.bnf.fr/fr/dublin-core>, consulté le 18 mai 2020].

Boldo D. 2003 : Fabrication de vraies ortho-images et correction des ombres, *Bulletin d'information scientifique et technique de l'IGN*, n° 74, p. 25-36.

Bolo A., De Muylder M., Font C. 2014 : De la tablette PC à la cartographie de terrain : exemple de méthodologie sur le chantier d'archéologie préventive de Noyon (Oise), in Costa L., Djindjian F., Giligny F. (dir.), *Actes des 3e journées d'informatique et Archéologie de Paris*, Florence, Archeologia e Calculatori (coll. Supplemento), p. 247-256. [URL : https://www.europeana.eu/fr/item/224/A_C_oai_Archive_sup_xml_795, consulté le 24 juin 2022].

Bordage F. 2019a : *Empreinte environnementale du numérique mondial*, report, Green IT, 40 p. [URL : https://www.greenit.fr/wp-content/uploads/2019/10/2019-10-GREENIT-etude_EENM-rapport-accessible.VF_.pdf, consulté le 11 septembre 2023].

Bordage F. 2019b : *Sobriété numérique : Les clés pour agir*, Paris, Buchet-Chastel, 208 p.

Bordage F., de Montenay L., Vergeynst O. 2021 : *Étude « iNUM » : impacts environnementaux du numérique en France*, Green IT, 21 p. [URL : <https://www.greenit.fr/wp-content/uploads/2021/02/2021-01-iNum-etude-impacts-numerique-France-rapport-0.8.pdf>, consulté le 17 mars 2021].

Borgman C. L. 2007 : *Scholarship in the Digital Age: Information, Infrastructure, and the Internet*, Cambridge, MA, USA, MIT Press, 360 p.

Boulétreau V., Habert B. 2014 : Les formats, in Sinatra M. E., Vitali-Rosati M. (dir.), *Pratiques de l'édition numérique*, Montréal, Presses de l'Université de Montréal [URL : <http://parcoursnumeriques-pum.ca/1-pratiques/https://www.parcoursnumeriques-pum.ca/1-pratiques/chapitre9.html>, consulté le 19 septembre 2023].

Bouriga S. 2020 : *Papier-crayon vs. écran-clavier : Effets sur le coût cognitif et sur la dynamique de la production de textes*, phdthesis, Université de Poitiers [URL : <https://theses.hal.science/tel-03795276>, consulté le 10 juillet 2023].

Bryas E., Bellan G., Speller A., Carpentier C. 2013 : Dolia, le système documentaire de l'Inrap, *Archéopages : archéologie & société*, 37, p. 72-81.

Bryas E., Carpentier C. 2012 : ACE Case Study: Inrap: Archival Preparation of the ArchéoDB field registration system, *Archaeology Data Service / Digital Antiquity* [URL : https://guides.archaeologydataservice.ac.uk/g2gp/DbSht_ACECS3-Inrap, consulté le 2 août 2022].

Buchsenschutz O., Batardy C. 2022 : *Bourges : un quartier de bijoutiers au Ve s. avant J.-C.*, *Bourges : un quartier de bijoutiers au Ve s. avant J.-C.* [URL : <https://chronocarto.eu/spip.php?article90&lang=fr>, consulté le 24 août 2023].

Buchsenschutz O., Cartereau M., Lambert G.-N. 1986 : Centrar : un système de gestion et de représentation graphique des données archéologiques, *Bulletin de la Société préhistorique française*, 83-10, p. 325-327.

Buchsenschutz O., Gruel K. 1994 : Le relevé de terrain en archéologie : le système Arkéoplan, *Histoire & Mesure*, 9-3, p. 231-237.

Buchsenschutz O., Richard H. 1996 : *L'environnement du Mont Beuvray*, Glux-en-Glenne, Centre archéologique européen du Mont Beuvray (coll. Bibracte, 1), 207 p.

Bulliot J.-G. 1870 : Fouilles de Bibracte, 1868 (suite), *Revue Archéologique*, XXI, p. 153-169.

Bulliot J.-G. 1883 : Fouilles du Mont-Beuvray. Le Parc aux Chevaux (suite). Maison n° 2, *Mémoires de la Société éduenne*, 12, p. 93-123.

Bulliot J.-G. 1899 : *Fouilles du Mont Beuvray (ancienne Bibracte) de 1867 à 1895*, Autun, 250 p.

Bulliot J.-G., Thiollier F., Thiollier N. 1899 : *Fouilles du Mont Beuvray. Ancienne Bibracte*, Saint-Étienne.

Burrough P. A. 1986 : *Principles of geographical information systems for land resources assessment*, Oxford, Clarendon Press (coll. Monographs on soil and resources survey, 12), 193 p. [URL : <https://archive.org/details/principlesofgeog00burr>, consulté le 26 juin 2021].

C3DC 2011 : *Guide pour la prise de vue photogrammétrique à destination de la plateforme C3DC*, Culture 3D Cloud, 35 p. [URL : http://www.tapenade.gamsau.archi.fr/Culture3D-Cloud_UsersGuideline.pdf, consulté le 9 décembre 2022].

Cacaly S. 1989 : Les banques de données de la recherche en archéologie: aperçu historique et problématique, *BRISES. Bulletin de recherches sur l'information en sciences économiques humaines et sociales*, 2-15, p. 147-153.

Cardon D. 2019 : *Culture numérique*, Paris, SciencesPo les presses (coll. Les petites humanités), 428 p.

CDL 2004 : *The UC Libraries' Digital Preservation Program : aims, overview, and initial priorities*, California Digital Library, 9 p. [URL : <https://libraries.universityofcalifornia.edu/wp-content/uploads/2012/11/newpresrep.pdf>, consulté le 22 octobre 2022].

Chaillou A. 2003 : *Nature, statut et traitements informatisés des données en archéologie : les enjeux des systèmes d'informations archéologiques*, Université Lumière - Lyon II, 270 p. [URL : <https://tel.archives-ouvertes.fr/tel-00137986>, consulté le 18 mai 2022].

Chaillou A. 2004 : Présentation de bdB, base de données utilisée à BIBRACTE, centre archéologique européen., *Cahier des thèmes transversaux ArScAn*, IV, p. 130-133.

Chang C.-Y., Tijus C., Zibetti E. 2015 : Les apprentissages à l'heure des technologies cognitives numériques, *Administration & Éducation*, 146-2, p. 91-98.

Chéné A., Réveillac G., Foliot P. 1999 : *La pratique de la photographie en archéologie*, Aix-en-Provence, Edisud (coll. Travaux du Centre Camille Jullian, 18), 143 p.

Clarke D. L. 1977 : *Spatial archaeology*, London, New York, San Francisco, Academic Press, 386 p.

CNRS 2019 : *Feuille de route du CNRS pour la science ouverte*, CNRS, 15 p.

cOAlition S 2021 : *Plan S - Macking full & immediate open access a reality*, cOAlition S, 3 p. [URL : https://www.coalition-s.org/wp-content/uploads/2021/03/Plan-S_profile_March2021.pdf, consulté le 23 avril 2023].

Collet P. 1996 : La photographie et l'archéologie : des chemins inverses, *Bulletin de Correspondance Hellénique*, 120-1, p. 325-344.

Conlon V. M. 1973 : *Camera techniques in archaeology*, London, J. Baker, 200 p. [URL : <http://archive.org/details/cameratechniques00conl>, consulté le 20 avril 2023].

Conolly J., Lake M. 2006 (1^{re} éd.) : *Geographical Information Systems in Archaeology*, Cambridge, Cambridge University Press (coll. Cambridge manuals in archaeology), 338 p. [URL : <https://www.cambridge.org/core/product/identifier/9780511807459/type/book>, consulté le 4 mai 2023].

Cookson M. B. 1954 : *Photography for archaeologists*, London : Parrish, 152 p. [URL : <http://archive.org/details/photographyforar0000cook>, consulté le 17 avril 2023].

Costa L., Desachy B. 2020 : L'approche de l'espace par les archéologues : des pratiques en évolution., *Histoire de la recherche contemporaine. La revue du Comité pour l'histoire du CNRS*, Tome IX-n°2, p. 142-152.

Dabas M., Music B., Laudrin F. 1997 : Prospections géophysiques sur le Mont Beuvray, in Guichard V. (dir.), *Rapport annuel d'activité scientifique 1997 du Centre archéologique européen du Mont Beuvray*, Glux-en-Glenne, Bibracte - Centre archéologique européen, p. 199-210.

Dabas M., Riahi Y. 2018 : Penser l'espace du terrain à la toile : d'Arkéoplan à Chronocarto, in Hiriart E., Genechesi J., Cicolani V., Martin S., Nieto-Pelletier S., Olmer F. (dir.), *Monnaies et archéologie en Europe celtique. Mélanges en l'honneur de Katherine Gruel*, Glux-en-Glenne, Bibracte (coll. Bibracte, 29), p. 75-80. [URL : <https://shs.hal.science/halshs-02068149>, consulté le 11 avril 2023].

DAC, BRC, NeSC 2005 : *Large-scale data sharing in the life sciences: data standards, incentives, barriers and funding models (The Joint Data Standards Study)*, The Digital Archiving Consultancy, The Bioinformatics Research Centre, University of Glasgow, The National e-Science Centre, 192 p. [URL : <http://hdl.handle.net/11343/28768>, consulté le 11 avril 2022].

Daston L., Galison P. L. 2012 : *Objectivité*, Dijon, les Presses du réel (coll. Fabula), 475 p.

De Ferrières M., Bouhot G., Bouillot R. 2015 : *Éléments de technologie pour comprendre la photographie : argentique & numérique*, Paris, Eyrolles, 352 p.

Déchelette J. 1908 : Rapport sur les fouilles du Mont-Beuvray, *Bulletin de la Société Française des Fouilles Archéologiques*, 2, p. 56-65.

Defer C. 2021 : La *domus* PC2, in Guichard V., *Rapport annuel 2020 du programme quadriennal de recherche 2017-2020 sur le Mont Beuvray*, Glux-en-Glenne, Bibracte EPCC, p. 147-168. [URL : <https://hal.science/BIBRACTE/halshs-02512035v1>, consulté le 3 octobre 2023].

Desachy B. 1990 : Le traitement formalisé des données en archéologie, *Brises - Bulletin de recherches sur l'information en sciences économiques, humaine et sociales*, 2-15, p. 21-36.

Desachy B. 2004 : Le sériographe EPPM : un outil informatisé de sériation graphique pour tableaux de comptages, *Revue archéologique de Picardie*, 3-1, p. 39-56.

Desachy B. 2008 : *De la formalisation du traitement des données stratigraphiques en archéologie de terrain*, Thèse, Université Panthéon-Sorbonne - Paris I [URL : <https://tel.archives-ouvertes.fr/tel-00406241>, consulté le 21 avril 2020].

Desachy B., Durost S., Giudicelli M., Guichard V. 2012 : Rapports scientifiques de fouilles programmées et publications de Bibracte. Réflexions, acquis et perspectives, *Les nouvelles de l'archéologie*, 130, p. 58-64.

Desprès-Lonnet M. 2013 : La photographie de travail dans les recherches en archéologie : un objet hybride entre prise de note informelle et compte rendu normé, *Sciences de la société*, 89, p. 74-94.

Dessus P. 2015 : Pour tout résoudre, cliquez ici, *Distances et médiations des savoirs. Distance and Mediation of Knowledge*, 9. [URL : <https://journals.openedition.org/dms/963>, consulté le 11 septembre 2023].

Deweirdt E. 2010 : *De l'analyse spatiale à la caractérisation de sites de la fin de l'âge du fer et du début de l'époque Gallo-Romaine dans le Nord et l'Est de la Gaule*, Thèse, Université de Gand, Université de Bourgogne, 377 p.

Di Cosmo R., Zacchioli S. 2017 : Software Heritage: Why and How to Preserve Software Source Code, in *iPRES 2017*, Kyoto, Japan, p. 1-10. [URL : <https://hal.archives-ouvertes.fr/hal-01590958>, consulté le 6 avril 2020].

Dietrich D., Adamus T., Miner A., Steinhart G. 2012 : De-Mystifying the Data Management Requirements of Research Funders, *Issues in Science and Technology Librarianship*. [URL : <https://journals.library.ualberta.ca/istl/index.php/istl/article/view/1556/1467>, consulté le 21 avril 2022].

DINUM 2015 : *Référentiel général d'interopérabilité - Standardiser, s'aligner et se focaliser pour échanger efficacement*, Direction Interministérielle du Numérique et du Système d'Information et de Communication de l'Etat, 84 p. [URL : <https://www.numerique.gouv.fr/publications/interopabilite/>, consulté le 18 février 2021].

Djindjian F. 2016 : Archéologie, de l'analogique au numérique : évolution technique ou révolution méthodologique ?, *Les nouvelles de l'archéologie*, 146, p. 6-11.

Djindjian F. 2017 : *L'archéologie : théorie, méthodes et reconstitutions*, Malakoff, Armand Colin, 612 p.

Donnelly M., Jones S. 2009 : *Data Management Plan Content Checklist. Draft Template for Consultation*, Data Curation Center, 5 p. [URL : https://www.dcc.ac.uk/sites/default/files/documents/resource/DMP/DMP_checklist_consultation-June-09.pdf, consulté le 22 novembre 2022].

Doorn P., Ronzino P. 2022 : ARIADNEplus Data Management Plan Tools, *ARIADNEplus* [URL : <https://vast-lab.org/dmp/index.html>, consulté le 28 juin 2022].

Ducomet G., Hesse A. 1985 : Compte rendu de la prospection électrique à la Caume Chaudron, in Peyre C., Buchsenschutz O., Beck F., Gruel K., Guillaumet J.-P., Duval A., Perichon R., Ducomet G., Hesse A., *Fouille du Mont Beuvray, campagne 1985*, Glux-en-Glenne, Bibracte - Centre archéologique européen, p. 104-111. [URL : <https://halshs.archives-ouvertes.fr/halshs-01299987/document>, consulté le 22 janvier 2018].

Dufay A. 1995 : La rue du Parc aux Chevaux, secteur 15, in Flouest J.-L., *Document final de synthèse, rapport triennal 1993-1995*, Glux-en-Glenne, Bibracte - Centre archéologique européen, p. 1-27. [URL : <https://hal.science/BIBRACTE/halshs-01299388v1>, consulté le 4 octobre 2023].

Ebrahim M. A.-B. 2003 : *Archeological photogrammetry and world heritage documentation*, 42 p. [URL : https://www.researchgate.net/publication/267037880_ARCHAEOLOGICAL_PHOTOGAMMETRY_AND_WORLD_HERITAGE_DOCUMENTATION?channel=doi&linkId=5442c1000cf2e6f0c0f93749&showFulltext=true, consulté le 16 février 2023].

Egels Y. 2011 : La photogrammétrie, principes généraux et terminologie, *Images et modèles 3D en milieux naturels*, 12-1, p. 41-50.

Ellul J. 2004 : *Le système technicien*, Paris, Le Cherche Midi, 337 p. [URL : <http://archive.org/details/lesystemetechnic0000ellu>, consulté le 9 août 2023].

ERC 2017 : *ERC Data Management Plan template*, European Research Council, 5 p. [URL : <https://erc.europa.eu/content/erc-data-management-plan-template>, consulté le 28 avril 2022].

ERC Scientific Council 2018 : *Open Research Data and Data Management Plans, Information for ERC grantees*, European Research Council, 23 p. [URL : <https://erc.europa.eu/content/open-research-data-and-data-management-plans-information-erc-grantees>, consulté le 28 avril 2022].

Esnault E. 2019 : Les nouvelles techniques de relevés appliquées à l'archéologie du bâti et à l'archéologie urbaine, in Eusèbe S., Nicolas T., Gouranton V., Gaugne R. (dir.), *Archéologie : imagerie numérique et 3D : actes du 3 e séminaire scientifique et technique de l'Inrap, 26-27 juin 2018, Rennes.*, Rennes, Open Edition [URL : <https://hal-inrap.archives-ouvertes.fr/hal-02020092>, consulté le 3 avril 2023].

Etalab 2017 : *LICENCE OUVERTE / OPEN LICENCE Version 2.0*, Etalab, 4 p. [URL : <https://www.etalab.gouv.fr/licence-ouverte-open-licence/>, consulté le 21 octobre 2023]

European Commission 2016 : *Guidelines on FAIR Data Management in Horizon 2020*, European Commission, 12 p. [URL : https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf, , consulté le 4 mai 2022].

Eusèbe S. 2019 : Imagerie numérique et représentation des données en archéologie, *In Situ. Revue des patrimoines*, 39. [URL : <https://journals.openedition.org/insitu/21467>, consulté le 2 août 2022].

Eusèbe S., Nicolas T., Gouranton V., Gagne R. (dir.) 2019 : *Archéologie : Imagerie numérique et 3D. Actes du 3e séminaire scientifique et technique de l'Inrap, 26-27 juin 2018, Rennes*, Rennes, Open Edition. [URL : <https://sstinrap.hypotheses.org/category/seminaires-st/imagerie-numerique-et-3d>, consulté le 27 janvier 2023].

Evans T., Gilham J. 2021 : *OASIS V User Manual*, OASIS, 30 p. [URL : https://docs.google.com/document/d/1vNnhFrgoURyOM5SLaKB_1JIHI6DjV4anykzVMKJPj_k/edit?usp=sharing, consulté le 10 août 2022].

Ferreboeuf H. 2018 : *LEAN ICT. Pour une sobriété numérique*, The Shift Project, 38 p. [URL : <https://theshiftproject.org/article/pour-une-sobriete-numerique-rapport-shift/>, consulté le 11 septembre 2023].

Ferreboeuf H., Efoui-Hess M., Verne X. 2021 : *Impact environnementale du Numérique : tendances à 5 ans et gouvernance de la 5G*, Note d'analyse, The Shift Project, 46 p. [URL : <https://theshiftproject.org/article/impact-environnemental-du-numerique-5g-nouvelle-etude-du-shift/>, consulté le 14 avril 2021].

Fochesato A. 2020 : *Aux origines de la ville européenne : technologie, typologie et fonction de l'architecture en bois dans l'habitat urbain des oppida celtiques (IIe-IIe s. av. J.-C.) : le cas de Bibracte, mont Beuvray (France)*, Thèse de doctorat, Bourgogne Franche-Comté, 457 p. [URL : <https://www.theses.fr/2020UBFCH002>, consulté le 19 juin 2023].

Francesconi A. 2023 : *BIMP. Batch Image Manipulation Plugin for GIMP*, *BIMP. Batch Image Manipulation Plugin for GIMP* [URL : <https://alessandrofrancesconi.it/projects/bimp/>, consulté le 31 mars 2023].

Friedt J.-M. 2023 : Exploitation des signaux de référence de navigation par satellite pour un positionnement centimétrique : RTKLIB fait appel à Centipède et l'IGN pour afficher dans QGIS, *Hackable*, 48, p. 110-115.

Gaime S. 2018 : L'imagerie 3D en archéologie : de l'enregistrement de terrain à l'aide aux relevés, in Eusèbe S., Nicolas T., Gouranton V., Gagne R. (dir.), *3ème séminaire scientifique et technique de l'Inrap*, Rennes [URL : <https://hal-inrap.archives-ouvertes.fr/hal-02124883>, consulté le 2 août 2022].

Gardin J. C. 1976 : *Code pour l'analyse des formes de poteries*, Paris, Éditions CNRS (coll. Analyse documentaire et calcul en archéologie), 116 p.

Gardin J.-C. 1979 : *Une archéologie théorique*, Paris, Hachette (coll. L'esprit critique), 339 p.

Garenne X. 1867 : *Bibracte*, Autun, L. Duployer, 228 p.

Getler M., Sisu D., Jones S., Miller K. 2014 : DMPonline Version 4.0: User-Led Innovation, *International Journal of Digital Curation*, 9-1, p. 193-219.

Giligny F. 2009 : De la fouille à l'interprétation : le traitement des données, in Demoule J.-P., Giligny F., Lehoërff A., Schnapp A., *Guide des méthodes de l'archéologie*, Paris, la Découverte (coll. Grands repères), p. 142-199.

Giligny F. 2011 : Informatique et archéologie : une révolution tranquille ?, in *Les historiens et l'informatique. Un métier à réinventer*, Rome, p. 189-198.

Giligny F., Desachy B. 2019 : Informatique et archéologie en France : les années 1980, *Les nouvelles de l'archéologie*, 157-158, p. 114-118.

Girod L. 2018 : *Improved measurements of cryospheric processes using advanced photogrammetry*, Thèse de doctorat, Universitetet i Oslo, 144 p. [URL : <https://www.duo.uio.no/handle/10852/62195>, consulté le 10 juin 2021].

Goudineau C. 2007 : *Regard sur la Gaule*, Arles, Actes Sud (coll. Babel, 836), 576 p.

Granier X., Chayani M., Abergel V., Benistant P., Bergerot L., Bernard J.-F., Bohbot H., Leroy Du Cardonnoy E., Cassen S., de Luca L., Dutailly B., Epaud F., Espinasse L., Eusèbe S., Flammin A., Fleury P., Grimaud V., Jeanson L., Joffres A., Kerouanton J.-L., Laroche F., Le Mort F., Lefèvre N., Lescop L., Madeleine S., Manséri H., Marlet O., Vergnieux R., Mathieu V., Mora P., Morineau C., Paitier H., Pamart A., Pouyet T., Quantin M., Rodier X., Sanz Laliberté S., Schmitt A., Sorin S., Tournon-Valiente S. 2019 : *Les recommandations du Consortium 3D SHS*, CNRS, 204 p. [URL : <https://hal.archives-ouvertes.fr/hal-01683842>, consulté le 23 juin 2022].

Greene K., Moore T. 2010 : *Archaeology: an introduction*, Milton Park, Abingdon, Oxon ; New York, Routledge, 394 p.

Gruel K., Buchsenschutz O. 1994 : De l'usage d'Arkéoplan pour l'enregistrement d'éléments en élévation : étude d'un mur du couvent du Mont Beuvray (Bourgogne), *Le médiéviste et l'ordinateur*, 29-1, p. 12-18.

Guillaumet J.-P. 1990 : *Base archéologique du Mont Beuvray, rapport sur l'opération Bibracte/Mont Beuvray de 1984 à 1989*, Bibracte - Centre archéologique européen, 190 p. [URL : <https://halshs.archives-ouvertes.fr/halshs-01299920/document>, consulté le 22 janvier 2018].

Guillaumet J.-P. 1996 : *Bibracte. Bibliographie et plans anciens*, Paris, Éditions de la Maison des Sciences et de l'Homme (coll. Documents d'Archéologie Française , 57), 170 p.

Guimarães N., Pádua L., Adão T., Hruška J., Peres E., Sousa J. J. 2020 : Viswebdrone: A web application for uav photogrammetry based on open-source software, *ISPRS International Journal of Geo-Information*, 9-11, p. 16.

Hantrais J. 2017 : *L'occupation du secteur des terrasses PC14 et PC15 de Bibracte à la fin de l'âge du Fer. Confrontation des données chronologiques, spatiales et fonctionnelles*, Mémoire de Master 2, Université de Franche-Comté.

Hardman C. 2009 : The Online Access to the Index of Archaeological Investigations (OASIS) Project: Facilitating Access to Archaeological Grey Literature in England and Scotland, *The Grey Journal*, 5-2, p. 76-82.

Harris E. C. 1979 : *Principles of archaeological stratigraphy*, London, Academic Press (coll. Studies in archaeological science), 136 p.

Harris E. C. 1989 : *Principles of archaeological stratigraphy*, London, Academic press, 2e éd. 170 p.

Historic England 2017 : *Photogrammetric Applications for Cultural Heritage. Guidance for Good Practice*, Historic England, 123 p. [URL : <https://historicengland.org.uk/images-books/publications/photogrammetric-applications-for-cultural-heritage/>, consulté le 9 décembre 2022].

Hulin G., Simon F.-X., Hatami N. 2020 : *Archaeological Geophysics Toolbox (AGT) : plugin de traitement des données géophysiques sous QGIS* [URL : https://geophinrap.hypotheses.org/plugin_qgis-agt, consulté le 11 mai 2023].

IA 2022 : *Issue 60 : Climate Change and Archaeology. EAC symposium proceedings*, Internet Archaeology [URL : <https://intarch.ac.uk/journal/issue60/index.html>, consulté le 29 août 2022].

Inrap 2007 : *Rapport d'activités 2006*, Inrap, 151 p. [URL : <https://www.inrap.fr/rapport-d-activite-9728>, consulté le 16 février 2023].

Inrap 2010 : *Rapport d'activités 2009*, Inrap, 199 p. [URL : <https://www.inrap.fr/rapport-d-activite-9728>, consulté le 16 février 2023].

Inrap 2012 : *Rapport d'activités 2011*, Inrap, 180 p. [URL : <https://www.inrap.fr/rapport-d-activite-9728>, consulté le 16 février 2023].

Inrap 2014 : *Rapport d'activités 2013*, Inrap, 81 p. [URL : <https://www.inrap.fr/rapport-d-activite-9728>, consulté le 16 février 2023].

Inrap 2015 : *Rapport d'activités 2014*, Inrap, 81 p. [URL : <https://www.inrap.fr/rapport-d-activite-9728>, consulté le 16 février 2023].

Inrap 2016 : *Rapport d'activités 2015*, Inrap, 89 p. [URL : <https://www.inrap.fr/rapport-d-activite-9728>, consulté le 16 février 2023].

Inrap 2017 : *Rapport d'activités 2016*, Inrap, 92 p. [URL : <https://www.inrap.fr/rapport-d-activite-9728>, consulté le 16 février 2023].

INSA 2021 : *Guide Science ouverte - Penser les futurs et agir avec sens*, INSA, 15 p. [URL : <https://hal-inrap.archives-ouvertes.fr/INSA-GROUPE/page/open-science-guide>, consulté le 13 juin 2022].

ISO 2009 : *ISO 15836:2009. Information and documentation — The Dublin Core metadata element set*, International Organization for Standardization, 5 p. [URL : <https://www.iso.org/standard/52142.html>, consulté le 24 avril 2023].

IWGDD 2009 : *Harnessing the Power of Digital Data for Science and Society*, Washington D. C., Interagency Working Group on Digital Data to the Committee on Science of the National Science and Technology Council [URL : <https://catalog.data.gov/dataset/harnessing-the-power-of-digital-data-for-science-and-society-report-of-the-interagency-wor>, consulté le 15 avril 2022].

Jarrige F. 2016 : *Technocritiques: du refus des machines à la contestation des technosciences*, Paris, la Découverte (coll. La Découverte Poche), 437 p.

Jeffrey S., Richards J., Ciravegna F., Waller S., Chapman S., Zhang Z. 2009 : The Archaeotools project: faceted classification and natural language processing in an archaeological context, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, p. 2507-2519.

Jones S. 2009 : *A report on the range of policies required for and related to digital curation*, Data Curation Center, 34 p. [URL : <https://era.ed.ac.uk/handle/1842/3371>, consulté le 25 mars 2022].

Jones S. 2011 : How to Develop a Data Management and Sharing Plan, *Data Curation Center* [URL : <https://www.dcc.ac.uk/guidance/how-guides/develop-data-plan>, consulté le 21 mars 2022].

JORF 2001 : Loi n° 2001-44 du 17 janvier 2001 relative à l'archéologie préventive [URL : <https://www.legifrance.gouv.fr/loda/id/JORFTEXT000000221337>, consulté le 19 juin 2023].

JORF 2004 : Loi n° 2004-575 du 21 juin 2004 pour la confiance dans l'économie numérique, 22 p. [URL : <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000000801164>, consulté le 3 mai 2021].

JORF 2016 : Loi n° 2016-1321 du 7 octobre 2016 pour une République numérique [URL : <https://www.legifrance.gouv.fr/jorf/id/JORFSCITA000033202935>, consulté le 10 juin 2022].

JORF 2017 : Décret n° 2017-638 du 27 avril 2017 relatif aux licences de réutilisation à titre gratuit des informations publiques et aux modalités de leur homologation [URL : <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000034502557/>, consulté le 17 mai 2022].

JOUE 2009 : Décision N° 922/2009/CE du Parlement Européen et du Conseil du 16 septembre 2009 concernant des solutions d'interopérabilité pour les administrations publiques européennes [URL : <https://eur-lex.europa.eu/eli/dec/2009/922/oj>, consulté le 8 juin 2022].

Kasser M., Egels Y. 2001 : *Photogrammétrie Numérique*, Paris (coll. ENSG-IGN), 379 p. [URL : https://www.academia.edu/20606086/Photogramm%C3%A9trie_Num%C3%A9rique, consulté le 2 février 2022].

Katsianis M., Tshipidis S., Kotsakis K., Kousoulakou A. 2008 : A 3D digital workflow for archaeological intra-site research using GIS, *Journal of Archaeological Science*, 35-3, p. 655-667.

Kitchin R. 2014 : *The data revolution: big data, open data, data infrastructures & their consequences*, Los Angeles, SAGE Publications, 222 p. [URL : <https://b-ok.cc/book/2476044/d0039c>, consulté le 18 juillet 2022].

Kitchin R., Dodge M. 2011 : *Code/Space: Software and Everyday Life*, Cambridge, MA, USA, MIT Press, 304 p.

Koehler A., Tufféry C. 2012 : Harmonisation des méthodes et outils pour l'information archéologique à l'Inrap : constats, enjeux et perspectives pour un établissement national, *Archeologia e Calcolatori*, Supplemento 3, p. 229.

La Salandra M., Miniello G., Nicotri S., Italiano A., Donvito G., Maggi G., Dellino P., Capolongo D. 2021 : Generating UAV high-resolution topographic data within a FOSS photogrammetric workflow using high-performance computing clusters, *International Journal of Applied Earth Observation and Geoinformation*, 105, p. 102-600.

Lacombe E., Reich G., Durost S., Girard J.-P., Rousset M. 2022 : Explorer la cité par le vocabulaire : le thésaurus et les graphes comme architecture heuristique au service de l'archéologie, *Management des Technologies Organisationnelles (2014-....)*, n° 14-1, p. 141-161.

Lees Perasso E., Vateau C., Domon F. 2023 : *Évaluation environnementale des équipements et infrastructures numériques en France*, Agence de l'Environnement et de la Maîtrise de l'Énergie, Autorité de régulation des communications électroniques, des postes et de la distribution de la presse, 132 p. [URL : <https://presse.ademe.fr/2023/03/impact-environnemental-du-numerique-en-2030-et-2050-lademe-et-larcep-publient-une-evaluation-prospective.html>, consulté le 11 septembre 2023].

Lehoërff A. 2009 : Le travail de terrain, in Demoule J.-P., Giligny F., Lehoërff A., Schnapp A., *Guide des méthodes de l'archéologie*, Paris, La Découverte (coll. Grands repères), p. 40-90.

Lepesant G. 2018 : *La transition énergétique face au défi des métaux critiques*, Ifri, 48 p. [URL : <https://www.ifri.org/fr/publications/etudes-de-lifri/transition-energetique-face-de-fi-metiaux-critiques>, consulté le 11 septembre 2023].

Limp F., Payne A., Simon K., Winters S., Cothren J. 2011 : Developing a 3-D Digital Heritage Ecosystem: from object to representation and the role of a virtual museum in the 21st century, *Internet Archaeology*, 30. [URL : http://intarch.ac.uk/journal/issue30/limp_index.html, consulté le 24 août 2022].

Lord P., Macdonald A. 2003 : *E-Science curation report: Data curation for e-Science in the UK : an audit to establish requirements for future curation and provision*, The Digital Archiving Consultancy Limited, 84 p. [URL : <https://www.digitalpreservation.gov/news/2004/e-ScienceReportFinal.pdf>, consulté le 22 novembre 2022].

Losier L.-M., Pouliot J., Fortin M. 2007 : 3D geometrical modeling of excavation units at the archaeological site of Tell 'Acharneh (Syria), *Journal of Archaeological Science*, 34-2, p. 272-288.

Lowe D. G. 2004 : Distinctive Image Features from Scale-Invariant Keypoints, *International Journal of Computer Vision*, 60-2, p. 91-110.

Lyon L. 2007 : *Dealing with Data : Roles, Rights, Responsibilities and Relationships*, UKOLN, 65 p. [URL : www.ukoln.ac.uk/ukoln/staff/e.j.lyon/reports/dealing_with_data_report-final.pdf, consulté le 22 novembre 2022].

Machado J. 2015 : Open data and open science, in Albagli S., Maciel M. L., Abdo A. H., *Open Science, Open Issues*, Brasilia, Rio de Janeiro, p. 295. [URL : <https://livroaberto.ibict.br/bitstream/1/1061/1/Open%20Science%20open%20issues%20Digital.pdf>, , consulté le 17 mars 2022].

Mahil A., Tremblay D.-G. 2017 : Théorie de l'acteur-réseau, in Bouchard F., Doray P., Prud'homme J. (dir.), *Sciences, technologies et sociétés de A à Z*, Montréal, Presses de l'Université de Montréal (coll. Thématique Sciences sociales), p. 234-237. [URL : <http://books.openedition.org/pum/4363>, consulté le 10 août 2023].

Marlet O., Baudoin B., Bernard L., Bézard L., Boissat R., Buard P.-Y., Halczuk A., Hivert F., Hélène J., Nouvel B., Rodier X., Rousset M., Scholtus L. 2022 : *Le Livre Blanc du consortium Mémoires des Archéologues et des Sites Archéologiques : Guide des bonnes pratiques numériques en archéologie*, Consortium MASA, 38 p. [URL : <https://halshs.archives-ouvertes.fr/halshs-03561376>, consulté le 23 mars 2022].

Marrone M., Peterlongo G., Pirina G. 2023 : Beyond the myths of digitalization: labor, space, and ecology in the digital age, in Carbonara E., Tagliaventi M. R. (dir.), *SMEs in the Digital Era: Opportunities and Challenges of the Digital Single Market*, Cheltenham, Elgar publishing, p. 84-99.

Martini C., Hamon E. 2016 : De la *domus* PC1 à la *domus* PC2, in Guichard V., *Rapport annuel 2016 du programme quadriennal de recherche 2013-2016 sur le Mont Beuvray*, Glux-en-Glenne, Bibracte EPCC, p. 207-232. [URL : <https://hal.science/BIBRACTE/halshs-01442242v1>, consulté le 3 octobre 2023].

Marx A., Rossenbach K. S., Bryas E. 2021 : Digital Archiving and Data Stewardship in French Archaeology, *Internet Archaeology*, 58. [URL : <http://intarch.ac.uk/journal/issue58/26/index.html>, consulté le 27 juillet 2022].

Marx K. 1908 : *Misère de la philosophie : réponse à la « Philosophie de la misère » de M. Proudhon (Nouvelle édition) / par Karl Marx ; avec une préface de Friedrich Engels*, Bibliothèque socialiste internationale, 300 p. [URL : <https://gallica.bnf.fr/ark:/12148/bpt-6k932177g>, consulté le 8 août 2023].

Matthews S. K. 1968 : *Photography in archaeology and art*, London, Baker, 161 p.

MESR 2022 : Commission d'enrichissement de la langue française. Vocabulaire de l'informatique, in *Bulletin officiel de l'enseignement supérieur et de la recherche*, Paris (coll. BO, 21) [URL : <https://www.enseignementsup-recherche.gouv.fr/fr/bo/22/Hebdo21/CTN-R2212247K.htm>, consulté le 8 juillet 2023].

MESRI 2018 : *Plan national pour la Science Ouverte*, Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation, 12 p. [URL : <https://www.ouvrirlascience.fr/plan-national-pour-la-science-ouverte>, consulté le 27 mars 2020].

MESRI 2021a : *Bilan du Plan national pour la science ouverte 2018-2021*, Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation, 33 p. [URL : <https://www.ouvrirlascience.fr/bilan-du-plan-national-pour-la-science-ouverte-2018-2021/>, , consulté le 12 mars 2022].

MESRI 2021b : *Deuxième Plan national pour la science ouverte. Généraliser la science ouverte en France en 2021-2024*, Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation, 30 p.

Meylan F. 2005 : *Les influences romaines dans l'architecture et l'urbanisme de l'oppidum de Bibracte (Mont Beuvray)*, Thèse de doctorat, Université de Bourgogne, Université de Lausanne, 243 p.

Milo P. 2014 : Prospections géophysiques au Champlain, à la Côme Chaudron, à PC 2, aux Vergers et à la Chaume, in Guichard V., *Rapport intermédiaire 2014 du programme quadriennal de recherche 2013-2016 sur le Mont Beuvray*, Glux-en-Glenne, Bibracte EPCC, p. 27-42. [URL : <https://hal.science/BIBRACTE/halshs-02512035v1>, consulté le 3 octobre 2023].

Mitcham J., Niven K., Richards J. 2010 : Archiving Archaeology: Introducing The Guides To Good Practice, in Rauber A., Kaiser R., Guenther R., Constantopoulos P. (dir.), *Proceedings of the 7th International Conference on Preservation of Digital Objects*, Vienne, Austrian Computer Society [URL : <https://phaidra.univie.ac.at/o:185505>, consulté le 9 août 2022].

Moore R., Hardman C., Xia L., Richards J. 2014 : ADS easy: an Automated e-archiving System for Archaeology, in Earl G., Sly T., Wheatley D., Romanowska I., Papadopoulos C., Murrieta-Flores P., Chrysanthi A. (dir.), *Archaeology in the Digital Era: Papers from the 40th Annual Conference of Computer Applications and Quantitative Methods in Archaeology (CAA), Southampton, 26-29 March 2012*, Amsterdam University Press, p. 299-306. [URL : <https://www.cambridge.org/core/books/archaeology-in-the-digital-era/ads-easy-an-automated-e-archiving-system-for-archaeology/0E4EC1CE672EEC1A08AD542897C71EC2>, consulté le 10 août 2022].

Moore T., Guichard V., Álvarez Sanchís J. 2020 : The place of archaeology in integrated cultural landscape management. A case study comparing landscapes with Iron Age oppida in England, France and Spain, *Journal of European landscapes.*, 1, p. 9-28.

Mora P., Espinasse L., Daniel F., Tournon-Valiente S. 2008 : *Carnyx Tintignac*, Archeo-vision [URL : <https://3d.humanities.science/CND3D/216007.d.2008>, consulté le 19 juillet 2023].

Morozov E. 2014 : *Pour tout résoudre, cliquez ici: l'aberration du solutionnisme technologique*, Limoges, Fyp éditions (coll. Innovation), 352 p.

Mounier P. 2010 : Manifeste des *Digital Humanities*, *Journal des anthropologues. Association française des anthropologues*, 122-123, p. 447-452.

Murray-Rust P., Neylon C., Pollock R., Willbanks J. 2010 : Panton Principles, Principles for Open Data in Science, *Panton Principles* [URL : <https://pantonprinciples.org/>, consulté le 3 avril 2020].

NERC 1996 : *NERC Data Policy Handbook*, Natural Environment Research Council, 21 p. [URL : <https://core.ac.uk/works/97275720>, consulté le 5 avril 2022].

NERC 1998 : *NERC Data Policy Handbook*, Natural Environment Research Council, 23 p. [URL : <https://core.ac.uk/works/97276362>, consulté le 5 avril 2022].

Neumann K. 2009 : Digital aerial cameras, in *Proceedings of the 24th International Cartographic Conference*, Santiago, Chili [URL : https://icaci.org/files/documents/ICC_proceedings/ICC2009/html/nonref/1_6.pdf, consulté le 9 décembre 2022].

Ngouajio M. L. R. 2015 : *Technologie topographique et calculs topométriques*, Support de cours, École Nationale Supérieure des Travaux Publics, 32 p. [URL : https://drive.google.com/file/d/1XwKtIc7eSbkP_E4pt58Aabp_GisrTr-5/view, consulté le 4 avril 2023].

Nicolucci F., Richards J. (dir.) 2019 : *The ARIADNE Impact*, Hungary, Archaeolingua Foundation, 201 p. [URL : <https://zenodo.org/record/4319058>, consulté le 25 août 2022].

NSB 2005 : *Long-lived digital data collections: enabling research and education in the 21st century*, National Science Foundation, 87 p. [URL : <https://www.nsf.gov/pubs/2005/nsb0540/>, consulté le 22 novembre 2022].

OECD 2004 : *Declaration on Acces to Research Data from Public Funding*, Organisation for Economic Co-operation and Development, 6 p. [URL : <https://legalinstruments.oecd.org/public/doc/157/157.en.pdf>, consulté le 22 novembre 2022].

Paul S., Ducher G., Jobard I., Latrache C.-H., Lenco M. 1997 : *Manuel terminologique didactique de télédétection et photogrammétrie: français-anglais*, Paris, Conseil international de la langue française. [URL : <https://fr.calameo.com/read/0009039478fa82b8cd21c>, consulté le 19 novembre 2022].

Pennock M. 2007 : Digital curation: A life-cycle approach to managing and preserving usable digital information, *Library & Archives*, 1, p. 3.

Pierrot-Deseilligny M., Clery I. 2011 : Apero, an open source bundle adjustment software for automatic calibration and orientation of set of images, in *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Trento, Italy, Copernicus GmbH, p. 269-276. [URL : <https://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XXXVIII-5-W16/269/2011/>, consulté le 23 septembre 2022].

Pierrot-Deseilligny M., Pappadimitriou N. 2006 : A multiresolution and optimization-based image matching approach: An application to surface reconstruction from spot5-hrs stereo imagery, in *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Ankara, Turquie, IAPRS [URL : <https://citeseerx.ist.psu.edu/viewdoc/citations;jsessionid=A193C909E4576A523016458FBEB6D050?doi=10.1.1.222.4955>, consulté le 23 septembre 2022].

Pierrot-Deseilligny M., Rupnik E., Muller J. M., Méneroux Y., Daakir M., Zhou Y., Rosu A.-M., Maillet G. 2022 : *MicMac, Apero, Pastis and Other Beverage in a Nutshell!*, IGNF, 479 p. [URL : <https://github.com/micmacIGN/Documentation/blob/341e9515b2844f-4445f5a73520a140db39d30b3f/DocMicMac.pdf>, consulté le 3 janvier 2022].

Pirot F., Varet-Vitu A. 2004 : Systèmes d'information géographique, archéologie et histoire, *Histoire & Mesure*, XIX-3/4, p. 222.

Pouyllau S., Capelli L., Minel J.-L., Bunel M., Sauret N., Baude O., Jouguet H., Busonera P., Desseigne A. 2021 : ISIDORE a 10 ans [URL : <https://humanum.hypotheses.org/6762>, consulté le 8 juillet 2023].

Rafowicz E. 2017 : La Commission de Topographie des Gaules (1858-1879) : structurer, encourager et contrôler le développement de l'archéologie nationale, *Organon*, 49, p. 155-182.

Reich G., Durost S., Girard J.-P. 2022 : Enjeux pour la mise en réseau et l'analyse des connaissances archéologiques, in *Terminologie & Ontologie : Théories et Applications. Actes de la conférence*, Chambéry, Université Savoie Mont Blanc, p. 31-50.

Renfrew C., Bahn P. G. 2004 : *Archaeology: theories, methods, and practice*, New York, Thames & Hudson, 656 p.

Reymonet N., Moysan M., Cartier A., Délémontez R. 2018 : *Réaliser un plan de gestion de données « FAIR » : modèle*, Université Paris Diderot, Université Sorbonne Paris Cité, Université Paris Descartes, 33 p. [URL : https://archivesic.ccsd.cnrs.fr/sic_01690547v2, consulté le 22 novembre 2022].

Richards J. 2002 : Digital Preservation and Access, *European Journal of Archaeology*, 5, p. 343-366.

Richards J. D. 1997 : Preservation and re-use of digital data: the role of the Archaeology Data Service, *Antiquity*, 71-274, p. 1057-1059.

Richards J. D. 2001 : Anglian and Anglo-Scandinavian Cottam: linking digital publication and archive, *Internet Archaeology*, 10. [URL : http://intarch.ac.uk/journal/issue10/richards_index.html, consulté le 24 août 2022].

Richards J. D. 2017 : Twenty Years Preserving Data: A View from the United Kingdom, *Advances in Archaeological Practice*, 5-3, p. 227-237.

Richards J. D. 2021 : Archiving Archaeological Data in the United Kingdom, *Internet Archaeology*, 58. [URL : <http://intarch.ac.uk/journal/issue58/21/index.html>, consulté le 8 août 2022].

Rodier X., Barge O., Saligny L., Nuninger L., Bertoncetto F. 2011a : *Information spatiale et archéologie*, Paris, Errance (coll. Archéologiques), 254 p.

Rodier X., Ciezar P., Moreau A. 2013 : Using GIS in French Rescue Archaeology. The Choice of Inrap: a Tool for Research at the Scale of Excavation, in *Fusion of Cultures*, Grenade, Espagne, Archeopress, p. 217-220. [URL : <https://shs.hal.science/halshs-00956293>, consulté le 23 février 2023].

Rodier X., Moreau A., Ciezar P. 2011b : SIG et archéologie préventive, le choix de l'Inrap : un outil d'aide à la recherche à l'échelle de la fouille, in *Information spatiale et archéologie*, Errance (coll. Collection Archéologiques), p. 33-38. [URL : <https://halshs.archives-ouvertes.fr/halshs-00956198>, consulté le 30 septembre 2022].

Roure R., Manseri H., Munos S., Py M. 2021 : Towards an Archaeological Information System: the evolution of Syslat, an archaeological data management software, in Djindjian F., Moscati P. (dir.), *Big Data and Archaeology*, Oxford, Archaeopress Archaeology, p. 62-70.

Rowley J. 2007 : The wisdom hierarchy: representations of the DIKW hierarchy, *Journal of Information Science*, 33-2, p. 163-180.

Rupnik E., Daakir M., Pierrot Deseilligny M. 2017 : MicMac – a free, open-source solution for photogrammetry, *Open Geospatial Data, Software and Standards*, 2-1, p. 14.

Samaan M. 2016 : *La photogrammétrie rapprochée au service de l'archéologie préventive*, Thèse de doctorat, Université Paris-Est, 137 p. [URL : <https://theses.hal.science/tel-01547923>, consulté le 22 mars 2023].

Sanchez C., Dabas M., Milo P., Quiquerez A., Thivet M. 2019 : Récolement de la documentation géophysique sur l'oppidum de Bibracte, in Guichard V. (dir.), *Rapport annuel 2019 du programme quadriennal de recherche 2017-2020 sur le Mont Beuvray*, Glux-en-Glenne, Bibracte - Centre archéologique européen, p. 61-66. [URL : <https://hal.science/BIBRACTE/halshs-02512035v1>, consulté le 25 mai 2023].

Schubert F. 1997 : Projet GéoTopoCart, in Guichard V. (dir.), *Rapport annuel d'activité scientifique 1997 du Centre archéologique européen du Mont Beuvray*, Glux-en-Glenne, Bibracte - Centre archéologique européen, p. 41-51.

Schubert F. 2000 : Projet GéoTopoCart (Géodésie, Topographie et Cartographie du Mont Beuvray), in Guichard V. (dir.), *Rapport annuel d'activité 2000 du Centre archéologique européen du Mont Beuvray*, Glux-en-Glenne, Bibracte - Centre archéologique européen, p. 45-50.

Schubert F., Hübner W., Krzystek P., Laudrin F. 2007 : Projet GéoTopoCart (Géodésie, Topographie et Cartographie du Mont Beuvray), in Guichard V. (dir.), *Programme triennal 2006-2008 de recherche sur le mont Beuvray, rapport intermédiaire 2007*, Glux-en-Glenne, Bibracte - Centre archéologique européen, p. 57-66.

Seng C., El-Hajaoui R., Font C., Guillemard T., Le Forestier C., Mercey F. 2020 : *La Nécropole Numérique - Compte-rendu d'activité - 2020*, report, Inrap, 36 p. [URL : <https://hal.archives-ouvertes.fr/hal-03620076>, consulté le 2 août 2022].

Shukla P. R., Skea J., Slade R., Al Khourdajie A., van Diemen R., McCollum D., Pathak M., Some S., Vyas P., Fradera R., Belkacemi M., Hasija A., Lisboa G., Luz S., Malley J. (dir.) 2022 : *Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.*, GIEC, 2029 p. [URL : <https://www.ipcc.ch/report/ar6/wg3/>, consulté le 3 octobre 2023].

Simmons H. C. 1969 : *Archaeological photography*, London, University of London, 138 p.

Smale N. A., Unsworth K., Denyer G., Magatova E., Barr D. 2020 : A Review of the History, Advocacy and Efficacy of Data Management Plans, *International Journal of Digital Curation*, 15-1, 30 p. [URL : <http://www.ijdc.net/article/view/525>, consulté le 14 janvier 2022].

Smith G. S. 1995 : Digital Orthophotography and GIS, in *ESRI 1995 User Conference Proceedings* [URL : <https://proceedings.esri.com/library/userconf/proc95/to150/p124.html>, consulté le 29 janvier 2022].

Suber P. 2012 : *Open access*, Cambridge, Mass (coll. MIT Press essential knowledge series), 242 p. [URL : <https://archive.org/details/9780262517638OpenAccess/mode/2up>].

Taillet R., Villain L., Febvre P. 2018 : *Dictionnaire de physique*, Bruxelles, De Boeck, s. v. Analogique, p. 976.

Tchernia A., Gardin J.-C., Morel J.-P., Gros P., Thébert Y., Guzzo P. G., Picard O., Ozanam D., Vincent B., Pietri C. 1986 : La publication en archéologie, *Mélanges de l'école française de Rome*, 98-1, p. 359-386.

Thibeault A., Villotte S. 2021 : *GRAVETT'OS, GRAVETT'OS*. [URL : <https://3d.humanities.science/CND3D/116903.d.2021>, consulté le 8 septembre 2023].

Trébuchet É., Cunault M., Salé P. 2017 : *Conférences sur l'archéologie. Exploitation des données de diagnostic et recherche à l'échelle des territoires : quelques réflexions sur les aspects documentaires et archivistiques. Exemples en région Centre-Val de Loire*, Caen [URL : <https://www.inrap.fr/exploitation-des-donnees-de-diagnostic-et-recherche-l-echelle-des-territoires-13229>, consulté le 1 août 2022].

Tricot A. 2020 : *Numérique et apprentissages scolaires : quelles fonctions pédagogiques bénéficient des apports du numérique ?*, report, Centre national d'étude des systèmes scolaires (Cnesco) ; Conservatoire national des arts et métiers (Cnam), 95 p. [URL : <https://cnam.hal.science/hal-03249545>, consulté le 10 juillet 2023].

Tudhope D., Binding C., Jeffrey S., May K., Vlachidis A. 2011 : A STELLAR role for knowledge organization systems in digital archaeology, *Bulletin of the American Society for Information Science and Technology*, 37-4, p. 15-18.

Tufféry C. 2017 : Ce que l'enregistrement archéologique pourrait signifier., *Les nouvelles de l'archéologie*, 149, p. 46-49.

Tufféry C. 2019 : Les compétences numériques en archéologie : un défi majeur et des risques de déni, *Interrogations ? Revue pluridisciplinaire de sciences humaines et sociales*, 28. [URL : <https://hal-inrap.archives-ouvertes.fr/hal-02380894>, consulté le 21 février 2023].

Tufféry C. 2022 : *Ce que le numérique fait à l'archéologie et aux archéologues. Contribution historiographique et épistémologique à l'étude des évolutions d'une discipline et de ses pratiques en France depuis les années 1970*, CY Cergy Paris Université, 277 p. [URL : <https://hal.science/tel-03936983>, consulté le 16 février 2023].

Tufféry C., Augry S. 2019 : Harmonisation de l'acquisition des données d'opérations d'archéologie préventive. Retours d'expériences et perspectives à partir de l'application EDarc, in Marinicia C., Guillet F., Laroche F., Velcin J. (dir.), *Atelier DAHLIA*, Metz, p. 21-27. [URL : http://dahlia.egc.asso.fr/atelier_DAHLIA2019_actes.pdf, consulté le 7 juillet 2023].

Tufféry C., Delvigne V., Fernandes P., Bressy-Léandri C. 2021 : À propos de quelques outils de collecte de données : réflexions sur les pratiques numériques en archéologie, *Humanités numériques*, 3. [URL : <https://journals.openedition.org/revuehn/1603>, consulté le 27 juillet 2022].

Tufféry C., Le Goff E., Boudry J., Nurra F. 2017 : Recours au CIDOC-CRM pour évaluer l'interopérabilité de données archéologiques de terrain très variées : présentation des premiers résultats des tests effectués par l'Inrap, in *Spatial Analysis and GEOMatics 2017*, Rouen, France, INSA de Rouen [URL : <https://hal.archives-ouvertes.fr/hal-01649750>, consulté le 1 août 2022].

Tyers P. 1996 : Roman amphoras in Britain, *Internet Archaeology*, 1. [URL : http://intarch.ac.uk/journal/issue1/tyers_index.html, consulté le 24 août 2022].

Van den Eynden V., Corti L., Woollard M., Bishop L., Horton L. 2011 : *Managing and sharing data. Best practice for researchers*, UK Data Archive, 33 p. [URL : <https://ukdata-service.ac.uk/media/622417/managingsharing.pdf>, consulté le 22 novembre 2022].

Verhoeven G. 2016 : Basics of Photography for Cultural Heritage Imaging, in Styliandis E., Remondino F., *3D Recording, Documentation and Management of Cultural Heritage*, Dunbeath, Whittles Publishing, p. 127-251.

Verriez Q. 2020a : *PCR Confluence : Plan de Gestion de Données initial*, UMR 6249 Chrono-environnement, UMR 5608 TRACES, UFC, 11 p. [URL : <https://dmp.opidor.fr/plans/5333/export.pdf>, consulté le 22 novembre 2022].

Verriez Q. 2020b : La *domus* PC2, in Guichard V., *Rapport annuel 2019 du programme quadriennal de recherche 2017-2020 sur le Mont Beuvray*, Glux-en-Glenne, Bibracte EPCC, p. 194-215. [URL : <https://hal.science/BIBRACTE/halshs-02512035v1>, consulté le 3 octobre 2023].

Verriez Q., Martini C. 2017 : La *domus* PC2, in Guichard V., *Rapport intermédiaire 2017 du programme quadriennal de recherche 2017-2020 sur le Mont Beuvray*, Glux-en-Glenne, Bibracte EPCC, p. 281-303. [URL : <https://shs.hal.science/halshs-01671478>, consulté le 3 octobre 2023].

Verriez Q., Martini-Picot C. 2018 : La *domus* PC2, in Guichard V., *Rapport intermédiaire 2018 du programme quadriennal de recherche 2017-2020 sur le Mont Beuvray*, Glux-en-Glenne, Bibracte EPCC, p. 227-248. [URL : <https://hal.science/BIBRACTE/halshs-02001397v1>, consulté le 3 octobre 2023].

Verriez Q., Tomasinelli A., Thivet M. 2023 : *Levé orthophotographique par photogrammétrie appliqué au patrimoine archéologique. Du choix des équipements aux paramètres de traitement par solution libre sous MicMac (IGN ENSG)*, Besançon, Presses universitaires de Franche-Comté (coll. Archéo Poche, 1), 110 p.

Vicente-Saez R., Martinez-Fuentes C. 2018 : Open Science now: A systematic literature review for an integrated definition, *Journal of Business Research*, 88, p. 428-436.

Wheatley D., Gillings M. 2002 : *Spatial technology and archaeology: the archaeological applications of GIS*, New York, Taylor & Francis, 269 p.

Wheeler M. 1954 : *Archaeology from the earth*, Oxford, Clarendon Press, 221 p. [URL : <http://archive.org/details/in.gov.ignca.1711>, consulté le 27 avril 2023].

Wilczek J. 2017 : Morphometrics of (especially ceramic) Celtic artefacts – new methods of acquisition, systematization and valorization of the past, in Wilczek J., Cannot A., Le Cozantet T., Remy J., Macháček J., Klápště J. (dir.), *Interdisciplinarité et nouvelles approches dans les recherches sur l'âge du Fer. Interdisciplinarity and New Approaches in the Research of the Iron Age. International Doctoral Conference 2015. Bibracte (France), 2015.*, Brno, Masarykova univerzita, p. 73-76. [URL : <https://digilib.phil.muni.cz/handle/11222.digilib/137961>, consulté le 25 mai 2023].

Wilczek J., Monna F., Jébrane A., Chazal C. L., Navarro N., Couette S., Smith C. C. 2018 : Computer-Assisted Orientation and Drawing of Archaeological Pottery, *Journal on Computing and Cultural Heritage*, 11-4, p. 1-17.

Wilkinson M. D., Dumontier M., Aalbersberg Ij. J., Appleton G., Axton M., Baak A., Blomberg N., Boiten J.-W., da Silva Santos L. B., Bourne P. E., Bouwman J., Brookes A. J., Clark T., Crosas M., Dillo I., Dumon O., Edmunds S., Evelo C. T., Finkers R., Gonzalez-Beltran A., Gray A. J. G., Groth P., Goble C., Grethe J. S., Heringa J., 't Hoen P. A. C., Hooft R., Kuhn T., Kok R., Kok J., Lusher S. J., Martone M. E., Mons A., Packer A. L., Persson B., Rocca-Serra P., Roos M., van Schaik R., Sansone S.-A., Schultes E., Sengstag T., Slater T., Strawn G., Swertz M. A., Thompson M., van der Lei J., van Mulligen E., Velterop J., Waagmeester A., Wittenburg P., Wolstencroft K., Zhao J., Mons B. 2016 : The FAIR Guiding Principles for scientific data management and stewardship, *Scientific Data*, 3, 9 p. [URL : <https://www.nature.com/articles/sdata201618>, consulté le 13 juillet 2022].

Winner L. 2002 : *La baleine et le réacteur : à la recherche de limites au temps de la haute technologie*, Paris, Descartes & Cie (coll. TechnoCité), 270 p.

Wise A., Miller P. 1997 : Why Metadata Matters in Archaeology, *Internet Archaeology*, 2. [URL : <https://intarch.ac.uk/journal/issue2/wise/toc.html>, consulté le 15 août 2022].



Titre : Rationaliser les pratiques numériques en archéologie : l'exemple des chantiers de fouilles de Bibracte

Mots clés : Science ouverte, Plan de Gestion de Données, drone, photogrammétrie, Système d'information Géographique, cycle de vie de la donnée

Résumé : Cette thèse explore la transition vers une archéologie numérique et ouverte. Elle se penche sur la faisabilité d'une fouille archéologique basée sur des logiciels libres, produisant des données transparentes, structurées et aux formats ouverts. Ces vingt dernières années correspondent en effet à une intensification majeure des usages de technologies numériques en archéologie, impactant la collecte, le traitement, la gestion, la conservation et la diffusion des données. L'approche de la science ouverte offre alors les solutions pour contrôler, exploiter et assurer la conservation de ces nouvelles données. Prenant pour cadre d'expérimentation quatre années de fouilles sur l'oppidum de Bibracte, ce travail examine ainsi les défis d'aligner la production de données sur les principes de la science ouverte, dès la phase de terrain. Il cherche également à sortir des considérations techniques pour éclairer les effets de l'archéologie numérique sur l'utilisateur et son environnement. Ce projet souhaite ainsi contribuer à la modernisation des méthodes archéologiques, en développant des pratiques numériques qui prennent en considération les objectifs des archéologues sur le terrain, en plus d'inscrire leur démarche dans un processus de maîtrise, de partage et de pérennisation de la connaissance archéologique.

Titre : Making digital practices in archaeology meaningful: the example of the Bibracte excavation sites

Keywords : Open science, Data Management Plan, drone, photogrammetry, Geographic Information System, data life cycle

Abstract : The thesis examines the move to open, digital archaeology. It investigates the feasibility of free software-based archaeological excavations producing transparent, structured data in open formats. Over the past two decades, the integration of digital technologies in the field of archaeology has increased considerably, affecting the collection, processing, management, preservation and dissemination of data. The open science approach offers solutions for the management, use and protection of this new type of data. This study uses a four-year excavation of the Bibracte Oppidum as a framework to test how open science principles can guide data production during fieldwork. Moreover, it aims to offer insights into the impacts of digital archaeology on its users and surroundings beyond technical concerns. The project aims to modernise archaeological methods by developing digital practices that consider fieldworkers' objectives and integrate their approach into a process of mastering, sharing and preserving archaeological knowledge.

