



HAL
open science

Enjeux communicationnels de la conception de dispositifs de médiation documentaire augmentée pour les herbiers numérisés

Lisa Chupin

► **To cite this version:**

Lisa Chupin. Enjeux communicationnels de la conception de dispositifs de médiation documentaire augmentée pour les herbiers numérisés. Sciences de l'information et de la communication. Conservatoire national des arts et métiers - CNAM, 2017. Français. NNT : 2017CNAM1141 . tel-01816645

HAL Id: tel-01816645

<https://theses.hal.science/tel-01816645>

Submitted on 15 Jun 2018

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.

École doctorale Abbé Grégoire

Laboratoire Dicen-IDF

Dispositifs d'information et de communication à l'ère numérique – Île-de-France

THÈSE DE DOCTORAT

présentée par : **Lisa CHUPIN**

soutenue le : 1^{er} décembre 2017

pour obtenir le grade de : **Docteur du Conservatoire National des Arts et Métiers**

Discipline / Spécialité : Sciences de l'information et de la communication

Enjeux communicationnels de la conception de dispositifs de médiation documentaire augmentée pour les herbiers numérisés

THÈSE dirigée par

M. ZACKLAD Manuel

Professeur du Cnam

RAPPORTEURS

Mme DESPRÉS-LONNET Marie
Mme MILLERAND Florence

Professeur des universités, Université de Lyon II
Professeur agrégée, Université de Québec à Montréal

PRÉSIDENT DU JURY

Mme SIMONNOT Brigitte

Professeur des universités, Université de Lorraine

EXAMINATEUR

Mme VIDAL Geneviève

Maître de conférences (HDR), Université de Paris XIII

Version du 5 juin 2018

Pour A., qui m'a dédié sa thèse il y a cinq ans. Puissent nos échanges, dons et contre-dons continuer à enrichir nos itinéraires aussi autonomes qu'essentiellement liés.

Remerciements

Mes recherches de thèse se sont fondées sur des observations de terrain réalisées auprès de multiples usagers et gestionnaires de collections d'histoire naturelle, choisis pour la plupart parmi les membres d'institutions partenaires de l'infrastructure e-ReColNat. Je commencerai donc par remercier les personnes à l'origine de ce projet, dont est partenaire le laboratoire Dicen-IDF au sein duquel j'ai effectué mes travaux. J'adresse d'abord mes remerciements à Manuel Zacklad, Professeur du Cnam, qui a défini, avec les autres porteurs du projet, les objectifs de recherche, de conception et de développement du laboratoire Dicen-IDF pour l'infrastructure e-ReColNat, et a encadré mes travaux pendant quatre ans. Mon travail s'est appuyé sur une précieuse coopération avec les ingénieurs responsables de la conception et chargés des développements des livrables du laboratoire, Gilles Bertin, Thomas Bottini, Benjamin Diemert, Dmitri Voitsekhovitch. Ce travail a été réalisé en collaboration avec le MNHN, et en particulier avec Marc Pignal, Directeur de l'infrastructure e-ReColNat, Julien Husson, chef de projet informatique (e-ReColNat), Simon Chagnoux (MNHN, DSI, responsable du Pôle Applications Scientifiques), et Eva Pérez (Direction des Collections, e-ReColNat).

Mes observations de terrain doivent beaucoup à l'accueil bienveillant que m'ont réservé de nombreuses personnes rencontrées lors des réunions de travail de l'infrastructure e-ReColNat, de visites de collections d'institutions partenaires, et de rencontres complémentaires plus ou moins formelles. Je pense en particulier au personnel et aux chercheurs du Muséum National d'Histoire Naturelle, en particulier Germinal Rouhan, Line Le Gall, Dario De Franceschi, Mélanie Tanrattana, Bernard Riera, Régine Vigne-Lebbe, Bruno Denetière, Cécile Aupic, Marc Jeanson, Véronique Andro, Gwenaël Lebras. Des rencontres très enrichissantes ont aussi eu lieu au sein de l'association Tela Botanica, notamment avec Véronique Schäfer et Louise Boulangeat, mais aussi au sein de l'herbier de l'université de Montpellier 2 notamment avec Caroline Loup, Joël Mathez et Peter A. Schäfer, et dans l'unité de recherche de l'AMAP (Cirad), en particulier avec Jérôme Münzinger et Jean-François Molino. À Lyon, je remercie aussi pour leur accueil les membres de la Société Linnéenne, les responsables des collections naturalistes du musée des Confluences, de l'herbier de l'université Claude Bernard et du Jardin Botanique, et en particulier Frédéric Danet qui a autorisé la diffusion d'images issues des collections dont il a la responsabilité. La collaboration avec des paléontologues comme Jérôme Thomas, Francis Duranthon, ou encore avec des mycologues comme Régis Courtecuisse et Jean-Philippe Rioult, a aussi été importante pour prendre la mesure de la diversité des collections et des méthodes de travail. Je ne pourrai remercier nommément toutes les personnes rencontrées sur le terrain, étudiants stagiaires déjà impliqués dans des sociétés savantes, techniciens participant aux activités d'informatisation, conservateurs et chercheurs. Mais quel que soit leur

âge et leur statut, en acceptant de me parler de leurs activités et de leurs passions, ces personnes m'ont apporté un précieux soutien.

Je tiens en outre à remercier les contributeurs de la plateforme « Les Herbonautes », dont la motivation fut communicative. Je pense en particulier à celles et ceux avec qui j'ai eu la chance d'échanger, de vive voix ou par mail, comme « Hélène », « Protea », « Coutu », « JCPogno », « Vital », « DMontagne », J. Peysson ou C. Delnatte, ainsi qu'aux animateurs qui ont bien voulu me parler de leur expérience.

Je remercie également les membres permanent·e·s et les doctorant·e·s du laboratoire Dicen-IDF qui, à l'occasion de séminaires ou de discussions amicales, ont contribué à orienter et enrichir ma réflexion. La mutualisation de retours d'expériences issus de différentes démarches participatives pour les sciences, les archives, les bibliothèques ou les musées a aussi été une aide précieuse. En plus des échanges avec les partenaires de l'infrastructure e-ReColNat — notamment avec Philippe Laroche (Agoralogie) et Tela Botanica — et avec des porteurs des projets participatifs du Muséum National d'Histoire Naturelle rencontrés notamment à l'occasion de l'Atelier de Recherche Participative coordonné par Frédérique Chlous, je mentionnerai les échanges avec Pauline Moirez (BNF), Raphaëlle Bats (ENSSIB) et Mathieu Andro (INRA). Le travail de recherche et de didactisation au service de formations innovantes à la médiation scientifique réalisé au Cnam avec le laboratoire HT2S, en particulier Loïc Petitgirard et Camille Paloque-Berges, a été lui aussi très stimulant. Ma réflexion doit aussi beaucoup aux conseils donnés par les relecteurs des articles que j'ai eu l'occasion de proposer à la publication, et par les discutants des colloques et journées doctorales auxquels j'ai pu participer grâce au soutien financier de l'infrastructure e-ReColNat et de l'école doctorale du Cnam (Abbé Grégoire).

Je souhaite aussi dire ma reconnaissance aux enseignant·e·s intervenu·e·s dans le master Documentation et Médiation des savoirs à l'Université de Lille 3, avec qui j'ai découvert des fondamentaux des sciences de l'information et de la communication, qui furent précieux pour amorcer mon travail de thèse. Ce dernier n'aurait pas été possible sans une culture en sciences humaines et sociales auxquelles de multiples professeurs ont contribué antérieurement, mais aussi des personnes de mon entourage, et en premier lieu mes parents, par leur goût pour les arts, les sciences, et le dialogue. Le soutien de mon conjoint a aussi été déterminant pour mener à bien ce travail. Son aide a été particulièrement précieuse pour relire et mettre en forme mon travail en LaTeX dans les délais impartis.

La présente version de ma thèse intègre des apports issus des échanges qui ont eu lieu lors de la soutenance. Je remercie Brigitte Simonnot qui a accepté de présider le jury, Florence Millerand et Marie-Després-Lonnet qui ont bien voulu rapporter ma thèse, et Geneviève Vidal, examinatrice. Je souhaite continuer à mettre à profit leurs conseils dans mes travaux ultérieurs.

Résumé analytique

Problématique

La présente thèse étudie les mutations de la communication médiatisée par les collections d'histoire naturelle dans le contexte de leur numérisation. Elle cherche à vérifier l'hypothèse selon laquelle des dispositifs numériques peuvent reconstituer et enrichir les médiations de la communication scientifique assurées par les herbiers, notamment en améliorant l'accès aux documents par différents publics et la coordination de l'activité d'acteurs aux compétences hétérogènes qui participent à la production, à l'exploitation et à la valorisation des collections. La thèse s'appuie sur une démarche de recherche-conception qui part de l'analyse de différentes bases de données d'herbiers numérisés et de leurs usages pour concevoir de nouvelles fonctionnalités favorisant l'exploitation des spécimens numérisés, et dessiner un modèle de médiation documentaire numérique augmentée amplifiant les caractéristiques et le rôle des herbiers dans la production des savoirs.

Les collections d'histoire naturelle ont servi de paradigme pour penser le rôle des artefacts matériels dans la production de données scientifiques standardisées par des acteurs aux compétences hétérogènes, y compris des amateurs (chapitre 2). Nous envisageons leur numérisation dans l'idée que ce même paradigme puisse être porteur de nouveaux modèles de dispositifs et pratiques de médiation, utiles pour valoriser d'autres corpus scientifiques et patrimoniaux, et conceptualiser les innovations stimulées par les technologies numériques déjà observées chez les professionnels de la médiation documentaire (Liquète *et al.*, 2010). Dans cette perspective, nous étudions les médiations de la communication scientifique nécessaires à développer les usages des spécimens numérisés à toutes les étapes où interviennent les collections : document issu de collectes parfois associées à des activités de loisir, la planche d'herbier devient une source d'information examinée au début de la recherche scientifique, puis le support des résultats d'une nouvelle expertise conservée en vue de sa réexploitation future ; elle acquiert avec le temps une valeur patrimoniale renseignant sur l'histoire des travaux scientifiques qu'elle a alimentés. Les collections d'histoire naturelle conduisent à adopter une conception de la communication scientifique qui intègre, en plus des interactions entre les acteurs de l'édition et les chercheurs, et aux côtés des médiations de l'écriture des publications scientifiques, celles qu'exercent les objets de collectes amassés par une grande diversité de contributeurs. Nous formulons la double hypothèse qu'il est possible de reconstituer sur des supports numériques les caractéristiques du dispositif de médiation de la communication scientifique que constitue l'herbier, mais aussi de l'enrichir par de nouvelles fonctionnalités qui amplifient son rôle. Le terme d'amplification est employé pour désigner le renforcement de propriétés déjà existantes dont les effets sont décuplés par un changement de support,

conformément à l'emploi que Dominique Boullier (2013) en a proposé pour penser les conséquences des technologies numériques et leurs effets, dans la continuité des travaux d'Élisabeth Eisenstein (1991).

Notre démarche nous amène à préciser la notion de « médiation documentaire augmentée », déjà employée notamment pour désigner les enjeux de l'introduction de ressources documentaires numériques en accompagnement des visiteurs dans l'espace des musées (Fabre, 2012). Nous désignons par ce terme les spécificités de la médiation documentaire observées dans le cas de l'herbier, avant sa numérisation, qui dépassent l'objectif de faciliter l'accès à l'information scientifique, pour assurer la coordination de l'activité d'acteurs aux statuts, compétences et centres d'intérêts hétérogènes. Ce modèle propre à l'herbier et aux multiples fonctions qu'il assure dans la production des connaissances conduit à penser les conditions de possibilité d'un dispositif numérique de médiation documentaire augmentée capable d'intégrer des objectifs d'exploitation scientifique et de valorisation patrimoniale tout en mobilisant des publics de spécialistes et non-spécialistes concourant à enrichir les collections (chapitre 3). L'étude des différentes bases de données reconstituant et augmentant sur des supports numériques les caractéristiques de la médiation documentaire propre aux herbiers, dans différents domaines (chapitre 4), puis parmi les sites utilisés par les chercheurs utilisant les collections d'histoire naturelle, permet de caractériser des dispositifs de médiation numériques qui pourraient être utilisés pour communiquer à des publics spécialistes ou non d'autres fonds d'archives ou de bibliothèques, dont l'herbier peut être rapproché.

Terrain d'observation et méthode d'analyse de la dimension communicationnelle des activités mobilisant les collections

L'augmentation des fonctionnalités des bases de données au-delà de la seule consultation des collections dépend de la volonté d'innovation des acteurs de la valorisation et de l'exploitation des collections. C'était l'un des objectifs de l'Infrastructure Nationale en Biologie Santé e-Recolnat, qui s'ajoutait à celui de numériser les collections détenues par les institutions partenaires (chapitre 5). Ce projet a conduit à la participation d'un laboratoire de Sciences de l'information et de la communication, Dicen-IDF, à l'infrastructure, avec une mission de recherche, conception et de développement d'applications destinées à la valorisation des collections numérisées. La conception s'est appuyée sur l'observation des pratiques d'utilisateurs de différentes spécialités et sur la recension de leurs attentes à l'égard des outils numériques. C'est sur cette recherche-action articulant des phases d'observation, de conception et de test des dispositifs développés que s'appuie notre étude.

Notre démarche consiste à décrire les spécificités des dispositifs de médiation documentaire donnant accès aux collections d'histoire naturelle en relation avec les pratiques de communication qui les mobilisent, à partir de l'observation de multiples activités utilisant les collections. Ces dernières ont été analysées à partir

d'une même théorisation de l'activité collective envisagée dans sa dimension communicationnelle, qui insiste sur les différents artefacts qu'elle participe à construire et qui la supportent et l'encadrent, la sémiotique des transactions coopératives élaborée par Manuel Zacklad. Cela nous amène à considérer les activités mobilisant les collections comme différentes formes d'interprétation de documents qui sont à l'origine de l'acquisition ou l'élaboration de nouvelles connaissances, et qui sont associées à une transformation des documents utilisés : prélèvements, recatégorisations, saisies de données... Cette double transformation — des connaissances et des documents — qui insiste sur le caractère productif de l'activité, est qualifiée de *transaction*. Nous analysons ainsi les activités d'informatisation, de conservation et d'exploitation scientifique des collections en distinguant les transactions qui les composent, et les dispositifs documentaires qu'elles mobilisent en fonction de l'articulation qu'ils opèrent entre collections numérisées ou non, et selon les formes de participation à l'enrichissement des collections qu'ils rendent possibles pour les usagers (chapitre 6). Cela nous permet d'identifier des gestes communs à différentes finalités et méthodologies scientifiques et de poser les bases d'une analyse fonctionnelle d'un dispositif de médiation susceptible d'accompagner les usages innovants des collections. Ce travail de conception, confronté aux retours des usagers sur les prototypes qui en sont issus, permet d'élaborer un nouveau modèle de médiation augmentant les fonctions documentaires d'un catalogue pour répondre aux besoins de la mobilisation des collections numérisées dans l'élaboration et la communication des connaissances scientifiques.

Communications et médiations à l'œuvre dans l'informatisation des collections

De la phase d'analyse des pratiques antérieure au travail de conception (chapitre 7), nous avons dégagé en premier lieu les renouvellements des activités documentaires prises en charge par les conservateurs et les gestionnaires. L'état de l'art des dispositifs documentaires existants a permis d'étudier les conditions de la centralisation des bases de données de collections, à partir d'un modèle de description de données commun inspiré du Dublin Core (Darwin Core), sans que n'existe de classification universelle stabilisée et consensuelle pour décrire la totalité des espèces. L'observation de la constitution des bases de données de spécimens peut être rapprochée du travail catalographique en réseau, qui suppose d'intégrer des métadonnées produites sur une même œuvre par différents acteurs de la chaîne du livre, de l'éditeur aux professionnels du dépôt légal et aux bibliothécaires qui intégreront à leur propre catalogue les notices produites en amont. Les activités de conservation s'inscrivent traditionnellement dans une perspective communicationnelle, puisqu'elles visent à faciliter la recherche de spécimens par l'utilisateur, qui est plus commode quand ils ont été enregistrés sous une dénomination actuellement valide. C'est d'autant plus manifeste que les conservateurs et les techniciens utilisent souvent des bases de données produites par d'autres, ce qui donne lieu à une progression concomitante des inventaires locaux. Les bases de données de spécimens, de plus en plus riches, peuvent faciliter et rendre accessible la recherche

des spécimens types (c'est-à-dire cités dans des publications scientifiques) contenus dans les collections, à un nombre croissant de gestionnaires : il suffit d'identifier des planches d'herbier issues d'une même récolte que le spécimen type déjà référencé pour signaler de nouveaux types. Si les conservateurs rencontrés soulignent l'importance d'un regard critique sur les résultats proposés par une base de données référençant des spécimens types, ces bases existantes permettent de mobiliser de nouveaux contributeurs à ces inventaires. L'intégration des chercheurs à ce travail documentaire reste dominé par des interactions avec les gestionnaires, même si certaines bases de données ouvrent des droits de correction aux usagers : ce type d'ouverture à la participation de nouveaux contributeurs apparaît comme un enjeu central pour l'élaboration de nouveaux modèles de médiation documentaire numérique capables de rendre soutenable pour les institutions scientifiques la numérisation de volumes chiffrés en millions de documents. Cet enjeu suppose d'appréhender l'activité scientifique mobilisant les spécimens numérisés et les modalités de production de données susceptibles d'être intégrées à la documentation des collections.

Médiations instrumentales de l'exploitation scientifiques des spécimens numérisés

L'étude des pratiques des scientifiques mobilisant les bases de données de spécimens montre leur intégration fréquente à l'environnement de travail des chercheurs, en premier lieu dans le but de faciliter l'accès à la planche d'herbier, grâce à un dispositif de médiation documentaire rendu plus performant. Des entretiens auprès des conservateurs et des chercheurs se dégagent le constat d'une plus grande autonomie de l'utilisateur dans ses recherches de spécimens, même si le conservateur garde un rôle essentiel pour compléter ses demandes. Alors que seules les liasses de planches d'herbier sont identifiées par une cote, ce qui oblige à les parcourir en entier pour sélectionner quelques spécimens qu'elles contiennent, la base de données fournit une URL correspondant au code-barres apposé sur chaque spécimen. Elle accroît la facilité d'accès et la permanence de la localisation du document — ou « référentialité », selon Manuel Zacklad (2015). Les informations des étiquettes fragmentées dans les multiples champs de saisie ainsi que le géoréférencement des spécimens ouvrent un accès aux données par visualisation et sélection cartographique. Les fonctions d'export facilitent la constitution d'un corpus correspondant à l'intérêt du chercheur, qui passe par un travail de vérification critique des données informatisées. Les pratiques scientifiques mobilisant les spécimens numérisés se distinguent par la valeur qu'elles leur accordent, entre simple intermédiaire facilitant l'accès aux spécimens physiques et utilisation du spécimen numérisé comme document de référence, qui se prête à de nouveaux traitements permettant d'extraire de l'objet de collection des données (relevé de caractères morphologiques, mesures par exemple).

Conçu pour encourager l'exploitation scientifique des spécimens numérisés comme documents de référence, le logiciel développé pour e-ReColNat par le laboratoire

Dicen-IDF intègre, aux fonctions de consultation, des outils d'annotation inspirés des pratiques développées par certains chercheurs individuellement selon leurs besoins (mobilisant conjointement tableur, logiciel de mesure manuelle ou d'analyse automatique de l'image). L'exploitation scientifique du document avec ce nouveau logiciel passe par la possibilité de prélever des mesures, et de les calibrer pour assurer leur fiabilité par rapport à celles qui pourraient être réalisées sur le document original (chapitre 8). Ces données quantitatives, prélevées sur la sélection de spécimens étudiés par le scientifique, sont ensuite consignées dans un tableau, et peuvent donner lieu à des traitements statistiques qui permettront de conclure en particulier à la nécessité de recatégoriser ou non les spécimens observés. L'intégration au même logiciel d'outils de regroupement progressif de spécimens par disposition libre dans l'espace de travail virtuel a vocation à associer prélèvement de mesure, observation de caractères, et proposition de nouvelle catégorisation – instrumentant ainsi les opérations essentielles du travail du systématique. Le passage entre observation et annotation d'un spécimen d'une part, et catégorisation d'autre part, montre la nécessité d'articuler deux échelles de médiation, d'une part celle qui s'inscrit dans le registre plus classique de la médiation documentaire, qui porte sur l'organisation des connaissances et le classement des documents issus des conclusions du chercheur, et qui pourra se traduire par des déplacements ultérieurs des spécimens dans les collections, et d'autre part celle qui porte sur la constitution de jeux de données se prêtant à une exploitation statistique. Les retours des usagers ont souligné le besoin de renforcer encore cette articulation entre médiation documentaire et médiation des données de recherche, dans le but d'accroître la précision de la référence au document de collection, ou à l'un de ses fragments sur lequel les données exposées dans une publication ont été prélevées.

Le travail de conception permet ainsi de préciser des caractéristiques d'un dispositif de médiation documentaire augmentée, articulant accès aux documents, gestion documentaire, production et archivage de données scientifiques, même si le logiciel effectivement développé ne répond pas encore à toutes les fonctions induites par les besoins dégagés. N'étant pas pour le moment associé à des interfaces de publication et partage des annotations réalisées par d'autres contributeurs, le logiciel développé n'intègre pas à la médiation documentaire l'archivage des données de la recherche – qui permettrait de revenir sur les mesures sur lesquelles s'est appuyée une publication : répondre à ce besoin reste encore une perspective de développement, au même titre que les possibilités d'intégration de corrections et d'enrichissements proposés par les usagers. D'autres exemples de bases de données étudiées au titre d'état de l'art permettent d'envisager des solutions techniques à cette ouverture sans excéder la disponibilité des techniciens susceptibles de procéder à l'intégration (chapitre 9). Si les bases de données ne transforment pas encore fondamentalement l'archivage des données et travaux de recherche mobilisant les collections, c'est en revanche la participation des amateurs au travail documentaire qui se trouve amplifiée, à partir de la plateforme de transcription participative « Les Herbonautes » (chapitre 10).

Communications et médiations en jeu dans la participation des internautes à l'informatisation des collections

Le dernier type de dispositif de médiation intervenant dans la production de données de recherche issues de l'informatisation des collections d'histoire naturelle de l'infrastructure e-ReColNat est la plateforme « Les Herbonautes ». Elle constitue un exemple de dispositif numérique amplifiant les caractéristiques de la médiation documentaire augmentée à partir de laquelle nous avons caractérisé l'herbier. En premier lieu, elle reconstitue un projet mobilisateur capable d'intégrer des amateurs avec différentes motivations dans la collecte de données standardisées scientifiquement utiles. Outre le fait qu'elle accroît massivement le nombre de contributeurs pouvant prendre part à la documentation des collections, elle rend possible une description standardisée de chaque planche perfectionnant les possibilités de recherche dans les collections. Elle offre ainsi un exemple de dispositif de médiation documentaire numérique qui permet l'organisation d'une activité collective associant néophytes et spécialistes par l'association d'interfaces de saisies ludiques, avec une confrontation automatisée des données issues des interprétations des étiquettes, et des forums de discussion permettant de justifier les choix. L'étude de la participation permet de caractériser le type de collectif mobilisé par ce dispositif : à la différence d'initiatives citoyennes dans le domaine de la surveillance de la biodiversité, dont la Ligue de Protection des Oiseaux est un exemple en France, la plateforme est une initiative institutionnelle, conçue comme une solution au problème de l'informatisation des collections de l'infrastructure e-ReColNat, tout comme les collections d'histoire naturelle résultent d'un projet scientifique construit en sorte de pouvoir intégrer la participation d'amateurs intéressés par la flore d'une région. Elle propose une forme de participation qui relève de « l'intelligence distribuée » selon la typologie des sciences participatives de Miku Haklay (2015) : la transcription repose en effet sur l'interprétation des documents, à la différence des logiques de contribution automatisée passant par l'utilisation de capteurs de surveillance environnementale.

Outre le fait qu'elle coordonne la réalisation d'une activité collective, la plateforme constitue un dispositif de médiation des savoirs scientifiques qui passe à la fois par la rédaction de multiples écrits didactiques nécessaires à accompagner la participation, et par des interactions entre contributeurs et animateurs. Elle conduit à développer une posture du médiateur scientifique qui interagit avec le public des internautes dans les espaces de discussion, et ouvre un espace d'échanges entre néophytes et chercheurs directement intéressés à l'activité des contributeurs, qui poursuit elle-même l'objectif d'accélérer la documentation de spécimens prioritairement utiles aux recherches. En permettant des apprentissages informels et communautaires autour des herbiers et des techniques documentaires qui s'y appliquent, et en accompagnant des amateurs dans l'utilisation de bases de données scientifiques de spécimens numérisés, la médiation documentaire reposant sur une réécriture de l'interface de saisie professionnelle dans un contexte scénarisé produit un effet de médiation des savoirs, essentielle à rendre compréhensible aux

néophytes des documents spécialisés mais aussi des savoirs documentaires nécessaires à réaliser les tâches proposées. On retrouve les trois catégories de savoirs dégagés par Patrick Fraysse (2015) dans ses études sur les dispositifs de médiation numérique patrimoniaux : premièrement, la plateforme permet d'acquérir des connaissances spécialisées sur les collections, mais en partant des connaissances des publics, qui sont prises en compte dans le choix des thématiques des missions. L'acquisition de savoirs sur les collections se fonde en second lieu sur « des habiletés techniques et numériques », et troisièmement sur des connaissances documentaires, qu'acquière aussi les contributeurs. Les motivations de certains contributeurs attestent de la découverte de l'herbier dans ses rapports au « monde des hommes qui l'ont produit, utilisé, codifié, embelli, voire au contraire saccagé ou détruit » (Davallon, 2006, p.123). Les commentaires laissés par les contributeurs n'attestent néanmoins que marginalement d'un regard porté sur les collections différent de celui que guident les consignes orientées par un objectif d'informatisation efficace – c'est toutefois le cas par exemple quand des tags sont apposés à certains spécimens pour signaler une collection historique par le nom d'un botaniste, ou en mentionnant sa signature. Si l'expression de médiation documentaire augmentée ne doit pas masquer les multiples registres et effets des médiations opérées par la plateforme, elle permet toutefois de souligner le rôle de l'objet de collection dans la mobilisation de la participation, et précisément de son caractère documentaire qui permet de faire le lien entre un spécimen, un territoire et une histoire dont atteste l'étiquette, qui joue un rôle essentiel dans l'intérêt porté par les contributeurs aux activités proposées.

Notre démarche de recherche-conception donne un point de vue original sur les médiations des formes de communication à l'œuvre dans l'informatisation ouverte aux internautes : l'étude des médiations est inscrite dans le processus de conception qui participe à les transformer, en particulier par le développement de nouveaux outils d'écriture collaborative destinés aux contributeurs. La démarche de recherche-conception trouve son rôle dans la vie de la plateforme du fait que ce ne sont pas les contributeurs qui concourent à son développement, et qu'une place apparaît donc pour une recherche contribuant à proposer des outils au service des différents profils d'utilisateurs selon les techniques de transcription expliquées lors des entretiens. La recherche-conception, par sa démarche itérative (Roy et Prévost, 2013, p.136 ; Parrini-Alemanno, 2004, p.58) qui conduit à travailler sur des données issues de deux versions principales de la plateforme, permet aussi de comparer l'efficacité de la communauté compte tenu de différents outils proposés. La participation de la recherche à la conception gagnerait à se prolonger par une implication dans l'animation pour développer un travail de catégorisation proposé à la communauté, étant donné le diagnostic posé d'un besoin d'accompagnement d'une utilisation coordonnée des tags.

L'activité des Herbonautes permet encore de dégager des défis que doivent relever les dispositifs documentaires ouverts à la participation. Les nouvelles médiations développées dans la plateforme conduisent de fait à une production de commentaires qui excède l'objectif de production de données standardisées. Leur valorisation suppose d'élaborer d'autres procédures d'intégration aux bases de données institutionnelles que celles qui sont utilisées dans les sciences participatives et les

plateformes de crowdsourcing fondées sur la standardisation des contributions en amont. L'intégration de ces contributions non standardisées pose le problème de l'articulation des interfaces distinctes intervenant dans l'informatisation et l'exploitation des collections au sein du réseau e-ReColNat (chapitre 11), qui recréent chacune un contexte d'usage adapté à certaines activités, mais ne permettent pas (encore) une vue d'ensemble des contributions concernant un même spécimen. Cette difficulté conduit à prendre la mesure de la complexité des médiations de la communication scientifique intégrant des ressources numériques : la centralisation des données pour faciliter l'accès à l'information scientifique n'est que la finalité et l'aboutissement d'un travail socialement distribué et techniquement réparti sur de multiples dispositifs ouvrant l'informatisation à des contributeurs de différents statuts.

Mots-clés : médiation documentaire ; recherche-action ; sciences participatives ; patrimoine scientifique et technique ; amateurs de science.

Summary

We describe evolving patterns of communication mediated by naturalistic collections in the context of their digitisation. Natural history collections have been a reference for theorizing the role played by material artifacts in the process of standardised data production lead by multiple and heterogenous actors, including amateurs. We study the digitisation of naturalistic collections with the aim of validating the hypothesis that their characteristics can be reproduced and enhanced with digital device. If so, digitised specimens could be used at each step of scientific communication involving uses of naturalistic collections : sometimes collected during leisure botany practices, specimens are a source of information used during the scientific study, and in the end they are the medium for publishing results of the study on a new label ; over time, they become a scientific heritage documenting the history of knowledge creation. The digital reproduction and enhancement of these functions of herbaria create a model for digital mediation enabling scientific collections to be considered as both scientific information and scientific heritage. This model could be used for other kinds of digitised collections.

The project of extending the functionalities of databases of digitised specimens can be lead only by curators and users. It was one of the goals of Recolnat, a national infrastructure in biology and health, in addition to digitisation of collections. In this aim, a lab in communication sciences, Dicen-IDF (Paris), has been part of the infrastructure. It had to design and develop digital tools for enhancing uses of digitised specimens. The designing process was grounded on the results of the observation of uses of digitalised herbaria and needs expressed by users working with different methods. We grounded our theoretical research on this inquiry involved in the design process in which we were taking part ; user tests of devices developed by the lab were the last part of our inquiry. We analysed all the activities using specimens with the same theoretical framework grounded on a conception of communication as a set of interactions transforming artifacts and creating new meanings and new forms of social organisation. Thereby we analysed different models of communication mediated by digitised collections, depending on the functionalities of databases and on the domain and skill set of users.

Digitised collections have been enhancing the numerous local efforts of inventorying collections by making available online the results of each of them. Researchers mainly communicate their results to curators in order to enrich collection data, because few databases integrate scientific annotations or corrections online. Scientific uses of digitised specimens are varying depending on the specialities : they can be used as a scientific reference, but are often used only for identifying the specimens owned by a herbarium. The software developed by the Dicen-IDF lab enhances scientific uses of digital specimens. Another software of the Recolnat project enables the traditional participation of amateurs in the curation of naturalistic collections by crowdsourcing online transcription of sheet labels. This

platform creates new forms of scientific mediation between animators promoting the understanding of scientific collections and volunteers learning the principles of documenting collections.

Keywords : Document and mediation, Participatory research, Citizen sciences, Scientific and technical heritage, amateurs of science

Table des matières

1. Introduction	1
I. La médiation documentaire augmentée à partir du modèle de l’herbier	7
2. L’herbier, un dispositif de médiation documentaire ouvert à la contribution des usagers de différents statuts	8
2.1. Penser les enjeux communicationnels des collections scientifiques	8
2.1.1. Le travail documentaire, une technique de centralisation des références?	8
2.1.2. Le traitement documentaire comme acte d’interprétation .	11
2.1.3. L’herbier paradigme d’un traitement documentaire multi-disciplinaire intégrant des amateurs	12
2.2. Les médiations de la communication scientifique mobilisant les collections	13
2.2.1. La communication scientifique appréhendée comme un ensemble de transactions	13
2.2.2. Caractériser les dispositifs, interprétations et collectifs à l’œuvre dans la communication mobilisant les collections .	16
2.3. Première caractérisation d’un modèle de médiation documentaire augmentée à l’aune de l’herbier	22
2.3.1. Enjeux de médiation de savoirs et de coordination de l’activité collective de la médiation documentaire	22
2.3.2. Hypothèse de recherche : un dispositif de médiation numérique peut reconstituer et augmenter les fonctions de l’herbier et amplifier son rôle dans la communication scientifique	24
3. Enjeux de la numérisation de l’herbier pour la médiation documentaire	26
3.1. Enjeux d’authentification du spécimen numérisé et conditions instrumentales de son exploitation scientifique	26
3.2. La reconstitution de la médiation documentaire des collections et ses enjeux	28
3.3. De la médiation documentaire à la médiation culturelle	29
3.3.1. De nouvelles complémentarités entre deux traditions professionnelles distinctes	29
3.3.2. Enjeux pour les médiations culturelle et scientifique de l’ouverture à la participation des activités documentaires . . .	30

3.4.	La médiation documentaire augmentée à la croisée de médiations instrumentales, organisationnelles et culturelles	32
4.	État de l’art des dispositifs de médiation documentaire reconstituant des fonctions de l’herbier : bases de données et collaboratoires, outils de lecture critique, et plateformes de crowdsourcing	35
4.1.	Des bases de données d’herbiers aux logiciels d’exploitation automatisée émergents	35
4.1.1.	Processus de numérisation, informatisation et conception de base de données d’herbier	36
4.1.2.	Logiciels et pratiques d’exploitation scientifique du spécimen numérisé	37
4.1.3.	Les collaboratoires	38
4.2.	Apports des dispositifs de médiation documentaire numérique de fonds patrimoniaux	39
4.3.	Des dispositifs de production de données standardisées multi-acteurs ouverts aux amateurs : les plateformes de crowdsourcing et de sciences participatives	41
II.	Méthodes d’observation et d’analyse du terrain : approche ethnographique et analyse quantitative fédérées par un objectif de conception	47
5.	Constituer e-ReColNat en terrain d’étude : mise en visibilité des conditions infrastructurelles de la communication médiatisée par les collections	48
5.1.	Des objectifs centrés sur le partage de données à grande échelle	48
5.2.	Partenaires et livrables	49
5.3.	Des réorganisations et des innovations dans la gestion des collections	51
5.4.	Un terrain pour appréhender les mutations de la communication	52
6.	Construire une démarche de recherche-action autour d’un objectif de conception d’application web	54
6.1.	Une démarche de recherche-action fédérée par un objectif de conception	54
6.2.	Opérationnaliser la sémiotique des transactions coopératives pour la conception	57
6.3.	Recueil et analyse des observations et des données	58
6.3.1.	Enquête ethnographique par entretiens et observations à l’échelle de l’infrastructure	58
6.3.2.	Analyse des dispositifs numériques et de l’activité en ligne	60
6.4.	Catégorisation des transactions à l’œuvre dans l’ensemble des activités observées sur différents supports	62

III. Mutations de la communication médiatisée par les collections scientifiques selon les dispositifs de médiation documentaire	65
7. La communication médiatisée par les bases de données de spécimens : émergence d'un travail d'inventaire en réseau renforçant les possibilités d'exploitation scientifique de l'image	66
7.1. Les caractéristiques documentaires de l'herbier amplifiées par les bases de données	67
7.1.1. La base de données d'herbier numérisés à l'aune de la typologie des médiations documentaires	67
7.1.2. Centralisation des données et travail d'articulation au sein du réseau botanique e-ReColNat	69
7.2. La communication médiatisée par les bases de données d'herbiers dans les activités de conservation	70
7.2.1. Un réseau documentaire morcelé	70
7.2.2. Diversité et évolution des méthodes de recherche de types dans le contexte de la constitution de bases de données	73
7.2.3. De l'inventaire à la mise en valeur de collections historiques : enrichir l'organisation des connaissances fondée sur la systématique dans la médiation documentaire numérique	74
7.2.4. Perspectives de l'analyse des activités de conservation pour la conception de dispositifs numériques de médiation documentaire augmentée	76
7.3. Le spécimen numérisé dans l'exploitation scientifique des collections	77
7.3.1. L'accès aux spécimens renouvelé par la base de données	77
7.3.2. Le spécimen numérisé dans les opérations d'identification	79
7.3.3. Le rôle du spécimen numérisé pour les systématiciens	80
7.3.4. Rôle du spécimen numérisé dans les recensions de données	81
7.4. Le statut documentaire du spécimen numérisé dans la communication médiatisée par les collections	83
8. Le laboratoire en ligne comme dispositif de médiation documentaire augmentée : enjeux pour la valeur documentaire du spécimen numérisé	88
8.1. Des spécificités documentaires de l'herbier aux spécifications d'un herbier numérique	88
8.2. Du projet de collaboratoire au laboratoire	91
8.2.1. Une politique scientifique privilégiant les fonctions d'aide à l'exploitation des spécimens en vue de la publication	91
8.2.2. Assistance à la constitution d'un corpus	92
8.2.3. Reconstituer la dimension spatiale de la catégorisation progressive	92
8.2.4. De la planche à la donnée, instrumentation du geste de prélèvement	95

8.3.	Retours sur le logiciel	96
8.3.1.	Sessions de test et démonstration	96
8.3.2.	Intégration du logiciel dans des pratiques scientifiques : retour d'un systématicien et d'une écologue	98
8.4.	Le laboratoire en ligne comme dispositif numérique de médiation documentaire augmentée	100
9.	Dispositifs numériques de médiation documentaire ouverts aux corrections des usagers	103
9.1.	Enjeux de l'intégration de corrections des usagers pour e-ReColNat	103
9.2.	Solutions d'intégration de contributions par les usagers des bases de données de spécimens numérisés	104
9.3.	Contributions non structurées validées par les conservateurs . . .	105
9.4.	Déterminations virtuelles et annotations sur une base de données enrichie	108
9.5.	Favoriser la contribution au service de l'informatisation des collections	110
10.	La plateforme de transcription participative comme dispositif de médiation documentaire contributif reconstituant la mobilisation d'amateurs dans la gestion des collections	112
10.1.	Amplification des spécificités de la médiation documentaire de l'herbier dans la plateforme participative	113
10.2.	Analyse des transactions au cœur de la transcription participative	116
10.2.1.	Un nouveau public pour les herbiers numérisés	116
10.2.2.	La transcription comme pratique informationnelle	119
10.3.	Rôle des transactions dialogiques dans la production de données standardisées	123
10.3.1.	Coordination de l'activité et apprentissages communautaires dans les forums de mission	125
10.3.2.	Les messages à l'origine des données valides	127
10.4.	Vers un dispositif de médiation documentaire contributif orienté dialogue valorisant les contributions non standardisées	130
10.5.	Médiation documentaire augmentée et médiation numérique du patrimoine à l'aune de la plateforme des Herbonautes	136
11.	Conclusion	139
	Bibliographie	146
IV.	Annexes	157
A.	Documents présentant l'infrastructure e-ReColNat	158
A.1.	Présentation synthétique du projet de recherche financé par l'ANR	158
A.2.	Page de présentation du réseau botanique e-ReColNat sur le site du MNHN	160

B. Tableau récapitulatif des tâches entreprises pour l'infrastructure e-Recolnat et des situations d'observations de l'enquête	162
C. Documents de conception internes du WP5	165
C.1. Synthèse de l'enquête sur les pratiques mobilisant les collections en préalable au travail de spécification avec les ingénieurs du WP5	165
C.2. Maquette de l'interface de transcription réalisée dans le cadre du travail de conception de la version 2 des Herbonautes	176
C.3. Maquette d'une version du laboratoire en ligne intégrant les contributions des Herbonautes et du logiciel de contrôle qualité des images	177
C.4. Échéancier prévisionnel des développements du Cnam concernant le collaboratoire	178
D. Réponses aux questionnaires détaillés envoyés à des contributeurs réguliers de la plateforme « Les Herbonautes »	180
E. Archivage des documents de l'enquête de terrain	196

Table des figures

2.1.	Chaîne des médiations matérielles (Latour, 2001, p.77)	10
2.2.	Planche d’herbier du jardin botanique de Lyon, scannée et mise en ligne sur Flickr	19
2.3.	Planche d’herbier du jardin botanique de Lyon, scannée et mise en ligne sur Flickr	20
5.1.	Schéma de l’organisation de l’infrastructure e-ReColNat en work-packages	50
5.2.	Cartographie dynamique des collections recensées par Tela Botanica (capture d’écran)	51
5.3.	Les institutions partenaires d’e-Recolnat	52
7.1.	Diagramme UML de la base de données Sonnerat au MNHN	85
7.2.	Cartographie des flux de données alimentant l’infrastructure e-ReColNat	86
7.3.	Diagramme UML de la base de données d’ethnobotanique au MNHN	87
8.1.	Panier de résultats de recherche du catalogue <i>Explore</i> exportable dans le laboratoire en ligne (capture d’écran)	93
8.2.	Espace de travail du laboratoire en ligne : placement manuel des images sur l’espace de travail (capture d’écran)	93
8.3.	Espace de travail du laboratoire en ligne : prélèvement d’une mesure taggée (capture d’écran)	94
8.4.	Espace de travail du laboratoire en ligne : mesure d’angle (capture d’écran)	94
8.5.	Espace de travail du laboratoire en ligne : calibrage de l’outil de mesure (capture d’écran)	94
9.1.	Formulaire d’annotation sur le portail Arctos (capture d’écran)	106
9.2.	Formulaire d’annotation typée et semi-structurée sur le site du <i>Consortium of California Herbaria</i> (capture d’écran)	106
9.3.	Exemple de commentaire laissé sur l’interface de consultation des spécimens de <i>Global Plants</i> (capture d’écran)	107
9.4.	Répartition par commentateur des commentaires laissés sur <i>Global Plants</i> entre 2011 et 2015	109
10.1.	Capture d’écran de l’interface de transcription de la version 1 de la plateforme « Les Herbonautes »	114
10.2.	« Page spécimen » consacrée à la planche n° P034067091 sur « Les Herbonautes » (capture d’écran)	116

10.3. Nuages de mots issus de l'analyse lexicale des forums de discussion sur « Les Herbonautes »	124
10.4. Page de la mission « Fleurs des champs » sur « Les Herbonautes » (capture d'écran)	125
10.5. « Page spécimen » consacrée à la planche n° P02115067 sur « Les Herbonautes » (capture d'écran)	126
10.6. Répartition des contributions annulées entre les spécimens complétés	127
10.7. Répartition du nombre de contributions selon la répartition des commentaires	128
10.8. Analyse des co-occurrences dans les messages associés aux spécimens, réalisée avec Iramuteq (Ratinaud, 2005)	128
10.9. Photomontage d'étiquettes d'un même botaniste réalisé par des contributeurs	131
10.10. Capture d'écran de la page de synthèse des discussions dans la version 2 des Herbonautes	131
10.11. Capture d'écran de l'interface de transcription de la version 2 des Herbonautes	132

1. Introduction

Aux temps de ses premiers développements, la photographie avait donné lieu à des initiatives visant à créer une « archive universelle » du patrimoine photographié à partir de ce nouveau « procédé d'enregistrement visuel » (Lautisser, 2009, p.4). De manière similaire, les possibilités de stockage et de mise à disposition de documents sur le web ont donné un nouvel essor à différentes utopies documentaires, comme celle d'offrir un aperçu synthétique de l'ensemble des connaissances disponibles dans une encyclopédie (Juanals, 2002), de donner accès à tous les ouvrages qui consignent les savoirs dans une bibliothèque, ou de conserver toutes les sources qui ont servi à les produire dans des archives. Les collections d'histoire naturelle sont issues d'un idéal de connaissance par description et classification de toutes les espèces ; il a pour corrélat une utopie documentaire de conservation dans un même espace des spécimens choisis pour être les plus représentatifs possibles de chaque espèce compte tenu de sa variabilité intraspécifique. Selon Geoffrey C. Bowker (2000, p.645), le « rêve panoptique » d'inventorier et classer la biodiversité, en la rendant visible depuis un point unique (Meyer, 2009), se trouve renforcé par les bases de données et les perspectives de mutualisation et de réexploitation qu'elles ouvrent, malgré les difficultés d'intégration de jeux de données issus d'approches différentes, systématique ou écologique par exemple (Bowker, 2000, p.672). La perspective de numériser les collections d'histoire naturelle pour les rendre accessibles sur Internet va dans le même sens. Elle a aussi pour enjeu la valorisation du patrimoine scientifique pour différents publics, y compris non spécialistes, comme c'est le cas d'autres programmes de numérisation souvent inscrits dans un idéal d'accès pour tous aux savoirs (Clavier et Paganelli, 2015).

Au-delà des volontés des politiques scientifiques qui soutiennent les plans de numérisation, des idéaux et des représentations qui les sous-tendent, notamment celle de la promesse d'une « e-science » perçue comme le moyen d'instaurer des collaborations scientifiques à grande échelle (Heaton et Millerand, 2013), l'objectif de notre recherche est de prendre la mesure des mutations des formes de communication médiatisées par les collections d'histoire naturelle dans le contexte de leur numérisation, en fonction des différents dispositifs développés pour donner accès aux spécimens numérisés, en permettre l'exploitation scientifique ou en enrichir la documentation. Nous abordons plus particulièrement le cas des collections de plantes séchées, les *herbiers*, qui se prêtent plus facilement à la numérisation en deux dimensions que les spécimens zoologiques ou fossiles, pour lesquels les procédés de capture d'images en trois dimensions sont de plus en plus privilégiés, d'après les retours d'expérience de paléontologues que nous avons rencontrés¹. Nous abordons ces mutations avec l'hypothèse qu'il est possible de reconstituer

1. les notes et verbatim d'entretiens mentionnés dans la présente thèse font l'objet d'un achivage en ligne – voir annexe 9

sur des supports numériques les caractéristiques des dispositifs de médiation documentaire qui caractérisent l'herbier et de les enrichir par de nouvelles fonctionnalités, en amplifiant ainsi le rôle des collections dans la communication scientifique. Le terme d'amplification est employé pour désigner le renforcement de propriétés déjà existantes dont les effets sont décuplés par un changement de support, conformément à l'emploi que Dominique Boullier (2013) en a proposé, dans la continuité des travaux d'Élisabeth Eisenstein (1991), pour penser les conséquences des technologies numériques. Étant donné les différentes fonctions que les collections d'histoire naturelle assurent dans la communication scientifique, nous les considérons comme un premier modèle permettant de définir une forme de médiation documentaire que nous qualifierons d'augmentée. Nous désignerons par ce terme les caractéristiques de la médiation documentaire observée dans le cas de l'herbier, qui dépasse l'objectif de faciliter l'accès à l'information scientifique, pour assurer la coordination des activités d'acteurs aux statuts, compétences et centres d'intérêts hétérogènes contribuant à enrichir les collections. Ce modèle propre à l'herbier et aux multiples fonctions qu'il assure dans la production des connaissances conduit à penser les conditions de possibilité d'une médiation numérique des collections qui soit capable d'intégrer des objectifs d'exploitation scientifique et de valorisation patrimoniale tout en mobilisant des publics spécialistes ou non pour concourir à enrichir la documentation des collections. L'étude de bases de données reconstituant et augmentant les fonctions des herbiers sur des supports numériques permettra de caractériser des dispositifs de médiation que nous qualifierons d'augmentés, et qui pourraient s'appliquer à d'autres fonds d'archives ou de bibliothèques, dont l'herbier peut être rapproché.

En tant que collection donnant accès aux sources utilisées dans les publications, l'herbier peut en effet être considéré comme une archive scientifique : sa numérisation pose la question de la fidélité du spécimen numérisé à l'original, condition de son exploitation. L'herbier se distingue ainsi des archives administratives et du traitement particulier qu'elles requièrent pour « préserver un ensemble documentaire dans son état originel, et non pas comme c'est le cas des bibliothèques ou des musées [pour] réorganiser l'ordre de la collection » (Müller, 2006). Si l'herbier peut être le lieu de dépôt de collections particulières remarquables, conservées séparément à cause de leur valeur historique², comme le sont les fonds d'archives personnelles de chercheurs, ses règles de conservation restent spécifiques à son statut de collection scientifique au service de l'élaboration de nouvelles connaissances. Cela permet de considérer les herbiers en tant qu'archives scientifiques comme un « système de données élaborées qui peuvent servir à la recherche elle-même et contribuer à [sa] réactivation », qui se distingue d'une « série documentaire à l'usage de l'historien » (Müller, 2006). C'est ainsi que les documents des collections peuvent être annotés, en partie dégradés pour des prélèvements au cours du travail,

2. Comme expliqué sur le site du MNHN, « les collections qualifiées d'«historiques» sont conservées séparément. Un herbier historique conserve son classement d'origine, traduisant l'état des connaissances d'une époque donnée. Il correspond au «système» de classification publié par leur «constituteur». Ainsi, l'herbier de Joseph Pitton de Tournefort, dès 1793, deviendra statutairement le premier herbier «historique». » – voir <http://www.mnhn.fr/fr/explorez/dossiers/orchidees-colombie-pas-humboldt-bonpland/collections-historiques-museum>

ou encore déplacés selon les nouvelles catégorisations qui en sont faites. Les herbiers répondent aussi aux fonctions d'une bibliothèque spécialisée : outre les dépôts de collections personnelles, ils doivent offrir un accès à l'information scientifique la plus diversifiée possible, chaque établissement cherchant par des systèmes d'échanges à constituer un fonds représentatif d'une diversité maximale d'espèces, avec parfois une spécialisation concernant certaines zones géographiques. En tant qu'ils réalisent une hybridation des fonctions de collection personnelle, d'archive et de bibliothèque, les herbiers constituent aussi un paradigme intéressant pour concevoir des dispositifs numériques associant ces différentes fonctions : ceci est précieux dans un contexte technique qui bouscule les périmètres traditionnels de l'action des archives, des bibliothèques et des centres de documentation (Chabin, 2012).

Les transformations de la communication scientifique dans le contexte de multiples programmes de numérisation constituent un sujet de questionnement pour les sciences de l'information et de la communication, dont les réponses dépendent intrinsèquement des initiatives des acteurs de la gestion des collections, ainsi que de l'investissement des usagers dans le développement de nouvelles méthodologies d'exploitation scientifique de l'image. Cet objet d'étude s'inscrit dans la perspective épistémologique proposée par Yves Jeanneret (2008) pour l'étude de la communication comme « science historique » (Passeron, 2006, p.80-85) dans laquelle « la prise en compte concrète des pratiques des hommes en situation est constitutive de l'objet » (Jeanneret, 2008, p.55). En l'occurrence, nous avons étudié la place des collections numérisées dans la communication scientifique à partir de leur intégration aux pratiques des conservateurs et des chercheurs des institutions partenaires de l'Infrastructure Nationale en Biologie Santé e-ReColNat³. Celle-ci a réuni, en vue de la numérisation et de la valorisation de leurs collections, le Muséum National d'Histoire Naturelle (MNHN) et différentes institutions universitaires et muséales françaises, ainsi que l'herbier de l'institut scientifique de Rabat, historiquement très lié à celui de Montpellier. Cette infrastructure était susceptible de donner lieu à des applications web de consultation et d'exploitation des collections particulièrement innovantes, étant donné que l'élaboration de dispositifs favorisant l'exploitation du spécimen numérisé dans la démarche scientifique était l'un de ses objectifs. Par rapport à d'autres études faites sur les opérations de numérisation d'herbiers (Heaton et Proulx, 2012; Heaton et Millerand, 2013), l'infrastructure était aussi originale par le choix d'une numérisation industrielle des collections, dans la continuité du processus initié pour l'herbier du MNHN en amont de la constitution de l'infrastructure. Ce processus s'accompagnait de la délocalisation et de la sous-traitance des tâches d'informatisation des spécimens numérisés : il en résultait une indexation partielle, que l'un des enjeux d'e-Recolnat était de compléter par le recours à une démarche qualifiée d'« externalisation vers la foule », selon la traduction généralement proposée du terme anglais « crowdsourcing » (Schenk et Guittard, 2012). Cette démarche s'applique à de multiples domaines, mais était ici susceptible de s'intégrer dans une tradition de sciences participatives ou de contribution d'amateurs au travail documentaire, usuelle dans la gestion des herbiers.

3. <https://www.recolnat.org/fr/>

Le double objectif d'instrumentation de l'exploitation des spécimens numérisés et de transcription participative a conduit à confier au laboratoire Dicen-IDF (EA 7339), spécialisé dans les « Dispositifs d'Information et de Communication à l'Ère Numérique », et au Conservatoire National des Arts et Métiers (Cnam) qui constitue l'une de ses tutelles, des missions de conception et de développement de nouvelles interfaces dédiées à l'utilisation scientifique des collections et à l'enrichissement de leur documentation. Nos recherches sur les nouvelles pratiques de communication médiatisées par les spécimens numérisés, selon les différents dispositifs y donnant accès, se sont appuyées sur une enquête que nous avons réalisée en vue de la conception des applications développées par le laboratoire. Cette démarche nous a conduits à observer les pratiques émergentes intégrant les spécimens numérisés, à partir des bases de données existantes et de la première version de la plateforme de transcription participative « Les Herbonautes »⁴, utilisée initialement pour enrichir la documentation des images issues de la numérisation de l'herbier du MNHN. Nous avons porté une attention particulière aux logiques d'innovation à l'œuvre chez les acteurs rencontrés et aux représentations qu'ils se faisaient des possibilités d'exploitation des spécimens numérisés. Étant donné l'imbrication du travail collectif de conception et de sa théorisation, facilitée par de multiples échanges au sein du laboratoire et lors d'événements scientifiques, cette thèse est rédigée à la première personne du pluriel. Ce choix d'énonciation correspond aussi à celui adopté pour les publications auxquelles nos travaux ont donné lieu (Chupin, 2014, 2015, 2016, 2017, 2018a, 2018b; Zacklad *et al.*, 2014; Zacklad et Chupin, 2015). Le singulier sera toutefois utilisé pour situer plus précisément les modalités de la participation de l'auteur de la présente thèse dans le projet.

La collaboration du laboratoire Dicen-IDF avec le MNHN a commencé dès 2010, dans le sillage du début de la numérisation de son herbier en 2009 : Manuel Zacklad, directeur du laboratoire Dicen-IDF, avait participé avec Marc Pignal (MNHN) et Philippe Laroche (Agoralogie), au montage d'un premier projet soumis à l'ANR, « Qualicontent », qui ne fut pas financé. Cette première réflexion a contribué à nourrir la construction du projet e-ReColNat, lauréat en 2012 de l'appel à projets « Infrastructures Nationales en Biologie et Santé » du Programme d'Investissements d'Avenir (PIA), financé pour cinq ans. J'ai participé au projet à partir du deuxième semestre 2013, suite à l'obtention du Capes en documentation, du master « Documentation et médiation des savoirs pour l'enseignement et la recherche » à l'Université Lille III, et d'un contrat d'attachée temporaire à l'enseignement et à la recherche à mi-temps au Cnam. J'ai choisi de commencer à m'impliquer dans le projet par l'observation des pratiques mobilisant les collections, dans le but de dégager les apports des outils numériques existants ainsi que les besoins et attentes qu'ils suscitaient, et de documenter un état initial de la communication mobilisant les collections, en amont des transformations qu'on pouvait attendre du processus de conception. L'enquête a donc été réalisée sans connaissance initiale sur les collections ou la botanique. J'ai pris part au travail de conception mené par les ingénieurs du laboratoire en relation avec les partenaires de l'infrastructure, à partir de septembre 2014, après avoir exploité mes observations pour établir des premiers axes de spécification; en 2015, j'ai participé aux échanges avec le pres-

4. <http://herbonautes.mnhn.fr/>

tataire choisi pour développer la version 2 de la plateforme « Les Herbonautes » ainsi qu'aux tests ; dans les mois qui ont suivi, j'ai contribué à recueillir les retours des usagers sur le logiciel développé par le Cnam, et j'ai proposé des spécifications complémentaires. La participation au travail de conception (dans des modalités précisées dans l'annexe B) visait à contribuer aux innovations dans l'utilisation de spécimens numérisés et à étudier leurs enjeux pour le renouvellement de l'activité scientifique envisagée dans sa dimension communicationnelle, et plus généralement pour les formes de communication mobilisant les collections. Cette démarche peut être considérée comme une « recherche-action », au sens où elle a contribué à « développer des stratégies permettant de soutenir des professionnels attachés au développement de leur métier » (Meyer, 2006).

Considérer les bases de données et applications permettant l'exploitation scientifique des spécimens numérisés comme un ressort de transformation de phénomènes communicationnels au cœur de l'élaboration des connaissances se fonde sur une conception de la communication donnant un rôle essentiel aux artefacts qui la rendent possibles. Nous commencerons par exposer les fondements théoriques de l'approche de l'herbier comme artefact de médiation documentaire constitutif de la communication entre les différents acteurs qui participent à son exploitation et à son enrichissement. Cette approche s'appuie sur l'analyse de leurs activités à partir de la mise en évidence des transactions qui les composent, définies selon l'approche de Manuel Zacklad (2015) comme des interactions produisant une transformation symbolique et artefactuelle au sein d'un collectif qu'elles participent à modérer (chapitre 2). Nous articulerons ensuite le sens de la médiation documentaire, entendue dans ce cadre d'analyse comme médiation artefactuelle de l'activité collective, avec celui de médiation des savoirs, qui s'attache à l'ensemble des dispositifs facilitant l'appropriation de l'information, et enfin avec celui de médiation culturelle et scientifique, nécessaire à comprendre la collaboration entre scientifiques et public non spécialiste. Nous serons alors en mesure de justifier l'hypothèse de la possibilité de reconstituer et d'augmenter les fonctions de l'herbier dans la communication scientifique par la conception d'un dispositif de médiation numérique combinant ces mêmes fonctions avec celles nécessaires à répondre à des usages en émergence des spécimens numérisés (chapitre 3). Nous étayerons notre hypothèse à partir d'un état de l'art des dispositifs consacrés à la production de données standardisées et mutualisées dans le domaine des sciences de la nature et du patrimoine, en particulier dans le cas où le travail documentaire est ouvert à des contributeurs non spécialistes (chapitre 4).

La suite de ce travail sera consacrée à l'analyse des formes de communication en jeu dans la production et l'exploitation des données des collections des partenaires de e-Recolnat, dans le contexte de la conception des nouvelles interfaces. Nous commencerons par présenter les spécificités de l'infrastructure (chapitre 5) ainsi que les formes prises par notre enquête, réalisée dans le contexte de la mobilisation de l'enquêteur aux côtés des acteurs du terrain dans un projet de conception. Nous préciserons les formes d'observation et de collectes de données selon les phases du projet, associant entretiens et observations en situation, analyse des données produites par les contributeurs, et étude des retours des usagers sur les logiciels développés dans le projet (chapitre 6). Nous exposerons dans une troisième partie

les différentes formes de communication médiatisées par les collections numérisées dans les différentes activités observées, selon les types de dispositifs utilisés : bases de données de consultation, laboratoire en ligne conçu dans le projet, bases ouvertes à des contributions scientifiques, ou à celles d'internautes non spécialistes. Nous présenterons en premier lieu les mutations des pratiques de conservation et d'inventaires locaux en développement concomitant, puis celles de l'exploitation scientifique des collections, en dégagant le rôle du spécimen numérisé, entre simple intermédiaire facilitant l'accès au spécimen original et document de référence (chapitre 7). Cette première analyse des transactions en jeu dans ces activités permettra d'en dépasser la diversité pour mettre à jour des gestes transverses d'interprétation, d'annotation et de traitement documentaires, ce qui servira à poser les bases de l'analyse fonctionnelle d'un modèle de dispositif numérique apte à reconstituer au mieux les fonctions de l'herbier (chapitre 8). Nous présenterons ensuite le logiciel effectivement développé, compte tenu de demandes complémentaires des usagers et de l'abandon de certaines fonctions induit par les contraintes de temps de développement : il constitue un modèle de base de données augmentée de fonctionnalités d'annotation, de tri et de mesure, au service de l'exploitation scientifique du spécimen, sans encore permettre l'archivage des données issues de ces opérations. Nous nous référerons à des dispositifs développés dans d'autres projets pour envisager des solutions à l'intégration et l'archivage de données des usagers des spécimens numérisés, qui ouvrent des perspectives de développement pour le futur de l'infrastructure (chapitre 9). Pour finir, nous présenterons le dispositif reconstituant l'intégration de contributeurs amateurs à l'informatisation des collections à travers une plateforme proposant la transcription des étiquettes des spécimens à des internautes non spécialistes. Nous serons attentifs aux formes de médiation scientifique qui s'exercent dans ce dispositif, mobilisant animateurs impliqués dans des activités de conservation et de recherche, et contributeurs réguliers qui s'en font les relais (chapitre 10). Ce nouveau contexte de lecture du spécimen, qui vient s'ajouter aux différents espaces de production de données de l'infrastructure, encore cloisonnés, ouvre la perspective de les coordonner d'avantage.

Première partie

La médiation documentaire augmentée à partir du modèle de l'herbier

2. L’herbier, un dispositif de médiation documentaire ouvert à la contribution des usagers de différents statuts

Nous commençons par dégager les caractéristiques de l’herbier comme document, collection, ou espace documentaire, dans l’idée qu’il peut constituer un modèle pour penser les enjeux du document et des collections numériques pour la production et la communication des savoirs scientifiques. Cela nous conduit à préciser la conception de la communication permettant d’appréhender l’activité de conservation et de recherche scientifique à partir des collections au niveau de ses enjeux communicationnels, et à la situer par rapport à d’autres approches du rôle de l’herbier dans la production des connaissances en histoire des sciences, que nous mobilisons également pour faire ressortir les caractéristiques de l’herbier comme modèle de dispositif documentaire.

2.1. Penser les enjeux communicationnels des collections scientifiques

2.1.1. Le travail documentaire, une technique de centralisation des références ?

Les collections d’histoire naturelle ont été prises pour paradigme du rôle du document et de la constitution de sa valeur documentaire dans l’activité scientifique, aussi bien dans les études de sciences et technologies (« Science and technology studies ») que dans les approches en sciences de l’information et de la communication cherchant à définir les nouveaux modes de « documentalité », ou manières de faire document (Zacklad, 2014), renouvelées par les supports numériques. L’approche de Suzanne Briet (1951), en ce qu’elle permet de penser la diversité des matérialités qui peuvent faire document au-delà des textes, intégrant aussi bien des objets de collections rapportés du terrain, a été reprise par Michael K. Buckland (1997) pour refonder une théorie du document tenant compte des transformations liées aux supports numériques. Suzanne Briet prend l’exemple de la constitution du spécimen d’histoire naturelle en document pour montrer le rôle du traitement et de l’espace documentaires dans la valeur de preuve acquise par l’objet collecté, qui lui confère son statut de document, selon la définition qu’elle reprend. Elle distingue ainsi l’objet observé sur le terrain, qui n’est pas encore document, ou la découverte scientifique (« une antilope d’une espèce nouvelle a été rencontrée en Afrique par un explorateur »), de l’inscription de l’antilope dans le catalogue du musée ; elle

mentionne aussi la « vêtue » de documents secondaires produits dans la communication de l'information scientifique, à partir du document primaire constitué par l'objet prélevé sur le terrain après qu'il ait été documenté. Le cas de l'herbier a aussi été mobilisé pour construire des apprentissages info-documentaires, et plus particulièrement la genèse du document à partir d'un objet observé et collecté sur le terrain : on le voit dans une séance pédagogique analysée par Cécile Gardiès (2014) et destinée à faire construire la distinction entre document par intention et par attribution théorisée par Jean Meyriat (1981), à partir de la différence entre la feuille observée sur le terrain (document par attribution en ce qu'elle peut être utilisée pour renseigner sur le milieu observé) et le statut qu'elle acquiert quand elle est intégrée à un herbier personnel (qui constitue un document par intention), ou encore à un centre documentaire rassemblant les herbiers des élèves.

La question de l'acquisition de la valeur documentaire d'un objet rejoint ainsi celle de la constitution de la référence, étudiée par Bruno Latour (2001) à partir d'une enquête ethnographique décrivant la démarche scientifique impliquée dans une étude sur l'évolution des sols selon leur composition, depuis le travail de terrain jusqu'au laboratoire. Faire référence par une inscription à un phénomène ou un terrain est la condition de possibilité de la production d'une connaissance à son sujet. L'observation du travail de collecte des échantillons et du suivi des opérations qu'ils subissent permet à Bruno Latour de décrire la production de différentes inscriptions et leur rôle dans le processus de production de nouvelles connaissances dont elles constituent les « médiations matérielles ». Au niveau de la collecte, Latour précise notamment l'importance de repères fixés sur le terrain qui permettront de rapporter précisément l'échantillon aux coordonnées du lieu d'où il provient. S'en suivent différentes étapes d'exploitation des premiers prélèvements qui débouchent sur la production d'autres inscriptions, de moins en moins analogues à la matière prélevée dont elles permettent toujours de parler, et de plus en plus faciles à mettre en relation les unes avec les autres. Si le spécimen prélevé constitue une réduction de l'information observée sur le terrain (par exemple, une telle sorte de terre se trouvait à tel endroit), la possibilité de le mettre en relation avec d'autres dans des conditions d'observation stabilisées et prolongées en laboratoire amplifie cette fois l'information portée par le spécimen (on a trouvé la même sorte de terre à tel endroit qu'à tel autre à tel moment). Outre l'amplification de l'information par comparaison des spécimens, la conservation des différentes inscriptions sur lesquelles s'est appuyée la théorisation progressive depuis l'observation du terrain permet aussi de revenir sur l'ensemble des étapes, de les vérifier, et de refaire le lien de la théorie et de la publication avec les matériaux du terrain. Cette conservation suppose un espace dédié qui joue lui aussi un rôle essentiel dans la production de nouvelles connaissances, puisqu'il facilite la mise en relation, le tri, l'amplification de l'information des spécimens, en plus de leur conservation.

Bruno Latour (1996) précise également le rôle de l'espace documentaire dans la production et l'exploitation des inscriptions pour produire des connaissances. Il le compare à un centre de calcul, qui instaure avec les périphéries d'où proviennent les matériaux du terrain un rapport de pouvoir, qui correspond à une relation de connaissant à connu. La capacité de production de connaissances du centre que constitue la bibliothèque, la collection, ou le laboratoire, tient à l'ensemble des

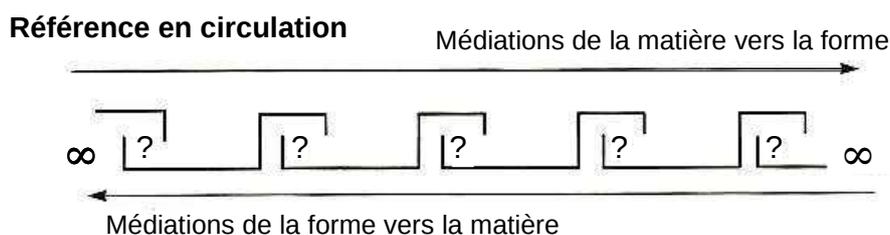


FIGURE 2.1. – Chaîne des médiations matérielles (Latour, 2001, p.77)

spécimens qui y sont accumulés, et que Latour décrit comme autant de connexions du centre documentaire avec le monde, connexions établies notamment grâce aux coordonnées de chaque prélèvement.

L'approche de Latour donne à la fois un modèle de compréhension du rôle de la collection dans la production scientifique au niveau de chaque document, en faisant droit à la matérialité du prélèvement, et de l'herbier comme espace documentaire, qui permet de mettre en relation avec la production scientifique les utopies documentaires cherchant à centraliser l'accès à tous les savoirs dans un même lieu panoptique. Dans la mesure où elle prend en compte la matérialité de l'inscription, elle permet de penser la numérisation comme nouvelle réduction de l'information, tout en interrogeant la capacité de la base de données donnant accès aux images à assumer de nouvelles fonctions d'amplification de l'information de chaque spécimen, susceptibles d'offrir des possibilités de « calcul » accrues par l'accès centralisé à un grand nombre d'images. La numérisation serait une nouvelle étape dans le double processus de réduction et d'amplification de l'information du terrain, donnant un degré de mobilité supplémentaire aux spécimens des collections, et permettant de convoquer des références aux spécimens observés sur le terrain plus nombreuses grâce à l'agrégation des données de différentes institutions. On penserait alors la collection scientifique numérisée à partir du modèle proposé par Bruno Latour (2011) pour la bibliothèque numérique, nouveau centre doté d'une puissance de calcul renouvelée, affranchi des contraintes spatiales et temporelles qui peuvent limiter l'accès aux documents.

Mais l'approche latourienne des inscriptions et des espaces consacrés à leur centralisation a fait l'objet d'une réappropriation critique en sciences de l'information et de la communication, qui se caractérise par le projet de prendre en compte les enjeux matériels de production et de communication de l'information scientifique, mais sans réduire à une problématique logistique les questions de l'accès à l'information et de l'appropriation qui en est faite. Yves Jeanneret (2008, pp.146-147 ; 2009) voit dans l'approche latourienne un risque de réduction de l'analyse de la communication et de l'appropriation des connaissances au transport matériel d'inscriptions mis en évidence par Bruno Latour (1996) : « des réseaux de transformation font parvenir aux centres de calcul, par une série de déplacements — réductions et amplifications — un nombre toujours plus grand d'inscriptions » (cité par Jeanneret, 2008, pp.146-147). La description latourienne des médiations matérielles de la production scientifique à partir de la métaphore d'une chaîne reliant

le terrain au laboratoire est critiquée par Yves Jeanneret du fait qu'elle ne rend pas compte des ruptures inhérentes au processus de communication, qui ont lieu à chaque appropriation dont fait l'objet le texte inscrit sur un support lorsqu'il est lu. En accord avec la conception de l'information développée par les fondateurs de l'approche des sciences de l'information en France, qui repose sur une « dialectique complexe entre la dimension intellectuelle de l'information comme relation et sa dimension matérielle en tant qu'inscription » (Jeanneret, 2008, p.80), la valeur informationnelle de la référence ne peut être décrite comme un rapport entre deux lieux qu'elle connecte, qui ferait disparaître la subjectivité du lecteur, car elle suppose un acte d'interprétation de ce rapport. Malgré les réserves et précautions qu'il exprime, Yves Jeanneret envisage la possibilité de lire le texte de Bruno Latour en accord avec sa propre perspective sur la communication, en « soulignant le rôle de la matérialité des gestes d'inscription », sans adhérer à la représentation d'« une chaîne logique de sujets et d'objets ». Et cette interprétation reste cohérente avec la schématisation produite par Latour des maillons de la chaîne des médiations matérielles qu'il décrit (Latour, 2001, p.77, reproduite en figure 2.1), dont on note le caractère discontinu : la distance laissée entre deux segments peut être vue comme celle de l'interprétation et de la libre réappropriation des inscriptions. Méthodologiquement, il en résulte la légitimité d'une étude des réappropriations successives des inscriptions « en suivant les étapes d'une chaîne d'objets et d'acteurs », à condition de montrer que « la circulation des objets conduit à leur redéfinition constante » (Jeanneret, 2008, p.104).

Yves Jeanneret propose une approche alternative de la circulation des objets matériels associés à la production de connaissances pour tenir compte des enjeux de réinterprétation des inscriptions : il s'agit de décrire un processus de construction des significations des objets culturels à mesure des interprétations qu'en font les acteurs qui les mobilisent dans différents lieux, contextes, « carrefours » (« trivia » latins) de la vie sociale — le terme de trivialité employé au sens étymologique désignant ce processus de circulation sociale et créatrice de sens. Les pratiques de traitement documentaire et plus généralement les « disciplines de l'archive » (Jeanneret, 2008, p.55) sont un objet d'intérêt en ce qu'elles constituent des terrains où se déploient de nouvelles interprétations corollaires de transformations matérielles subies par les objets à chacune des opérations de classement ou de déplacement qui participent à en construire le sens.

2.1.2. Le traitement documentaire comme acte d'interprétation

On peut rattacher à cette même perspective communicationnelle adoptée sur le traitement documentaire l'approche qu'en propose Marie Després-Lonnet (2014a), à partir d'un parallèle qu'elle établit entre les opérations documentaires et le travail à l'œuvre dans la constitution des collections d'histoire naturelle. Elle remobilise avec un regard critique l'approche de Bruno Latour des collectes naturalistes (pp.107-108), en soulignant l'importance des choix interprétatifs, aussi bien au moment de la sélection des objets sur le terrain, que de l'attribution d'une place dans

la collection. Selon elle, le traitement documentaire peut être considéré comme une sélection des documents à conserver, qui subissent d'une « réduction » de leur signification lors des opérations de condensation et d'indexation, avant que ne soit « amplifiée » leur valeur informative par leur mise en relation avec d'autres documents portant sur le même sujet. Cette amplification n'est possible que parce qu'a été donné au document, par rapport aux autres qui composent la collection, un « sens, en cohérence avec le projet qui a conduit à leur acquisition, » et qui n'est pas neutre : « la qualité du traitement documentaire n'est pas liée à sa neutralité, mais bien à sa capacité à répondre aux attentes d'un collectif réuni autour d'un projet commun » (p.96).

La prise en compte des phénomènes de communication en jeu dans les opérations documentaires permet aussi de prendre en compte la dimension collective du processus de constitution et de documentation des collections. À ce titre, l'approche de Marie Després-Lonnet rend compte, dans le cas de la base de données photographiques du site archéologique de Bibracte, de la conception d'un dispositif documentaire numérique permettant de répondre aux besoins de différents spécialistes ayant chacun leur point de vue sur les objets : elle y observe l'accumulation dans une même notice de champs de description qui correspondent aux points de vue des multiples professions travaillant sur le site archéologique. Prendre en compte les enjeux communicationnels de la production des collections par des acteurs de spécialités hétérogènes est particulièrement important dans le cas de l'herbier du fait que des scientifiques de différentes disciplines participent eux-mêmes à certaines phases du travail documentaire, qui intègrent encore en plus des contributeurs amateurs. Les études de sciences ont particulièrement théorisé la problématique de l'appropriation de matériaux collectés pendant les recherches par des scientifiques aux compétences et motivations variées. Nous les mobilisons pour préciser le processus documentaire dans le contexte collaboratif et intrinsèquement lié à l'activité scientifique propre à l'herbier.

2.1.3. L'herbier paradigme d'un traitement documentaire multi-disciplinaire intégrant des amateurs

En sociologie des sciences, le spécimen est décrit comme l'un des objets auquel s'applique un travail « d'équipement » lui permettant de circuler « dans un espace d'échange entre acteurs plus ou moins hétérogènes », qui se sont préalablement « accordés sur les éléments qu'il convient d'ajouter » (Vinck, 2009). Le processus décrit suppose une négociation en amont de la constitution des échantillons. En cela, elle se distingue de l'approche qu'ont développée Susan L. Star et James R. Griesemer (1989) d'un processus de catégorisation itératif, sans négociation préalable, à partir de l'histoire de la collection du muséum de zoologie vertébrée de l'université de Californie. Ils décrivent la constitution d'un édifice documentaire à partir de l'implication de contributeurs aux motivations hétérogènes, rendant possible la réalisation d'un projet scientifique requérant de vastes collectes largement distribuées dans l'espace. Ils mentionnent le rôle d'objets-frontières qui, par la multiplicité des interprétations auxquelles ils se prêtent, parviennent à rassembler largement des acteurs aux motivations variées dans un but commun (par

exemple l'intérêt pour une région, plus que pour une approche scientifique de sa flore). Ils enclenchent un processus progressif de catégorisation, à mesure des travaux mobilisant les objets. L'objet-frontière, constitué d'un ensemble d'artefacts intervenant dans le travail collectif, existe « entre plusieurs mondes sociaux où il est mal structuré », peut donner lieu ensuite à une interprétation plus spécialisée, correspondant aux objectifs d'un groupe qui le mobilise localement (Star, 2010, p.22). Cette approche explique la réexploitation des documents scientifiques par une activité de réinterprétation et réappropriation des artefacts selon le projet propre aux groupes qui se l'approprient. On peut rapprocher ce processus de la circulation de l'objet culturel créatrice de sens à mesure des interprétations qui en sont faites, qui a été décrite par Yves Jeanneret comme trivialité (2008). L'élaboration progressive des catégories de la taxonomie se comprendrait ainsi comme un processus de trivialité contrôlée, qui se caractériserait par des réinterprétations encadrées par des normes disciplinaires, et productrices de nomenclatures consensuelles (par exemple la classification APG III utilisée au MNHN, et produite par le groupe de botanistes « Angiosperms Phylogeny Group » à partir de l'association de techniques moléculaires et de l'analyse cladistique). Différentes réinterprétations des documents d'herbier restent possibles dans d'autres communautés scientifiques, avec une approche historique par exemple, indépendante de la catégorisation par cette classification phylogénétique.

De ces différentes études ressort la nécessité de prendre en compte les enjeux communicationnels du travail documentaire, dont ne rend pas compte une approche par la seule centralisation des inscriptions. Nous cherchons à comprendre à présent l'intégration du traitement documentaire dans l'ensemble des activités de production et de communication de connaissances qui mobilisent le document d'herbier — principalement en systématique (l'approche qui vise à décrire et classer les espèces vivantes) et en écologie (qui prend davantage en compte les interactions entre espèces en lien avec le milieu). Cela revient à décrire la communication scientifique des disciplines qui s'appuient sur les herbiers, en prenant en compte les enjeux communicationnels du travail documentaire intégré à l'activité des chercheurs. Nous mobilisons pour cela une analyse des formes de communication à l'œuvre dans la constitution et l'exploitation des collections scientifiques, centrée sur l'examen de l'activité des différents acteurs qui y participent et des pratiques documentaires qui sont les leurs.

2.2. Les médiations de la communication scientifique mobilisant les collections

2.2.1. La communication scientifique appréhendée comme un ensemble de transactions

Après avoir montré les enjeux communicationnels à l'œuvre dans la constitution et la réappropriation des spécimens des collections, nous les situons dans « le système de communication scientifique » entendu comme l'ensemble des médiations

« entre tous les acteurs présents et futurs de la recherche » (Beaudry, 2011, p.21). Les collections sont utilisées dans la production, la diffusion, et la conservation qu'assure ce système, que Guylaine Beaudry analyse en huit fonctions : rédaction, évaluation par les pairs, édition, fabrication, diffusion et distribution auprès des lecteurs, appropriation, et conservation (p.31). Seules les fonctions d'édition ne concernent pas les spécimens. Les collections permettent de stocker et de rendre accessible les documents examinés en amont d'une étude aux côtés des publications. L'information lue ou prélevée sur le spécimen est la base de l'élaboration des connaissances, que nous entendrons comme résultat d'un « travail productif des sujets sur eux-mêmes pour s'approprier des idées ou des méthodes » (Jeanneret, 2000), et qui se fonde en l'occurrence sur l'examen du spécimen. L'accès aux spécimens est également essentiel à l'évaluation des publications. Les collections participent ainsi à la légitimation des connaissances, et à leur intégration au terme du processus de validation et de publication dans des « ensembles organisés de connaissances cumulées et durables » qui définissent les savoirs selon Jean Meyriat (1981), et dont Yves Jeanneret (2000, p.85) précise qu'ils sont « reconnus par une société », où ils « reçoivent un statut répertorié ». Outre les fonctions recensées par Guylaine Beaudry, les collections sont aussi le vecteur de la communication des savoirs scientifiques à des publics non spécialistes, puisque les musées d'histoire naturelle ont servi à la diffusion des connaissances sur le vivant, et ont intégré des amateurs à leur constitution — à différents degrés selon les époques toutefois : l'intégration des pratiques amateurs a été rendue plus difficile à partir de l'adoption de la taxonomie d'Antoine-Laurent de Jussieu (Rasse, 2013), alors que les dispositifs numériques de collecte favorisent au contraire de nouvelles modalités de participation (Charvolin, 2004; Heaton *et al.*, 2011). Nous présentons à présent un cadre d'analyse permettant de prendre en compte les spécificités de la communication à l'œuvre dans les activités mobilisant les collections, activités caractérisées par des dispositifs documentaires visant à faciliter l'accès à l'information scientifique et enrichis collectivement par des contributeurs de statuts, compétences et centres d'intérêts variés.

Pour décrire les formes de communication à l'œuvre dans les activités de conservation et d'exploitation des spécimens et les traitements documentaires dont elles sont indissociables, nous mobilisons le cadre théorique de la sémiotique des transactions coopératives, conçu par Manuel Zacklad (2013) pour appréhender les activités collectives dans leur dimension communicationnelle, et faire droit au rôle des différents artefacts sémiotiques (c'est-à-dire pouvant faire l'objet d'une interprétation) qu'elles utilisent et participent à construire. Au-delà de leurs finalités respectives, les activités des usagers de l'herbier ont une dimension collective, en ce qu'elles partagent l'objectif commun d'intégrer les matériaux du terrain et les expertises aux collections existantes. La sémiotique des transactions coopératives (STC) est particulièrement adaptée pour rendre compte de l'activité scientifique qui participe à transformer les collections qu'elle mobilise, puisque l'herbier, régulièrement restauré, complété de nouvelles étiquettes qui sont ajoutées aux spécimens en conclusion d'un examen, réorganisé selon un nouveau classement, se caractérise par les transformations qu'il peut subir lors de son examen contrairement au document d'archive administrative utilisé comme source historique.

La STC appréhende justement l'activité intellectuelle comme une transformation double, qui opère au niveau symbolique, par l'élaboration de nouvelles significations ou l'acquisition de connaissances, et, de manière indissociablement liée, au niveau des artefacts, qui sont modifiés au cours de l'activité (par exemple par de nouvelles annotations). Manuel Zacklad nomme *transaction* cette transformation réciproque, qui peut être le fait d'un seul sujet utilisant un artefact (par exemple un livre, ou un outil d'écriture permettant d'objectiver l'évolution de sa pensée), ou un collectif produisant ou transformant différents artefacts à mesure que ses membres élaborent des connaissances. Par rapport à la notion d'interaction, la transaction insiste aussi sur son efficacité, en soulignant les transformations qu'elle produit, qu'elles soient matérielles, symboliques, ou même sociales (dans le cas de la transformation des relations entre les parties prenantes impliquées dans l'action). Grâce à cette approche, nous pourrions étudier le processus de « trivialité » spécifique au cas de l'herbier, en montrant comment, à mesure de la circulation sociale du document de collection et des interprétations qui en sont faites, sont générées des connaissances scientifiques par les acteurs et collectifs qui les mobilisent, corrélativement aux transformations qu'ils leur apportent ; et comment la valeur patrimoniale de document qu'il y a obligation de conserver (Davallon, 2006) émerge également de ce processus.

La STC permet de caractériser les différentes valeurs que prend le document dans l'activité, selon qu'il est considéré comme l'aboutissement d'une performance, ou utilisé dans différentes sous-transactions intermédiaires comme une ressource pour atteindre un objectif. Cette distinction permet de préciser la notion de document par intention pour prendre en compte différentes intentions constitutives des interprétations dont fait l'objet l'herbier selon les phases du processus d'action collective dans lequel il s'inscrit. En effet, l'herbier est bien considéré comme une fin ou un moyen, selon qu'il est déposé au terme d'une recherche pour conservation dans une institution, ou consulté au titre de l'état de l'art ou pour prélever des données sur lesquelles s'appuieront de nouvelles analyses. La prise en compte de ces réappropriations du document avec différentes significations qui lui sont données est importante pour comprendre la communication scientifique mobilisant les mêmes documents d'herbier d'étude en étude, et penser la remobilisation du document dans une finalité différente de l'intention de signification de son auteur — par exemple quand un usage historique est fait d'un document initialement produit dans une perspective scientifique.

La notion de médiation employée dans l'expression d'artefact médiateur désigne dans la STC une relation de co-constitution des supports et des productions cognitives ou sociales, en cohérence avec différentes théories de la communication médiatisée qui partagent l'attention portée au rôle essentiel des intermédiaires techniques dans la construction des significations, à l'opposé de l'idée de supports de l'information transparents à son sens. Ces théories partagent une conception de la communication, selon un « modèle de la médiation », comme « mise en œuvre d'un élément tiers rendant possible l'échange social, alors même que les univers de la production et de la réception sont *a priori* par nature disjoints » (Davallon, 2004, p.53). Elles s'opposent à un modèle linéaire de la communication comme transmission de l'information entre émetteurs et récepteurs, qui n'insisterait pas sur le rôle

des multiples intermédiaires, techniques et humains, nécessaire à permettre leur compréhension. La prise en compte du rôle des médiations matérielles de la communication s'inscrit aussi dans la continuité de l'emploi de la notion de médiation en philosophie, notamment dans l'ontologie hégélienne, qui décrit les processus de subjectivation à partir de la description du rôle qu'y acquièrent les objets et notamment les outils.

La STC fournit un cadre d'analyse commun pour appréhender, à travers les flux de transactions qui les composent, les activités des acteurs qui participent aux différentes étapes de la communication scientifique. Elle conduit ainsi à décrire la communication scientifique comme un ensemble de transactions coordonnées, qui produisent de nouveaux documents dans le cadre d'observations, de collectes, ou de publications, à mesure que les connaissances sont élaborées. Elle permet de questionner en particulier la complémentarité des transactions sur différents supports. Nous précisons à présent les principes d'analyse de l'activité issus de la STC, à partir desquels on dégagera des types de transactions mobilisant les documents des collections pour comprendre les formes de communication qu'elles médiatisent.

2.2.2. Caractériser les dispositifs, interprétations et collectifs à l'œuvre dans la communication mobilisant les collections

Nous emploierons le terme d'*activité* pour désigner les opérations entreprises par les usagers, observées selon la description qu'ils en font. Nous utiliserons le terme de *pratiques*, et notamment de pratiques informationnelles, ou de pratiques de lecture et d'écriture, pour dégager, à partir des observations des activités singulières, des gestes et des manières d'agir socialement partagées et constituées en objet d'étude dans la littérature scientifique. La description des pratiques procédera d'une recatégorisation des observations des activités telles que décrites par les acteurs dans les entretiens ; les catégories d'artefact documentaire, de transaction et de collectif propres à la STC jouent le rôle de « grille de structuration des données issues de l'enquête » (Le Marec et Babou, 2003) transformant la « précatégorisation » issue des discours des acteurs. Nous n'avons pas retenu pour notre analyse des activités observées la division « diachronique » de la réalisation de la performance (Zacklad, 2013), mais nous nous appuyons sur les dimensions d'analyse de l'activité collective comme ensemble de transactions et sous-transactions, comprises comme interactions socialement et symboliquement productives médiatisées par des artefacts et des dispositifs qui participent à les coordonner.

En effet, la STC propose de mettre en relation les caractéristiques des artefacts documentaires, au niveau de leur matérialité, avec les dimensions d'analyse sociale et sémiotique, c'est-à-dire « celle dans laquelle prédomine l'échange avec un tiers (composante sociale dominante) et celle dans laquelle prédomine l'échange avec soi » (Zacklad, 2007). De ce point de vue, elle ouvre des perspectives d'analyse similaires aux trois dimensions proposées pour l'analyse du document numérique par le collectif réuni sous le pseudonyme Roger T. Pédaque : l'artefact, dans sa

dimension matérielle d'objet anthropologique, est mis en relation avec les enjeux sociaux du médium, et « le contenu ou le texte » objet d'interprétation (Salaün, 2007). L'étiquette de l'herbier, envisagée au niveau de sa matérialité, comme annotation accolée sur la planche aux côtés du spécimen, pourra ainsi se caractériser au niveau de l'interprétation qu'elle rend possible (avec la prise en considération d'une expertise passée), qui passe par la maîtrise d'un vocabulaire normé nécessaire à donner du sens à l'identification taxonomique du document ; elle pourra aussi être envisagée dans sa fonction de coordination, l'emploi d'une norme taxonomique traduisant l'inscription de l'activité dans une discipline et la volonté de faciliter l'intégration du document à un ensemble institutionnel plus vaste. La prise en compte des formes de collectifs mobilisés à partir du *dispositif*, constitué de l'ensemble des artefacts intervenant dans les transactions, permet de justifier l'emploi de ce terme en référence à la signification que lui a donnée Foucault, en soulignant les formes de pouvoir qu'il configure, au-delà de son utilisation fréquente dans le sens d'une « logique d'agencement », selon la recension de Violaine Appel et Thomas Heller (2010). Ainsi, outre le fait que le terme souligne la diversité d'éléments, relevant plus ou moins directement du discours, qui composent les artefacts intervenant dans la communication, nous employons le terme en faisant référence à la capacité de l'ensemble des artefacts mobilisés par le collectif à configurer les relations entre acteurs qui le mobilisent, et notamment des inégalités de droits de contribution et de fixation des règles de l'action collective. En plus de fournir ce cadre d'analyse de la communication à partir de l'activité et des supports qu'elle mobilise à des fins d'interprétation et de coordination, la STC fournit des typologies pour caractériser les dispositifs et collectifs impliqués par les activités collectives. Nous nous en inspirons à présent pour mettre en évidence les spécificités de l'herbier et pouvoir les comparer ensuite avec celles des nouveaux dispositifs de médiation numérique, en envisageant les formes de collectifs impliqués dans les deux cas.

Dans les termes de la STC, le document d'herbier est constitué de l'enregistrement d'une observation (le spécimen séché), qui a fait l'objet d'une documentarisation, c'est-à-dire d'un traitement permettant « la réutilisation du support dans le cadre de transactions ultérieures, de la personne avec elle-même ou avec d'autres personnes » (Zacklad, 2007), en permettant « au document de circuler à travers l'espace, le temps, les communautés d'interprétation », en l'occurrence celle des écologues et des systématiciens, mais aussi celle des historiens, et encore à travers les expositions pour lesquelles les spécimens peuvent être prêtés. La documentarisation du spécimen comprend des éléments standardisés obligatoirement présents, à savoir les informations sur la collecte, et la détermination de l'espèce du spécimen quand elle a été identifiée, consignée dans une étiquette collée sur la *planche d'herbier* (feuille de papier, généralement de format A3, sur laquelle le spécimen est également épinglé ou attaché, par collage de bandelettes de papier). S'ils sont essentiels à l'interprétation scientifique du spécimen, ces éléments font aussi partie de la « documentarisation externe » du spécimen, en ce qu'ils permettent de le classer par rapport à d'autres : la classification taxonomique, jointe à des critères géographiques, régit souvent, comme c'est le cas au MNHN, le plan de classement dans l'espace de collections. Peuvent s'ajouter facultativement un dessin, un ex-

trait de la description qui a été faite du spécimen lors de sa première publication (protologue), des informations manuscrites sur le contexte d'observation ; un numéro peut correspondre également à un paragraphe d'un carnet de récolte comportant d'autres détails. L'organisation des connaissances qui régit les étiquettes a la particularité d'être issue de l'activité scientifique. La nouvelle identification d'un spécimen correspond à une opération de redocumentarisation, au sens où elle se traduit par l'ajout d'une nouvelle étiquette, associée à une nouvelle signification attribuée au document. La typologie des régimes de documentalité, qui sont autant de manière de « faire document », permet de préciser encore les spécificités de l'herbier.

Les planches d'herbier entreposées dans des liasses et des « compactus » supposent des recherches permettant de sélectionner les spécimens adéquats aux besoins. En ce sens, ils relèvent du régime de la « publication documentaire » supposant une demande ou une recherche de la part de l'utilisateur avant d'être consultée, à la différence des documents « poussés » vers lui, ou s'inscrivant dans le régime de la « circulation documentaire » (Zacklad, 2015). La diversité des étiquettes apposées à une même planche donne à l'herbier le caractère de document *fragmenté*, dans lequel « les sous-transactions et les productions sémiotiques associées sont mécaniquement articulées les unes aux autres par le biais d'un système de numérotation, d'un sous-titrage, d'un ensemble d'attributs particuliers visant à expliciter le statut des fragments dans le document et leur relation avec les fragments environnants ». Dans le cas de l'herbier, le collage des différentes étiquettes est le moyen d'annoter le spécimen et d'organiser différentes annotations explicitant les résultats scientifiques dont il a été l'objet (voir figures 2.2 et 2.3). Le document d'herbier présente un caractère inachevé, de nouvelles déterminations pouvant lui être apposées dans l'avenir. La notion de fragmentation s'applique aussi aux collections de l'herbier général, composées de planches séparées les unes des autres, permettant des réagencement selon l'évolution des catégorisations scientifiques et des besoins des usagers ; certains herbiers, plus rares, sont reliés, ou conservés dans l'ordre des liasses ou boîtes du collecteur, quand ils présentent une valeur historique.

Ces caractéristiques de fragmentation et d'inachèvement rapprochent l'herbier des documents constitués par et pour l'action d'un collectif (Zacklad, 2004), et qu'on peut caractériser selon les formes de contribution qui les constituent : étant donné le format standardisé des inscriptions des étiquettes selon les règles du code de nomenclature, il s'agit de contributions « orientées données ». Étant donné le caractère manuscrit des étiquettes, la dénomination « orientée données » ne doit pas sous-entendre qu'il s'agirait de données standardisées, puisque les inscriptions ne peuvent faire l'objet immédiat de calcul et d'exploitation informatisée. Mais leur forme déjà fragmentée et normalisée souligne l'accointance de l'herbier avec le document numérique.

Nous appliquons la typologie des formes d'action collective aux relations entre les membres du collectif participant à la production et à l'exploitation documentaire des collections. Parmi les situations d'activité collective, Manuel Zacklad distingue la *coordination*, qui ne suppose pas d'interactions entre membres, des situa-



FIGURE 2.2. – Planche d’herbier du jardin botanique de Lyon, scannée et mise en ligne sur [Flickr](#). Elle compte, en plus de l’étiquette de la collecte (en bas à gauche), une annotation signée et datée suggérant un statut probable de type.

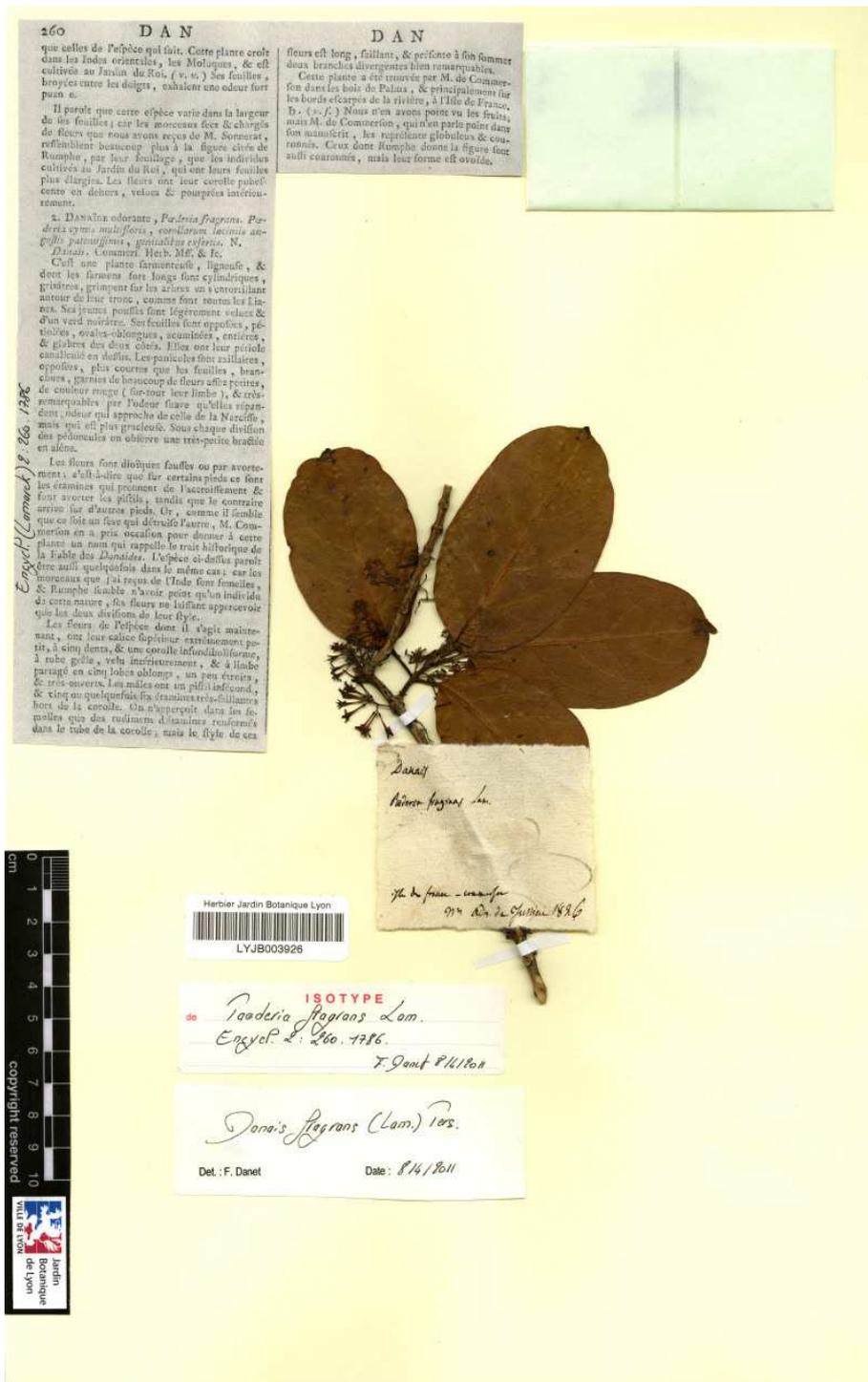


FIGURE 2.3. – Planche d’herbier du jardin botanique de Lyon, scannée et mise en ligne sur [Flickr](#). En plus de l’étiquette originale, et de la nouvelle qui signale le statut de type du spécimen, a été ajouté un extrait de l’*Encyclopédie de la botanique* de Jean-Baptiste Lamarck mentionnant le spécimen.

tions de *coopération*, dans lesquelles existe un objectif partagé qui nécessite d'entrer en interaction pour s'accorder sur les manières de l'atteindre et de se coordonner (Zacklad, 2015). Si on trouve des activités coopératives dans certains projets de recherche ou de gestion des collections, l'enrichissement et l'exploitation des collections tient davantage à des pratiques de dépôt et de consultation, « fortement asynchrones », dont la dimension collective tient au respect de normes documentaires communes, mais ne suppose pas de coopération entre usagers des collections. Toutefois, l'édifice des collections nécessite également des activités d'inventaire essentielles à permettre sa réexploitation, et la consultation des collections passe également par des interactions avec les conservateurs investis dans cette connaissance des collections. On peut considérer ces tâches, parfois réalisées par des bénévoles, comme un « travail d'articulation », caractérisé par Anselm Strauss comme « ce qui fait en sorte que le travail peut être accompli » (Strauss, 1988), et comportant notamment des activités de coordination souvent peu visibles. La notion est appliquée par Florence Millerand (2012) au travail de documentation et de constitution des bases de données de recherche en écologie ; elle est également mobilisée pour désigner les tâches à l'œuvre dans l'informatisation des collections d'herbier par Lorna Heaton et Serge Proulx (2012). On peut en rapprocher également les activités de nomenclature, spécifiquement consacrées à la mise en cohérence des dénominations employées par les uns et les autres, qui sont elles aussi essentielles à la cohérence de l'édifice des collections ; cependant, à la différence du sens classique de « travail d'articulation », elles donnent lieu à des publications et font ainsi partie de la production scientifique visible. Différents rôles apparaissent donc dans l'enrichissement et la conservation des collections, même s'ils peuvent selon les moments être dévolus aux mêmes personnes, les tâches de conservation étant réparties entre conservateurs et usagers, et en particulier certains d'entre eux — par exemple, les enseignants-chercheurs du MNHN ont statutairement des responsabilités de conservation d'un secteur des collections.

Les observations des activités utilisant l'herbier nous amèneront à préciser les transactions qu'elles mettent en œuvre au niveau des formes d'interprétation impliquées dans la lecture du document scientifique. Selon les cas, la notion de transaction sera précisée à partir de théorisations des formes d'interprétation de documents inhérentes aux activités intellectuelles, particulièrement dans les situations de recherche d'information, notamment collaborative. On privilégiera dans cette perspective la référence à des modèles de la recherche d'information qui peuvent être pensés comme la mise en œuvre d'activités d'interprétation associées à la transformation d'artefacts, parfois dans un contexte collectif. C'est en particulier le cas du modèle élaboré par Andrew Dillon (1996), qui souligne, parmi les compétences nécessaires à la recherche d'information, le rôle des « Manipulation Skills and Facilities » — « l'aptitude et l'habileté à manipuler », et notamment à naviguer (Kembellec, 2012, p.126). On retrouve aussi dans ce modèle la notion d'objectif de la recherche, qui correspond à la finalité de la transaction, et deux notions d'interprétation du document, la prévisualisation permettant de localiser l'information, et la lecture à proprement parler. Pour appréhender sa dimension collective, nous convoquerons également l'approche de la recherche d'information comme un dialogue différé (Tricot et Comtat, 2012), et plus précisément « un double dialogue,

entre les chercheurs d'une part, entre les chercheurs et le document d'autre part ». Le dialogue conduit à distinguer pour les participants des rôles, qui apparaissent dans la plupart des situations de recherche d'information collaborative observées : on pourra bien appréhender ce dialogue comme une succession de transactions, au sein d'un collectif éphémère fédéré autour de l'objectif de recherche. La STC guidera l'analyse des observations des activités mobilisant les collections ; elle permettra de préciser les caractéristiques documentaire de l'herbier et des formes de communication qu'il organise, entre acteurs participant à une production documentaire coordonnée, qui configure leurs pratiques d'écriture selon les normes nécessaires à faciliter leur archivage et leur intégration aux ensembles existants. Elle permettra d'organiser les résultats de nos observations en sorte de caractériser les spécificités des médiations intervenant dans l'élaboration des connaissances et la conservation des collections, selon que les bases de données de spécimens sont ou non utilisées. Elle rendra possible une comparaison des dispositifs de médiation documentaire constitués par les bases de données de spécimens d'une part, et l'herbier non numérisé d'autre part, qu'elle permet dès à présent de caractériser.

2.3. Première caractérisation d'un modèle de médiation documentaire augmentée à l'aune de l'herbier

2.3.1. Enjeux de médiation de savoirs et de coordination de l'activité collective de la médiation documentaire

Mobiliser la STC pour appréhender la constitution collective de l'herbier demande de penser l'articulation de plusieurs conceptions de la médiation documentaire, qui sont souvent dissociées : Manuel Zacklad (2015) distingue par exemple son approche de la médiation documentaire considérée dans la perspective d'une médiation des savoirs (Liquète *et al*, 2010) et des acceptions de la médiation numérique qui s'inscrivent également dans cette perspective. Nous précisons ces différentes approches avant de montrer leur complémentarité pour penser la médiation documentaire propre à l'herbier, dont nous postulons l'intérêt pour concevoir des dispositifs de médiation numérique jouant un rôle de mobilisation de publics hétérogènes dans l'enrichissement des fonds numérisés.

Premièrement, dans la tradition des métiers de la documentation, la médiation documentaire désigne l'ensemble des dispositifs facilitant l'accès à l'information scientifique, et questionne les conditions de son appropriation. Elle conduit à une conception de la documentation centrée sur l'accompagnement de l'utilisateur et l'étude des conditions de l'appropriation de l'information en vue de l'élaboration ou de l'acquisition de connaissances : « en se centrant sur la médiation, il ne s'agit donc pas de penser la circulation, ni les interactions, mais réellement d'accompagner la transformation de l'information par le récepteur » (Liquète *et al*, 2010). Cette approche conduit ensuite à considérer la documentation comme une

médiation des savoirs (Liquète *et al.*, 2010), qui passe par des actions visant à faciliter l'appropriation de l'information. Le cas de l'herbier conduit à identifier le rôle des artefacts documentaires dans l'élaboration ou l'acquisition de connaissances, au sein de démarches scientifique, d'enseignement, ou de diffusion de savoirs à un public non spécialiste. C'est ainsi que les étiquettes, éléments essentiels pour identifier un objet au sein d'une collection, ou les *compactus* (étagères regroupant les spécimens en liasses), peuvent être vues aussi bien comme un moyen de trier pour retrouver les documents que pour apprendre ou augmenter les connaissances produites par comparaison.

Rendre possible la participation d'amateurs au travail de conservation ou aux collectes suppose encore dans le cas de l'herbier de rendre possible la communication des documents à un public spécialiste ou non : c'est là un geste essentiel de la médiation documentaire qui consiste à établir « un système de communication adapté à l'utilisateur » (Liquète *et al.*, 2010), quel que soit son statut. Dans cette perspective, les collections peuvent se voir inscrites dans d'autres dispositifs construits à partir de la médiation documentaire, comme les dispositifs pédagogiques visant à favoriser l'acquisition de savoirs par l'observation des spécimens, les visites de collections universitaires à l'université de Toulouse ayant montré l'importance des spécimens dans certains enseignements de biologie. L'exposition de certains objets des collections demande encore de penser l'articulation entre médiation documentaire et muséale, dont l'objectif de « faciliter le travail d'appropriation de connaissance » (Frayse, 2015) rejoint directement la fonction de médiation des savoirs.

Le dispositif de médiation documentaire propre à l'herbier se distingue encore par son imbrication dans l'ensemble des supports d'écriture mobilisés par les chercheurs : par le système d'étiquetage, les mêmes documents ayant servi de support d'analyse et de prélèvement sont réintégrés aux collections. On peut préciser ces spécificités en reprenant la métaphore de la « chaîne lectoriale » (Bottini, 2010, p.19), en gardant toutefois les mêmes précautions que celles formulées par Yves Jeanneret (2008) à l'égard de cette image, qui conduisent à souligner les opérations d'interprétation associées à chacune de ces étapes : dans le cas de l'herbier, on constate la présence des documents des collections aussi bien dans les opérations de lecture, d'annotation et de tri de la phase d'analyse, qu'en conclusion de la phase de la synthèse, par la publication des données associées aux publications – même si d'autres supports complémentaires s'ajoutent aux spécimens dans le travail personnel du chercheur aboutissant à une publication.

À cette conception déjà complexe de la médiation documentaire comme médiation des savoirs, articulée avec d'autres médiations des espaces d'enseignement, d'exposition ou du travail scientifique, l'herbier demande d'articuler le rôle des dispositifs de médiation documentaire dans l'organisation collective. Manuel Zacklad envisage les dispositifs de médiation dans leur capacité à servir de support à l'activité collective, certains d'entre eux constituant des documents par leur pérennité et leur capacité à pouvoir être identifiés et réutilisés dans un processus d'action collective. L'analyse des transactions permet une approche des enjeux organisationnels des dispositifs documentaires en décrivant leur caractéristiques en

relation avec celles des collectifs qui les mobilisent (Zacklad et Chupin, 2015). Le cas de l'herbier permet de préciser l'articulation entre ces approches de la médiation documentaire, centrées sur une problématique d'accès et d'appropriation de l'information d'une part, et de coordination de l'activité collective d'autre part, en faisant ressortir la fonction des dispositifs de recherche et d'accès aux documents dans la coordination de l'activité collective : l'organisation des connaissances définie à partir des normes taxonomiques régit non seulement le classement dans l'espace des collections pour faciliter l'accès aux spécimens selon les besoins scientifiques, mais aussi les normes d'écriture sur les étiquettes, ce qui permet d'ouvrir aux usagers l'archivage des conclusions de leurs examens, consignées dans une nouvelle étiquette et rendues ainsi disponibles pour ceux qui s'y intéresseront ultérieurement. Le cas de l'herbier apporte donc un éclairage spécifique sur les relations entre techniques documentaires et production des savoirs dans un contexte collectif. Nous convoquerons conjointement ces différentes approches pour appréhender les médiations de la communication scientifique mobilisant les collections, en prenant en compte à la fois les conditions de leur production par de multiples acteurs, de l'interprétation des documents selon les dispositifs qui y donnent accès, dans les espaces de recherche ou d'exposition.

2.3.2. Hypothèse de recherche : un dispositif de médiation numérique peut reconstituer et augmenter les fonctions de l'herbier et amplifier son rôle dans la communication scientifique

L'examen des spécificités des médiations documentaires constitutives des collections ainsi que de leur rôle dans la communication scientifique conduit à une conceptualisation qui complexifie la notion de médiation documentaire, qu'il faut appréhender en faisant droit à ses aspects matériels (médiation instrumentale), à ses enjeux organisationnels, et symboliques (médiation des savoirs scientifiques pour différents publics). La caractérisation des spécificités de l'herbier comme dispositif de médiation documentaire produit collectivement et essentiel à l'activité scientifique donne la mesure du défi que représente sa numérisation si on cherche à ne pas appauvrir, mais au contraire à enrichir le travail avec les spécimens : il s'agit de reconstituer et de développer davantage avec les images les opérations dont l'herbier est le support, y compris celles qui sont pris en charge par des amateurs. L'herbier constitue en cela un modèle pour dépasser une conception de la médiation numérique répondant à la seule fonction de mise à disposition des documents. L'objectif de l'infrastructure e-ReColNat de favoriser les usages scientifiques de l'image se traduit dans l'hypothèse de la possibilité d'une médiation numérique reconstituant les fonctions de l'herbier. Cette hypothèse a pour corollaire la définition d'un dispositif numérique de médiation documentaire capable d'assurer non seulement un accès facilité aux documents, mais aussi la coordination de l'activité collective participant à leur enrichissement.

Pour vérifier cette hypothèse, nous commencerons par préciser l'esquisse d'un premier modèle d'une médiation numérique des collections, défini pour le moment

d'une manière fonctionnelle, en examinant notamment des expériences ayant contribué à reconstituer certaines fonctions d'annotation collaborative pouvant correspondre à celles de l'herbier. Nous revenons donc à présent sur des théories et expériences de numérisation permettant de dégager les conditions de possibilité de la réalisation de notre hypothèse. Nous préciserons également ses enjeux théoriques concernant l'articulation de différentes traditions professionnelles de médiation (documentaire, culturelle et scientifique), que les dispositifs donnant accès aux fonds numérisés relie encore plus étroitement.

3. Enjeux de la numérisation de l'herbier pour la médiation documentaire

Prendre la mesure des conditions de possibilité de la reconstitution et de l'augmentation des fonctions de l'herbier sur supports numériques conduit à mobiliser conjointement des théorisations élaborées pour penser la numérisation de documents scientifiques, contributifs, mais aussi historiques et patrimoniaux. Ces théories partagent en commun une appréhension riche de la numérisation, qui ouvre des perspectives sur ses enjeux communicationnels, à l'opposé d'une conception qui en ferait une simple conversion sans conséquence pour l'interprétation des documents. Nous revenons d'abord sur une théorisation de la numérisation envisagée dans ses relations avec la préservation et la lisibilité des documents, qui fonde la légitimité d'un travail de conception visant à reconstituer les conditions de la lecture et les médiations matérielles du travail scientifique ; nous envisagerons ensuite les conditions nécessaires à recréer l'accès à l'information scientifique et l'appropriation des fonds numérisés par différents publics. Ces différents enjeux seront ceux que nous envisagerons pour la numérisation de l'herbier, et qui permettront de préciser les caractéristiques d'une médiation numérique capable d'amplifier le rôle des collections dans la production des savoirs.

3.1. Enjeux d'authentification du spécimen numérisé et conditions instrumentales de son exploitation scientifique

Les distinctions établies par Bruno Bachimont (2001) pour envisager la numérisation et ses enjeux sur l'interprétation du document permettent de préciser les conditions de possibilité de l'exploitation scientifique de l'herbier numérisé. Nous appliquerons à l'herbier l'analyse des transformations du document papier subi par le numérique : avant la numérisation, le support et la forme d'enregistrement, qui sont respectivement le « support que l'on charge de préserver et conserver le contenu inscrit » (par exemple la page du livre), et la modalité de l'inscription sur le support (les caractères respectant un certain format), sont confondus avec le support et la forme d'appropriation (le livre comme objet matériel facilement consultable rendant possible sa lecture, et le texte comme objet d'interprétation). La numérisation de l'herbier introduit de nouveaux intermédiaires entre le support et la forme d'enregistrement (respectivement, la feuille de papier et le spécimen séché considérés dans leur fonction de stockage de l'information d'une part, et les caractères de l'étiquette et du spécimen, alphabétiques et morphologiques, d'autre

part), et ses formes physique et sémiotique d'appropriation (respectivement, l'ensemble des pixels affichés à l'écran, et l'image constitué par le spécimen numérisé dans sa capacité à être identifié à partir de ses caractères morphologiques, ainsi que les étiquettes destinées à être interprétées). La prise en compte des nouvelles médiations techniques à l'œuvre dans l'enregistrement sur support numérique sous la forme d'un fichier image (qui constitue le support d'enregistrement du spécimen numérisé), et sa restitution sous une forme d'appropriation reconnaissable (l'image du spécimen) se traduit par l'élaboration d'une procédure visant à garantir et manifester l'authenticité de la forme d'appropriation du spécimen numérisé, par rapport à celle du spécimen papier. Ce contrôle est essentiel pour conférer à l'herbier numérisé une valeur de preuve au moins en partie comparable à celle du spécimen initial : « la légitimité d'une archive repose sur son authenticité, c'est-à-dire sa capacité à pouvoir prouver que le document conservé est un témoin vérac et fidèle du document tel qu'il a été produit selon ses conditions propres de production et de réception » (Bachimont, 2001, p.10). Dans le cas de l'herbier, il s'agit en particulier de prouver le respect de la taille du document initial, nécessaire à l'exploitation morphométrique du spécimen numérisé, c'est-à-dire procédant par relevé de mesures. Cela se traduit dans la plupart des bases de données d'herbiers par un indicateur de taille du spécimen. Dans le cas du projet « Global Plants Initiative » ont été observées une « échelle de calibrage des couleurs et une règle métrique sur la page de fond » qui constituent un « mécanisme de compensation propre au format numérique », « l'affichage de couleurs pouvant varier d'un moniteur à l'autre » (Heaton et Millerand, 2013) ; on retrouve dans les spécimens issus de la chaîne de numérisation utilisée par le MNHN et e-ReColNat « la règle métrique [offrant] un repère en cas de zoomage ». Malgré ces moyens de contrôle, il reste que le spécimen pouvait donner lieu à des interprétations et des résultats plus riches que ceux qui sont possibles à partir de la prise de vue de sa surface. La transformation du support d'enregistrement de l'inscription dans le cas de l'herbier va avec un amoindrissement de la richesse informationnelle de la forme d'appropriation, l'information étant constituée non seulement d'un « regard » porté sur le spécimen (Jeanneret, 2000), comme dans le cas du texte, mais parfois de manipulations comme la palpation de la texture, la réhydratation et la dissection, la constitution de lames destinées au microscope à partir de prélèvements, d'après les témoignages des chercheurs rencontrés au MNHN.

La question des nouvelles formes d'exploitation de l'image ne saurait être tranchée *a priori*, mais demande d'observer les pratiques effectives et la manière dont se constitue une tradition d'interprétation (Bachimont, 2001) de l'image d'herbier qui invente et intègre progressivement les opérations que le spécimen numérisé rend possible. La liste, liée aux opérations de tri et de catégorisation de la « raison classificatoire » (Bachimont, 2004, p.104) associée au support papier, correspond à des gestes fondamentaux opérés dans le travail de systématique qui procède en recatégorisant les spécimens. En suivant l'interprétation développée par Bruno Bachimont des enjeux cognitifs des supports numériques, la question de l'exploitation scientifique du spécimen numérisé pourra dépendre du développement de programmes *ad hoc*, qui rendront possibles l'exploitation automatisée des listes, ce qui constitue l'opération cardinale de la « raison computationnelle ». L'utilisation

des spécimens numérisés dépendra aussi des nouvelles façons d'en « instrumenter la lecture », qui rejailliront sur les prérequis des logiciels visant à en favoriser la possibilité (Bachimont, 2004, p.197).

3.2. La reconstitution de la médiation documentaire des collections et ses enjeux

Les nouvelles formes d'exploitation des documents sont également conditionnées par la possibilité de les retrouver parmi les millions de documents uniques des collections, ce qui suppose un travail de documentarisation des images issues de la numérisation, reconstituant celui qui a été réalisé depuis plusieurs siècles au moyen des étiquettes (comme exposé dans le chapitre précédent). La constitution d'une base de données constitue une réelle valeur ajoutée pour l'herbier qui ne disposait pas de catalogue avant la numérisation. La reconstitution de dispositifs de médiation documentaire permettant l'accès aux documents numériques suppose un investissement dans de nouveaux traitements; il s'agit donc d'une re-documentarisation des collections, au sens de nouveau traitement documentaire (Salaün, 2007). Ces nouveaux traitements documentaires supposent aussi de « construire ou développer des techniques de gestion documentaire » (Salaün, 2007), comme l'illustre l'élaboration du standard de description du Darwin Core, élaboré pour le partage des données sur la biodiversité, créé par le Biodiversity Information Standards (TDWG). Lorna Heaton et Serge Proulx (2012) ont montré l'importance des normes internationales, différentes des habitudes locales, dans les activités d'informatisation des collections.

Au nombre des enjeux de ces nouvelles techniques documentaires, évoqués par Jean-Michel Salaün comme les conséquences « plus larges » de la redocumentarisation dans le contexte de la multiplication des documents numériques, il faut compter, dans le cas des collections d'histoire naturelle, l'élaboration d'un modèle économique rendant soutenable le nouvel effort documentaire. Les techniques documentaires du web mettent dans la position de *contributeurs* des millions d'utilisateurs d'Internet, étant donné que leurs traces de navigation « sont la source de la création de la valeur économique des géants » (Proulx, 2014, p.21). Mais les outils numériques facilitant l'agrégation des productions et l'activité collective nourrissent plus précisément « la contribution en tant que forme sociale ». Celle-ci désigne la mobilisation d'internautes à des réalisations collectives diverses, fondées sur les relations horizontales, associées à « un univers d'attentes normatives communément partagées », parfois prolongé par un sentiment d'appartenance à une communauté, et qui requièrent de leur part une participation souvent « modeste et anonyme » (Proulx, 2014, p.21). C'est cette forme sociale, capable de s'insérer dans l'économie marchande, qui est aussi susceptible de rendre soutenables les défis de la numérisation et de l'informatisation des collections scientifiques et patrimoniales. C'est aussi à elle que nous nous référerons pour en préciser les enjeux dans le cas des collections d'histoire naturelle, en parlant d'internautes contributeurs pour désigner les usagers participant à enrichir la documentation des collections

d'herbiers numérisés. L'externalisation vers la foule (crowdsourcing), appliquée non plus dans une perspective d'accroissement de la valeur ajoutée de l'entreprise, mais comme une contribution volontaire bénévole visant à favoriser l'accès à l'information et au patrimoine scientifique, a été envisagée dans cette perspective.

3.3. De la médiation documentaire à la médiation culturelle

3.3.1. De nouvelles complémentarités entre deux traditions professionnelles distinctes

Outre des enjeux d'accès aux documents scientifiques, la numérisation et la mise en ligne des collections a été pensée comme un moyen de constituer une opération de médiation du patrimoine et de la culture scientifique pour de nouveaux publics. Dans le cas de l'herbier, cela revient à envisager un dispositif numérique capable de reconstituer les modalités d'exposition des collections des muséums dans un but de transmission de connaissances taxonomiques à des non-spécialistes (Rasse, 2013). L'exposition numérique des collections pourrait même être perçue comme l'un des moyens d'enrayer la crise qu'a connue la muséologie fondée sur une présentation des collections régie par la taxonomie. Ce projet d'ouverture des fonds spécialisés à des publics diversifiés a souvent été associé aux entreprises de numérisation, comme l'avait déjà montré Marie Després-Lonnet (2000) dans le cas de la réalisation d'une base de données iconographique par un centre de recherche en histoire, qui avait donné lieu dans une seconde phase à un projet d'ouverture de la base à « toute personne intéressée par l'histoire du Nord ». Cette tendance a de plus en plus conduit à intégrer des éléments de la documentation des collections comme parties prenantes d'une exposition, parfois virtuelle, et donc intégrées dans des fonctions de médiation culturelle (Després-Lonnet, 2014b). Comprendre les enjeux de cette démarche suppose de revenir sur les fondements théoriques des différents types de médiation en jeu dans ces projets.

Selon Marie Després-Lonnet (2014b), la documentation est réalisée en vue « d'exigences d'inventaire et de connaissance érudite des collections », souvent internes à l'institution, et se distingue des actions destinées à faire découvrir un patrimoine à des publics caractérisés par leur diversité (qui relèvent davantage de l'objectif de la médiation culturelle). Isabelle Fabre (2012) rappelle la spécificité du point de vue adopté pour décrire les objets des collections par les documentalistes, qui se démarque des points de vue disciplinaires par son effort d'objectivité et de normalisation. Documentation et médiation culturelle ne s'opposent pas, mais agissent en complémentarité, comme en témoigne le rôle des étiquettes et de différents écrits d'exposition réalisés par des acteurs de la médiation culturelle à partir de la documentation, qui en constitue un fondement nécessaire. Selon Julie Desjardins et Daniel Jacobi (1992), l'étiquette est constitutive du sens de l'objet dans la collection : « avec la disparition de son étiquette, l'unité de la collection est privée tout à

la fois, de sa mémoire, de sa place dans l'inventaire et en définitive c'est sa signification qui devient incertaine ». L'étiquette, constitutive de la valeur documentaire de l'objet, devient aussi, après des réécritures plus ou moins importantes, un élément essentiel dans le dispositif de l'exposition, assurant une « médiation entre l'objet et le visiteur » et contribuant « à l'élaboration du sens ». Cette complémentarité est renforcée par les formats numériques des productions documentaires, qui facilitent leur mise en ligne à destination des publics des musées. Marie Després-Lonnet (2014a) a montré comment des expositions virtuelles pouvaient s'appuyer sur un parcours de navigation proposé à partir de notices de bases de données patrimoniales, le visiteur étant redirigé sur les notices catalographiques quand il consulte les œuvres présentées au sein du parcours. Isabelle Fabre (2012) décrit, par l'observation de l'activité du service de documentation d'un musée d'art, comment un blog complète des actions classiques de composition de dossiers documentaires ; la médiation documentaire s'y prolonge, plus librement, par des « médiations cognitives, médiations actives » et « créatives ». Les dispositifs numériques renforcent les proximités entre deux démarches qui ont en commun l'objectif de « faire accéder un public à des œuvres (ou des savoirs) », en « construisant une interface entre ces deux univers étrangers l'un à l'autre (celui du public et celui, disons, de l'objet culturel) dans le but précisément de permettre une appropriation du second par le premier », même si les modalités d'action diffèrent : ces objectifs décrits par Jean Davallon (2004) pour la médiation culturelle peuvent aussi correspondre à des finalités de médiation documentaires, mais ce sont les moyens d'action qui les séparent (avec pour la médiation documentaire une spécialisation sur les traitements organisant les documents dans l'espace, les différents outils de recherche documentaire et d'accompagnement de l'utilisateur dans leur utilisation).

Ces nouvelles formes d'exploitation de la documentation, plus directement accessibles aux visiteurs grâce à leur support numérique, supposent toutefois un travail spécifique pour surmonter la distance qui peut exister entre leur point de vue et celui des professionnels de la documentation. De même que l'étiquette de musée est issue d'une réécriture plus ou moins importante des informations issues de l'inventaire (Desjardins et Jabobi, 1992), utiliser la mise en ligne du catalogue pour donner accès au patrimoine numérisé suppose un travail de médiation permettant de dépasser l'obstacle constitué par les langages documentaires professionnels pour ceux qui les ignorent (Després-Lonnet, 2000), par exemple en organisant des parcours de navigation. L'intégration des publics à certaines tâches intervenant dans la gestion des collections est un moyen de renforcer encore davantage les liens entre fonctions culturelles et informationnelles des produits documentaires.

3.3.2. Enjeux pour les médiations culturelle et scientifique de l'ouverture à la participation des activités documentaires

De fait, de nombreux projets participatifs ont été développés dans l'idée que l'ouverture à la participation des activités de traitement documentaire pouvait être un moyen de renouveler la médiation culturelle en exploitant de nouvelles possibilités d'interactions avec les publics. L'enjeu de la participation des internautes à la

documentarisation de fonds numérisés, en plus de la recherche de « l'efficacité documentaire » (Casemajor Loustau, 2011), entre dans une perspective de médiation culturelle si on prend la mesure de ses effets sur la mobilisation de nouveaux publics : Mia Ridge (2013) signale l'intérêt que présentent, pour tisser des liens forts entre l'institution et ses usagers, des dispositifs qui proposent des tâches évolutives, orientées vers la réalisation de projets proposés par l'institution détentrice des collections ou en association avec une équipe de recherche en vue de leur exploitation. De même, dans le cas des archives, Pauline Moirez (2012) évoque un souci des professionnels d'organiser une « véritable dissémination massive des fonds d'archives, et d'initier un travail scientifique collaboratif sur celles-ci, à distance », plutôt que d'attirer le public en salle de lecture, ce qui différencie leur action de médiation de celle des musées.

L'objectif d'intégration des contributeurs non spécialistes pose en de nouveaux termes la question des conditions de la communication de documents provenant de bases de données conçues dans des logiques professionnelles à l'égard d'un public non spécialiste, puisque c'est sur ces codes documentaires professionnels ou disciplinaires que doit porter l'action de communication adaptée à ce nouveau public. La mise à disposition de fonds spécialisés ouvre déjà la perspective d'étudier leur appropriation par de nouveaux publics, et de « mesurer les conséquences d'un déplacement de la définition des publics visés, du spécialiste vers le grand public », notamment dans les formes de perception et de catégorisation spontanément produites des fonds mis en ligne (Després-Lonnet, 2000). Dans le cas des dispositifs participatifs, les outils d'écriture proposés aux internautes en vue de l'intégration de leurs contributions aux bases de données des collections sont susceptibles de modifier le statut traditionnel du public, puisque selon Joëlle Le Marec (2005, p.148), « ne rien inscrire est une des caractéristiques du public en tant qu'instance de réception ». Si l'on cherche à situer les modalités de communication des savoirs à l'œuvre dans les projets de documentarisation participative par rapport à des modèles de communication des sciences, on peut constater une situation de communication inversée par rapport aux formes de médiation des savoirs scientifiques qui portent traditionnellement sur des connaissances déjà produites et validées par les chercheurs (Jacobi, 2007), qu'il s'agisse de les transmettre ou d'en faire discuter les enjeux pour la société. Participer à des tâches documentaires conduit les internautes à produire une partie des données qui constitueront le catalogue des collections, alors que c'est notamment sur lui que s'appuient les concepteurs d'exposition. Les revues de littérature sur les sciences participatives et les formes de participation qu'elles encouragent (Houllier *et al.* 2016 ; Haklay, 2013) sont susceptibles de fournir des directions d'analyse pour cette nouvelle mobilisation d'un public non-spécialiste dans une production documentaire concourant à enrichir un fonds destiné à la recherche. Le renversement à attendre dans la position du contributeur, par rapport à la position du public dans les modèles de communication des sciences reposant sur une posture de vulgarisation, reste donc à questionner, selon les modalités de participation proposées. Le renouvellement des formes de médiation culturelle et scientifique opérées dans les démarches documentaires ouvertes à la participation dépend des modalités de participation des contributeurs à la production des savoirs, qu'il s'agisse, selon la typologie de Muki Haklay (2015)

d'un simple travail de collecte (qualifié de « crowdsourcing »), de la participation à l'interprétation des données (« intelligence distribuée »), ou de la participation à la définition des problèmes (sciences participatives), ou encore d'une intégration des contributeurs citoyens à toutes les phases de la recherche. Les formes d'encadrement de la participation utiles à produire des données conformes aux normes institutionnelles et exploitables pourraient concourir à reconstituer dans les plateformes participatives les caractéristiques d'une médiation scientifique fondée sur la transmission d'un savoir validé par l'institution à ses publics. Dans le cas de l'herbier, les formes de participation à la démarche scientifique proposées par la plateforme (transmission permettant de participer aux collectes, ou intégration plus ambitieuse des contributeurs dans les ajustements en jeu dans la plateforme) pourront en plus être confrontées avec celles que prend la participation des bénévoles dans la constitution des collections.

3.4. La médiation documentaire augmentée à la croisée de médiations instrumentales, organisationnelles et culturelles

Ces différentes théorisations et expériences de numérisation permettent de préciser les questionnements multiformes que soulève l'hypothèse de la possibilité de reconstituer ou d'augmenter, à partir de dispositifs de médiation numériques, les fonctions de l'herbier dans la communication scientifique, à savoir : la médiation instrumentale de l'activité scientifique (avec un rôle de support de prélèvement et d'exploitation de l'information en vue de la production de nouvelles connaissances); la médiation documentaire envisagée comme facilitation de la mise à disposition et l'appropriation de l'information scientifique, qui suppose dans notre cas de penser de nouveaux modèles de financement des opérations d'indexation; et des formes de médiation culturelle et scientifique nécessaires à pouvoir intégrer des contributeurs à des tâches de collecte ou de conservation des collections, ou à faire valoir l'intérêt des collections pour des publics non spécialistes. L'hypothèse d'une médiation numérique capable de recomposer ces fonctions et registres de médiation correspondant à des traditions professionnelles distinctes, mais qui sont déjà exercées conjointement par l'herbier, est aussi fondée sur d'autres expériences attestant de la possibilité de ces rapprochements, comme on l'a vu avec les exemples de réinvestissement des dispositifs documentaires dans des démarches culturelles. On peut ainsi caractériser le type de médiation numérique que nous cherchons à concevoir pour l'herbier par la recherche d'une solution unique à ces différents défis : comment garantir les conditions de fiabilité de l'image à la planche d'herbier ? Comment rendre le travail documentaire soutenable économiquement ? Quelles démarches de médiation culturelle et scientifique inspirées des sciences participatives sont susceptibles de constituer une solution économique au problème de l'informatisation des collections, mais aussi une forme de valorisation pour un public non spécialiste ?

Nous qualifierons de médiation documentaire augmentée un dispositif capable de répondre à ces enjeux multiples, pour faire droit à l'idée que les fonctions d'ac-

cès aux documents scientifiques restent au centre du dispositif, la documentation des images étant un préalable aux possibilités de collaboration, de centralisation accrue, de tri ou de visualisation des spécimens développées en plus dans certaines bases de données. Nous reprenons la notion d'augmentation de la médiation au champ de la muséologie où la notion est appliquée aux renouvellements de la médiation culturelle intégrant des dispositifs numériques – par exemple dans le cas du développement d'une application de la Cité nationale de l'histoire de l'immigration, qui permet à partir d'un dispositif « mobile, interactif et personnalisé » de « personnaliser la médiation institutionnelle » (Flicoteaux et Kinoshita, 2014, p.327). Cette notion de médiation augmentée a aussi été utilisée pour caractériser le rôle de la documentation dans la production de nouvelles médiations numériques au musée, qui peuvent par ailleurs utiliser des technologies de l'univers des jeux et de la réalité augmentée (Fabre, 2012, p.83). En outre, l'idée d'augmentation de la médiation documentaire s'inscrit dans les efforts des professionnels de la documentation faits pour penser et « reconfigurer » leurs pratiques en mobilisant les technologies numériques dans la perspective d'accompagner les usagers pour faciliter l'appropriation de l'information (Liquète *et al.*, 2010).

Au-delà des efforts et volontés d'innovation mobilisant les technologies numériques au profit d'un renouvellement des actions de médiation, évaluer un dispositif de médiation documentaire qui prétend « augmenter » son rôle demande de mesurer les effets qu'il induit, puisque la médiation est une « activité productive et créative qui consiste à intervenir sur le cours de la communication en lui apportant une dimension nouvelle » (Jeanneret, 2014, p.13). Il s'agit donc d'évaluer les effets potentiels « d'amplification » produits par les différentes bases de données de spécimens sur la communication scientifique mobilisant les collections. La notion d'amplification avait été utilisée par Élisabeth Eisenstein (1991) pour penser les effets de l'imprimerie sur la vie culturelle, notamment sur certaines théories religieuses ou scientifiques. Dominique Boullier (2015) a proposé de reprendre la notion qu'il préfère à celle de révolution numérique pour souligner qu'un effet devenu prégnant avec le développement de nouveaux supports pouvait déjà exister avec des enjeux moindres avant le changement technologique. Cet emploi de la notion d'amplification¹ permet dans notre cas de penser synthétiquement les continuités et les innovations liées à la mobilisation de supports numériques dans les activités scientifiques. Elle donne les moyens de surmonter le risque de réduction de la compréhension des enjeux de la numérisation des collections qui existe si on les pense à partir d'un seul objectif de reconstitution des fonctions des spécimens non numérisés. Penser les conditions de la reconstitution des fonctions de l'herbier sur support numérique est en effet utile pour orienter le début du travail de conception, mais le questionnement théorique et pratique sur les nouvelles possibilités d'exploitation des spécimens numérisés ne saurait être restreint à cette dimension.

Dans le cas de l'herbier, évaluer d'éventuelles amplifications des formes de communication médiatisée par les collections conduit à questionner les nouveaux échanges et réappropriations des documents par les multiples acteurs participant à leur

1. Notion dont Florence Millerand nous a suggéré l'utilisation lors de la soutenance de la présente thèse

constitution, ainsi que les changements des méthodes scientifiques utilisées. La sémiotique des transactions coopératives (Zacklad, 2013) permettra de comparer les dispositifs destinés à opérer une médiation documentaire augmentée par rapport à ceux qui restent centrés sur des fonctions comparables à celles des catalogues de type OPAC. L'herbier non numérisé constituera notre première référence pour penser les apports effectifs des dispositifs numériques de médiation documentaire augmentée dans les pratiques qui les mobilisent, étant donné le travail scientifique possible à partir des spécimens déjà réduits à une surface en deux dimensions et en cela comparables aux images, et compte tenu des fonctions de coordination du travail documentaire de multiples contributeurs qu'exercent les collections. Mais afin de prendre la mesure des spécificités des dispositifs faits pour reconstituer ou augmenter en ligne les fonctions de l'herbier, et pour pouvoir montrer l'originalité des médiations qu'ils exercent, nous nous penchons d'abord sur des outils existants conçus pour répondre à des objectifs en partie similaires à ceux de l'herbier.

4. État de l'art des dispositifs de médiation documentaire reconstituant des fonctions de l'herbier : bases de données et collaboratoires, outils de lecture critique, et plateformes de crowdsourcing

Après avoir identifié les principaux enjeux de la numérisation de l'herbier et les différents registres de médiation qu'ils demandent de prendre en compte, nous revenons sur des dispositifs qui reconstituent des fonctions similaires dans la communication des savoirs scientifiques. Ces dispositifs illustrent les formes que peuvent prendre une médiation documentaire numérique augmentée par des fonctions facilitant l'exploitation scientifique des documents, et la contribution des usagers à l'enrichissement de la documentation. Cet état de l'art intervenait en amont de la conception d'applications pour l'exploitation des spécimens numérisés d'e-ReColNat : les fonctionnalités développées pour d'autres jeux de données scientifiques ou documents patrimoniaux étaient susceptibles d'être utilisées pour l'herbier. Au plan théorique de l'étude des mutations de la communication associées aux dispositifs d'herbiers numérisés, cette recension permettra de situer les outils développés dans l'infrastructure par rapport aux bases de données existantes, et de comparer les retours d'expériences des usagers de l'infrastructure avec ceux issus d'autres projets de valorisation numérique de corpus. Nous présenterons différents types de dispositifs, susceptibles de reconstituer chacun certains aspects de la médiation opérée par l'herbier. Outre la littérature consacrée aux bases de données d'herbiers numérisés et aux logiciels qui renouvellent l'utilisation des images qui en sont issues, nous mentionnerons des projets d'humanités numériques ayant vocation à faciliter l'exploitation d'images patrimoniales, qui pourraient être repris pour valoriser les collections d'herbier en faisant droit à leur dimension historique ; les dispositifs de sciences participatives et de crowdsourcing ont été considérés plus particulièrement sous l'angle des outils d'écriture et de coordination des productions individuelles, dans l'optique d'alimenter la réflexion sur la version 2 de la plateforme des Herbonautes.

4.1. Des bases de données d'herbiers aux logiciels d'exploitation automatisée émergents

Le recensement de la littérature sur les herbiers à partir de la base de données bibliographique du *Web of Science*, qui couvre aussi bien les domaines de la com-

munication, des sciences de l'information, et de la conservation des collections, que ceux de l'écologie et des études sur les plantes (*plant sciences*) ou sur la biodiversité, montre le dynamisme des projets de numérisation d'herbiers. Lors du bilan conclusif de nos recherches bibliographiques, nous avons recensé, en octobre 2017, 47 publications¹, dont 13 en 2016-2017; le premier résultat remonte à 2003, et seulement 7 sont antérieurs à 2010; 19 résultats sont classés en « *plant science* » et 15 en écologie, et les deux tiers proviennent des États Unis et du Brésil (impliqué dans le projet d'herbier virtuel « Re flora », sur lequel nous reviendrons au chapitre 8). Une explication avancée par Pamela S. Soltis (2017) de cet intérêt croissant réside dans l'atteinte d'une masse critique de spécimens mis en ligne, qui laisse à présent envisager des changements significatifs dans les méthodes scientifiques, par exemple avec la perspective de croiser des données des collections avec les données climatiques ou génétiques.

4.1.1. Processus de numérisation, informatisation et conception de base de données d'herbier

Une première catégorie de publications est constituée de la présentation de différents processus de numérisation ou « workflows » destinés à favoriser ces initiatives (Nelson *et al.*, 2015), compte tenu des spécificités de certaines collections; par exemple les petites collections (Harris et Marsico, 2017) ont donné lieu à une estimation du temps des opérations de numérisation dans le cas où elles sont prises en charge par des étudiants, en vue d'aider d'autres institutions pauvres en personnel à s'y engager. La présentation des programmes de numérisation insiste aussi sur les résultats qu'ils produisent sur la connaissance des collections (Sreekumar, 2017); la description des données issues de la numérisation de l'herbier du MNHN et des différentes catégories de spécimens détenus a donné lieu à un « *data paper* » (Le Bras *et al.*, 2017), publication fondée sur la présentation d'un jeu de données et du dépôt dans lequel elles sont consultables, en vue de valoriser le travail de mise à disposition de données et de favoriser leur réappropriation.

Le processus de numérisation a aussi été étudié dans ses relations avec l'informatisation des étiquettes. Les études sur la constitution des bases de données ont souligné les recherches d'information à l'œuvre dans cette activité, qui s'ajoutent aux opérations moins intellectuelles de scan du spécimen et de saisie des informations une fois comprises et vérifiées. Ces recherches d'information sont particulièrement importantes dans le cas où seule une partie des spécimens est numérisée, à savoir les spécimens « types » ayant été cités comme échantillons de référence dans une publication pour décrire une espèce. La valeur ajoutée de la connaissance des collections associée à la constitution des bases de données de spécimens types numérisés est le produit des « recherches bibliographiques » et du patient travail de « déchiffrement d'une écriture cursive, souvent vieille de plus de cent

1. Il s'agit des résultats de la requête suivante :
TOPIC : ("virtual herbari *") OR TOPIC : ("digital herbari *") OR TOPIC : (digitization herbari *) AND TOPIC : (image NEAR/5 herbari*)?
Timespan : All years.

ans, sur du papier jauni, et utilisant une encre ayant tendance à s'effacer » (Heaton et Proulx, 2012). Le précédent conservateur de l'herbier de l'université de Montpellier, Peter A. Schäfer (2004), précise les modalités des recherches de types, selon les situations qui les suscitent : dans le cadre d'un programme de numérisation qui procède liasse par liasse, mais aussi d'une demande d'un scientifique réalisant une monographie sur un genre ou une espèce, avec lequel s'engage une réelle coopération, par exemple par mail, ou pour déterminer les spécimens types, le travail sur l'œuvre d'un auteur bien représenté à l'herbier. Le processus d'informatisation varie encore quand il a lieu dans une plateforme de transcription participative, comme « Notes from Nature »², issue d'une collaboration entre les musées d'histoire naturelle de Londres et de l'université du Colorado et le SERNEC ; les outils de communication de la plateforme et la procédure de validation (à partir de trois réponses concordantes issues de contributeurs distincts), ainsi que la procédure de restitution des jeux de données validés et des erreurs à l'institution détentrice du fonds (Hill *et al.*, 2012), ont pu en particulier être confrontés avec le processus à l'œuvre dans la plateforme analogue du MNHN, « Les Herbonautes ».

Une enquête a été menée sur les attentes des usagers concernant les bases de données pour le jardin botanique belge de Meise par une agence prestataire de développement numérique, en vue de l'élaboration d'un nouveau portail et de la priorisation des besoins. Les différents profils d'usagers interrogés correspondaient à ceux que nous avons identifiés dans notre enquête, ce qui la rend comparable. De l'étude est ressortie l'importance des données liées à d'autres bases, en particulier bibliographiques ; la qualité de transcription à mettre en balance avec la quantité de spécimens informatisés amènent à considérer l'utilité de solutions d'intégration de corrections par les usagers. Les mesures sur l'image sont considérées comme un fonction complémentaire intéressante. On trouve une bonne partie de ces fonctions réalisées sur le portail de consultation des types informatisés dans le programme « Global Plants Initiative », qui se démarque par les liens établis entre publications et spécimens, grâce à des opérations corollaires de numérisation de fonds de publications scientifiques ; l'ajout de tags et commentaires aux spécimens est également possible, ainsi que l'utilisation d'outils de mesures calibrées, mais ces fonctions avancées ne sont disponibles que sur abonnement payant.

4.1.2. Logiciels et pratiques d'exploitation scientifique du spécimen numérisé

Des études récentes montrent une attention portée à de nouveaux défis de l'exploitation automatisée des images. L'accès aux données sur terminal mobile devient une préoccupation (Maya-Lastra, 2016). Des clés de détermination utilisant des images ont été élaborées (Kirchoff, 2011), dans lesquelles les caractères à observer pour distinguer deux espèces sont décrits visuellement, et non par une explication textuelle. L'identification automatique d'un spécimen d'herbier grâce à un processus d'apprentissage automatique (*machine learning*) « entraîné » sur un large échantillon de spécimens numérisés provenant de plusieurs institutions a

2. <https://www.notesfromnature.org/>

donné des résultats prometteurs (Carranza-Rojas *et al.*, 2017), alors que cette méthode était antérieurement appliquée à des photographies de plantes vivantes et à des feuilles (Unger *et al.*, 2016; Corney, 2012). Le partage d'annotations d'images, sur lequel nous reviendrons au chapitre 8, a également été abordé récemment (Surbier, 2017).

Ces études montrent en particulier les nouvelles avancées que connaît l'exploitation scientifique du spécimen numérisé, qu'on peut comparer aux conclusions que Christine Hine (2013) a tirées quelques années auparavant d'une étude des conséquences de la numérisation des collections sur le travail des systématiciens, propre à un contexte de politiques scientifiques favorable à ces projets. À partir de 22 réponses obtenues suite à un appel à témoignages envoyé sur une liste de diffusion internationale de systématique concernant le travail fait avec les images d'herbiers, Christine Hine a montré la prégnance d'une « culture matérielle », qui reste aujourd'hui marquée par la reproduction en ligne du travail sur l'herbier papier. Il n'est pas question de substituer les images aux collections existantes (Hine, 2008, p.117) ; cependant, des pratiques faisant une place à l'image dans la recherche sont en émergence avec — outre l'accès facilité aux collections — l'identification de spécimens en ligne, mais aussi quelques expériences de sollicitation de l'expertise de chercheurs distants concernant la détermination de l'espèce d'un spécimen à partir de sa photographie. Alors que l'identification à partir d'une image fut initialement un sujet de plaisanterie sur la liste de discussion étudiée, elle se développe aujourd'hui, y compris de manière automatisée, même si les techniques et pratiques sont en émergence.

4.1.3. Les collaboratoires

Nous avons également étudié les rapprochements possibles entre les dispositifs de l'infrastructure e-ReColNat et les *collaboratoires*, environnements informatiques visant à favoriser les collaborations scientifiques — « *computer-supported scientific collaboration environments, often called collaboratory* » (Bos *et al.*, 2007, p.653). L'un des livrables confiés au laboratoire Dicen-IDF avait initialement été désigné par ce terme, avant de devenir « laboratoire en ligne ». Alors que l'habitude de collecter des spécimens en plusieurs exemplaires et de les déposer auprès de différentes institutions avait vocation à faciliter le travail des spécialistes à l'échelle régionale de l'université, ou à celle de la société savante locale ou de centres nationaux plus importants, le collaboratoire est pensé comme un moyen d'utiliser les technologies de l'information et de la communication en vue d'un travail scientifique organisé à l'échelle internationale : l'objectif est de permettre de travailler ensemble malgré la distance grâce à des outils et ressources partagées (Glasner, 1996). La constitution de bases de données d'herbiers se rapproche des collaboratoires construits autour d'un objectif principal de partage de données, qualifiés de « *community data systems* », dans une typologie réalisée à partir de 212 études de cas recensés dans 9 champs disciplinaires dans le cadre du projet SOC (*Science of collaboratories*) (Bos *et al.*, 2007, p.660). Elle catégorise les collaboratoires selon leurs objectifs et les technologies et formes d'organisation qu'ils mobilisent. Ce type de collaboratoires, qui reposent davantage sur l'agrégation de données que

sur la co-création, est à l'origine d'une production de standards de description et d'outils de visualisation de données (p.661); ils doit surmonter la difficulté d'une motivation insuffisante à déposer des données, mentionnée également dans le cas de la base de données d'histoire naturelle du muséum du Luxembourg (Meyer, 2009).

Outre cette typologie de dispositifs et formes de collaboration, la littérature sur les collaboratoires a fourni des exemples de « bureaux virtuels », qui dans le processus de conception étaient susceptibles d'inspirer le laboratoire en ligne de l'infrastructure e-ReCoNat, et avec lesquels l'outil effectivement développé peut être comparé. L'exemple du logiciel CORE2000 développé par le *Pacific Northwest National Laboratory* (PNNL) associe fonctions de communication variées (discussion vidéo, tableau partagé, outils de vote, messagerie instantanée) et outils plus spécialisés d'analyse automatisée d'image (Olson *et al.*, 2008, p.125). Un exemple dans le domaine de l'étude de l'atmosphère permet une campagne d'observations partagée, intégrant différents outils de visualisation d'observations à des fonctions de messagerie instantanées (p.157). Les outils de communication synchrone ont été développés davantage que les solutions de gestion documentaire dans les exemples évoqués : étant donné l'importance des documents des collections dans les pratiques scientifiques imprégnées par la culture matérielle des collections (Hine, 2013), nous avons cherché des complémentarités entre ce type de bureau partagé et les logiciels développés pour la lecture critique dans le cadre de projets d'humanités numériques. Cela nous conduit à élargir notre état de l'art, alors que la littérature sur les herbiers numérisés est dominée par la perspective de l'exploitation scientifique et du recensement des collections.

4.2. Apports des dispositifs de médiation documentaire numérique de fonds patrimoniaux

La problématique des conditions de l'exploitation scientifique de l'image traverse d'autres disciplines ayant de plus en plus recours à la numérisation des documents sur lesquels elles s'appuient. En 2000, Marie Després-Lonnet signalait, dans le contexte de l'étude de la base de données du projet LIBRIS, que « les bases de données ne suscitent pas l'intérêt ni la confiance de l'ensemble de la communauté historienne ». Le projet de reconstituer les conditions scientifiques de l'exploitation de l'image d'un objet numérisé s'inscrit dans le champ de questionnements communs aux humanités numériques fédérés par l'objectif de « s'approprier les technologies informatiques offertes par la numérisation de documents » et d'« exploiter des possibilités de recherche inédites avec l'accès simplifié à des sources issues de la construction de nouvelles bases de données », qui demande corrélativement d'acquérir de nouvelles « méthodes d'analyse et de recherche » (Broudoux et Kembellec, 2017, pp.20-21). Les projets d'humanités numériques ont donné lieu à de multiples exemples de valorisation de corpus, dont certains modules auraient

pu s'appliquer à l'herbier, même si l'évolution du projet et les contraintes de développement n'ont pu donner lieu à leur intégration actuellement : identification et description des entités nommées présentes dans les métadonnées des collections pour intégrer automatiquement des enrichissements sur les lieux des collectes ou la biographie des collecteurs ; exploitation des possibilités de visualisation des documents par navigation géographique des enregistrements géolocalisés ou chronologiques, ou encore par une « frise chronologique permettant l'affichage et l'interrogation des contenus d'une base de données » (Broudoux et Kembellec, 2017, p.21).

Pour compléter l'approche des bureaux virtuels dans les laboratoires, nous nous sommes intéressés aux logiciels intégrant un ensemble de fonctionnalités destinées à l'interprétation critique d'images, qui ont aussi donné lieu à des interfaces susceptibles d'être adaptées aux besoins de l'infrastructure e-ReColNat. Par exemple, l'interface permettant la « fragmentation d'une entité matérielle graphique au sein d'un espace libre », développée par Thomas Bottini (2010, p.210) a été une référence dans le contexte de la conception. L'exemple de la bibliothèque numérique de la BNF, « Gallica », a retenu notre attention en ce que l'un de ses récents développements a été consacré à une interface de lecture, visant à améliorer les possibilités d'exploitation personnalisée des documents en ligne. Gallica s'était également dotée, le temps d'un projet de recherche qui n'a pu être pérennisé, d'une plateforme de correction des transcriptions automatisées par OCR, également envisagée pour faciliter l'informatisation des étiquettes d'herbier, et fournit à ce titre des retours d'expérience utiles pour notre cas (BNF, 2015).

Avec son système d'étiquettes ajoutées successivement, l'herbier constitue un modèle sur papier des « fonctionnalités d'annotation », « qu'il est devenu commun pour un projet d'humanités numériques de mettre à disposition du lectorat » (Broudoux et Kembellec, 2017, p.27). Les nouvelles solutions d'annotation pourraient reconstituer et revisiter le modèle de l'herbier pour l'appliquer aux collections numérisées. Si les systèmes de « tags » peuvent correspondre à la notion d'étiquette dans la dimension graphique qu'ils donnent à l'annotation, ils introduisent une organisation des connaissances sans vocabulaire contrôlé qui ne correspond pas à la pratique systématique, mais qui peut toutefois accompagner des usages individuels de l'image. Les systèmes d'annotation développés pour des corpus audiovisuels nous ont intéressés par leur souci de s'adapter à la dimension sensori-motrice du geste d'annotation, dans l'optique de prendre en compte l'inscription des pratiques mobilisant les collections dans la « culture matérielle » dans laquelle elles se sont constituées. Les outils d'annotation dans un contexte collaboratif ont été en particulier développés dans des projets d'archives participatives et de sciences citoyennes mobilisant des plateformes de crowdsourcing, qu'il était essentiel de confronter avec la plateforme des Herbonautes en vue de la réflexion sur les évolutions à lui apporter.

4.3. Des dispositifs de production de données standardisées multi-acteurs ouverts aux amateurs : les plateformes de crowdsourcing et de sciences participatives

Les plateformes de crowdsourcing ont notamment été développées comme nouvelle modalité de réalisation des tâches documentaires, sous le terme de « *crowd-curation* ». Elles constituent des dispositifs partageant des points communs avec l'herbier, dans la mesure où elles donnent accès à un corpus documentaire, concerné par les tâches proposées, et ouvrent la participation à la production de données standardisées sans condition de compétence particulière. Les mêmes outils d'écriture y sont appliqués aussi bien aux sciences de la nature qu'à des corpus patrimoniaux, et sont *a fortiori* susceptibles de s'appliquer à l'herbier.

Nous employons le terme de plateforme pour désigner un dispositif modulable utilisé pour différents programmes de tâches, à partir d'un ensemble d'outils communs. Nous avons analysé ces plateformes à partir d'un point de vue communicationnel, mais sans le réduire aux techniques de mobilisation les plus efficaces des contributeurs. Nous avons pour cela analysé des formes de communication à l'œuvre dans les activités d'interprétation des documents et les opérations d'écriture associées, en tenant compte de leur dimension collective, inscrite comme c'est le cas pour l'herbier dans les normes qui régissent ces écritures. Le but de cette analyse était de déterminer les caractéristiques de dispositifs de médiation documentaire ouverts à la participation de contributeurs non spécialistes. La littérature sur ces plateformes permet de préciser ces caractéristiques, au niveau des modèles de validation de données produites par de multiples contributeurs, des techniques de communication à adopter pour motiver la communauté, et du rôle des espaces de discussion permettant de soutenir l'activité collective sur la plateforme. Notons que l'analyse communicationnelle que nous proposons pour les plateformes de crowdsourcing concerne le cas bien précis de l'adaptation des principes et techniques du crowdsourcing aux objectifs de valorisation de corpus scientifiques et patrimoniaux (Dunn et Hedges, 2014), fondée sur la participation bénévole d'un nombre réduit de contributeurs (ce qui conduit à parler de « *community sourcing* », selon la recension de Mathieu Andro, 2016). Ces plateformes se prêtent à des stratégies de coopération et suscitent des interactions entre les contributeurs, ouvrant une approche de la communication qu'elles médiatisent différente de celle développée dans le cas de dispositifs agrégeant les traces de contributions d'un grand nombre de contributeurs, parfois même sans que les usagers en aient conscience, dans des formes de crowdsourcing implicite (Andro, 2016). Olivier Sarrouy (2014) propose justement, pour appréhender les dispositifs exploitant la contribution des foules, de remettre en cause le cadre d'analyse des interactions, pour prendre en compte des effets de dissolution du « *sujet de l'énonciation* » (p.276) dans une masse de traces agrégées par un « *dispositif de coordination de l'action des foules* » (p.284), qui devient alors le principal élément d'analyse. Les plateformes de ce type organisent alors une « *communication comme abandon de*

traces », en « soustrayant » les traces, « par divers procédés techniques, à leurs auteurs ».

Le cas qui nous intéresse demande au contraire de prendre en compte les interactions personnelles dont ces plateformes sont le lieu, sans pour autant négliger le dispositif de coordination dans lequel elles s'inscrivent. On peut caractériser avec la sémiotique des transactions coopératives les dispositifs impliqués par un répertoire d'action plus large, qui étend certaines pratiques d'écriture intervenant dans le travail documentaire hors de la sphère des chercheurs et documentalistes qui travaillent à la constitution de bases de données ou à la documentation des corpus. Les plateformes associent dispositifs d'écriture contributifs orientés données (par des tâches d'écritures fragmentées et contrôlées) et orientés dialogue (étant donné les espaces de discussion ouverts). Les outils d'écriture proposés distinguent les plateformes de crowdsourcing d'autres formes de participation du public dans les institutions culturelles, étudiées notamment par Geneviève Vidal (2010), à partir de l'étude des « traces d'usages » et des « usages inscrits » dans des « services web 2.0 » inclus à 60 sites web de musées. Les productions destinées à s'insérer dans des bases de données institutionnelles relèvent du registre de la *contribution*, qui « apporte des contenus nouveaux, produits de façon collective, par la communauté des usagers », renégociant ainsi parfois l'autorité de l'institution qui intègre leurs productions. Les formes de contribution aux tâches documentaires, contrôlées, se distinguent toutefois des créations libres réalisées par les publics. La contribution peut être considérée comme une « rationalisation » des « interventions des internautes » à partir d'une « régulation de ces dernières fondée sur une sophistication technique et éditoriale des médiations numériques » (Vidal, 2012). Elle demande de réfléchir aux manières de ne pas compromettre l'autorité de l'institution. C'est ainsi qu'elle est souvent organisée sur un site autonome, distinct de la base de données institutionnelle même s'il l'alimente. De même, Nathalie Casemajor Loustau (2011) souligne dans son étude des dispositifs collaboratifs de « Bibliothèque et Archives Canada » la différence entre les apports intégrés au site institutionnel, communiqués uniquement par mail et faisant l'objet d'une validation, et les commentaires, plus libres, laissés sur les collections mises en ligne sur Flickr : « le fait de déplacer l'activité collaborative des usagers sur un site autre que celui de l'institution patrimoniale atténue quelque peu la tension entre, d'une part, la volonté d'offrir aux internautes des espaces d'expression et, d'autre part, le besoin de contrôler cette expression publique ». Dans le domaine naturaliste, étudiant la participation sur le site d'un muséum régional, Bernadette Dufrêne et Madjid Ihadjadene (2010) montrent que si des gestes « péri-professionnels » sont confiés aux internautes, c'est sous le contrôle de modérateurs professionnels dont les internautes s'approprient les pratiques.

L'ouverture des corpus à la contribution passe par l'encadrement de la participation par des outils d'écriture adéquats, développés en adaptant au contributeur néophyte des dispositifs d'écriture utilisés pour la recherche et l'édition scientifique (comme la transcription selon le standard XML TEI dans le site « Transcribe Bentham », projet du *University College* de Londres visant à transcrire des papiers du philosophe Jeremy Bentham³). Ce type d'outil se caractérise par la prégnance

3. <http://blogs.ucl.ac.uk/transcribe-bentham/>

des critères de légitimité des institutions qui les ont développés. En effet, la régulation des écrits produits dans ces plateformes est pensée en amont des projets selon des critères scientifiques ou liés à la gestion des collections. Outre l'enjeu du contrôle des contributions, la fragmentation des tâches d'écriture correspond aussi à l'objectif de mobiliser des contributeurs ponctuellement, tout en ouvrant la possibilité de susciter leur engagement croissant (Ridge, 2013). Selon Caroline Haythornthwaite (2009), l'ouverture à la foule est associée à une régulation hiérarchique des contributions, prise en charge par le dispositif qui instaure par exemple des rôles prédéfinis, ce qui permet de solliciter moins activement les contributeurs que dans les formes de collaborations organisées selon un modèle communautaire, dont la régulation est prise en charge par les contributeurs eux-mêmes. Les différentes formes de régulation des contributions se traduisent sur les outils d'écriture : dans un modèle communautaire comme Wikipédia, des pages de discussion sont dédiées à la validation des contributions selon des règles qui font elles-mêmes l'objet de débats. Gilles Sahut (2014) l'a montré en retraçant la genèse de la règle de citation des sources dans la communauté. Les règles d'écriture sont un élément essentiel de la régulation des communautés impliquées dans la production de connaissances et dépendent des objectifs et des évolutions de leur projet. Elles conduisent à une « régulation collective des écritures particulièrement originale », observée dans Wikipedia, « où chacun est appelé à être non seulement écrivain, mais aussi contrôleur des écritures des autres » (Cardon et Levrel, 2009).

Toutefois, les deux modèles peuvent être hybridés (Haythornthwaite, 2009). Cela se traduit au niveau des outils d'écriture, par lesquels nous appréhendons les formes de communication médiatisées sur ces plateformes, qui ne sont plus soumis à la fragmentation et à la validation selon des critères institutionnels. Des forums de discussions existent dans l'une des principales plateformes proposant des projets de science participative, « Zooniverse » (Owens, 2014, p.273). Une étude quantitative de l'utilisation de ces forums dans dix projets de Zooniverse (Luczak-Roesch *et al.*, 2014) a montré qu'ils étaient utilisés principalement par certains contributeurs, caractérisés plutôt par un profil que par un volume de contributions. Une autre étude sur un nombre plus restreint de programmes de Zooniverse a approfondi qualitativement les usages de ces forums, et fait ressortir leur importance pour rendre visibles les pratiques des contributeurs et favoriser les apprentissages informels, dans un contexte où les contributions individuelles sont masquées (Mugar *et al.*, 2014) : ces espaces sont un ingrédient essentiel à l'intégration des nouveaux contributeurs, à qui ils offrent une forme de retour (« *feedback* ») sur la justesse de leurs contributions. C'est aussi le constat fait par Pauline Moirez (2012) à propos du rôle de l'animation dans le cadre du projet « PhotoNormandie » : « Le succès du projet repose sur la constitution et l'animation dynamique d'un groupe de contributeurs qui ont largement utilisé les fonctionnalités sociales de Flickr pour échanger et commenter. Ici, le réseau social a eu un rôle de levier sur l'opération de crowdsourcing ». Au rang des projets mobilisant les réseaux sociaux existants pour communiquer avec les usagers, elle cite également l'exemple de « Transcribe Bentham », « dont la communauté des transcripateurs est recrutée et animée via un blog et un fil Twitter particulièrement dynamiques » (Moirez, 2012).

L'étude des enjeux communicationnels des plateformes a aussi porté sur l'éva-

luation de l'efficacité des messages adressés aux contributeurs, susceptibles d'éviter l'arrêt de la contribution s'ils sont envoyés au bon moment pour anticiper la démotivation (Lee *et al.*, 2017). Le design de la plateforme a été lui aussi abordé : selon Donelle McKinley (2013), citée dans l'état de l'art de Mathieu Andro (2016, p.220) sur les plateformes de crowdsourcing dans les bibliothèques, « un site de crowdsourcing doit avoir une page d'accueil qui décrit le projet et invite à la participation des bénévoles, et d'autres pages Web pour instruire les bénévoles sur l'exécution des tâches. Les internautes doivent avoir une idée claire de pourquoi ils sont là et qu'est-ce qu'ils ont à faire ». Donelle McKinley (2013, p.24) précise sept prérequis pour l'interface de crowdsourcing. Ils concernent en particulier la minimisation des efforts des usagers, la facilité de la compréhension des tâches par les nouveaux contributeurs ayant une habitude moyenne de navigation sur le web, l'intégration des tâches au processus de recherche qu'elles supposent, ainsi que la production de données structurées et contrôlées, mais ouvertes à une certaine flexibilité. C'est davantage vers la recherche de la mobilisation d'une dynamique communautaire, par des outils adaptés permettant de structurer les traces des contributions, que s'orientera le travail sur les outils d'écriture au service de l'animation de la communauté des Herbonautes.

La question des formes de mobilisation des contributeurs par l'animation est souvent associée à la compréhension des motivations de la participation, qui orientent aussi les choix à faire pour l'animation. Les projets patrimoniaux se caractérisent par une motivation liée intrinsèquement aux documents proposés. Trevor Owens (2014) rapporte l'expérience de transcripteurs de journaux de guerre mis en ligne par la bibliothèque de l'université de l'Iowa qui « suivent l'histoire racontée dans les manuscrits et s'engagent dans l'histoire ou sont motivés par la pensée de recherches futures » (p.278). On retrouve dans le document d'archive, selon Pauline Moirez (2012), de quoi alimenter une source de motivation propre à la nature du document, qu'il s'agisse de sa valeur de document historique ou administratif. La nature des documents proposés permet de mobiliser des publics spécifiquement intéressés, comme les généalogistes « de plus en plus jeunes et férus de technologies numériques », qui « étendent leurs centres d'intérêt largement au-delà d'une simple généalogie "génétique" consistant à remonter un arbre généalogique, et s'intéressent de plus en plus aux sociétés et territoires dans lesquels leurs ancêtres ont vécu » (Moirez, 2012).

Les retours d'expérience attestant de l'intérêt suscité par les fonds proposés aux contributeurs ouvrent pour la médiation culturelle une perspective nouvelle, fondée sur un projet d'enrichissement d'un dispositif de médiation documentaire adapté à des usagers non spécialistes. La reconstitution de l'ensemble des fonctions de l'herbier suppose de conjointre les objectifs des plateformes de crowdsourcing et de ceux instrumentant la collaboration entre professionnels : les collaboratoires et les différents corpus patrimoniaux mis en ligne dans la perspective de renouveler les modalités de leur étude scientifique constituent autant d'exemples de dispositifs qui peuvent inspirer un herbier numérique. Penser l'application de ces outils du web collaboratif à la valorisation des collections a été l'un des objectifs de l'infrastructure e-ReColNat. Nous étudions maintenant les dispositifs mobilisés par les usagers de l'infrastructure, en incluant les applications produites en son

sein. Nous nous demanderons en quelle mesure ils reconfigurent les spécificités de la médiation documentaire opérée par l'herbier, indissociable de la médiation du travail scientifique dans sa dimension collective et ouverte aux amateurs. Nous procéderons pour cela en étudiant les formes de communication qu'ils médiatisent au niveau de l'activité des acteurs qui les mobilisent aux côtés des collections non numérisées.

Deuxième partie

**Méthodes d'observation et
d'analyse du terrain : approche
ethnographique et analyse
quantitative fédérées par un
objectif de conception**

5. Constituer e-ReColNat en terrain d'étude : mise en visibilité des conditions infrastructurelles de la communication médiatisée par les collections

Les exemples d'outils d'écriture recensés dans l'état de l'art, capables de reproduire la production de données standardisées ouverte à l'enrichissement par les usagers y compris non spécialistes, avaient vocation à inspirer le travail de conception et à être confrontés aux dispositifs développés dans le cadre de l'infrastructure e-ReColNat. Nous précisons à présent la méthodologie qui nous permettra d'appréhender ces dispositifs et leurs enjeux pour les formes de communication impliquées dans les activités de recherche, de conservation et de gestion des collections, traditionnellement réparties entre professionnels et amateurs. Nous commençons par présenter l'infrastructure e-ReColNat pour justifier le fait de l'avoir considérée comme un terrain porteur de potentielles innovations dans les pratiques mobilisant les collections. Nous reviendrons ensuite sur la démarche mise en place pour appréhender les formes de communication médiatisées par les collections dans le contexte d'un processus de numérisation industrielle des collections, et compte tenu de l'implication des acteurs dans une démarche de recueil de besoins.

5.1. Des objectifs centrés sur le partage de données à grande échelle

E-ReColNat est une Infrastructure Nationale en Biologie-Santé, financée par le ministère de l'enseignement supérieur et de la recherche (par l'intermédiaire de l'ANR), au titre des investissements d'avenir, du 1er janvier 2013 au 31 décembre 2019 et coordonné par le MNHN. D'après la présentation du projet mise en ligne par le Ministère de l'Enseignement et de la Recherche lors de la communication des résultats de la sélection des projets élus au titre des investissements d'avenir (voir annexe A.1), son objectif est la « valorisation de 350 ans de collections d'histoire naturelle » à travers la numérisation de collections publiques de France (ainsi que de l'herbier de Rabat, historiquement en étroites relations avec celui de l'université de Montpellier). Les objectifs de l'infrastructure sont en premier lieu scientifiques, avec quatre disciplines mentionnées, « Systématique, Botanique, Zoologie, Écologie », et un centrage particulier sur la systématique, discipline de plusieurs responsables du montage du projet puis de la coordination de l'infrastructure. La

perspective du projet pour le « système de la recherche » s'inscrit dans une problématique d'accès à l'information scientifique, celle de favoriser les collaborations scientifiques de grande échelle grâce aux environnements numériques de travail, propres aux laboratoires, en l'occurrence par la valorisation des données des collections : « L'ensemble de ces données pourra être agrégé à l'ensemble mondial et contribuer ainsi aux modélisations de la biodiversité à toutes les échelles. Avec l'infrastructure E-RECOLNAT l'ensemble de la communauté mondiale des systématiciens aura accès à cette infrastructure unique, qui leur permettra de répondre à certaines questions à distance ». Ces objectifs concernent la mutualisation de données à une échelle comparable à celle du projet canadien Canadensys étudié par Florence Millerand et Lorna Heaton, qui est un nœud régional qui alimente la base de données du GBIF, l'instance internationale ayant vocation à centraliser les données sur la biodiversité. La réponse à l'évaluation à mi-parcours de 2017 de l'infrastructure e-ReColNat rappelle l'objectif de centraliser « sur une plateforme unique les images et les données de la totalité des collections naturalistes françaises (100 millions de spécimens). Ces données proviennent de bases existantes et interoperables et des programmes de numérisation qui ont lieu au niveau national ». Elles sont centralisées à partir d'une procédure de « moissonnage » (intégration régulière et automatisée de données par une base centrale) des bases de données des collections des différentes institutions partenaires, qui transmettent aussi leurs données au GBIF.

L'infrastructure a aussi un objectif d'innovation logicielle, qui visait initialement à favoriser le travail à distance dans des sessions individuelles et partageables d'un « laboratoire » en ligne. C'est ce qui permet de la rapprocher non seulement des laboratoires centrés sur le partage de données, dont relèvent les « *community data systems* » (Bos *et al.*, 2007), mais également de laboratoires orientés vers la mutualisation d'outils. Elle poursuit aussi des buts de partage de connaissances, qui s'étendent non pas à des activités d'enseignement comme certains des cas étudiés dans la typologie, mais à l'élargissement de la maîtrise des techniques de documentarisation des collections par de nouveaux publics. L'accès libre et gratuit aux données de la recherche n'est pas un enjeu réservé aux scientifiques : les retombées de l'infrastructure pour les citoyens passent surtout par leur participation « en partie active » (annexe A.1) à l'actualisation des données des collections. L'accès à ces données est présenté comme un enjeu économique pour l'« expertise environnementale » et la création d'emplois d'informaticiens nécessaires au développement de l'infrastructure.

5.2. Partenaires et livrables

L'infrastructure s'est organisée autour de différentes équipes, « *workpackages* » (WP), définies chacune par les missions et livrables qui lui étaient attribués. Sans recenser l'ensemble des tâches exercées par chacun des WP, qui devaient être déclarées sur une plateforme de communication interne au projet, nous présentons l'activité des différentes équipes telle que nous l'avons appréhendée, à partir des relations entretenues par notre équipe avec les autres WP.

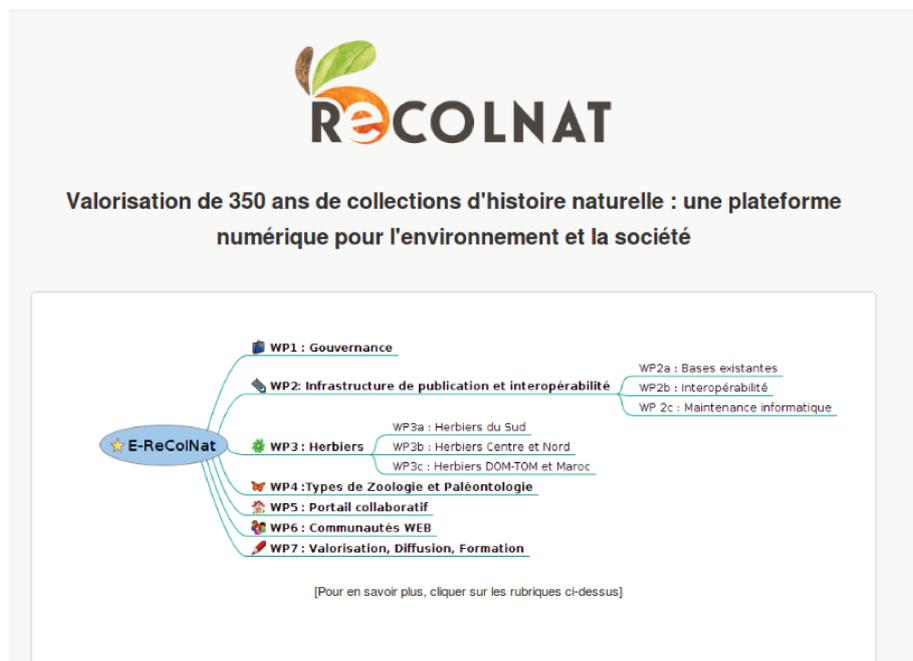


FIGURE 5.1. – Schéma de l’organisation de l’infrastructure e-ReColNat en workpackages. Capture d’écran du premier site du projet (2013-2015), antérieur au déploiement du portail e-Recolnat.

Notre enquête s’inscrivait en amont et en accompagnement du travail de conception du WP5, chargé de la constitution d’applications ayant vocation à s’intégrer à la plateforme centrale de l’infrastructure, dont le nombre et les objectifs ont évolué au cours du projet. Ce travail de conception au sein du WP5 a conduit à interagir avec le WP1, chargé de la coordination de l’infrastructure, et avec le WP2, chargé notamment du modèle et des développements de la base centralisant les données régionales, de l’application de consultation des données de la base (moteur de recherche et visualisation cartographique), et d’une application de saisie de données proposée aux institutions ne pratiquant pas encore l’informatisation. L’un des ingénieurs du WP5 a davantage interagi avec le WP3 chargé de l’organisation du processus de numérisation industrielle, en vue de l’élaboration du cahier des charges d’un logiciel de contrôle de la qualité de la numérisation, capable de repérer des lots d’images défectueux à partir de la vérification d’un échantillon des images produites par la chaîne de numérisation. Conçu pour être ouvert à la participation des internautes, ce logiciel sera finalement utilisé seulement par le personnel des collections, en raison d’un refus tardif de la part de certaines institutions partenaires de déléguer cette mission à des amateurs. L’enquête sur les usages de la plateforme participative « Les Herbonautes » nous a conduits à interagir également avec le WP6, chargé de l’animation de la plateforme et coordonné par l’association Tela Botanica (Heaton *et al.*, 2011) qui fédère un réseau de botanique francophone (comptant près de 40 000 inscrits sur le site web de l’association en 2017). Les échanges avec Tela Botanica nous ont aussi conduits à assister à des réunions liées au travail de recensement des collections, piloté également par le



FIGURE 5.2. – Cartographie dynamique des collections recensées par Tela Botanica (capture d'écran)

réseau de botanique, d'abord en Languedoc-Roussillon (Mathieu et Mathez, 2004) puis finalement intégré aux objectifs de e-ReColNat. Le recensement avait pour mission de développer l'inventaire dans de nouvelles régions, à partir d'un accompagnement méthodologique (Boulangeat, 2014). Ce recensement s'est traduit par l'intégration dans le portail de e-ReColNat du widget de l'application développée par Tela Botanica de signalement de collections, le catalogue CoEl (figure 5.2). L'infrastructure e-ReColNat s'appuie ainsi sur les collectifs existants ; comme dans le cas de « Global Plants Initiative » (Heaton *et al.*, 2012), ce type de programme apporte un financement complémentaire réinvesti dans la communication et le développement informatique essentiels aux activités du réseau de botanique.

5.3. Des réorganisations et des innovations dans la gestion des collections

À la différence de l'informatisation et de la numérisation réalisées en interne par le personnel des collections, le processus industriel choisi pour les collections de l'infrastructure demande au personnel de travailler avec de nouveaux partenaires, les prestataires choisis au terme d'appels d'offres pour le chantier de numérisation, auxquels s'ajoutent les partenaires et prestataires impliqués dans la conception et le développement des logiciels.

La centralisation des données s'est faite à partir de la constitution d'un réseau réunissant les institutions détenant des collections et acceptant de mutualiser leurs



FIGURE 5.3. – Les institutions partenaires d'e-Recolnat

données dans la base du MNHN, « Sonnerat ». Les gestionnaires des collections numérisées dans l'infrastructure étaient pour beaucoup impliqués en amont dans le réseau des herbiers de France, devenu réseau « ReColnat », littéralement « Réseau des collections naturalistes », au cours du projet. Mais les différentes institutions étaient inégalement avancées dans l'informatisation de leurs collections : l'herbier du muséum avait déjà été largement numérisé, le travail avait été commencé localement dans d'autres institutions, notamment de la participation au projet « Global Plants Initiative ». Certaines ont rejoint l'infrastructure pendant la durée du projet, sur la base de financements complémentaires indépendants. La constitution du réseau en vue de la centralisation des données et de la numérisation des collections supposait l'implication de tous les partenaires dans des choix de numérisation communs, malgré les différentes politiques de valorisation des collections locales et les différentes règles de conservation selon le statut des collections (différentes pour un musée de France, le MNHN, ou une université). Il fallait surmonter des intérêts antagonistes (préférer une numérisation locale afin d'obtenir un financement pour la gestion de sa collection, plutôt que de participer à un projet de numérisation industriel), ou la crainte pour les partenaires de disparaître dans un réseau national rendant surtout visible l'institution coordinatrice, aspects qui sont ressortis lors de certains de nos entretiens.

5.4. Un terrain pour appréhender les mutations de la communication

La constitution de l'infrastructure, par les nouveaux partenaires et contraintes qu'elle implique pour le personnel des collections compte tenu des décisions du réseau, par les opérations d'inventaire et les opérations de numérisation qu'elle fédère, et par les données et applications qu'elle centralise, a généré des transformations dans les activités de gestion des collections et de nouveaux outils logiciels pour leur exploitation. La constitution du réseau et des nouveaux outils constitue un moment qui rend visible l'*infrastructure* des collections, entendue dans le sens des artefacts matériels qui structurent les pratiques tout en étant déterminée par les choix des acteurs (Star et Ruhleder, 2010). Elles deviennent invisibles quand l'activité prend un caractère routinier, mais les crises et dysfonctionnements rappellent leur existence. La constitution de l'infrastructure e-ReColNat constitue une

perturbation, rendant visible l'ensemble des artefacts et formes de coordination du travail des acteurs participant à leur gestion. Le travail de conception impliquant les partenaires de l'infrastructure a créé des situations privilégiées pour observer les enjeux communicationnels de ces mutations. Ce contexte de perturbation a suscité l'investissement des usagers des collections dans notre enquête et dans le travail de conception visant à produire de nouveaux dispositifs permettant l'exploitation des données centralisées et les activités nécessaires à les enrichir. Nous avons observé à partir de ces situations les mutations de la communication médiatisée par les documents des collections, appréhendée à travers l'élaboration de nouvelles connaissances sur la base de l'interprétation des documents et des différents prélèvements et écrits, plus ou moins standardisés, associés à ces opérations. Ainsi, à partir des situations générées par l'infrastructure e-ReColNat en cours de constitution, qui constitue par conséquent notre principal terrain de recherche, nous avons appréhendé les formes de communication médiatisées par les collections d'herbiers dans le contexte de la constitution d'un réseau d'institutions impliquées dans leur numérisation, dont les objectifs étaient aussi l'innovation logicielle. L'infrastructure a généré de multiples réorganisations et échanges entre les partenaires, qui ne constituent pas directement notre objet d'étude, mais sont le contexte d'observation des activités de conservation et de recherche dont nous appréhendons les enjeux communicationnels. Nous revenons à présent sur les spécificités de notre enquête construite autour d'un objectif de conception logicielle constituant l'un des livrables d'une infrastructure de recherche en cours de développement, et sur ses rapports avec une démarche de recherche-action.

6. Construire une démarche de recherche-action autour d'un objectif de conception d'application web

6.1. Une démarche de recherche-action fédérée par un objectif de conception

L'objectif de conception a fédéré une démarche qu'on peut rapprocher de la recherche-action, attribuée à Kurt Lewin, et qui a fait l'objet de multiples développements dans différentes disciplines, notamment en communication des organisations, dans l'optique de « développer des synergies entre la recherche académique et la recherche appliquée liée à une demande sociale » (Meyer, 2006). En l'occurrence, les programmes de numérisation génèrent de la part des institutions culturelles et scientifiques qui s'y engagent une demande d'élaboration d'applications favorisant l'usage scientifique des documents et la participation d'internautes non spécialistes aux tâches documentaires. Cette demande sociale génère un besoin d'accompagnement du changement, dont le chercheur peut être partie prenante, en tant qu'« agent du changement » (Meyer, 2006). L'une des expériences ayant conduit Kurt Lewin à mettre en place une démarche de recherche-action était une question de diffusion de l'information de santé publique, avec le constat que l'implication de membres de la population dans la démarche facilitait l'appropriation de l'information. Dans notre cas, le travail de conception peut être vu comme un accompagnement du changement des pratiques scientifiques mobilisant les collections : notre intervention vise à favoriser l'intégration du spécimen numérisé aux méthodes scientifiques conformément aux objectifs des politiques favorables aux programmes de numérisation (Hine, 2008), et à accompagner l'intégration des usagers à la correction des données des étiquettes informatisées des spécimens. La démarche de conception, en partant du recueil des besoins et de l'observation des pratiques des usagers, et en intégrant des ateliers de test des prototypes, constitue « un processus de construction collective d'une représentation, qui passe notamment par une étape de mise en commun des perceptions, de confrontation, de négociation et de délibération entre différentes subjectivités » (Meyer, 2006). Elle a pour enjeu de favoriser le développement d'un statut de document scientifique accordé au spécimen numérisé, qui passe par son exploitation à des fins de production de connaissances et non de simple intermédiaire d'accès aux collections. La participation du chercheur au travail de conception, à partir de ses observations et avec un rôle de synthèse des besoins exprimés par la diversité des usagers rencontrés, le conduit aussi à proposer des spécifications fonctionnelles. Cela fait bien du

chercheur un « coacteur » qui « intervien[t] pour répondre aux besoins et aux préoccupations qu'il partage avec les personnes provenant des milieux concernés » (Meyer, 2006).

La démarche de conception que nous avons suivie correspond aussi à une progression itérative associée à la recherche-action (Roy et Prévost, 2013, p.136 ; Parrini-Alemanno, 2004, p.58). Elle conduit en premier lieu à recueillir les besoins directement exprimés par l'utilisateur sur la base des dispositifs existants, mais aussi à observer les pratiques sur l'herbier papier et à analyser les raisons du non-usage de l'image, afin de définir en termes de spécifications fonctionnelles les conditions de la reconstitution des gestes opérés sur les spécimens dans des dispositifs numériques. Cette première phase d'observation des pratiques construit une représentation du statut du spécimen numérisé dans l'ensemble des documents mobilisés, et permet de caractériser les pratiques d'accès et d'exploitation de l'information qui le mobilisent par rapport à celles qui utilisent d'autres documents. Cette phase correspond à l'élaboration des connaissances, décrite dans la démarche de recherche-action comme construction « d'un point de vue sur la communication interne de cette organisation et des préconisations pour l'organisation grâce au tri des données sur leur critère de récurrence et richesse de sens » (Parrini-Alemanno, 2004).

Cette première phase d'enquête à partir des multiples dispositifs existants s'est prolongée par un travail de conception. Il s'est appuyé sur la synthèse des besoins exprimés par les usagers dans le cadre des entretiens et questionnaires (voir annexe C.1). La conception a porté sur la recherche de solutions visant à dépasser les limites rencontrées dans l'exploitation de l'image avec les dispositifs existants. La première phase d'élaboration de spécifications fonctionnelles à laquelle nous avons participé s'est suivie par le travail des ingénieurs du WP5 pour déterminer les solutions techniques susceptibles de répondre au mieux à l'ensemble des spécifications (voir annexe B). Pour la version 2 du site « Les Herbonautes », le cahier des charges élaboré à partir de l'enquête d'usages et des discussions avec les animateurs a été repris par un prestataire extérieur de développement web. Pour le laboratoire en ligne, les spécifications ont connu une étape supplémentaire, celle des ateliers de test pendant lesquels les retours des usagers ont été collectés pour déterminer des ajustements nécessaires.

Les retours des usagers sur l'utilisation du laboratoire en ligne, comme l'analyse des données produites à partir des outils de catégorisation de la version 2 de la plateforme des Herbonautes, enrichissent les usages du spécimen numérisé par rapport à la situation initiale. Les nouveaux logiciels incluant les retours des usagers permettent d'observer les pratiques renouvelées par les dispositifs conçus dans le projet, qui sont susceptibles de rendre visibles des innovations méthodologiques dans l'exploitation scientifique de l'image, qui n'ont pu se développer auparavant que marginalement, faute de dispositif adapté. Le processus de conception enrichit le questionnement des nouvelles formes de communication médiatisées par les spécimens numérisés, dont l'exploitation scientifique suppose l'élaboration de dispositifs de médiation numérique adéquats que l'enquête du chercheur contribue à définir. La recherche-action fédérée par l'objectif de conception est adaptée pour appréhender le développement de pratiques émergentes et les mutations des

formes de communication médiatisée par des dispositifs documentaires en cours de transformation. Elle conduit à appréhender le renouvellement des dispositifs de médiation documentaire non pas comme un phénomène extérieur aux pratiques mobilisant les collections, mais intrinsèquement lié à la manière dont les acteurs envisagent les enjeux du numérique pour leurs pratiques et acceptent de s'investir dans un travail de conception généré par le contexte de l'infrastructure de recherche.

La participation du chercheur au travail de conception de dispositifs de médiation dont il étudie également les enjeux communicationnels peut aussi être légitimée par les travaux épistémologiques de Joëlle Le Marec (2002) destinés à prendre en compte les situations de communication générées par les enquêtes de terrain : si on considère la participation à la conception au même titre que d'autres formes d'interactions de l'enquêteur avec les acteurs du terrain, cette implication n'est pas vue comme un biais, comme ce serait le cas dans une perspective de neutralisation des interactions du chercheur avec les personnes auprès desquelles il mène son enquête. Il s'agit au contraire d'inclure à la description des formes de communication étudiées celles qui sont générées par les interactions suscitées par l'enquête, et en l'occurrence par le travail de conception. Cette approche nous a conduits à considérer l'implication des acteurs de l'infrastructure e-ReColNat dans le processus d'élaboration d'outils innovants et les interactions qui en résultent avec de nouveaux partenaires, y compris des chercheurs en sciences de l'information et de la communication, comme un premier niveau de mutation de la communication scientifique induit par le contexte de la numérisation des spécimens.

Cette posture de recherche-conception est aussi susceptible de donner un éclairage original sur les médiations de la communication scientifique mobilisant les collections dans le contexte de leur numérisation, puisque la participation du chercheur à la conception de nouvelles fonctionnalités des bases de données l'inclut dans l'ensemble des médiations en cours d'évolution qu'il étudie. Le travail de conception s'inscrit en effet dans l'étude des médiations qui rendent possibles les phénomènes communicationnels : il peut en lui-même être considéré comme une action de médiation définie selon Yves Jeanneret (2014, p.13) comme « activité productive et créative qui consiste à intervenir sur le cours de la communication en lui apportant une dimension nouvelle », et qui conduit « le chercheur à interroger sa propre place dans la circulation sociale des savoirs » (Jeanneret, 2008, p.266). Une démarche de recherche-conception qui appréhende les médiations en jeu dans la production de connaissances à partir des spécimens numérisés vise délibérément à influencer les formes de communication observées dans la production des connaissances, en favorisant leur développement par une meilleure instrumentation de l'exploitation des spécimens numérisés. Pour prendre en compte notre propre place dans les phénomènes communicationnels observés, nous avons cherché à mettre nos observations en relation avec notre rôle dans le projet, et les situations d'entretiens orientés par la collecte de besoins que cette fonction justifiait. Nous avons pour cela utilisé les notions de la sémiotique des transactions coopératives, afin de disposer d'un cadre commun à la description des interactions du chercheur avec les enquêtés, au déroulement du projet et aux activités observées.

6.2. Opérationnaliser la sémiotique des transactions coopératives pour la conception

Le double statut accordé aux artefacts dans la STC – comme moyen et finalité, selon les étapes du flux de transactions qui composent une activité – permet de rendre compte du processus de conception itératif associé à l’élaboration de connaissances théoriques qui a lieu dans la démarche de recherche-action : dans ce processus, le prototype issu de l’analyse fonctionnelle peut être considéré comme l’aboutissement de la première phase des observations, mais il est réinvesti comme nouvel artefact d’enquête permettant de recueillir les réactions des usagers lors des démonstrations. Ces réactions enrichissent les premières observations faites à partir des logiciels existants, et permettent de préciser les besoins et de suggérer de nouvelles pistes d’innovation méthodologique par la manipulation des prototypes.

La notion de transaction s’est avérée opératoire pour qualifier nos interactions avec les enquêtés, en soulignant leur caractère « productif », au sens où la rencontre produit de multiples artefacts. Aux enregistrements sonores, photographies, captures vidéos et aux notes prises pendant ou après les entretiens se sont ajoutés les documents et objets qui nous ont été donnés pendant l’enquête. Deuxièmement, l’enquête a produit des réflexions entrant dans le processus de conception, consignés dans des documents internes à l’équipe du WP5 : synthèses de notes de réunions et de collectes de besoin (voir annexe B), décrivant le répertoire d’acteurs impliqués dans l’utilisation des collections, et mettant en relation leurs besoins selon leur statut et domaine ; documents intermédiaires permettant de cheminer progressivement vers la définition de besoins et de spécifications fonctionnelles.

La notion de transaction médiatisée est aussi opératoire pour guider les observations dans un sens permettant d’aboutir à une analyse de besoins et à l’élaboration de spécifications fonctionnelles. Nous sommes ainsi partis de l’observation de l’activité centrée sur le rapport à un artefact. L’analyse des différentes activités observées se laisse tout d’abord grouper selon les métiers, orientés plutôt sur des activités de conservation ou de recherche – même si les deux types d’activités peuvent être liées, avec des recherches sur l’histoire des collections qui sont un préalable à leur exploitation scientifique. Ensuite, chaque activité mobilisant différents dispositifs est décomposée comme un ensemble de transactions, que l’on regroupera selon les formes d’interprétation qui s’y déploient, les formes de production de données ou de documents, et les interactions qu’elles supposent entre acteurs de différents statuts. Le cadre commun de l’analyse des transactions à l’œuvre dans les activités conduit à éclater la logique professionnelle pour rapprocher les formes similaires d’utilisation des documents et des dispositifs, transverses à différentes professions et finalités. Outre son utilité pour la conception, cette analyse permet de reconstituer la circulation d’un document à mesure des différentes interprétations et annotations dont il fait l’objet, ce qui permet de faire ressortir la complémentarité des activités qui le mobilisent. Les formes de communication en jeu dans l’infrastructure e-ReColNat et dans la mobilisation des collections dans le contexte de leur numérisation sont ainsi appréhendées à travers ces ensembles de transactions et l’étude de leur complémentarité. Les différentes transactions

peuvent encore être confrontées selon la variable du rôle qu'elles accordent au spécimen numérisé : c'est une manière d'appréhender la valeur documentaire qui lui est conférée, à partir des usages, selon les différentes activités et les habitudes de travail. Nous précisons à présent les modalités de collecte de données de l'enquête réalisée autour de cet objectif de conception et orientée à l'analyse de l'activité par la sémiotique des transactions coopératives.

6.3. Recueil et analyse des observations et des données

6.3.1. Enquête ethnographique par entretiens et observations à l'échelle de l'infrastructure

Notre enquête s'est fondée sur les observations utiles au recueil des besoins, et plus généralement sur le travail de conception des applications, étant donné l'entrée en relation qu'il impliquait avec les usagers. Elle relève de la recherche qualitative, même si elle sera articulée à des méthodes quantitatives à certaines phases, pour appréhender l'activité en ligne en complémentarité des observations.

Compte tenu des premiers objectifs attendus des livrables du WP5, une enquête sur les activités mobilisant les collections dans leur diversité a été réalisée, en incluant les problématiques de valorisation patrimoniale — alors que celles-ci sont finalement passées au second plan dans la priorisation des fonctionnalités à développer, conformément aux objectifs de l'infrastructure de favoriser le travail de la communauté des systématiciens. Ce sont les pratiques scientifiques qui ont été analysées de manière plus approfondie, au vu de la direction prise par le projet, mais le champ initial de l'enquête ouvert aux activités de valorisation pour différents publics était aussi cohérent avec la diversité des missions assumées par les acteurs rencontrés. Outre le travail des conservateurs qui s'investissent sur l'exposition de spécimens ainsi que sur la présentation de leurs collections, notamment lors des journées du patrimoine ou de la fête de la science, le travail scientifique s'accompagne d'actions destinées au public scolaire, et de la rédaction d'ouvrages de vulgarisation.

L'observation des différentes activités a demandé d'adopter plusieurs techniques selon les phases du projet : entretiens réalisés en préalable aux spécifications fonctionnelles des logiciels ; observations lors de réunions de projet ; entretien de groupe et observation des tests des prototypes lors des ateliers de test organisés par les ingénieurs du WP5. S'est ajoutée à ces observations des pratiques l'étude des sites web et logiciels existants utilisés par les usagers des collections (bases de données sans possibilité d'enrichissement, sauf l'ouverture à des commentaires pour certaines ; plateformes proposant la transcription des informations des étiquettes associées aux spécimens à des contributeurs non spécialistes), ainsi que l'analyse des traces d'usages disponibles dans les sites ouverts à la contribution.

Le travail de conception participait à configurer les observations de la première phase de l'enquête, puisque l'entretien était contextualisé par le rôle qu'il avait

dans le cadre d'une enquête préalable aux développements des livrables du WP5 pour l'infrastructure e-ReColNat, au moment de la prise de contact par mail, au début de l'entretien et dans les questions de relance concernant les attentes sur les outils numériques de l'infrastructure, ou l'analyse des besoins à partir des limites rencontrées dans les dispositifs utilisés. L'orientation de l'entretien par la démarche de recueil des attentes et de définition des besoins a permis d'inclure les représentations de l'utilisateur sur son utilisation des outils numériques et ses projections quant à leurs enjeux pour sa discipline de recherche. Les entretiens semi-directifs étaient orientés par une première demande de description des activités réalisées avec les herbiers physiques et numérisés, et par des questions de relance, pour préciser chacune des activités selon son objectif, les documents et outils utilisés, ainsi que les interactions avec d'autres personnes. Des relances étaient faites sur l'intérêt et les limites des outils numériques, les idées d'améliorations qu'ils suscitent, et les attentes par rapport à l'infrastructure. L'intérêt des différentes bases de données d'herbiers utilisées et leurs limites justifiant le recours aux collections physiques étaient questionnés, se prolongeant dès que possible par des démonstrations de leur utilisation, sur le poste du bureau dans lequel avait lieu généralement l'entretien. En plus de ces démonstrations, portant notamment sur l'utilisation des référentiels taxonomiques et des bases de données de collections numérisées, trois situations d'observation nous ont permis de mieux cerner les opérations centrales pour le travail avec les collections. La recherche de spécimens types mentionnés dans des publications a été observée en participant à une journée de formation organisée par Tela Botanica et le personnel ayant participé à la numérisation financée par la « Global Plants Initiative » ; nous avons assisté à l'identification d'un spécimen (la détermination de l'espèce à laquelle il appartient) à partir d'une image de base de données, par le chef du projet lors d'une entrevue à l'occasion d'une démonstration du logiciel Xper, ainsi qu'à la mesure de caractères morphologiques à partir d'images, lors d'une séance de travail de deux chercheurs au MNHN. Les situations d'observations ont été analysées à partir des notes prises pendant et après les entretiens et observations avec les documents rapports (photographies, films), directement exploitées pour dresser un recueil des besoins. Certains entretiens ont fait l'objet d'un retour plus approfondi et d'une transcription partielle, mais l'ensemble des tâches impliquées par la participation au projet de conception, s'ajoutant aux multiples demandes de communication sur le projet, n'ont pas permis pour le moment de procéder à des transcriptions entières.

Le choix des chercheurs avec qui les entretiens ont été réalisés a été élaboré à partir de la demande, adressée au coordinateur du projet, d'une dizaine de contacts de chercheurs travaillant dans des disciplines variées, afin d'établir un premier panorama des usages des herbiers et des attentes liées aux outils numériques. Ce premier groupe était centré sur des chercheurs MNHN et des partenaires de Montpellier. Des réunions du projet (réunion de lancement, réunions en vue de la définition d'un module de contrôle qualité, réunion de coordination entre plusieurs *workpackages*, conseils scientifiques...) ont donné lieu à des prises de contacts et entretiens complémentaires, notamment avec des chercheurs, des conservateurs de collections, des responsables de conservatoires botaniques nationaux, des salariés du réseau Tela Botanica, le responsable de l'inventaire des types de paléonto-

logie, des informaticiens du MNHN. Cela a permis de mieux appréhender le travail des différents WP, et en particulier des différents métiers qu'impliquait la constitution de l'infrastructure e-ReColNat. L'évolution des attentes du conseil scientifique et du coordinateur du projet a demandé d'autres entretiens, auprès d'usagers travaillant avec le logiciel Xper, développé par un laboratoire du muséum, pour lequel le coordinateur souhaitait envisager l'intégration aux outils d'e-ReColNat. Se sont encore ajoutés d'autres entretiens organisés à l'occasion de déplacements occasionnés par des événements scientifiques : deux colloques au Maroc ont permis la visite de l'herbier de l'Institut scientifique de Rabat, partenaire d'e-ReColNat ; un colloque à Grenoble a été l'occasion d'organiser un entretien avec une conservatrice de l'herbier de Lyon qui a rejoint e-ReColNat quelques mois plus tard ; une journée doctorale à Lille a donné lieu à un entretien complémentaire avec un mycologue conservateur de l'herbier de l'université de Lille 2 ; une journée d'étude en amont du colloque de l'Acfas organisé à Montréal par Florence Millerand, Lorna Heaton et Patricia Dias Da Silva a été l'occasion de rencontrer le personnel de l'herbier Marie-Victorin, et de profiter du dernier jour du déplacement au Canada pour faire une démonstration des applications développées par e-ReColNat et recueillir les retours des bénévoles, du conservateur, et de quelques chercheurs qui s'étaient réunis pour l'occasion. Ces entretiens complémentaires ont été précieux pour appréhender le projet avec le point de vue des personnels des institutions qui n'en étaient pas partenaires, voire avaient initialement refusé d'y participer. Les personnes extérieures au réseau étaient notamment intéressées par les logiciels développés dans le cadre du projet, sans avoir nécessairement d'intérêt particulier pour les collections d'e-Recolnat. Les déplacements personnels et les visites occasionnelles de muséums ont aussi été exploités pour enrichir la connaissance d'un monde initialement complètement étranger à toute l'équipe de conception du Cnam, qui supposait d'être apprivoisé progressivement – à l'échelle d'un réseau, de ses dynamiques régionales, non réductibles aux pratiques d'une unique institution.

Les analyses de la présente thèse se sont fondées sur les matériaux traités au fur et à mesure de l'enquête au service de l'objectif opérationnel de conception. Les comptes-rendus d'entretiens et d'observations, les documents et objets qui nous ont été confiés sur le terrain font l'objet d'un travail d'archivage à partir des services de la TGIR Huma-Num (Nakala et Nakalona, voir annexe E). Cela constitue un prolongement des réflexions développées dans cette thèse sur les conditions de l'archivage des données de recherche mobilisant les herbiers (chapitre 8) dans la perspective de leur application à la conservation des données d'une enquête ethnographique.

6.3.2. Analyse des dispositifs numériques et de l'activité en ligne

Se sont ajoutés à ces observations l'analyse des sites web et des logiciels existants mentionnés par les usagers des collections pendant les observations. La prise en compte des principaux dispositifs utilisés par les acteurs pour consulter et exploiter les données a conduit à analyser d'autres sites que ceux développés par

l'infrastructure e-Recolnat, notamment des bases locales des partenaires, ou des bases internationales. Les bases de données ouvertes à la contribution ont donné lieu à une analyse de données produites par les usagers. Nous avons plus particulièrement analysé deux sites intéressants par leurs possibilités d'intégration de commentaires et d'annotations, la base du « Global Plants Initiative », à laquelle plusieurs collections du réseau ont participé (Paris, Rabat, Montpellier notamment), ainsi que l'herbier virtuel du Brésil, avec lequel le MNHN avait un partenariat. Afin d'en permettre la comparaison, l'analyse des dispositifs a été réalisée à partir des catégorisations que nous avons utilisées pour l'herbier et issues de la sémiotique des transactions coopératives : régimes de diffusion, de fragmentation, de référentialité, et de contribution.

Les dispositifs ouverts à la contribution se sont prêtés à une étude de l'activité en ligne, par une analyse des contributions, complémentaire de celle observée en entretiens. L'analyse des commentaires issus du site « Global Plants Initiative » a été réalisée à partir d'une structuration par des expressions régulières du corpus récupéré sur le site — préalable à la réalisation de statistiques descriptives. La plupart de nos analyses de la participation en ligne ont été consacrées à la plateforme des Herbonautes, qui avait aussi été abordée dans les entretiens avec les chercheurs qui y participaient comme contributeurs ou animateurs.

L'étude de la participation générée par la plateforme avait commencé par un questionnaire en ligne (ayant obtenu 126 réponses de contributeurs et 76 d'internautes non contributeurs, entre janvier 2014 et novembre 2015), ainsi que des entretiens téléphoniques, entrevues ou questionnaires plus approfondis (selon les préférences des enquêtés) avec une dizaine de contributeurs réguliers. Ils visaient à mieux appréhender les motivations et les techniques de lecture et de recherche d'information mobilisées. Nous avons ensuite analysé l'ensemble des inscriptions nécessaires à produire les transcriptions validées. Elles comprennent l'ensemble des saisies proposées pour chaque étiquette, par tous les contributeurs y ayant participé, y compris les annulations et modifications toujours enregistrées, ainsi que les commentaires associés aux missions ou aux spécimens. S'y ajoutent aussi le texte présentant chaque mission et les messages d'animation (alertes) de la plateforme. Ces données sont pour la plupart invisibles pour les contributeurs (chacun ne voit que ses propres transcriptions et celles qui ont été validées, ainsi que les commentaires des espaces de mission, même si dans la version 2 du site une page récapitule l'ensemble des discussions récentes sur la plateforme). Elles ont été analysées à partir d'un export des données anonymisées de la plateforme, visualisées dans un logiciel de gestion de bases de données et examinées par le biais de requêtes SQL. L'extraction des données de la plateforme (base de données Oracle), réalisée par l'un des responsables informatique du MNHN, avait elle-même supposé un travail de conversion. Cette conversion a été effectuée par l'un des partenaires du WP5, directeur de la société Agoralogie, pour les données enregistrées de novembre 2012 à septembre 2015, puis par un des ingénieurs du WP5 pour le dernier jeu de données issu de la version 2 de la plateforme, qui suivait un format différent, de janvier 2016 à juin 2017.

L'analyse des espaces de discussion, mise en relation avec l'ensemble des données à l'origine des transcriptions, a été importante pour avancer dans la compré-

hension des ressorts de la production de données, ainsi que des enjeux de l'animation. Le site « Les Herbonautes » comporte un nombre important de commentaires, relativement au total des transcriptions réalisées : de novembre 2012 à septembre 2015, on compte 36 496 commentaires pour 125 808 spécimens proposés à la transcription, dont 77 919 complets ou signalés inutilisables. En comparaison, on compte 13 046 commentaires sur le site homologue de la plateforme Zooniverse, « Notes From Nature », pour un million de transcriptions environ d'avril 2013 à août 2015 — d'après un corpus de commentaires récupéré lors de la participation à une université d'été à Oxford, animée par des ingénieurs de Zooniverse en juillet 2015. La plateforme du réseau e-ReColNat s'est donc prêtée à l'étude de la fonction de ces commentaires, préalable nécessaire à la réflexion sur les outils de communication de la version 2 du site.

La taille du corpus des commentaires a conduit à envisager une étude lexicale automatisée avec le logiciel Iramuteq (Ratinaud, 2005), accompagnée de la lecture d'un ensemble de commentaires sélectionnés à partir de l'export de la base de données dans les discussions comportant plus de 15 messages, ainsi que dans les commentaires les plus longs et les plus courts. S'y est ajoutée, dans le cours de l'étude, la lecture de commentaires comportant des mots-clés pour lesquels l'analyse en cours soulevait l'intérêt. Les discussions de trois forums de mission ont aussi été prises en compte. L'annotation manuelle des messages associée à la lecture — avec les logiciels Glozz (Mathet et Widlöcher, 2011) et Analec (Landragin, 2011) utilisés en complémentarité — a facilité le repérage de fils de discussion dans les forums et la détermination de régularités dans les thématiques et les fonctions des messages par rapport à l'activité de transcription ; elle a constitué une aide à la lecture, dans une approche qualitative, sans pouvoir être systématisée faute de temps, ce qui aurait permis d'envisager une exploitation statistique des catégories dégagées. Dans une toute autre direction d'utilisation de ces logiciels, que nous ne faisons que suggérer, la littérature développée autour des applications de Glozz à des campagnes d'annotation collaboratives, pour effectuer des mesures d'accord inter-annotateurs, ouvre des perspectives d'application à l'étude de la production de transcriptions concordantes dans les plateformes de crowdsourcing, mais nous n'avons à ce jour réalisé que des calculs sur l'efficacité documentaire de la plateforme à partir du nombre de contributions nécessaires à valider les données d'un spécimen.

6.4. Catégorisation des transactions à l'œuvre dans l'ensemble des activités observées sur différents supports

L'ensemble des observations, y compris les retours sur les prototypes, a donné lieu au recensement des types de transactions à l'œuvre dans les activités mobilisant les collections en lien avec différentes bases de données. La sémiotique des transactions coopératives a été utilisée comme cadre des observations de l'enquête,

avec la double visée théorique et pratique de comprendre les formes de communication médiatisée par les collections, et de produire une analyse fonctionnelle du dispositif numérique à concevoir. Une fois l'ensemble des transactions recensées et regroupées selon les formes d'interprétation, d'écriture et d'interaction qu'elles mobilisent, nous avons choisi de présenter les transactions en les rassemblant selon le type de dispositif qui les encadre. Cela permet de montrer les complémentarités d'utilisation d'un même dispositif, et donc de reconstituer les formes de communication auxquelles il donne lieu. Nous avons ainsi été conduits à distinguer les transactions médiatisées par les bases de données de spécimens numérisés impliquées dans les activités des conservateurs, des gestionnaires de collections et des chercheurs. Nous distinguons ensuite les transactions inscrites dans la conception et le développement du laboratoire en ligne, destiné à l'exploitation scientifique du spécimen. Comme cette application n'a pas intégré de fonctions d'enrichissement de la base des collections à mesure des travaux de recherche, nous compléterons l'analyse des dispositifs numériques augmentant la médiation documentaire par l'ouverture des corrections aux usagers avec un état de l'art des herbiers numériques intégrant ces fonctions. Le dernier ensemble de transactions se distingue par les nouveaux publics qu'il implique, aux côtés d'animateurs professionnels, mobilisés par une plateforme adaptant les principes de l'externalisation vers la foule (crowdsourcing) aux tâches de transcription d'étiquettes d'herbier. Les bases de données utilisées pour la recherche et la conservation sont aussi utilisées par les usagers de cette plateforme, mais dans un contexte défini par la plateforme de transcription qui les encadre, et qui constitue donc le dispositif de référence justifiant notre regroupement.

Troisième partie

Mutations de la communication médiatisée par les collections scientifiques selon les dispositifs de médiation documentaire

7. La communication médiatisée par les bases de données de spécimens : émergence d'un travail d'inventaire en réseau renforçant les possibilités d'exploitation scientifique de l'image

Nous rendons compte à présent des mutations de la communication médiatisée par les collections dans le contexte de leur numérisation à partir de l'analyse des transactions impliquant les acteurs de l'infrastructure e-ReColNat et les spécimens des institutions partenaires. Nous commençons par analyser dans ce chapitre les transactions à l'œuvre dans les activités de conservation ou d'exploitation scientifique des collections, supportées par les bases de données de spécimens en articulation avec les collections physiques. Ces activités partagent une structure commune qui permet de les appréhender à partir du cadre transactionnel issu de la sémiotique des transactions coopératives, comme différentes interprétations de planches d'herbiers, numérisées ou non, qui conduisent à produire de nouvelles annotations ou de nouveaux documents. Les activités d'informatisation se distinguent par le fait qu'elles se traduisent par une production de données issues de la saisie standardisée des valeurs des différents champs décrivant le spécimen dans la base de données, et utilisent comme outils d'écriture les interfaces de saisie standardisées, avec pour objectif final de l'activité l'enrichissement de la collection. L'exploitation scientifique des spécimens est moins directement orientée à cette perspective d'intégration de données aux collections et d'enrichissement de leur documentation, même si certaines études y contribuent. Et les outils d'écriture associés à la consultation des collections à des fins de recherche restent plus personnels à chaque chercheur. Avant de présenter les transactions impliquées dans ces deux types d'activités, nous rappellerons que le contexte dans lequel nous les avons observées est marqué par la constitution d'une base centralisant les données des institutions partenaires d'e-ReColNat. Nous comparons la base de données de collections numérisées avec le dispositif de médiation documentaire de l'herbier, que nous avons décrit notamment comme un système d'archivage et de mise à disposition de documents scientifiques coordonnant une activité de production documentaire collective d'acteurs aux compétences hétérogènes, régi par un système d'organisation des connaissances directement issu des résultats scientifiques, mais compatible avec les multiples regards sur les collections, y compris l'intérêt patrimonial qui peut leur être porté.

7.1. Les caractéristiques documentaires de l’herbier amplifiées par les bases de données

7.1.1. La base de données d’herbier numérisés à l’aune de la typologie des médiations documentaires

Les bases de données de spécimens numérisés représentent le principal dispositif utilisé dans la communication scientifique — de l’élaboration des connaissances à l’archivage des sources sur lesquelles elles se sont appuyées, jusqu’à la communication des documents à d’autres personnes. Elles sont aussi consultées par les usagers de plateformes de transcription participative, mais dans un contexte et avec des motivations suffisamment différentes pour que nous nous y intéressions dans un chapitre distinct (chapitre 10). On peut dégager des éléments communs aux caractéristiques de l’herbier et à la définition de la base de données par la Commission ministérielle de terminologie informatique, comme un « ensemble de données organisées, en vue de son utilisation, par des programmes correspondant à des applications distinctes et de manière à faciliter l’évolution indépendante des données et des programmes » reprise par l’ADBS (2004). En effet, la structuration des données comme le classement des planches d’herbier est la condition de leur utilisation indépendante, selon des combinaisons définies par les besoins de l’usager. Il s’agit dans les deux cas de faciliter de multiples opérations et usages des données ou documents, conformément à la nécessité de pouvoir réutiliser les matériaux au travers des études, soulignée par Susan Star et James Griesemer (1989). Les bases de données issues de l’informatisation de chaque planche d’herbier permettent d’exploiter directement les données documentaires saisies pour chaque spécimen. La numérisation se traduit par des bases de données comportant en plus une table de données qui regroupent les métadonnées de l’image associées au spécimen, produites au moment de la numérisation (table IMAGES dans le schéma de la base Sonnerat dans la figure 7.1, à distinguer d’autres documents qui peuvent être associés au spécimen dans la médiathèque du MNHN).

La base de données perfectionne les possibilités de recherche dans les collections. Avant le travail de numérisation n’existe souvent dans l’herbier qu’un plan de classement référençant la place des liasses dans les compactus, ce qui conduit à devoir parcourir toute une liasse pour localiser un document recherché. La base de données accroît la précision de la « référentialité » de la planche d’herbier (Zacklad, 2015), en fournissant une URL unique à chaque spécimen numérisé, qui inclut le numéro du code-barres, afin de pouvoir citer la référence dans une publication (Hyam *et al.*, 2012). En termes de régime d’autonomie, qui distingue les documents « qui ne sont accessibles qu’en mode connecté [des] documents autonomes qui peuvent être dupliqués sur le terminal et édités en mode déconnecté » (Zacklad, 2015), le recours au téléchargement de planches classées selon leur espèce dans les ordinateurs individuels montre que l’autonomie du document est retrouvée. On peut même dire que l’autonomie du spécimen numérisé est accrue, étant donné la facilité à accéder à un document par rapport à la procédure de prêt nécessaire à pouvoir consulter les spécimens non numérisés. L’accès au spécimen est facilité

en dehors des heures d'ouverture des collections et, pour les usagers distants, des périodes où ils peuvent se rendre sur place.

Comme dans le cas des bases patrimoniales (Després-Lonnet, 2000, p.28), les annotations du spécimen acquièrent elles-mêmes un caractère plus structuré, puisque, même si les étiquettes suivent souvent une structuration en champs prédéfinis pour toute la collection, l'informatisation conduit à « morceler le texte portant sur un objet » (c'est-à-dire les inscriptions de l'étiquette dans notre cas), « pour le ventiler selon des rubriques ou des champs » (Brochu et Davallon, 1998), qui sont plus nombreux dans le cas de la base de données que sur l'étiquette de la planche d'herbier. Les éléments de description des étiquettes saisis dans la base sont standardisés et actualisés selon les normes d'organisation des connaissances correspondant aux classifications botaniques régies par un consortium (APG III). Cela rend possible une visualisation des données de la collection, synthétisant la répartition historique ou géographique des spécimens. En revanche, dans les bases de données ne sont pas toujours saisies toutes les informations des étiquettes, notamment les notes de terrain manuscrites.

On retrouve dans la base de données la possibilité de produire une nouvelle sélection de spécimens par la formulation des requêtes, qui remplacent la sélection de spécimens dans les liasses directement sorties des *compactus*. Si le mobilier des collections est configurable selon les besoins de l'utilisateur et si la facilité à déplacer les planches d'une liasse (non attachées, contrairement aux pages d'un livre) pour composer un corpus de travail permet déjà de parler d'une quasi-interactivité du document d'herbier, celle-ci est accrue par la base de données numérique. L'interactivité du document numérique permet aussi d'exploiter les liens hypertextuels pour faciliter l'accès aux spécimens à rapprocher, par exemple à cause de leur auteur commun ou de leur appartenance à une même espèce.

Une caractéristique importante de l'herbier est le régime de contribution qu'il rend possible. Dans l'herbier, comme dans le document pour l'action décrit par Manuel Zacklad (2015), il n'existe pas « une séparation tranchée et asymétrique entre les participants à la transaction, auteurs d'un côté et lecteurs de l'autre, et une dissociation de la temporalité de leurs activités » : chaque chercheur ou conservateur est alternativement dans la position d'examiner des spécimens et de les modifier, ou d'en créer de nouveaux. En revanche, il existe une asymétrie entre, d'une part, les acteurs de la conservation des collections détenant des droits de contribution dans la base de données, d'autorisation de visite ou de prêt, et d'autre part, les scientifiques qui ajoutent de nouvelles étiquettes après autorisation des conservateurs. La base de données reconstitue cette asymétrie au niveau des droits de saisie, qui ne sont pas ouverts aux usagers mais seulement à l'ensemble des gestionnaires du réseau e-ReColNat. L'ajout d'une étiquette sur le spécimen est conditionné par le respect d'une procédure de dépôt, auprès du personnel de l'herbier, des spécimens à annoter, pour que soit opérée une nouvelle saisie dans la base de données. L'ouverture des droits de contribution a été une demande importante, notamment de la part des conservateurs souhaitant apporter leur expertise sur les collections d'autres institutions, lorsqu'elles contiennent des documents produits par des auteurs dont ils détiennent une part importante de la production.

7.1.2. Centralisation des données et travail d'articulation au sein du réseau botanique e-ReColNat

Nous analysons dans le présent chapitre les formes de communication médiatisées par la consultation et l'enrichissement des bases de données, à l'œuvre dans le travail scientifique et la gestion des collections. Les transactions qui les composent, à mesure de l'informatisation et la production de nouvelles connaissances, s'appuient sur un travail préalable de conception et de développement de bases de données en réseau. Nous précisons donc pour commencer les transformations introduites par la constitution de l'infrastructure au niveau des réseaux informatiques : elles s'ajoutent aux modifications subies par les lieux de conservation, les artefacts matériels et les objets des collections.

L'infrastructure e-ReColNat suppose un nouveau travail de conception, développement et maintenance de bases de données en réseau, qui se superpose à celui qui a déjà lieu au sein de certaines des institutions partenaires. À l'échelle du MNHN a déjà été décrite, dans la présentation du réseau des collections e-ReColNat¹ (voir annexe A.2) une veille sur les besoins des usagers, et la collaboration des gestionnaires et chercheurs avec la direction du service de l'informatique, qui permet la mise à jour de la base de données des collections, Sonnerat, notamment avec l'ajout de champs correspondant aux besoins de certaines spécialités (comme les données sonores associées à la prononciation des noms vernaculaires des espèces). Les logiciels de gestion de collections, JACIM au MNHN, évoluent aussi selon les besoins (prise en compte de demandes de prêt et d'impression d'étiquettes) – le choix de ces logiciels de gestion de collections ayant été laissé à la responsabilité des institutions, avec toutefois la création d'un label e-ReColNat pour les logiciels interopérables avec la base du réseau. Outre la conception de l'architecture des données et de l'infrastructure informatique permettant leur conservation et leur mise à jour, la préservation de la cohérence des données malgré les multiples gestionnaires participant à l'informatisation et à la numérisation des spécimens (à raison de 80 spécimens par jour en moyenne) passe par un travail de vérification des données. D'après l'un de nos entretiens avec un responsable de gestion de données de l'infrastructure e-ReColNat, des corrections étaient déjà nécessaires, au sein du réseau des Herbiers de France qui regroupait déjà des institutions intégrant leur données à la base Sonnerat, pour pallier les incohérences de saisie entre institutions. Le même type de travail était réalisé au MNHN dans la perspective de la correction des données de spécimens types transmises dans le cadre du projet Global Plants Initiative.

L'objectif de centralisation et de gestion en réseau des collections e-ReColNat est passé par la conception et le développement d'une nouvelle base de données, livrable du WP2, conçue avec le souci de respecter l'indépendance de la gestion des collections par les partenaires et la diversité des moyens et de l'état d'avancement de l'informatisation. Nous avons observé le travail de conception de la base de données par la documentation et les communications qui en étaient faites au WP5 à mesure de l'avancée du projet. Comme illustré en figure 7.2, afin d'intégrer les données de collections locales, des applications et procédures ont été mises en place

1. <http://www.mnhn.fr/fr/collections/programmes-partenariats/reseau-herbiers-france>

pour convertir les données des bases des institutions partenaires dans le format développé par le Gbif (Darwin Core), et les rendre interopérables avec la base e-ReColNat où elles sont intégrées. Les données des bases locales sont ainsi moissonnées à partir d'un logiciel (IPT) qui réalise leur conversion (mapping) au standard international du Darwin Core (GBIF). Avant l'intégration à la base e-ReColNat, les données sont soumises à une vérification de leur qualité réalisée à l'INPN, instance chargée de l'inventaire du patrimoine naturel national et de la mise à jour du référentiel qui permet de le décrire (TAXREF) : selon le bilan d'activité de 2014 du WP2 rédigé par Julien Husson, « l'ensemble est ensuite analysé afin, d'une part, d'être introduit dans le système d'information de l'INPN, et d'autre part, d'ajouter de l'information en se basant sur les référentiels géographiques et taxonomiques de l'INPN. » L'infrastructure prévoit encore d'intégrer l'expertise de l'INPN pour identifier des espèces sensibles, pour lesquelles les données ne seront que partiellement communiquées. Enfin, la réutilisation de données de l'infrastructure à d'autres fins que celle de la consultation de la base de données via l'interface de consultation Explore², et notamment son exploitation dans les autres applications développées au sein de l'infrastructure, est ensuite rendue possible par une API³. Nous questionnons maintenant les enjeux de l'architecture des bases de données en réseau de l'infrastructure e-ReColNat pour les opérations d'informatisation et le travail de conservation des collections, envisagés dans leur dimension communicationnelle, qui sont aussi essentiels au processus de centralisation des données.

7.2. La communication médiatisée par les bases de données d'herbiers dans les activités de conservation

7.2.1. Un réseau documentaire morcelé

Les activités de conservation et de gestion des collections ont en commun des pratiques de recherche d'information sur les spécimens des collections dont les résultats sont consignés dans un catalogue informatisé de plus en plus souvent mis en ligne. Pour chaque spécimen est recherché l'ensemble des données sur la récolte, les publications dans lesquelles le spécimen a été cité (et le statut de type qui en découle le cas échéant), et les noms valides (selon les classifications actuelles) sous lesquels le spécimen est informatisé. Ces recherches ont lieu dans différents contextes : mise à jour régulière réalisée par les gestionnaires chargés d'identifier les publications écrites à partir des spécimens de leurs collections ; réponse à une demande de prêt ; programme de numérisation accélérant l'informatisation par de nouveaux recrutements. Comme l'expliquait une conservatrice à Lyon, ces opérations sont réalisées plus ou moins régulièrement, selon les disponibilités du personnel affecté à l'herbier, souvent en partie bénévole. Les demandes de prêt

2. <https://explore.recolnat.org/>

3. <https://api.recolnat.org/erecolnat/apidoc/#>

sont l'occasion d'une collaboration entre chercheurs et conservateurs ; cette collaboration bénéficie à la connaissance des collections, puisque l'établissement d'une monographie, « une liste exhaustive des noms donnés à un taxon », conduit un scientifique à « cherche[r] les publications » et « trouver les échantillons étudiés par les auteurs qui l'ont précédé dans les différents herbiers » (Schäfer, 2004, p.115), en collaboration avec le personnel.

Nous appréhendons la dimension communicationnelle de ces activités de conservation, dans les différents contextes où elles se déploient, en les décrivant à partir d'une structure transactionnelle qui leur est commune : elles interprètent les documents des collections non numérisées, avant de saisir les données dans une forme plus standardisée. Ces activités hybrident les environnements documentaires papier et numérique, puisque les opérations de saisie et de recherche d'informations complémentaires dans d'autres bases se fondent sur la lecture des étiquettes du spécimen papier, et parfois des carnets de récoltes non numérisés. La collection qu'il s'agit de valoriser constitue bien la finalité de ces activités de conservation, mais elle est tout autant l'artefact sémiotique (porteur d'information) transformé par l'opération d'interprétation et de standardisation (avec l'apposition d'un code-barres). La connaissance des collections progresse à mesure de ces transformations, qui concourent à produire un catalogue et à décrire et reclasser les documents. Les opérations de numérisation dynamisent ce processus essentiel à la conservation des collections ; il correspond à celui, décrit par Jeanneret (2008, p.58) pour penser l'archive, « par lequel tout ce qui a été produit dans la culture », et dans le cas qui nous concerne dans la connaissance des espèces, « est perpétuellement repris et transformé ».

La diversité de ces activités d'inventaire et d'informatisation entraîne leur progrès simultané et morcelé, sans empêcher que ces activités soient fortement liées les unes aux autres. Leur dimension communicationnelle ressort d'autant plus dans le contexte de bases de données en cours d'enrichissement simultané, du fait que les conservateurs et techniciens utilisent souvent des bases produites par d'autres. Une image déjà identifiée pourra être utilisée pour renseigner une planche moins complètement indexée, à partir du moment où il a été établi par certains indices (le nom du collecteur, la date de la collecte par exemple) que les deux spécimens proviennent de la même récolte. La recherche dans d'autres collections est aussi le moyen, pour qui connaît la facture des étiquettes d'un botaniste qui l'intéresse, de retrouver d'autres planches de sa collection personnelle. Les recherches dans les bases de différentes institutions sont d'autant plus fécondes que les collections se constituent par agrégation de différents herbiers, détenus par des botanistes ayant eux-mêmes procédé de leur vivant à de nombreux échanges de spécimens avec leurs pairs (Pérez et Pignal, 2013). Ces collaborations, corollaires aux échanges scientifiques qui ont présidé aux progrès des disciplines et à la constitution des collections, sont d'autant plus utiles que les connaissances en botanique sont très spécialisées et liées à un domaine d'étude, centré sur certains genres de plantes ou certains botanistes. Il est alors essentiel de pouvoir mobiliser les compétences réparties à l'échelle internationale, entre botanistes professionnels ou passionnés érudits, pour reconstituer les informations relatives à certaines collections éclatées

entre de multiples institutions. Pour faciliter cette collaboration dans la connaissance des collections, des répertoires internationaux de ressources se sont développés, tels que « *The Plant List* » (répertoire de tous les noms d'espèces existantes, mis à jour régulièrement)⁴, « *Tropicos* » (base de données précisant l'origine et la littérature associés aux noms d'espèces)⁵, et l'index des botanistes de l'université de Harvard⁶.

La dimension collective du travail documentaire nécessaire à rendre les collections plus accessibles à partir des bases de données s'exerce aussi au niveau du recensement des collections, organisé en France par Tela Botanica sur la base d'une incitation des acteurs locaux détenant des collections à les décrire par le biais d'un formulaire standardisé qui alimente le Catalogue des collections en ligne de Tela Botanica⁷. Le même principe existe à l'échelle internationale sur l'*Index Herbariorum*.

Cette dimension de travail en réseau s'ajoute à la dimension communicationnelle intrinsèque des activités de conservation qui se fondent sur l'anticipation des recherches des usagers, que facilite l'actualisation par les gestionnaires de collections des dénominations associées aux spécimens en cohérence avec les conventions actuellement valides. Les activités de conservation sont conditionnées par l'accès à l'information scientifique qu'elles concourent elles-mêmes à faciliter. Cela rend le libre accès aux données des collections particulièrement important pour favoriser le recensement et la connaissance des herbiers. Ces activités peuvent être rapprochées du travail des « réseaux documentaires », essentiels à la réalisation du « projet fondateur de la documentation – transmettre au plus grand nombre l'ensemble des connaissances produites » (Chartron, 2013), qui se sont développés à mesure des progrès de la documentation et de l'accès à l'information scientifique. Toutefois, par rapport aux initiatives des bibliothèques, l'unicité de chaque document d'herbier – caractéristique qu'il partage avec les documents d'archive – limite la collaboration à une logique de réutilisation du travail déjà réalisé. Certains projets, comme la réalisation d'un inventaire mondial des herbiers ou la mutualisation des bases de données taxonomiques permettant de tenir à jour la liste de tous les taxons (noms d'espèces) historiquement employés et leurs relations, relèvent davantage formes de collaboration proches de celle des réseaux documentaires conçus comme « ensemble[s] d'entités documentaires qui s'associent selon des critères variés pour collaborer à des objectifs communs, comme par exemple un catalogue collectif, une banque de données établie par plusieurs producteurs, un dispositif de prêt entre bibliothèques, un réseau de lecture publique » (Chartron, 2013). Au-delà de ces caractéristiques communes, nous précisons les principales opérations observées entrant dans les activités de conservation – recherche de types, inventaire – qui ont suscité des besoins spécifiques dans l'utilisation de l'image.

4. <http://www.theplantlist.org/>

5. <http://www.tropicos.org/>

6. http://kiki.huh.harvard.edu/databases/botanist_index.html

7. <http://www.tela-botanica.org/eflore/coel/appli/Coel.html>

7.2.2. Diversité et évolution des méthodes de recherche de types dans le contexte de la constitution de bases de données

L'activité de recherche de types fournit un bon exemple de reconfiguration du travail de conservation sous l'effet de la mise en ligne de nouvelles ressources. À Montpellier, le conservateur est revenu sur l'histoire des projets de numérisation, en soulignant que le travail initial de recherche de types avait justement permis d'obtenir un financement pour la participation au programme de numérisation des types de la « Global Plants Initiative », en s'appuyant sur la connaissance de la richesse de la collection en types, estimée d'après ceux qui y avaient déjà été identifiés. Les opérations de recherche de types à l'herbier ont eu lieu initialement à partir de la comparaison des publications, des carnets de récolte et des collections détenues sur place : il s'agissait d'identifier les spécimens correspondant à ceux mentionnés dans les publications. À présent, les types déjà en ligne simplifient ce travail de recherche, qui peut être conduit indépendamment des carnets et de la détention d'une œuvre complète d'un auteur : il suffit de pouvoir vérifier que le spécimen d'une collection provienne de la même récolte que les types déjà inventoriés pour qu'on puisse lui attribuer un statut de type (et plus précisément, d'isotype – ensemble des spécimens collectés en même temps que le type). La comparaison se fonde essentiellement sur l'étiquette, et peut donc se faire à distance sur la seule image de la planche. La recherche facilitée permet d'ouvrir à davantage de conservateurs cette activité de recherche de types, et de faire d'autant plus fructifier le travail documentaire morcelé des différentes initiatives d'informatisation. Cette démarche d'ouverture par la formation aux méthodes de recherche de types était l'un des objectifs de la journée de formation organisée par Tela Botanica en octobre 2014, pour diffuser les pratiques développées dans le cadre du projet « Global Plants Initiative » au profit de la connaissance du patrimoine naturaliste. La journée fut pour nous l'occasion d'observer la démarche et l'environnement de travail de la recherche de types dans la perspective de son instrumentation dans les applications d'e-ReColNat. La navigation sur le web facilite une consultation alternée de ressources répertoriées en favoris, aux côtés du spécimen papier : d'autres spécimens numérisés susceptibles de provenir d'une même récolte, sur lesquels il est possible de zoomer, et qu'on peut aussi retourner pour mieux déchiffrer les étiquettes ; bases de données biographiques sur les récolteurs ; bases de données taxonomiques donnant les références des publications citant les types ayant servi à décrire un taxon ; bases de données bibliographiques permettant de consulter ces publications pour vérifier si les données mentionnées sur la collecte correspondent à celles du spécimen observé.

L'accès aux types numérisés déjà répertoriés ne rend pas pour autant cette activité triviale, tant elle s'avère complexe : le conservateur de l'herbier de l'Université de Montpellier et son prédécesseur insistent sur le fait que les bases de données déjà constituées peuvent être erronées, ce qui exclut d'en faire une utilisation sans regard critique. D'autres informations concernant la collection dont le spécimen est issu (en particulier les pratiques de *lectotypification* – identification des types – propres à l'auteur d'un herbier) peuvent amener à apporter des modifications

aux statuts de types déclarés dans les bases de données. Il est parfois nécessaire de reconstituer des types, quand l'auteur ne s'est pas conformé aux règles en vigueur aujourd'hui. Prendre la mesure de la complexité de cette activité conduit à envisager des bases de données ouvertes à la contribution dialogique. De fait, le précédent conservateur investi dans la recherche de types à Montpellier est utilisateur régulier des fonctions de commentaire de la plateforme « Global Plants » – et c'est le seul commentateur régulier dans ce cas parmi les personnes ayant participé à notre enquête. Nous étudierons dans le chapitre 9 les spécificités des formes de communication médiatisées par ces bases ouvertes aux commentaires, peu nombreuses en comparaison des bases de données conçues essentiellement dans une logique de diffusion de l'information, auxquelles nous nous intéressons pour le moment.

7.2.3. De l'inventaire à la mise en valeur de collections historiques : enrichir l'organisation des connaissances fondée sur la systématique dans la médiation documentaire numérique

Nous avons recensé dans plusieurs herbiers des activités d'étude et d'inventaire des spécimens d'un collecteur. Ces activités sont parfois une activité amateur, comme nous l'avons constaté suite à la réunion des coordinateurs du réseau « Sauvages de ma Rue » à Tela Botanica en 2014. Nous y avons en effet rencontré l'auteure de l'inventaire amateur d'une espèce décrite et recensée par un botaniste local, Gaston-Sacaze (Albert-Llorca et Garreta, 2013); cette tâche a conduit la botaniste amateur à consulter dans ses loisirs différents herbiers et bases de données pour répertorier les différents spécimens de l'espèce en question. Ces opérations d'inventaire se traduisent souvent par un projet de valorisation destiné à communiquer le fruit des recherches sur une collection, son histoire, les territoires auxquelles elle est attachée, pour un public spécialiste ou non. Elles donnent lieu à des publications comportant un nombre d'éléments communs (Hoff *et al.*, 2009), que nous avons recensés dans la perspective de penser leur transposition dans un site web valorisant une collection numérique historique : éléments et illustrations bibliographiques, description de la collection, photographies d'étiquettes et de planches remarquables, liste des publications et localités mentionnées dans la collection. Elles donnent aussi lieu à des expositions, parfois en ligne, présentant une collection particulière avec quelques illustrations. Elles offrent en outre une contextualisation parfois nécessaire au travail scientifique, comme nous l'a expliqué une chercheuse en écologie à l'origine de l'étude de l'herbier Rivals, à l'occasion de sa restitution au muséum de Toulouse après une longue période de prêt au MNHN. Le travail sur la collection, en relation avec les descendants du collecteur principal, a donné lieu à une présentation en ligne⁸ et à une exposition à l'Université Paul Sabatier, revenant notamment sur le lien entre les récoltes et la seconde guerre mondiale.

8. <http://www.ecolab.omp.eu/ACTUALITES/Expo-herbier-Reunion>

Ces perspectives historiques sur les collections, et les travaux valorisant leur ancrage dans un territoire et une histoire locale, questionnent les dispositifs de médiation documentaire dans leur capacité à rendre possible ces regards sur les collections, correspondant à un autre point de vue que celui de la systématique. Si certains sites créent des espaces dédiés à la présentation de collections particulières (par exemple sur le site de l'herbier de Strasbourg⁹), ces projets sont parfois confrontés à l'inadéquation de leurs objectifs avec le projet de communication qui préside aux bases de données scientifiques d'herbiers numérisés. On le voit dans le cas de l'informatisation des collections d'ethnobotanique, qui ont demandé une structure de données spécifique permettant d'inclure d'autres informations collectées sur le terrain (voir figure 7.3). Les travaux de nomenclature (qui étudient les relations entre les noms accordés aux espèces à travers l'histoire) se heurtent aussi aux systèmes d'organisation des connaissances des bases de données régies par les conventions internationales de systématique : la valorisation des classifications et taxons très anciens trouve ses limites dans les bases de données qui décrivent les spécimens seulement selon les dénominations présentement utilisées (sous la forme d'un binôme genre espèce, qui n'a pas toujours été utilisé). Ces projets portent à leur maximum la tension décrite par Lorna Heaton et Serge Proulx (2012) entre les enjeux locaux de valorisation et la base conçue pour offrir des données standardisées, susceptibles d'être comparées entre elles, et par là même « délocalisées » des pratiques de gestion des collections locales, et *a fortiori* des herbiers historiques dans lesquels elles avaient initialement leur place. Cette tension met en évidence le besoin d'une réflexion sur l'articulation entre différents systèmes d'organisation des connaissances dans la médiation documentaire numérique, afin de rendre accessibles pour les pratiques systématiques tous les spécimens des collections, y compris historiques ou ethnobotaniques, sans empêcher d'autres formes de valorisation. Ainsi la base de données d'e-ReColNat a-t-elle été construite (par le WP2 dirigé par Julien Husson) en veillant à prendre en compte le besoin d'ajouter à un spécimen de nouvelles déterminations sans écraser celles qui ont été réalisées antérieurement. Le site donnant accès à l'herbier numérisé de Jean-Baptiste Lamarck¹⁰ donne aussi un exemple d'une articulation de deux choix d'organisation des connaissances, puisqu'on peut consulter l'herbier en naviguant selon les liasses et taxons historiques, tout en ayant accès par un lien hypertexte à la fiche de chaque spécimen dans la base de données du MNHN qui permet de retrouver le référentiel systématique actuel. Les entretiens avec des conservateurs investis dans des activités de nomenclature ont donné lieu à une amorce de solution, en envisageant de créer des catégorisations de spécimens et de les présenter dans une arborescence définie par l'utilisateur et régissant le plan d'un parcours de navigation. Plusieurs conservateurs ont aussi exprimé le souci d'avoir dans e-ReColNat un espace de présentation personnalisée librement éditable de leurs collections — demande dont porte la trace la page consacrée aux partenaires du site e-ReColNat¹¹.

La réflexion sur l'enrichissement des principes de description des spécimens

9. <https://herbier.unistra.fr/herbier-de-strasbourg/herbier-general/herbier-du-lycee-fustel-de-coulange/>
10. <http://www.lamarck.cnrs.fr/herbier.php>
11. <https://www.recolnat.org/en/nos-partenaires>

pourrait aussi bénéficier à des travaux de systématique. Nous l'avons déduit à partir de l'intérêt pour la notion de *récolte* (ensemble des spécimens collectés à une même date au même endroit par un même récolteur), dont nous avons pris la mesure à l'occasion d'une journée d'étude sur la valorisation scientifique des herbiers : un chercheur en systématique avait donné l'exemple de l'intérêt de consulter les spécimens d'une même récolte pour avoir une indication de l'abondance de l'espèce sur une station. Aujourd'hui, les spécimens d'une même récolte se suivent en général dans la liasse, et partagent donc un numéro d'inventaire proche. L'entité récolte est formalisée dans les bases de données (du MNHN, comme on le voit en figures 7.3 et 7.1), et dans la base d'e-ReColNat¹². Mais un conservateur du MNHN en cryptogamie (étude des plantes sans fleurs, comme les mousses), investi dans la recherche de types, suggérerait d'exploiter davantage cette table, afin de donner accès d'un clic aux spécimens d'une même récolte, pouvant partager le cas échéant le caractère de spécimen isotype. Outre le rassemblement facilité de types, l'exploitation accrue du concept de récolte déjà implémentée dans la structure des données laisse envisager des visualisations de l'évolution chronologique des spécimens et espèces récoltés sur une même station correspondant à une géolocalisation unique.

7.2.4. Perspectives de l'analyse des activités de conservation pour la conception de dispositifs numériques de médiation documentaire augmentée

L'analyse des activités de conservation a permis de faire ressortir la place importante qu'elles donnent à l'image, médium d'un travail d'informatisation qui progresse de façon parallèle. Ces différentes activités dégagent des besoins communs et plus spécifiques, en particulier liés à la prise en compte de la saisie des informations relatives aux spécimens comme résultat d'une activité complexe d'interprétation de documents des collections – activité qui suppose une connaissance complète de leur histoire, sujette à remise en question et contestation. Des outils d'écriture non standardisée sont d'autant plus pertinents que ce travail fait droit aux coopérations internationales, entre conservateurs, chercheurs, et gestionnaires des différents programmes d'informatisation amenés à réutiliser les informations issues de précédents inventaires. La prise en compte de la nécessité d'acquérir une connaissance des collections locales pour les informatiser conduit à envisager de reconstituer les outils de mémorisation observés chez certains gestionnaires. Ainsi, la présence d'un wiki utilisé par le personnel informatisant les collections avait été mise en évidence à Montpellier ; au MNHN, nous avons vu la réalisation par une gestionnaire d'un cahier de signatures et d'étiquettes pour faciliter l'informatisation. Nous avons observé également dans les herbiers les modalités de la communication avec les visiteurs, qui remplissent un cahier (à Rabat) ou un tableau informatisé (à l'herbier du jardin botanique de Lyon) retraçant les objectifs de toutes les visites. Les bases de données pourraient s'inscrire également dans ces relations, à condition d'envisager des comptes personnels associés aux demandes de prêt des collections physiques, facilitant les demandes de retour et le signalement des publications auxquels donnent lieu les spécimens.

12. <https://api.recolnat.org/erecolnat/apidoc/#>

Après avoir identifié les formes de transaction auxquelles donnent lieu les bases de données de spécimens pour les conservateurs et gestionnaires de collections, dans les étapes de constitution d'inventaires et d'informatisation des collections, mais aussi de recherche nomenclaturale et d'analyse de collections historiques, nous dégageons celles qui émanent des activités visant à produire de nouvelles connaissances dans différentes disciplines scientifiques.

7.3. Le spécimen numérisé dans l'exploitation scientifique des collections

Certaines des opérations réalisées par les conservateurs, comme la nomenclature ou la recherche de types, constituent également des étapes de l'activité des chercheurs. Par exemple, la « révision d'un genre » pour déterminer de nouvelles catégorisations d'espèces sur la base d'observations plus approfondies, par exemple croisées aux données génomiques, commencera par une revue de la littérature et une recherche des types, dans la perspective d'apporter une éventuelle critique aux descriptions et catégorisations antérieures. Nous nous intéressons à présent aux opérations plus spécifiques entrant dans l'activité de recherche scientifique mobilisant les collections, qui s'insèrent dans la communication scientifique au niveau de l'appropriation de la littérature existante et de l'élaboration de nouvelles connaissances par examen de spécimens existants, couplées parfois à de nouvelles récoltes. Nous identifierons le rôle des bases de données dans ces opérations parmi les différents outils de consultation et d'écriture mobilisés pour travailler avec les spécimens, et le statut donné au spécimen numérisé. Pour appréhender la dimension communicationnelle de l'activité scientifique mobilisant les collections, nous conservons le même cadre transactionnel facilitant la comparaison entre l'exploitation des spécimens et des bases de données dans des travaux de conservation ou de production scientifique. Les transactions correspondant aux pratiques de recherche d'un document et d'accès aux spécimens ont été repérées dans toutes les activités; d'autres transactions sont plus spécifiques à certains objectifs scientifiques — nous les aborderons dans un second temps.

7.3.1. L'accès aux spécimens renouvelé par la base de données

Le premier ensemble de transactions en jeu dans les activités scientifiques mobilisant les bases de données de spécimens est constitué par les opérations d'écriture, de lecture, et les interactions intervenant dans l'accès aux collections, papier ou numériques; ces transactions sont communes aux activités qu'on distinguera ensuite par le statut qu'elles confèrent au spécimen. La recherche de documents dans la base de données de la collection peut s'analyser comme une sous-transaction qui, par la formulation d'une requête, aboutit à la composition automatique d'un ensemble de résultats rassemblant des unités documentaires fragmentées (les planches). Ces unités documentaires vont faire l'objet de tris et de filtres pour constituer un sous-corpus d'étude répondant au besoin du chercheur,

qui pourra être téléchargé ; l'opération sera répétée dans différentes bases. Nous avons observé des pratiques de téléchargement qui conduisent à élaborer des collections personnelles d'images issues de requêtes dans les catalogues en ligne des différentes institutions à l'échelle internationale.

L'accès facilité depuis la base de données ne se substitue pas à l'accès dans les collections physiques, et rencontre ses propres difficultés, en particulier quand les principes de classement des collections ne correspondent pas à la logique de l'étude. L'efficacité de la recherche avec la base du MNHN est maximale pour les recherches par nom d'espèce, car ce sont les seules données précisément saisies au moment de la numérisation industrielle, aux côtés de l'indication d'une zone géographique large. Les recherches portant sur les spécimens récoltés dans une même localité, ou par un même botaniste, obligent à consulter les étiquettes d'un grand nombre d'images jusqu'à arriver aux spécimens cherchés. Imparfaitement informatisée, la base de données de spécimens issus d'un processus de numérisation industrielle ne résout pas entièrement la difficulté qui existe déjà avec les collections papier, qui doivent adopter un ordre de rangement fixe, attribué selon les besoins des disciplines et les questionnements de recherche privilégiés (ainsi, au MNHN, le classement de l'herbier par espèce s'est substitué lors du chantier de numérisation en 2012 au classement géographique, qui obéissait aux études de flores régionales moins dynamiques aujourd'hui). Mais quand elles décrivent suffisamment précisément chaque spécimen, les bases de données permettent d'affiner la recherche par des filtres ; la visualisation cartographique des spécimens géolocalisés permet aussi de rassembler les spécimens d'une zone géographique d'intérêt à partir d'une sélection dans une carte de répartition. Les possibilités de filtrage, en particulier par type, sont même susceptibles de transformer le choix des spécimens consultés. En effet, au MNHN, la consultation d'un spécimen type identifié par son numéro d'inventaire ne peut se faire qu'à partir de la recherche dans toute la liasse correspondante. Selon l'un des systématiciens impliqué dans la conservation des collections, cet état de fait présente des aspects positifs, en évitant de créer un biais lié à une présélection incomplète des spécimens types, et de réduire la variabilité de l'espèce aux caractéristiques des types. L'obtention de résultats de recherche très précis par filtrages successifs dans la base de données pourrait ainsi limiter les échantillons consultés – hypothèse qui demanderait davantage de recul sur les pratiques pour être confirmée.

De même, la relation entre usagers et conservateurs peut être affectée. Les usagers interrogés apprécient la plus large autonomie que leur donne la recherche dans les collections numérisées par rapport à la formulation d'une demande au personnel de l'herbier, qui, d'après l'une des chercheuses du MNHN, ne permet pas toujours de « savoir ce qu'ils ont », surtout dans les cas où l'herbier souffre d'un manque de personnel. Réciproquement, les personnes chargées des prêts ont attesté recevoir des demandes de visiteurs de plus en plus précises. Toutefois, il reste souhaitable selon eux que la consultation de la base se poursuive par un échange avec les conservateurs, dont la connaissance locale de la collection peut s'avérer essentielle pour enrichir la sélection de spécimens faite par l'utilisateur.

Malgré l'avantage qu'elle représente, la recherche dans la base de données pour accéder aux spécimens n'est pas généralisée. Les usages que les chercheurs en font

dépendent de leur appartenance institutionnelle et de l'accès qu'ils ont aux collections physiques. Certains chercheurs du MNHN utilisent peu les bases de données pour leurs propres activités scientifiques. Ils préfèrent parcourir les « compactus » (rayonnages d'herbiers). L'un de ces chercheurs rapporte comment il est parfois détourné de son objectif de recherche d'un spécimen précis par une anomalie de classement, qu'il corrigera, à moins qu'elle ne le mette sur la piste d'une découverte (espèce non encore décrite par exemple). Pour les chercheurs du MNHN privilégiant l'accès direct aux compactus, les bases de données du muséum servent principalement à répondre aux besoins des visiteurs auxquels leurs missions de conservation les conduisent à répondre, mais ils utilisent en revanche les bases d'autres collections importantes au plan international, notamment pour identifier et localiser les types d'une espèce. Au contraire, les chercheurs extérieurs au MNHN ont évoqué la collection numérisée comme un outil consulté fréquemment, en complémentarité d'autres bases internationales.

La sélection de spécimens pertinents pour les besoins de l'utilisateur ouvre sur l'étude de leur interprétation. À la différence des activités d'inventaire de collections des conservateurs, dans de multiples cas l'image ne suffit pas à procéder aux analyses auxquelles les chercheurs souhaitent procéder. La consultation de la collection numérisée est alors utilisée seulement comme une pré-visite, ou pour préparer la demande de prêt. Nous analysons à présent les différentes opérations qui font suite à la consultation de la base de données.

7.3.2. Le spécimen numérisé dans les opérations d'identification

L'identification est l'opération fondamentale à l'œuvre dans la constitution et la mobilisation des collections. Elle consiste à associer à un spécimen une espèce répertoriée, ou au moins un genre. Elle a lieu juste après la collecte, avant le dépôt en herbier, et est très importante pour constituer le document, puisque c'est sur cette base que sera attribuée à la planche d'herbier une place par rapport aux autres planches de la collection (à moins que le spécimen n'ait pas pu être identifié, et qu'il reste alors dans les spécimens « indéterminés » de chaque zone géographique). L'herbier intervient lui-même dans l'identification — c'est le premier des 72 usages de l'herbier recensés par Vicky Funk (2003) : « *discover or confirm the identity of a plant or determine that it is new to science* ».

L'identification peut se faire de différentes façons, qui varient selon les manières d'intégrer l'herbier et l'image à la démarche. À moins qu'on ne reconnaisse immédiatement une espèce connue, l'identification peut passer par l'observation méthodique d'un ensemble de critères significatifs dans un ordre préconisé par une *clé de détermination*, organisée comme une liste arborescente de questions permettant de reconnaître une espèce donnée. Pour les flores qui n'ont pas encore été entièrement décrites, notamment pour lesquelles on ne dispose pas de clé permettant de discriminer l'ensemble des espèces, on se réfère à l'herbier, en comparant l'échantillon du terrain avec des spécimens susceptibles de lui ressembler. Cette opération

nous a été décrite par un botaniste, ancien conservateur d'herbier, comme un travail d'enquête solitaire et ardu, mais essentiel au plaisir du travail en herbier. En cas d'échec, une demande d'aide à un collègue spécialiste peut être faite. Un autre chercheur signale que la photographie du spécimen facilite alors l'échange.

Nous avons assisté à deux situations d'identification mobilisant une image. Dans les deux cas, les spécimens identifiés à partir des images en ligne étaient destinés à rejoindre l'herbier institutionnel. L'identification suivait une démarche de comparaison du spécimen à identifier avec une série d'images du genre auquel on avait rattaché en première analyse le spécimen. Les images étaient regroupées, dans un cas dans un dossier spécifique constitué à partir de spécimens numérisés de différentes collections, et dans l'autre cas par la consultation dans la base de données du MNHN. Elle passait par l'observation des caractères morphologiques significatifs du nouveau spécimen et des images consultées, facilitée par les fonctions de zoom. L'un des systématiciens signalait tout de même l'inadéquation de la recherche dans la base de données pour l'identification : constat important pour le travail de conception, il faudrait pouvoir réaliser une recherche par filtrage successif correspondant à des caractères observés.

Ces pratiques d'identification au moyen d'images en ligne d'un spécimen physique destiné à rejoindre les collections nuance l'idée d'une dégradation des informations de l'image, qui entraînerait un statut nécessairement subordonné au document physique : le rapport est ici renversé puisque l'image devient un document de référence pour l'identification de nouveaux spécimens adjoints à la collection. On ne saurait néanmoins généraliser le statut de l'image dans la démarche, puisqu'il dépend des espèces et des spécificités des caractères à examiner pour les identifier. Nous revenons à présent sur des démarches scientifiques qui intègrent ces opérations d'identification et de consultation des bases de données au sein d'objectifs spécifiques supposant des opérations complémentaires.

7.3.3. Le rôle du spécimen numérisé pour les systématiciens

Les activités de taxonomie visent à identifier de nouvelles espèces, et à décrire plus précisément les espèces d'un genre, par exemple en tenant compte de leur variété géographique. Elles contribuent à préciser les relations entre les différentes espèces d'un genre, relations sur lesquelles reposent les classifications. Elles comportent bien une dimension transactionnelle, avec la mobilisation critique des documents produits par d'autres, réexaminés et reclassés, qui conduit à produire de nouvelles publications, des redéterminations dans de nouvelles étiquettes sur les spécimens réidentifiés, mais aussi à redéfinir les catégorisations de l'organisation des connaissances dans les classifications. Ces activités comprennent les gestes précédemment décrits de constitution de corpus – à partir de sélections dans les bases de données, complétées par les interactions avec les gestionnaires de collections de différentes institutions – et de lecture critique de l'étiquette et vérification de l'identification. Mais les activités de systématique se distinguent des méthodes utilisées pour les identifications, étant donné l'objectif de révision des critères distinctifs d'une espèce : il s'agit de produire de nouvelles clés de détermination, et

non d'appliquer les clés existantes. Pour cela, les caractères sont souvent observés au microscope (qui donnera lieu à la prise de clichés cités dans les publications), ou à partir des résultats d'une analyse moléculaire du spécimen, mis en relation avec les caractères morphologiques. Cela suppose des prélèvements, et rend la consultation du spécimen original obligatoire, parfois par l'intermédiaire du conservateur qui enverra des fragments pour les analyses. Certains procédés de morphométrie, en ce qu'ils portent sur des caractères internes d'une fleur séchée par exemple, ne peuvent se faire à l'image, demandant d'abord une réhydratation du spécimen puis une décomposition – c'est là le troisième usage des herbiers recensé par Vicky Funk (2003), « *provide material for making morphological measurements (taxonomy, systematics)* ». La limitation de l'utilisation de la base à un outil d'accès ou de pré-visite s'explique dans ces cas par l'impossibilité de procéder, à partir de la simple image, aux examens qui seraient nécessaires pour répondre à certaines questions de recherche. Mais l'analyse de ces cas est utile pour préciser les fonctions d'une médiation documentaire numérique adaptée, qui soit par exemple capable d'associer à un spécimen les résultats d'analyses, photographies ou mesures prélevées sur lui, parfois à partir d'une dégradation de certains fragments.

Malgré ces limites, les opérations de systématique ont donné lieu à l'observation de la mobilisation de l'image, au terme d'un travail de deux chercheurs du MNHN ayant conduit à décrire de nouvelles espèces. Les spécimens avaient déjà été recatégorisés, et des mesures devaient être prélevées pour compléter la description des deux espèces. La consultation des images était une aide à la phase finale de la rédaction de la publication scientifique, où il suffisait de compléter la description de l'espèce par la description de caractères discriminants, et la mesure de traits de caractères distinctifs. Procéder aux mesures à l'écran sur le spécimen numérisé a facilité cette étape de rédaction, par le rassemblement de l'article en cours, des spécimens numérisés et du tableau de prélèvement sur un même poste de travail informatique. La prise de mesure était réalisée par l'utilisation simultanée d'un tableur et d'un logiciel de mesure d'images intégrant la fonction de calibrage, permettant de s'assurer de la fiabilité de la mesure par rapport à la taille du spécimen original. La mise au point de cette combinaison de logiciels pour exploiter l'image était particulièrement intéressante dans la perspective de l'instrumentation de l'exploitation scientifique de l'image. À ce titre, le travail de mesure des images que nous avons observé était toutefois ralenti par l'absence d'outil de calibrage et de mesure intégré à la base du MNHN. Le tableau des caractères observés se retrouve dans d'autres activités, qui mobilisent les spécimens originaux : nous l'avons retrouvé dans les démarches scientifiques procédant par inventaire d'occurrences d'espèces dans des zones déterminées.

7.3.4. Rôle du spécimen numérisé dans les recensions de données

Si l'inventaire d'une collection est central dans les activités de conservation, on retrouve dans les activités scientifiques le geste de relevé de données standardisées à partir de la vérification de chaque spécimen, par l'observation des caractères morphologiques et la lecture critique des étiquettes. À l'échelle de notre enquête, nous

avons observé des démarches d'inventaire pour décrire la flore encore incomplètement connue d'une région (la Nouvelle-Calédonie) ou en connaître l'évolution, et réaliser un suivi des espèces d'arbre présentes sur une parcelle. L'inventaire peut être réalisé dans le cadre d'une surveillance écologique essentielle à l'établissement des politiques de conservation de l'environnement. Nous avons également rencontré des acteurs réalisant des inventaires pour les Conservatoires Botaniques Nationaux, qui nourrissent l'Inventaire National du Patrimoine Naturel, et utilisent les herbiers pour recueillir des données et les intégrer avec celles issues d'observations récentes recueillies avec le concours de différents amateurs, étudiants, et professionnels qui y travaillent.

Dans ces différentes démarches, l'image n'était pas mobilisée, puisque la consultation se faisait justement à l'herbier afin de pouvoir examiner chaque spécimen papier, dont il s'agissait de valider les données. La vérification des coordonnées de géolocalisation, signalées comme parfois erronées, est particulièrement importante et constitutive du travail d'inventaire, au même titre que celle de l'identification de l'espèce. C'est aussi la consultation des spécimens non encore informatisés qui justifie le déplacement dans un herbier par rapport à la seule consultation des bases en ligne. Une fois la donnée d'occurrence validée, elle est consignée, dans la base de données d'inventaire pour les Conservatoires Botaniques Nationaux, et dans un tableau adapté aux besoins de l'étude (et parfois partagé entre plusieurs personnes) pour les chercheurs. Les examens de spécimens pendant l'inventaire conduisent à enrichir la collection de nouvelles déterminations, notamment de plantes encore indéterminées en provenance de la zone étudiée par l'inventaire. Pourtant, il n'y a pas toujours d'intégration des données synthétisées dans l'inventaire à la base de données de l'institution détentrice des collections qui ont été inventoriées, en particulier à partir de la transcription des étiquettes.

Dans un entretien à propos de l'inventaire de la flore de Vanoise – réalisé à partir de spécimens du principal herbier détenant des collections du territoire concerné, à Genève, mais sans consultation de bases de données complémentaires –, une utilisation de la base de données du MNHN a été évoquée, suite à une question de relance, comme un moyen d'actualiser l'inventaire plusieurs années après sa constitution. S'est aussi dégagé de l'entretien l'intérêt pour une solution qui permettrait plus facilement d'intégrer des données issues d'un examen d'inventaire à la base de données institutionnelle.

L'inventaire fait donc une place assez restreinte à l'image : elle n'est pas le document principal utilisé. Les représentations cartographiques issues des bases de données sont un point de départ pour la démarche de l'inventaire, qui consiste à vérifier chaque donnée qu'il utilise, qu'elle vienne d'une planche papier ou d'une base de données. Une fois que toutes les données ont été validées, la visualisation cartographique peut aussi constituer un point d'aboutissement du travail. Les démarches d'inventaire ont aussi fait ressortir l'intérêt pour des outils permettant de sélectionner géographiquement un ensemble de spécimens, mais en faisant correspondre la zone géographique d'intérêt aux parcelles des inventaires. Il s'agirait donc d'intégrer, à la structure des données de la base, différents référentiels géographiques rendant possibles des sélections par « maille » correspondant aux normes des principaux programmes d'inventaire.

7.4. Le statut documentaire du spécimen numérisé dans la communication médiatisée par les collections

À partir de l'étude des activités mobilisant les collections et les bases de données existantes, nous pouvons à présent caractériser les formes de la communication scientifique dans le contexte de leur numérisation, en présentant l'ensemble articulé des transactions complémentaires réparties sur différents supports et intégrant différenciellement l'image, à l'œuvre dans l'accès, l'appropriation de l'information scientifique, la publication, et l'archivage de nouveaux documents. L'utilisation de la base de données de spécimens intervient dans les phases de recherche documentaire communes aux activités de conservation et d'exploitation scientifique des collections, qui en cela se recoupent. Ces recherches ne sont cependant qu'une première phase de travail pour le systématiseur, qui utilise plutôt les bases de données déjà réalisées (même si ses recherches peuvent l'amener à les compléter), alors que la constitution des bases de données de spécimens est au centre des activités de conservation. La production des connaissances à partir des spécimens a pour condition une collaboration entre conservateurs et chercheurs. Celle-ci peut être une coopération, par l'échange d'informations permettant de déterminer les spécimens pertinents pour l'étude du chercheur, ce qui amène à trouver directement dans la collection un type (Schäfer, 2004), ou à anticiper ces demandes et à réaliser ces recherches en amont des demandes, comme nous l'avons observé notamment à Lyon. Le travail d'actualisation des termes des collections, observé plus fréquemment, s'inscrit aussi dans cette collaboration anticipée. L'accès facilité à l'information scientifique est le fruit de ces recherches. Pour les conservateurs comme pour les chercheurs, la base de données constitue un dispositif d'accès aux documents plus performant ; outre les possibilités de requêtes, dépendant du degré de complétude de l'informatisation, la base contient des informations qui ne sont pas directement déchiffrables sur la planche, facilitant les liens vers les publications associées. Elle permet d'atténuer en partie une prépondérance des normes documentaires, fondées sur la systématique, dans les modalités d'accès aux spécimens, rangés par espèces dans le bâtiment des collections ; elle n'efface cependant pas entièrement cette prépondérance, étant donné que lesdites normes régissent une partie des champs de description des spécimens.

Ce type de collaboration entre recherche et conservation suppose des outils adéquats. Il s'agit de faire le lien entre les outils d'écriture utilisés par les conservateurs et ceux utilisés par les chercheurs. Alors que la phase de recherche et d'interprétation des documents d'herbier aboutit dans la conservation à la saisie de données standardisées, et parfois à un travail de valorisation d'un corpus historique visant à préserver l'organisation des connaissances sur laquelle il est construit, les activités scientifiques se caractérisent par l'élaboration individuelle de documents facilitant la structuration des prélèvements de données (tableaux), rassemblant les observations qualitatives (caractères) et quantitatives (mesures). Certains des éléments de ces tableaux ont vocation à être cités dans une publication, et constituent à ce titre des données de la recherche associées aux documents des collections scientifiques. Certains prélèvements d'herbier sont d'ailleurs archivés dans des bases de

données génomiques — la base Genbank a été particulièrement mentionnée dans les entretiens.

Les différentes démarches rencontrées ne font pas une place égale à l'exploitation de l'image et des bases de données. Cette diversité est un obstacle pour qualifier le statut documentaire de l'image *a priori*, à partir de l'étude de sa « documentalité », ou manière de faire document, envisagée par les seules caractéristiques du dispositif (fragmentation, référentialité, autonomie, rédaction et contribution) ; nous avons donc été conduits à questionner la valeur documentaire du spécimen numérisé à partir de ses « conditions de recevabilité », « en tant que document », en suivant l'approche proposée par Marie Després-Lonnet (2014) dans son étude de la constitution de la valeur documentaire des photographies de fouille en archéologie, qui deviennent document « lorsqu'elles sont mises au service d'un projet au sein duquel elles jouent ce rôle ».

La constitution de bases de données a permis de constituer l'image en document jouant un rôle de conservation et de diffusion de l'information scientifique, en reconstituant des fonctions d'accès plus efficaces, mais aussi de support d'observation des étiquettes, et dans certains cas des caractères morphologiques et de prélèvement de données. Quand bien même l'image est utilisée comme simple outil d'accès au spécimen papier, elle est le support d'une première lecture critique en vue de sélectionner des spécimens, puisqu'elle induit de nouvelles pratiques de tri d'images, en complémentarité avec la consultation des collections papier et les échanges avec les conservateurs. Le spécimen numérisé est aussi un intermédiaire facilitant le lien entre le spécimen et les étapes du travail de prélèvement de données et de rédaction. Il peut être utilisé comme un élément enrichissant la notice du catalogue, et donc comme un document secondaire. Mais les opérations qui le mobilisent pour des prélèvements qui seront cités des publications, ou pour attester de l'identification d'un nouveau spécimen qui rejoindra les collections, lui donnent un statut de document primaire, sans qu'il soit question qu'il soit substitué aux originaux des collections. En outre, si elle constitue bien une référence documentaire, l'image du spécimen ne reconstitue pas dans e-ReColNat la fonction de support d'annotation apposée par le scientifique lui-même et donc de recatégorisation du spécimen à mesure des résultats des nouvelles études, cette opération restant réalisée à partir d'échanges avec les conservateurs. La base de données n'est un dispositif contributif que pour les gestionnaires qui possèdent les droits de saisie, et qui peuvent enrichir leurs notices des informations trouvées sur d'autres bases, dans une logique de travail en réseau encore morcellé et émergent.

Les fonctions et statuts accordés à l'image selon les étapes de la démarche scientifique, les acteurs et les activités supposent des réponses appropriées en termes de fonctionnalités logicielles. L'objectif de conception est de contribuer à faire évoluer l'image vers un document de travail intégré à l'ensemble des étapes de la communication scientifique, entendue dans un sens large (Beaudry, 2011) incluant l'élaboration des connaissances. L'observation des principaux usages et gestes observés chez le maximum d'acteurs dans les opérations de conservation, d'inventaire historique ou écologique, et d'exploitation systématique des collections a permis de dresser une analyse des fonctions auxquelles l'herbier numérique devrait idéalement pouvoir répondre, que nous abordons au chapitre 8.

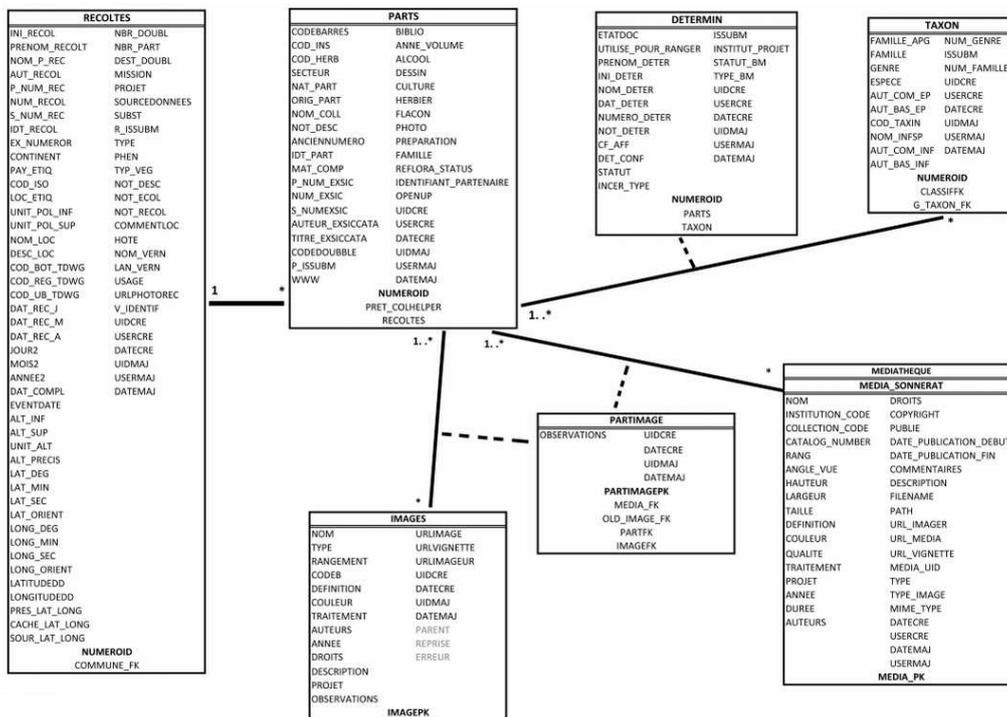
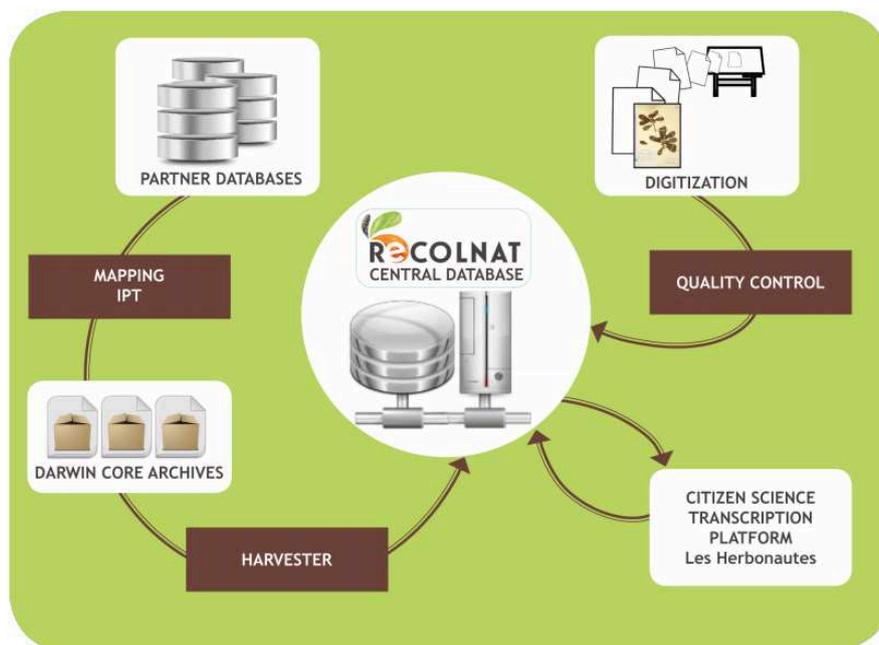


FIGURE 7.1. – Diagramme UML de la base de données Sonnerat au MNHN. Document communiqué suite à un entretien avec un gestionnaire de données du MNHN et de l'infrastructure e-ReColNat.



a. Schéma du flux des données entrant – infrastructure e-ReColNat 2014

FIGURE 7.2. – Cartographie des flux de données alimentant l'infrastructure e-ReColNat. Extrait du rapport d'activité de 2014 du WP2. Auteur : Julien Husson.

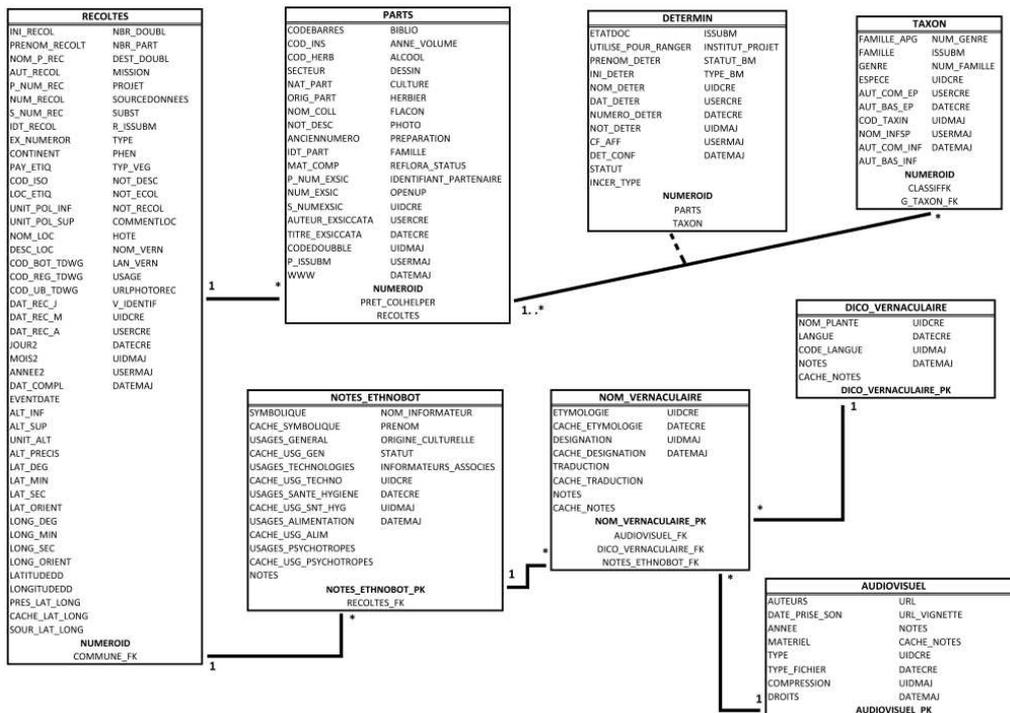


FIGURE 7.3. – Diagramme UML de la base de données d’ethnobotanique au MNHN. Document communiqué suite à un entretien avec un gestionnaire de données du MNHN et de l’infrastructure e-ReColNat.

8. Le laboratoire en ligne comme dispositif de médiation documentaire augmentée : enjeux pour la valeur documentaire du spécimen numérisé

8.1. Des spécificités documentaires de l'herbier aux spécifications d'un herbier numérique

L'observation des formes de communication médiatisées par l'herbier a permis de rendre compte des modalités de la mobilisation des spécimens numérisés aux différentes étapes de la communication scientifique – conservation, mise à disposition à l'utilisateur, exploitation du document scientifique et, plus marginalement dans nos observations, ouverture des collections à des publics non spécialistes. Les principaux gestes observés dans ces différentes activités, d'inventaire historique ou écologique et de systématique, ont permis de dresser une analyse des fonctions auxquelles l'herbier numérique devrait idéalement pouvoir répondre. Répertorier les gestes revient aussi à déterminer des équivalents pour les différents outils utilisés pour travailler avec l'herbier : le compactus, permettant la recherche des spécimens pertinents ; les tables de travail et d'observation au microscope situées à proximité des compactus ; les chariots et portulaires permettant d'apporter à ces tables les documents sélectionnés pour l'étude, parfois avec l'aide du personnel de l'herbier ; les outils d'observation disponibles sur ces tables (principalement microscopie et poste informatique relié à Internet) ; et les outils d'écriture personnels du chercheur, en premier lieu son ordinateur portable, le tableau dans lequel il consigne souvent les caractères prélevés, et les ouvrages empruntés à la bibliothèque.

Cette analyse constitue un premier modèle de ce que peut être un dispositif numérique susceptible de favoriser la reconstitution des fonctions de l'herbier. Ce modèle traduit notre hypothèse, fondée sur les observations de l'acquisition d'une valeur documentaire du spécimen numérisé, selon laquelle il est possible de reconstituer par l'image les fonctions des collections dans la démarche scientifique. Mais il est déjà aussi la première étape d'un travail de conception qui contribue à réaliser cette hypothèse. Le modèle et les prototypes qui en découleront deviendront ensuite un moyen de questionner de nouveau la même hypothèse, cette fois en observant l'appropriation dont feront l'objet les logiciels récemment développés pour l'infrastructure.

À partir de l'analyse des activités observées en transactions, il s'agit de procéder à une recomposition des opérations d'interprétation et d'écriture, pour déterminer des fonctions applicatives nécessaires à les rendre possibles. En classant

les différentes possibilités de transformation des supports, d'échelle d'examen des documents, de relation établie entre eux, et de documentarisation des annotations ou des productions documentaires impliquées par la transaction, nous avons défini différentes fonctions auxquelles le logiciel doit répondre, que nous détaillons ci-dessous.

RECHERCHER. La base de données est un outil de plus en plus utilisé pour l'accès aux spécimens par une grande diversité d'utilisateurs. Leurs différents objectifs et spécialités supposent une multiplicité de critères de recherches (par exemple par caractères visibles : je sélectionne les spécimens dont les racines, les fleurs, les bourgeons, etc., sont visibles). Cela peut passer par des possibilités de *tags* apposés aux spécimens par l'utilisateur, et intégrés aux filtres de recherche. À partir de la plateforme de transcription participative augmentée d'outils de classement d'images, l'intégration de catégorisations réalisées par des amateurs peut être envisagée comme une solution pour enrichir la description des spécimens.

L'efficacité des recherches peut aussi être accrue par l'introduction de référentiels taxonomiques, permettant d'étendre les résultats de recherche au-delà des taxons entrés dans le moteur de recherche à l'ensemble des taxons synonymes (c'est-à-dire employés à d'autres périodes pour décrire la même espèce correspondante). Le choix des référentiels à privilégier pose toutefois le problème d'un accord au sein de l'infrastructure, en l'absence de référentiel unifié et valide internationalement, qui permettrait de décrire l'ensemble des espèces représentées dans les collections, qui incluent aussi des espèces animales.

La recherche dans les spécimens doit pouvoir porter également sur les annotations faites sur les spécimens par différents contributeurs ayant accepté la publication de leurs données (recherche par *tags*, ou dans le texte des commentaires associés aux spécimens). Aux fonctions d'annotation et de catégorisation d'images décrites ci-après doivent donc correspondre des fonctions de recherche dans les annotations, ou de navigation dans les catégorisations créées.

Étant donné que la recherche conduit soit à la consultation et à l'exploitation des images en ligne, soit à des demandes d'accès ou de prêt, les fonctionnalités de recherche doivent pouvoir s'articuler non seulement avec les logiciels d'exploitation de l'image, mais aussi avec ceux de gestion des prêts et de contact avec le personnel de gestion, en plus des indications de localisation dans les collections des institutions.

OBSERVER. La visualisation des spécimens doit s'intégrer à différentes phases de l'activité : prévisualisation lors de la consultation des résultats d'une requête ; examen d'un spécimen ; examen de plusieurs spécimens comparés. Le spécimen doit pouvoir être examiné au niveau d'un fragment d'image, d'une image entière, ou d'un ensemble d'images comparées. Des éléments de l'image peuvent demander à être retournés (en cas d'étiquette à l'envers).

ANNOTER. Utiliser l'image comme support d'observation suppose de compléter la visualisation par des possibilités d'annotation de différentes natures, permettant d'identifier des caractères remarquables ou de prélever des mesures sur un fragment de l'image, mais aussi de constituer un regroupement de différentes images.

Une annotation doit pouvoir porter sur un point, ou une zone de l'image. Il peut s'agir de désigner l'un des échantillons présents sur la planche dans les cas où elle en contient plusieurs, un système de coordonnées étant notamment nécessaire pour localiser les observations sur les images issues de préparations microscopiques contenant un grand nombre de spécimens distincts.

LIER. Il doit être possible d'associer à un spécimen des documents complémentaires, en particulier des photographies au microscope ou des résultats d'analyses réalisées sur des fragments du spécimen papier. Il doit être possible de signaler par des liens hypertextes des ressources méritant d'être consultées en même temps que le spécimen.

CLASSER ET METTRE EN SÉRIE. L'activité de systématique, visant à établir différentes espèces, demande en particulier de prévoir des regroupements progressifs d'images, à partir de caractères communs relevés par des annotations. L'opération de classement peut être conçue comme l'apposition d'une catégorie préalablement définie à une planche. Il doit donc être possible d'associer à un spécimen un commentaire, une étiquette virtuelle indiquant une nouvelle détermination, un tag en vocabulaire libre ou contrôlé.

Des relations doivent pouvoir être définies entre différentes catégories de spécimens, organisées en une arborescence (ce qui correspond à la structure d'une classification).

Des fragments d'images doivent pouvoir être découpés, annotés, regroupés ensemble (mis en série), et constitués en unité documentaire : une URL doit pouvoir être créée pour constituer une référence unique à cet ensemble constitué par l'utilisateur. Un exemple d'utilisation serait la création d'une collection d'étiquettes significatives pour un récolteur.

EXPORTER. Étant donné l'utilisation d'une multiplicité d'outils de statistiques, de visualisation de données, ou de traitement d'images dans le travail scientifique, un export des annotations doit être prévu.

ENREGISTRER, PARTAGER ET PUBLIER. Il faut pouvoir mémoriser les annotations et catégorisations réalisées sur un ensemble de spécimens dans une optique de travail personnel, mais aussi les partager à un groupe de travail, et pour certaines d'entre elles les rendre publiques (avec une URL pérenne).

Il doit être possible de consulter toutes les annotations rendues publiques sur un spécimen.

EXPOSER ET NAVIGUER. Conformément à l'utilisation du document d'herbier dans des expositions, une catégorisation produite dans le cadre du laboratoire peut servir pour fonder un parcours de navigation dans une sélection d'images de la collection – par exemple à des fins d'exposition virtuelle. Il s'agirait de pouvoir présenter les spécimens selon une taxonomie historique ou un classement différent de celui de la systématique.

CONTRIBUER. Nous entendons ici par contribution l'ajout d'annotations standardisées susceptibles d'être intégrées à la base de données institutionnelle e-ReColNat, et qui suivent donc son format, à la différence des libres annotations des spécimens dans l'espace de travail. Ces fonctions d'intégration de données, susceptibles de faire entrer l'herbier numérique dans le régime contributif du dispositif de médiation documentaire (Zacklad, 2015), supposent de définir des droits de saisie pour les usagers, et non seulement les gestionnaires. Cette forme d'accréditation doit reconstituer l'autorisation du conservateur de travailler sur des spécimens, sur place ou par prêt, et d'y apposer de nouvelles annotations.

CONTACTER. Le format séché des spécimens les destine à être envoyés à des correspondants et échangés. Les comptes usagers pourraient servir à faciliter les contacts avec les conservateurs, les demandes de prêts, le rappel des prêts en cours, et le signalement des publications mentionnant les spécimens consultés ou prêtés. Les profils remplis par les usagers pourraient servir à ouvrir des droits de contribution. Ils pourraient être exploités dans des fonctions de contact, qui pourraient recommander des usagers ayant travaillé sur les mêmes spécimens que soi.

Compte tenu de l'ensemble des activités observées, d'après les besoins collectés et l'observation des pratiques mobilisant des images, l'herbier numérique devrait pouvoir répondre à cet ensemble de fonctionnalités. Parmi celles-ci, ont été retenues comme prioritaires celles qui concernent l'observation de l'image accompagnée de prise de notes et de prélèvement. Les dimensions de publication, d'exposition, et d'intégration aux collections de données standardisées produites par l'utilisateur n'ont donné lieu à aucun développement. Nous revenons à présent sur les choix ayant présidé au dispositif effectivement développé, qui ont transformé le modèle d'herbier numérique produit par les applications de l'infrastructure e-ReColNat, avant d'envisager la manière dont ce modèle a pu s'insérer dans les pratiques de quelques usagers, et d'évoquer pour conclure ce chapitre la façon dont ces pratiques des usagers testeurs ont contribué à redéfinir certaines fonctionnalités.

8.2. Du projet de collaboratoire au laboratoire

8.2.1. Une politique scientifique privilégiant les fonctions d'aide à l'exploitation des spécimens en vue de la publication

Le conseil scientifique d'e-ReColNat a fait le choix, pour le laboratoire en ligne, de prioriser le développement d'un logiciel « mono-utilisateur » – c'est-à-dire non

collaboratif, car ne permettant pas le partage des données entre usagers –, centré sur l'objectif de faciliter les publications à partir des données de l'infrastructure, qui soit opérationnel et présentable lors de démonstrations à l'échelle internationale, pour accroître sa visibilité. Ce choix, lié à la politique scientifique d'e-ReColNat, rejoint aussi une contrainte d'efficacité de développement, ainsi que les besoins prioritaires pour les principaux chercheurs impliqués dans la coordination de l'infrastructure. Un consensus ressortait en effet sur des fonctions d'exploitation de l'image essentielles dans un ensemble d'opérations communes à différentes perspectives de travail, réalisées le plus souvent en solitaire, et également fondamentales dans des situations de travail à plusieurs. Notre enquête en amont soulignait aussi l'importance de ces gestes, mais elle avait additionnellement recensé des fonctions essentielles à l'articulation des activités de conservateur et de chercheur, donnant plus de poids à la dimension d'intégration de données aux bases institutionnelles. Le modèle issu de l'enquête proposait aussi des fonctions de catégorisation utilisables dans une perspective d'exposition virtuelle, conformément à la mission de médiation du patrimoine pour des publics non spécialistes. La priorisation qui est devenue consensuelle s'appuie sur le principe selon lequel l'exploitation de l'image scientifique conditionne l'intérêt d'un archivage des données produites par les usagers, qui n'a d'intérêt que si ces derniers ont atteint un nombre et une régularité de contribution suffisants.

8.2.2. Assistance à la constitution d'un corpus

Un premier ensemble de fonctions concerne la constitution du corpus de travail (« set »), qui pourra être annoté et spatialisé. Une sélection de résultats d'une requête formulée dans le moteur de recherche développé par le WP2 peut être placée dans un panier et exportée vers le laboratoire (figure 8.1). Des images qui n'appartiennent pas à la base e-ReColNat peuvent être importées grâce à une procédure plus complexe, reposant sur la création d'un fichier contenant la liste des URL des images à importer. Une fonction de pré-tri a été développée à la demande des usagers, fonction qu'ils souhaitent encore voir enrichie de critères de tri plus nombreux, voire individualisables, pour trier les spécimens d'un même set selon des critères choisis. Ces fonctions de sélection, import et tri permettent donc de constituer un ou plusieurs dossiers constituant un premier classement d'images utiles à la sélection, qui pourront être eux-mêmes divisés dans les phases ultérieures du travail.

8.2.3. Reconstituer la dimension spatiale de la catégorisation progressive

Le logiciel développé correspond à un projet d'instrumentation de l'étape d'observation et de prélèvement d'information sur le spécimen, se traduisant par des prises de notes et des annotations, essentielles à élaborer des hypothèses de catégorisation nouvelles. Elle correspond à une instrumentation de la lecture critique associant analyse et synthèse, au cœur de la production de connaissances mobilisant

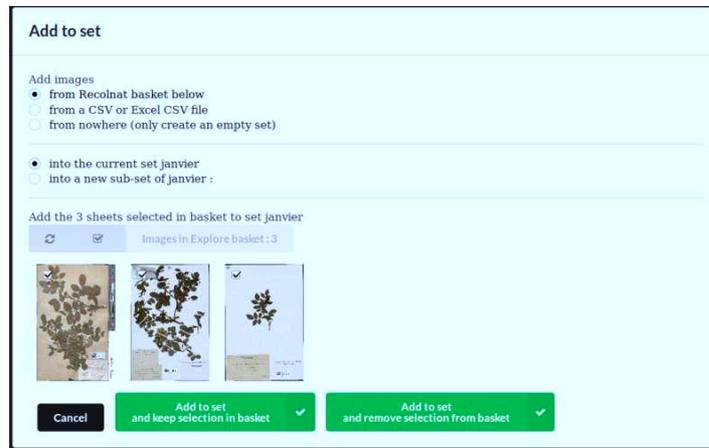


FIGURE 8.1. – Panier de résultats de recherche du catalogue *Explore* exportable dans le laboratoire en ligne (capture d'écran)

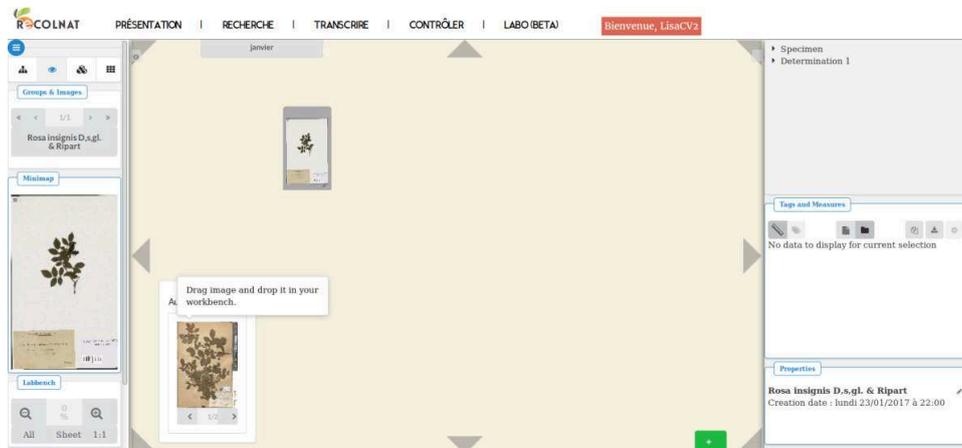


FIGURE 8.2. – Espace de travail du laboratoire en ligne : placement manuel des images sur l'espace de travail (capture d'écran). Dans toutes les captures, sauf indication contraire, le panneau latéral de gauche montre la vue entière du spécimen et celui de droite récapitule les mesures prélevées.

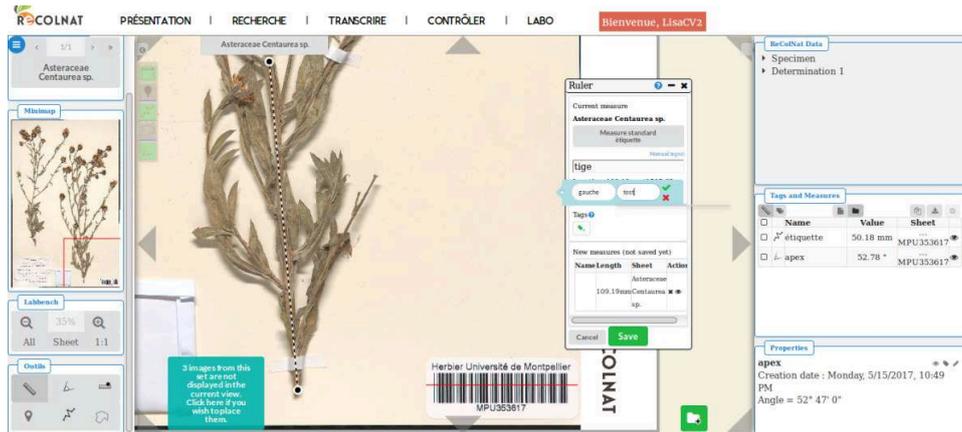


FIGURE 8.3. – Espace de travail du laboratoire en ligne : prélèvement d'une mesure taggée (capture d'écran). Le panneau latéral de gauche montre les outils disponibles.



FIGURE 8.4. – Espace de travail du laboratoire en ligne : mesure d'angle (capture d'écran)



FIGURE 8.5. – Espace de travail du laboratoire en ligne : calibrage de l'outil de mesure (capture d'écran)

des documents (Bottini, 2017). Dans les termes des types de régime de documentalité avec lesquels nous avons décrit l'herbier, le laboratoire en ligne se démarque de la base de données par les outils d'annotation non standardisée et non spatialisée qu'il propose, qui ne sont en revanche pas contributifs, puisque les données produites dans la session ne peuvent être partagées ou publiées. Décrire cette application demande de qualifier en particulier la dimension spatialement ancrée des annotations et fonctions de catégorisation qu'elle propose : le set correspondant à la sélection initiale d'images, qui sera enrichi des annotations faites sur l'ensemble d'images qu'il contient, peut être divisé en sous-dossiers (« sous-sets »); les images du set peuvent être librement déplacées pour produire des regroupements émergeant à mesure des observations (figure 8.2). Les tags apposés à une annotation (figure 8.3) ou à une planche permettent de conjointre l'annotation et la catégorisation spatiale, en apposant par exemple un même ensemble de tags à un groupe d'images, encadrées alors d'une couleur identique. La laboratoire en ligne instrumente ainsi le geste essentiel de la systématique, celui de la catégorisation progressive réalisée à mesure des observations et annotations, sur la base de caractères distinctifs justifiant le regroupement de spécimens. La spatialisation des annotations reproduit la possibilité de répartir sur le bureau en différentes piles les spécimens examinés, selon les caractères observés. Le zoom fluide reconstitue le geste de manipulation des différentes planches rassemblées sur un bureau de travail, examinées les unes après les autres, ou comparées : il permet de passer d'un détail d'une planche à la visualisation de l'ensemble des spécimens du set répartis en catégories dépaçables à mesure des observations. L'observation zoomée d'un spécimen est toujours associée à la présentation de la miniature de la planche dans le panneau de gauche (figure 8.3), pour en donner une vision globale quel que soit le niveau de zoom, ainsi qu'au jeu de données déjà prélevées dans le set (consignées dans le panneau de droite).

8.2.4. De la planche à la donnée, instrumentation du geste de prélèvement

À part les tags qui peuvent s'appliquer à une planche entière, les annotations réalisées dans le laboratoire ont la caractéristique d'être ancrées sur l'image, et se traduisent par un marquage du spécimen selon l'outil choisi (dans le panneau de gauche), qui correspond soit à un point ou une zone d'intérêt, soit à un segment ou à un angle mesuré (figure 8.4). Pour être valides, les mesures doivent être précédées d'une action de calibrage appliqué sur la planche sur laquelle est effectué le prélèvement (figure 8.5).

Ce type d'annotation, qui porte une valeur (celle de la mesure réalisée) est le moyen de prélever une donnée sur l'image. Les tags permettent de repérer dans l'espace un spécimen par une couleur, mais ils peuvent aussi porter sur chacune des annotations (mesure, point ou zone d'intérêt, tag du spécimen entier), chaque mesure pouvant se voir associer également un attribut et une valeur (par exemple « couleur : jaune »). Chaque annotation avec ses étiquettes et tags est rapportée sur le panneau latéral (de droite), dans un tableau pouvant faire l'objet d'une exportation. Les possibilités d'export des mesures réalisées permettent l'articulation du

travail d'observation de caractères et d'annotation dans l'espace de travail avec des travaux statistiques réalisés à partir des mesures prélevées avec d'autres logiciels. Le tableau permettant de visualiser les mesures au moment où elles sont réalisées est l'amorce de ce qui est était prévu initialement comme une vue tabulaire, permettant de passer du travail d'observation à une logique statistique ou à une vue synthétique. La présence d'un tableau de synthèse des prélèvements de mesures et d'observations, constatée dans de multiples entretiens, demandait de reconstituer une vue, même minimale, associant des perspectives sur les spécimens orientées documents et données. Nous montrons à présent comment les situations de démonstrations et de tests ont permis de questionner la pertinence pour le travail scientifique de l'association de ces différentes fonctions de regroupements spatialisés, d'annotations valuées et localisées, exportables et applicables à une sélection de spécimens de la base ou d'un corpus importé.

8.3. Retours sur le logiciel

8.3.1. Sessions de test et démonstration

Les premiers retours sur le logiciel ont été observés lors de situations de tests donnant à voir les réactions que suscitait sa prise en main. Ces ateliers de tests réunissaient des usagers correspondant à la diversité des métiers des collections, de la conservation, de la valorisation patrimoniale ou de la recherche, sans faire toutefois de place aux amateurs, avec qui les interactions sont envisagées dans le contexte de la plateforme participative qui sera présentée au chapitre 10. Pour chaque atelier, après un tour de table sur les attentes liées au projet e-ReColNat et aux nouveaux outils d'exploitation de l'image, qui était une occasion de confirmer et d'enrichir ponctuellement les résultats de l'enquête en amont de la conception, et suite à la démonstration par l'ingénieur de développement de l'utilisation de chaque fonction disponible, les usagers testeurs utilisaient l'outil en autonomie, avec l'aide ponctuelle des ingénieurs.

Les tours de table des ateliers ont fait ressortir que les fonctions d'analyse automatique d'image étaient un horizon de travail important, plus que ne l'avait souligné l'enquête. Certains chercheurs se projettent en outre dans une utilisation du spécimen numérisé facilitée par le recours à des programmes participatifs pour pré-catégoriser les images. Si nous avons bien recensé ce besoin, notamment souligné par un paléontologue du MNHN pour l'aide à l'exploitation des fossiles, nous avons moins approfondi ce champ de réflexion dans le travail de conception, puisqu'il ne correspondait pas directement au domaine d'expertise du laboratoire (mais plutôt à celui de l'un des partenaires de l'infrastructure e-ReColNat à Montpellier, le Cirad, à l'origine de l'application Plantnet d'aide à l'identification de photos). Sans entamer le consensus pour une première version mono-utilisateur du laboratoire, des tours de tables est aussi ressorti l'intérêt du travail collaboratif pour intégrer les étudiants dans des programmes de recherche portant sur des identifications.

L'atelier de test a permis d'observer la durée réelle nécessaire à la prise en main du logiciel — elle est longue : les questions des usagers concernant certaines fonctions ont émergé sur toute la durée de l'atelier. De fait, une bonne part des remarques a concerné l'ergonomie, les demandes allant dans le sens d'une homogénéisation des pratiques de navigation dans l'application avec les conventions habituelles d'expérience utilisateur (par exemple, possibilité d'utilisation de raccourcis dans la saisie du texte des annotations, curseur changeant de forme selon l'outil sélectionné, zones de sélection polygonales dont la clôture est effectuée automatiquement à partir des points tracés). Elles ont aussi fait ressortir une difficulté à visualiser simultanément des images d'échelles différentes : la comparaison de mesures après étalonnage reste en l'état la seule garantie contre les distorsions.

Les ateliers ont encore précisé la description de certains des gestes mobilisant l'image. Les annotations doivent ainsi pouvoir être numérotées comme une série cohérente. S'est aussi exprimée une reconnaissance de l'apport du logiciel pour introduire dans la méthodologie du chercheur un nouveau degré de précision dans le travail de relevé de données : lors d'une démonstration au Canada après un colloque de l'Acfas organisé par Florence Millerand et Lorna Heaton, une chercheuse en systématique nous disait son enthousiasme de pouvoir repérer le point où a été prélevé la mesure, ce qui se substituera à la pratique, contraire aux règles de conservation, de noter d'une minuscule marque au crayon sur la planche le niveau où une mesure a été réalisée. Plusieurs usagers se sont projetés dans des fonctions de publication et d'archivage des travaux réalisés dans la session personnelle. La session de travail comportant des mesures pourrait être associée à une publication, sous forme de données complémentaire en permettant la vérification, comme le consigne l'un des comptes-rendus d'atelier des ingénieurs : « une demande fonctionnelle importante concerne la possibilité via une URL d'accéder à la présentation comme preuve dans un papier par exemple d'une étude scientifique faite par un chercheur (plusieurs à l'avenir) dans le laboratoire. Ce serait donc une preuve interactive » (au sens où il serait possible de zoomer sur la session, vérifier la mesure d'un segment tracé sur l'image, etc.). Cela suppose un archivage pérenne, garanti par l'institution de l'herbier au même titre que les types déposés, constituant une extension des missions de l'herbier de rang international à des missions de conservation des données de la recherche.

La fenêtre regroupant les données associées à l'image dans la base institutionnelle a soulevé la question de la possibilité d'éditer directement ces données depuis l'espace de travail, dans le cas où des erreurs sont déterminées. Cette demande souligne la pertinence de la fonction de médiation documentaire caractéristique de l'herbier, qui intègre l'utilisateur à l'enrichissement des données documentaires. Étant donné que le logiciel développé par le WP5 est mono-utilisateur et centré sur des annotations non standardisées, c'est plutôt au niveau de l'interface d'édition et de saisie des données livrée par le WP2, et pour le moment réservée aux gestionnaires des collections du réseau, que serait envisagée à l'avenir une ouverture du « répertoire des participants » aux usagers contribuant à la production de données standardisées sur les spécimens visant à corriger ou enrichir les données de collecte ou les déterminations scientifiques.

8.3.2. Intégration du logiciel dans des pratiques scientifiques : retour d'un systématien et d'une écologue

Au-delà de la seule prise en main et des projections allant de pair observées dans les ateliers, deux chercheurs ont fait une utilisation plus approfondie du logiciel. Le chercheur en systématique, coordinateur du projet, a pu utiliser la fonction de mesure soulignée comme un besoin essentiel. De la même façon que lors des ateliers de tests, il a été constaté que des obstacles ergonomiques, en particulier le temps de chargement des sessions de travail composées d'un set de plusieurs images et d'annotations, ont empêché de prendre dès le départ la pleine mesure des possibilités d'exploitation de l'image. Le temps de chargement restait problématique même en exploitant la possibilité d'ouvrir l'espace de travail en chargeant les images par set de travail à la demande, avec une option d'affichage évitant de charger les mesures ; ce problème est accru par le choix posé au début du projet, en concertation avec les coordinateurs, d'une application web plutôt que d'une application native (« client lourd ») pouvant fonctionner indépendamment du serveur.

Malgré ces obstacles, le laboratoire en ligne a donné lieu à la réalisation de plusieurs centaines de mesures, réalisées pour certaines parallèlement à d'autres effectuées sur les spécimens physiques par un second chercheur, avec un tableau de synthèse commun, non intégré à l'espace de travail en ligne. La pratique des mesures a confirmé le besoin d'inclure des fonctions de calcul à la synthèse tabulaire des mesures (volet de droite), afin de fournir à chaque nouvelle annotation des indicateurs statistiques sur les valeurs prélevées (tels que moyenne ou écart-type), associées à un même tag, un même numéro d'inventaire, ou un même nom d'espèce (taxon). Il s'agit d'optimiser le nombre de prélèvements nécessaires à obtenir la moyenne de la taille du caractère observé. Cette demande a confirmé l'intérêt d'une vue orientée données sur le corpus de spécimens annotés et catégorisés constituant le set, sous la forme d'un tableau récapitulatif des prélèvements réalisés, selon les spécimens et les catégories leur ayant été associées, en vue de faciliter leur exploitation statistique. Un enjeu similaire de faciliter un passage fluide entre le contexte d'observation du spécimen et celui de l'analyse des données prélevées est ressorti du travail du second usager testeur.

L'utilisation du laboratoire a été envisagée dans le cadre de l'élaboration d'une méthodologie de thèse en écologie (« Adaptation de la morphologie foliaire à travers les collections d'herbier : applications aux assemblages fossiles », thèse de Mélanie Tanrattana sous la direction de Dario De Franceschi et d'Anaïs Boura). L'étude demandait d'effectuer des relevés de mesures de caractères foliaires sur un grand nombre de spécimens appartenant à quelques espèces prises pour référence, sur des fossiles de feuilles et des feuilles mises en herbier, afin de noter d'éventuelles évolutions associées à des changements climatiques. Une première orientation méthodologique de l'étude supposait un grand nombre de spécimens à mesurer, faisant préférer l'utilisation de logiciels d'analyse d'images automatisée (qui calculent les dimensions de surfaces et de segments repérés automatiquement sur un corpus d'images de feuilles). Une seconde orientation méthodologique, qui avait réduit le nombre d'espèces de référence et de spécimens à mesurer, a permis

d'envisager l'utilisation du laboratoire en ligne plus facilement, pour articuler aux mesures prélevées automatiquement les prélèvements qui ne pouvaient être réalisés par les logiciels d'analyse d'images, à savoir les mesures d'angles. Finalement, compte tenu de difficultés ergonomiques, c'est encore un autre logiciel proposant des fonctions équivalentes qui a été retenu pour les mesures d'angles, en association avec les logiciels d'analyse automatisée, — mais aussi avec le laboratoire en ligne, finalement utilisé essentiellement comme outil de sélection et de pré-tri des spécimens pertinents selon leur espèce. L'obstacle ergonomique majeur, outre les lenteurs de l'application, venait d'une difficulté à distinguer visuellement un spécimen dans l'espace de travail du laboratoire par son numéro d'inventaire, qui servait de référence dans les mesures réalisées avec les autres logiciels : l'encadré de chaque spécimen avait été construit à partir nom d'espèce, et la correction supposait une modification non triviale de la structure de données.

Cet obstacle a fait apparaître l'importance des fonctions de repérage des spécimens appelés à circuler entre différents logiciels et environnements de travail dans le cadre d'études hybridant les supports et les échelles de traitement (observation d'un spécimen ou traitement d'un jeu de données issu de nombreuses observations). Dans la méthodologie de cette thèse en écologie, le travail sur le spécimen numérisé articule différentes techniques de prélèvement et mesure d'images, manuelles ou automatisées, en complémentarité avec les mesures des fossiles non numérisés. Ce travail opère un passage d'une multitude de documents à la constitution d'un jeu de données traité ensuite statistiquement pour faire ressortir des différences significatives dans la taille des caractères mesurés. Le spécimen est mobilisé à différentes reprises au sein d'un flux de transactions utilisant chacune des instruments et logiciels spécifiques pour prélever des mesures. On peut décrire de la même manière le cas observé en systématique dans les opérations de mesures sur différents supports, mesures destinées à être analysées ensuite dans un tableau synthétique. Pour le dispositif de médiation documentaire, l'enjeu est d'intégrer à la fois la gestion du document et celle de la donnée qui lui est associée. En l'occurrence, la donnée se définit, à partir du contexte du travail sur l'herbier, comme la mesure ou l'annotation marquant la présence d'un caractère, issues d'une première phase d'observation, et restant contextualisées à l'unité documentaire qui permet de les rattacher au contexte de la collecte et à l'espèce concernée. L'enjeu de l'articulation du travail sur les documents et les données s'est vu confirmer par la volonté, apparue au cours du projet, d'utiliser d'autres logiciels en complémentarité avec le laboratoire en ligne, qui permettraient à la fois des fonctions de création de référentiels, laissant envisager un raffinement des fonctions de tri des documents, et une exploitation statistique des données qui leur sont prélevées.

Le logiciel Xper développé par le laboratoire du LIS, spécialisé dans le développement de la Systématique Assistée par Ordinateur, en étroite relation avec le MNHN, permet de créer des modèles de description de données et des vocabulaires contrôlés, qui peuvent être appliqués à des images de spécimens. À partir d'un ensemble de données décrites à partir d'un modèle créé, le logiciel propose des catégorisations (cluster), à partir de l'identification automatique des traits de caractères communs statistiquement significatifs, et fournit une assistance à la construction de clés d'identification. À partir d'un modèle de données et de clés construites

correspondant à la flore de la région concernée, il peut être utilisé sur le terrain comme une application d'aide à l'identification. La réflexion sur les convergences entre le laboratoire en ligne et ce logiciel a donné lieu à des propositions de scénarios qui s'intègrent à la fois dans une problématique de gestion documentaire et d'exploitation des données. En effet, les modèles de description produits sous Xper pourraient être intégrés dans le laboratoire, pour annoter des images selon le vocabulaire et la structure produite en amont. Corrélativement, une campagne d'annotation d'un lot d'images selon un modèle de description de données pourrait être suivie de traitements statistiques proposant des regroupements de spécimens selon des traits morphologiques significativement récurrents. La description de caractères observés dans certains spécimens, ou associés en amont à une espèce, permettrait de rechercher des spécimens dans le moteur de recherche de la base e-ReColNat par de nouveaux critères, correspondant à des caractères cette fois, conformément aux besoins de l'identification en ligne. Dans une perspective de valorisation patrimoniale, le logiciel pourrait être utilisé pour construire un parcours dans les collections selon un ordre déterminé correspondant à un modèle de description de données définie en amont appliqué à une sélection de spécimens à exposer. Si ces projets n'ont pas abouti dans la durée étudiée de l'infrastructure, ils confirment les enjeux de la circulation du spécimen associée à l'élaboration des connaissances dans différents contextes interprétatifs, de la collection régie par la taxonomie à la base de données personnelle ou partagée par un groupe de travail, régie par un modèle de description de données spécifique.

Les attentes complémentaires des usagers ont mis en évidence l'enjeu pour le laboratoire en ligne de s'intégrer au flux de données et documents exploités dans différents logiciels au cours de la démarche scientifique. Il s'agit de faciliter le passage de l'observation du document au traitement des données qui en ont été prélevées. Au niveau logiciel, cela suppose de prévoir en amont les conditions de l'interopérabilité entre le laboratoire en ligne, les autres logiciels de l'infrastructure et ceux susceptibles de s'y intégrer dans un avenir proche, ainsi que les logiciels utilisés par les usagers. Plus généralement, cela conduit à penser un prolongement de la médiation documentaire par une médiation de la donnée, posant la question des modalités d'archivage des annotations et prélèvements faits sur le document.

8.4. Le laboratoire en ligne comme dispositif numérique de médiation documentaire augmentée

Le laboratoire en ligne a permis de préciser les modalités que pouvaient prendre, dans le cas de l'herbier, un dispositif destiné à faciliter l'exploitation scientifique du document, en augmentant les bases de données de consultation des collections par des fonctionnalités d'assistance à la constitution d'un corpus, de mesure, d'annotation non standardisées, ainsi que d'export de toutes les données produites par ces différentes opérations. L'application intègre un flux de transactions essentielles à

l'activité de lecture manuelle de l'image. Elle rend possible une partie de l'exploitation de l'image, celle qui concerne l'observation et le prélèvement de mesures semi-automatisé, qui s'articulent sur d'autres étapes de la démarche scientifique passant par l'étude morphologiques, microscopique et moléculaire des spécimens non numérisés, et l'analyse automatisée des images. L'ergonomie et lenteur liée à la lourdeur des images restent les principaux obstacles à la fluidité qu'on pourrait attendre entre différents contextes d'étude. Le choix technologique de la réalisation d'une application web s'est avéré pour le moment moins utile que prévu, étant donné l'absence de fonctions collaboratives. Développer les fonctions collaboratives à l'avenir serait essentiel à ce qu'on puisse rapprocher le laboratoire en ligne des bureaux virtuels développés pour des « laboratoires », évoqués dans l'état de l'art (chapitre 4), centrés sur des de communication. Mais l'instrumentation d'opérations d'herméneutique numérique proposées dans le laboratoire en ligne s'inscrit dans les problématiques informationnelles impliquées dans les laboratoires, avec la question commune de l'archivage de long terme des données de la recherche, en l'occurrence dans notre cas des jeux de données produites dans les sessions individuelles de travail (Bos *et al.*, 2007, p.387).

Nous pouvons caractériser, à partir de l'ensemble des retours des usagers, le type de dispositif de médiation documentaire numérique augmentée que constitue le laboratoire en ligne, par rapport aux modèles que représentent les autres bases de données, et la formalisation des attentes des usagers rencontrés. Alors que toutes les fonctions nécessaires à reconstituer ou prolonger celles de l'herbier n'ont pu être développées dans le temps du projet, les retours des usagers ont validé la pertinence de la plupart d'entre elles, en soulignant en particulier l'importance de l'archivage des données produites dans le laboratoire en ligne, et l'attente confirmée de l'ouverture de la session de travail avec d'autres contributeurs, à plus long terme. En ce qui concerne les fonctions de valorisation de l'herbier dans d'autres perspectives que la systématique et l'écologie, elles n'ont pas été retenues prioritaires, compte tenu de la politique scientifique de l'infrastructure, sans être invalidées pour autant – et ce d'autant que la plateforme de transcription participative abordée au chapitre 10 répond à cette perspective, sous un angle qui diffère des expositions virtuelles prévues en cohérence avec les pratiques observées chez les conservateurs.

Malgré les difficultés d'utilisation qui en ont limité la portée et empêchent de tirer des conclusions sur les transformations des pratiques du laboratoire en ligne, en production dans sa version actuelle depuis avril 2017, la conception du laboratoire en ligne, au travers des retours et des demandes complémentaires des usagers, a questionné la place que pouvait avoir l'image dans les pratiques scientifiques, quand étaient mis à disposition des outils facilitant leur exploitation au-delà de la consultation et de l'export de données proposés par les bases de données classiques. Le laboratoire en ligne assoit le statut documentaire de l'image, comme « forme d'appropriation » du spécimen numérisé (Bachimont, 2004) qui facilite le prélèvement de données essentielles à l'élaboration de connaissances, qui seront traitées ou directement citées dans les diagnoses (descriptions d'espèces) des publications. Cette approche conduit à concevoir la valeur documentaire non seulement par le prisme des usages et situations qui la constituent, mais également des

dispositifs qui la rendent possible, en cohérence avec la conception de la médiation instrumentale de l'interprétation du document, notamment présente dans les travaux de Bruno Bachimont sur l'ingénierie des connaissances que nous avons rappelés au chapitre 3. Le statut du spécimen numérisé dépend du système de médiation documentaire qui permet son archivage pérenne, et son évolution passe par la capacité des herbiers numérisés à archiver également des données ou à faire le lien vers des dépôts spécifiques. Si l'archivage des données produites dans le laboratoire n'a pas trouvé de solution dans le projet, en l'absence du développement d'une fonction minimale de publication d'un jeu de données correspondant à une URL pérenne garantie par l'infrastructure, l'intégration de données standardisées au titre de correction de la description des spécimens numérisés paraît réalisable plus facilement à courte échelle, compte tenu des différentes solutions déjà en place dans plusieurs institutions. Nous revenons à présent sur les réflexions que cette fonction a suscitées dans l'infrastructure.

9. Dispositifs numériques de médiation documentaire ouverts aux corrections des usagers

9.1. Enjeux de l'intégration de corrections des usagers pour e-ReColNat

Favoriser la participation des usagers à la correction et à l'actualisation des données des collections est un enjeu pour l'infrastructure, même s'il n'a pas pu faire l'objet de développement logiciel à ce jour : il s'agit d'intégrer l'identification progressive des spécimens types, et plus généralement de favoriser la dynamique de travail en réseau pour enrichir la qualité des données de toutes les collections de l'infrastructure. Nous avons recueilli des demandes allant en ce sens de la part de plusieurs conservateurs (notamment à Strasbourg et Lyon), qui étaient amenés à transcrire des étiquettes des collections du MNHN dans le cadre des activités concernant leurs propres collections (pour les récolteurs détenus en commun), sans pouvoir faire facilement bénéficier le MNHN des données standardisées issues de leurs efforts. Ce besoin a conduit à se pencher sur différentes bases de données ouvertes à la contribution des usagers dans le cas des herbiers, au titre de la réalisation d'un état de l'art susceptible d'éclairer les choix futurs de l'infrastructure ; ces solutions développées par d'autres institutions et projets de recherche définissent différents modèles de dispositifs de médiation documentaire ouverts à la contribution (Vidal, 2010), qui se distingue d'autres formes de participation par la perspective commune d'intégration des données aux bases institutionnelles – même si dans les faits, cet objectif n'est pas toujours atteint. Comme c'est le cas des outils d'annotation spatialisée du laboratoire, de tels dispositifs peuvent être utiles pour penser une médiation documentaire augmentée pour d'autres corpus, notamment dans la perspective d'ouverture à la participation des archives (Moirez, 2012).

Le coordinateur de l'infrastructure juge important de permettre des corrections en ligne, pour profiter des mises à jour résultant de la fréquentation du site par des spécialistes des collections. Mais la question des modifications à autoriser sans passer par la consultation du spécimen physique ne fait pas l'objet de réponses concordantes chez les chercheurs et les conservateurs rencontrés. Il n'y a pas encore de politique commune aux institutions, ni de point de vue unanime sur la valeur documentaire du spécimen numérisé par rapport à la planche d'herbier papier ; par conséquent, il n'y en a pas non plus sur les pratiques de conservation à adopter concernant l'intégration d'enrichissements de la base de données par des usagers sans qu'ils aient consulté les spécimens originaux. Certains chercheurs

pensent qu'il faut accepter une discordance entre la planche d'herbier et sa notice dans la base, pour utiliser au mieux les moyens limités consacrés au travail de conservation des collections, et éviter de devoir réintégrer de nouvelles précisions sur les planches. Cette solution suppose que la consultation de la base de données pour accéder aux spécimens soit systématisée, alors même que cela ne correspond pas à la pratique de consultation des collections : nous avons mentionné au chapitre 7 les chercheurs qui vont directement aux compactus, au moins occasionnellement. La planche d'herbier perdrait aussi son indépendance par rapport à la base de données, qui contiendrait l'ensemble des déterminations. C'est contre cette discordance du document papier et de la base de données que s'élèvent d'autres chercheurs, attachés à une complétude de la planche d'herbier — pourtant déjà associée à des publications ou des carnets de récolte. Selon eux, la planche d'herbier a jusqu'à présent tiré sa valeur de la présence de l'ensemble des déterminations réalisées à son égard, consignées dans les étiquettes. Le document leur paraîtrait incomplet s'ils se voyaient obligés de passer par le catalogue pour chercher les déterminations faites à propos de la planche. Cette différence de point de vue recoupe en partie des différences d'habitudes d'accès aux collections, entre les usagers distants qui passent systématiquement par la base, et profiteraient des annotations virtuelles, et ceux rattachés à l'institution, qui peuvent aller directement travailler dans l'herbier, et craignent de manquer ainsi une partie des informations sur le spécimen.

À cette première divergence, s'ajoute celle de la valeur informative accordée au spécimen numérisé, et la question de savoir s'il peut fonder une redétermination, nouvelle identification intégrée à la base de données, susceptible d'entraîner un déplacement du spécimen vers une autre liasse, au même titre que le permet la consultation du spécimen papier — sachant que nous avons observé que des déterminations de nouveaux spécimens intégrés à l'herbier sont déjà actuellement réalisées à partir de la seule consultation d'images, sans que cela ne soit signalé d'une quelconque manière dans la base de données ou sur l'étiquette. Face à cette absence de consensus, une recension des solutions d'intégration de données produites par les usagers à partir de la seule consultation des spécimens numérisés a été proposée en préalable à la conception d'un système de correction des données par les usagers pour l'infrastructure.

9.2. Solutions d'intégration de contributions par les usagers des bases de données de spécimens numérisés

Au-delà de sa fonction dans le projet, l'enjeu de cet état de l'art est d'étudier les formes d'ouverture de la médiation documentaire numérique à la participation, dans le respect de l'autorité de l'institution, et de prendre la mesure de ses enjeux pour les formes de communication entre acteurs de la gestion et de l'exploitation scientifique des collections, qui s'ajoutent à leurs interactions dans l'espace des

collections. La notion de transaction permet bien de décrire l'action de commentaire ou de correction de données réalisée par l'utilisateur, qui interprète un spécimen numérisé avant de proposer une nouvelle valeur pour l'un des champs de la notice du spécimen ; à cette première transaction répond la validation de la contribution – sauf dans les cas où le système prévoit une alternative à ce retour du conservateur vers l'utilisateur. En reprenant le cadre d'analyse utilisé pour décrire l'herbier, nous distinguons les différents dispositifs selon les formes de contributions qu'ils permettent (structurées, limitées à des champs de formulaires ou libres), selon les « répertoires de participants » auxquels ils sont ouverts (nécessité ou non d'une inscription pour contribuer), et enfin selon les interactions qu'elles suscitent entre conservateurs et gestionnaires, qui dépendent des modalités d'intégration des contributions au dispositif documentaire (simple ajout en commentaire d'une notice ou donnée de la base, pouvant être retrouvée à partir du moteur de recherche des collections en ligne). La consultation des différentes collections répertoriées à l'échelle mondiale dans l'Index Herbariorum¹ permet d'appréhender la diversité des portails donnant accès aux bases de données des collections, plus ou moins propices à l'intégration de contributions produites par les usagers. Nous prenons aussi en compte les contributions lorsqu'il a été possible d'analyser les données produites par les contributeurs, et les retours auxquels certains de ces sites ont donné lieu chez les usagers rencontrés dans notre enquête.

9.3. Contributions non structurées validées par les conservateurs

La première solution pour intégrer des actualisations et des corrections venant des usagers passe par un contact avec l'institution, via le courrier électronique ou un formulaire en ligne, pour suggérer des modifications ; cette solution est disponible pour toutes les collections mises en ligne. L'incitation à faire remonter les données peut passer par un bouton intégré aux notices consultées dans la base de données. On trouve par exemple ces fonctions de contact/modification dans le portail Arctos² (figure 9.1), qui regroupe les données de 84 collections naturalistes américaines, ou dans « l'Herbier Virtuel A. de Saint-Hilaire » (Romaniuc *et al.*, 2013). Une solution de contributions plus structurées, proposée sur le site de l'herbier de Californie³, vise explicitement à accélérer le travail d'intégration des données par les conservateurs en proposant une catégorisation d'annotation, un champ de saisie de coordonnées géolocalisées et un champ de commentaires pour justifier la modification (figure 9.2).

Le formulaire de l'herbier de Californie s'affiche lors de la consultation d'un spécimen après une requête dans la base de données. La modification effective des données reste soumise à modération par le personnel de l'herbier. Les intitulés donnés aux commentaires montrent qu'il s'agit, pour les deux tiers des 3000 commentaires qui avaient été laissés sur le site en 2015, de corriger une localité géographique.

1. <http://sweetgum.nybg.org/science/ih/>

2. <http://arctos.database.museum/SpecimenSearch.cfm>

3. *Consortium of California Herbaria*, <http://ucjeps.berkeley.edu/consortium/about.htm>

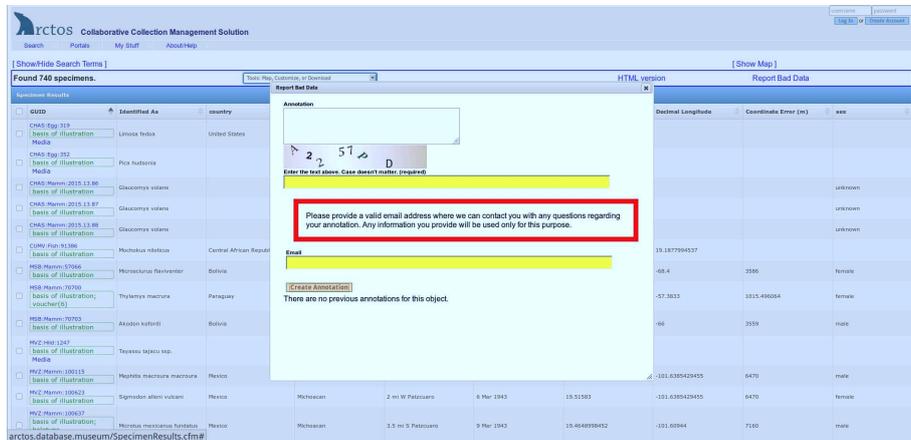


FIGURE 9.1. – Formulaire d’annotation sur le portail Arctos (capture d’écran)

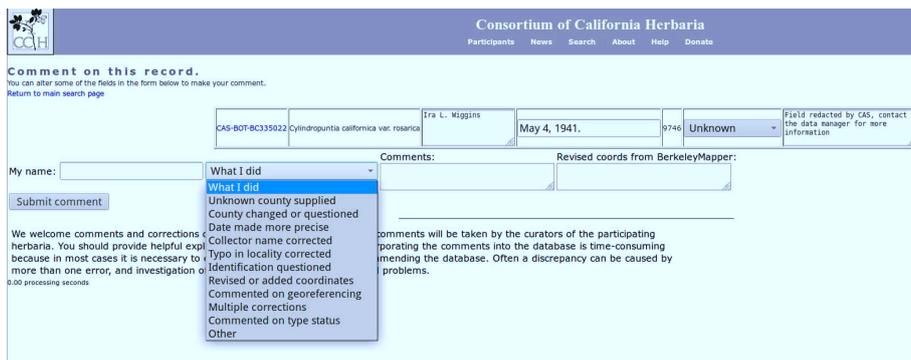


FIGURE 9.2. – Formulaire d’annotation typée et semi-structurée sur le site du Consortium of California Herbaria (capture d’écran)



[View Larger](#)

Open Viewer

PDF

Collection

Partner Page

Isotype of *Pithecellobium coripatense* Rusby [family FABACEAE]

Herbarium	Museo Botánico, Facultad de Ciencias Exactas, Físicas y Naturales, Universidad Nacional de Córdoba (CORD), CORD00002936
Collection	Herbarium Specimens
Resource Type	Specimens
Collector	Bang, M., #2176
Collection Date	1894/05/06
Locality	Coripata.
Country	Bolivia (Bolivia)
Identifications	Isotype of <i>Pithecellobium coripatense</i> Rusby [family FABACEAE] (stored under name);
Notes	Cfr.: Bull. New York Bot. Gard. 4(14):349. 1907.
Pages	1
Data last modified	2011-01-28

View Edit History

What the Plant Community is Saying

1 Comment JSTOR Global Plants [Login](#)

Sort by Best [Share](#) [Favorite](#)

Join the discussion...

Sandy Knapp · 2 months ago

 *Solanum sericeum* var. *strigillosum* is now a synonym of *Solanum salicifolium* Phil.- see <http://phytokeys.pensoft.net/a...>

[^](#) | [v](#) · [Reply](#) · [Share](#)

ALSO ON JSTOR GLOBAL PLANTS

Holotype of *Calcitrapa patibilcensis* Kunth
[family ASTERACEAE]

1 comment · 18 days ago

[disqus_VdDiuZtK1o](#) — The collection number is 3768.

WHAT'S THIS?

***Campuloclinium macrocephalum* (Less) DC.**
[family ASTERACEAE]

1 comment · 18 days ago

[disqus_VdDiuZtK1o](#) — I have studied other duplicates of Gardner 66 and would suggest that this barcode is more applicable to the ...

FIGURE 9.3. – Exemple de commentaire laissé sur l’interface de consultation des spécimens de *Global Plants* (capture d’écran)

De cette solution, qui suppose un travail d'intégration de la part des conservateurs, explicitement soulignée comme coûteuse en temps sur le formulaire de l'herbier de Californie, se démarque la contribution par commentaire des notices sur le site *Global Plants*. Le site utilise l'agrégateur de commentaires Disqus pour afficher les commentaires associés à chaque notice. Il affiche l'identité des commentateurs et met en valeur leur activité récente. Il n'est pas prévu de visualiser facilement l'ensemble de ces commentaires, ou de faire des recherches dans le texte, si ce n'est par la consultation du profil des commentateurs sur Disqus. Les commentaires sont, pour 85% d'entre eux, composés d'un message unique associé à un spécimen (voir figure 9.3), sachant que cette proportion est encore sous-estimée du fait des modalités du traitement des commentaires auquel nous avons procédé – les commentaires associés aux spécimens enregistrés sous un même nom sont regroupés. Les réponses données à un commentaire sont majoritairement celles du personnel confirmant la modification de la base, puisque l'internaute ne peut faire lui-même la mise à jour. À part des discussions extrêmement ponctuelles entre usagers et membres du programme *Global Plants* – nous avons notamment relevé une discussion à propos des désagréments générés par l'accès payant aux images de grande taille – la plupart des commentaires mentionnent une précision ou une correction à réaliser, à propos du contexte de la collecte ou du statut de type d'un spécimen. Ce constat montre que les commentaires rejouent la logique d'agrégation de données issue d'un travail individuel sans engager une démarche coopérative – à la différence de la dynamique générée par la plateforme ouverte aux amateurs que nous étudierons au prochain suivant. Si le commentaire n'est pas modéré sur le site « *Global Plants* », les contributions n'affectent pas pour autant l'autorité du personnel de l'herbier, qui autorise de fait les modifications en les reportant dans sa base ; et le changement est d'autant moins important qu'il arrive déjà, selon le témoignage d'un chargé de collections du MNHN, que les tableaux de données issus de l'examen des spécimens par les visiteurs pour leurs besoins soient intégrées à celles du Muséum. La solution de la séparation de la base institutionnelle et de l'espace de commentaire est le moyen pour l'institution de conserver son rôle de validation des données.

Le nombre de commentaires laissés sur *Global Plants* atteste des potentialités de l'intégration des contributions des usagers pour améliorer la qualité de la documentation des collections. Mais le site montre aussi la difficulté d'engager cette dynamique, puisque l'utilisation des commentaires est majoritairement le fait de quelques professionnels des collections particulièrement investis. En témoigne la répartition du nombre de commentaires par contributeurs (figure 9.4). Les trois quarts des interventions sont le fait d'un seul commentateur, à l'époque membre du personnel d'un herbier finlandais. Les commentateurs les plus investis appartiennent à des institutions détentrices de collections.

9.4. Déterminations virtuelles et annotations sur une base de données enrichie

D'autres sites ont élaboré une solution déchargeant en partie le personnel des collections du travail de validation et de saisie manuelle des corrections apportées

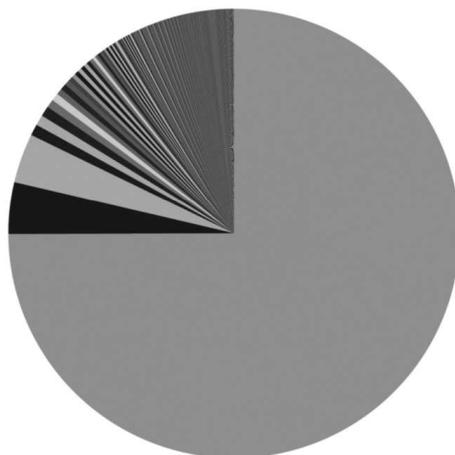


FIGURE 9.4. – Répartition par commentateur des commentaires laissés sur *Global Plants* entre 2011 et 2015

par les usagers, parfois trop important par rapport aux moyens humains limités de la plupart des herbiers.

Nous avons relevé le cas remarquable de l’herbier virtuel Reflora⁴, qui a d’entrée de jeu été pensé dans une perspective de travail à distance avec le spécimen numérisé comme alternative au déplacement dans l’herbier. Le site est en effet associé à un programme de numérisation des collections concernant le territoire du Brésil détenues dans de multiples institutions à l’échelle internationale, notamment celles du MNHN. Il s’agit de faciliter le travail des chercheurs brésiliens sur la flore du pays en diminuant les frais de déplacement dans les institutions qui les détiennent. Pour accélérer l’inventaire des spécimens brésiliens des collections, des possibilités de détermination en ligne sont proposés, après une demande d’inscription par mail formulée auprès des responsables du programme Reflora. La délégation des droits de contribution est fondée sur une autorisation du visiteur souhaitant travailler dans un herbier à apposer des étiquettes aux spécimens. Cette procédure reproduit la procédure d’admission d’un visiteur dans un herbier, qui repose sur la confiance établie par le gestionnaire des collections dans sa relation avec lui, en partie liée aux marques de sa compétence – profession, notoriété ou publications. Les déterminations faites en ligne sont explicitement signalées comme telles, ce qui permet une traçabilité et éventuellement une attention critique accrue : un peu plus de 6000 déterminations virtuelles (sur les deux millions que contient l’herbier virtuel) ont été effectuées depuis la création du site.

Une dernière solution remarquable mise en place pour rendre possibles et exploitables les contributions, sans que le conservateur ne soit forcément disponible pour les valider, est celle mise en place dans le projet de recherche Annosys. Elle repose sur une base de données accueillant les corrections, qui est couplée aux bases institutionnelles qui l’alimentent, mais distincte d’elles. L’ensemble des commentaires est consultable dans une interface dédiée, avec possibilité de s’abonner aux commentaires de spécimens choisis. Les conservateurs ont les moyens d’apporter

4. <http://reflora.jbrj.gov.br/>

leur validation, facultativement, et le cas échéant, le système facilite aussi la réintégration des contributions vers leur propre base. La compatibilité des structures de données des deux bases permet en effet au gestionnaire d'autoriser l'intégration par un simple clic.

La solution d'Annosys, par sa facilité à s'intégrer à des portails de consultation des collections déjà existants, puisqu'elle est proposée par le GBIF (*Global Biodiversity Information Facility*, infrastructure internationale de centralisation de données sur la biodiversité), fournit un système abouti correspondant au besoin d'e-ReColNat, et plusieurs acteurs de la coordination et de la gestion informatique de l'infrastructure sont favorables à son intégration dans le portail. Ce pourrait donc à terme être une solution retenue pour l'infrastructure. Les retours sur cette solution attestent toutefois d'une difficulté à inciter la contribution (Suhrbier, 2017), qui est aussi celle repérée dans les laboratoires fondés sur l'enrichissement collectif d'un jeu de données partagées (Bos *et al.*, 2007). Ces difficultés ne sont pas nouvelles à l'herbier, car elles correspondent aussi à la difficulté, parfois exprimée par les gestionnaires, d'obtenir des retours de la part des usagers sur les publications auxquelles ont donné lieu leurs spécimens.

9.5. Favoriser la contribution au service de l'informatisation des collections

La diversité des solutions mises en place pour reconstituer la possibilité des corrections venant des usagers atteste de la reconnaissance de son utilité, pour alimenter les logiques de travail en réseau qui émergent au-delà de la diversité des bases de données de spécimens. Les contributions et commentaires laissés permettent de documenter les relations entre conservateurs de différentes institutions et la participation des usagers à l'enrichissement des données, alors que l'on a vu au chapitre 7 que les inventaires progressaient simultanément, sans échanges directs entre conservateurs, et sans que la consultation de multiples bases nécessaires à procéder aux nouveaux travaux d'informatisation ne laisse de trace permettant de valoriser la logique d'un réseau documentaire médiatisé par les bases de données en émergence.

Outre le problème de l'intégration des commentaires qui suppose une procédure de validation coûteuse en temps, qui peut trouver une alternative dans les bases enrichies couplées aux bases institutionnelles, l'ouverture des bases de données d'herbiers numérisés aux corrections des usagers suscite la difficulté d'entretenir la dynamique de la contribution : les principaux contributeurs sont des personnes déjà très impliquées dans le travail d'informatisation et d'inventaire des collections. Favoriser ces logiques émergentes de contribution ouvre la perspective d'utiliser la médiation documentaire numérique pour renforcer les liens entre le personnel des collections et leurs usagers, ce qui constitue un réel besoin étant donné que la question de la communication des résultats de recherche publiés à partir des spécimens, ainsi que le suivi des retours de prêt, constituent déjà un

souci des conservateurs et des gestionnaires, indépendamment des opérations de numérisation et d'informatisation des collections.

En termes de dispositif, une solution de couplage entre base institutionnelle et base ouverte aux enrichissements des contributeurs paraît s'imposer dans les différents projets, y compris dans l'infrastructure e-ReColNat. Elles varient les unes des autres selon les procédures de réintégration de données dans les bases institutionnelles. Elles rejoignent les solutions analysées dans le domaine patrimonial pour préserver l'autorité institutionnelle sur les collections sans compromettre la dynamique de participation (Casemajor Loustau, 2011).

En outre, la médiation documentaire numérique participative ne fonctionne pas sans la reconstitution d'une communication destinée aux usagers pour les inciter à contribuer. Le projet brésilien suscite davantage de contributions car il s'inscrit dans un projet scientifique partagé.

Les dispositifs ouverts à la contribution des internautes non spécialistes recréent cette incitation à la participation, qui semble manquer à certains dispositifs destinés aux professionnels, et suscitent de fait une participation plus importante, susceptible d'apporter un réel avantage au travail documentaire d'informatisation, et de renouveler les formes de communication sur lesquelles il repose.

10. La plateforme de transcription participative comme dispositif de médiation documentaire contributif reconstituant la mobilisation d'amateurs dans la gestion des collections

Nous avons jusqu'à présent appréhendé les nouvelles médiations de la communication mobilisant les collections en caractérisant le rôle des bases de données de spécimens dans les usages scientifiques et la conservation des herbiers : nous avons distingué certaines bases de données selon la manière dont elles reconstituent ou même renforcent la capacité des dispositifs de médiation documentaire des collections d'histoire naturelle à rendre possible conjointement l'accès aux documents et leur exploitation, ainsi que l'archivage d'une partie des données issues des travaux de recherche de multiples contributeurs — par le dépôt de nouveaux spécimens ou l'ajout d'annotations complémentaires sur les planches. Nous avons analysé les formes de communication générées par les bases de données multiformes qui coexistent, en particulier à travers la réappropriation des informations des spécimens numérisés mis en ligne, dans le travail d'inventaire comme dans leur exploitation scientifique. Les dispositifs envisagés dans les chapitres précédents concernaient des opérations prises en charge majoritairement par des professionnels. Nous nous intéressons à présent au dispositif développé pour ouvrir le travail d'informatisation des collections à des internautes non spécialistes. Nous comparerons plus précisément les formes de médiation documentaire en jeu dans l'herbier et dans la nouvelle plateforme, et en particulier les modalités d'intégration d'amateurs au travail documentaire qu'elles rendent possibles. Nous questionnerons les spécificités des formes de communication qui en résultent en précisant les caractéristiques des transactions impliquant les internautes contributeurs, notamment au niveau des pratiques de lecture et d'écriture qui se développent sur la plateforme.

10.1. Amplification des spécificités de la médiation documentaire de l'herbier dans la plateforme participative

La plateforme « Les Herbonautes »¹ reconstitue un nouveau projet mobilisateur capable d'intégrer des amateurs avec différentes motivations dans la collecte de données standardisées scientifiquement utiles. En ce sens, on peut considérer qu'elle réalise une amplification de la caractéristique essentielle de l'herbier dégagée par Susan Leigh Star et James Griesemer (1989), que nous avons reformulée comme la capacité d'un dispositif à associer d'une part la médiation organisationnelle de la mobilisation d'acteurs de compétences hétérogènes dans les collectes, et d'autre part la médiation documentaire permettant l'accès aux documents collectés par des chercheurs et amateurs de spécialités variées.

La plateforme renouvelle le fonctionnement de la médiation proprement documentaire des collections, par « la mise en place d'un système de communication adapté à l'utilisateur » (Liquète *et al.*, 2010), qui doit en l'occurrence s'adapter à l'absence de connaissance des collections chez les internautes découvrant la plateforme. C'est à ce niveau des modalités d'accès aux documents scientifiques repensées que se joue la première étape de l'effet de médiation exercé par la plateforme, qui permettra de rapprocher le public novice de la culture spécialisée des usagers des collections, et ainsi de « créer un lien et de concilier deux choses jusque-là non rassemblées pour établir une communication et un accès à l'information » (Liquète *et al.*, 2010). La plateforme organise en effet de nouvelles modalités d'accès aux spécimens à informatiser : les planches à transcrire sont sélectionnées parmi un lot défini en amont qui suit une cohérence thématique et définit une « mission », et sont réparties automatiquement entre contributeurs. Les spécimens numérisés sont ainsi distribués à chaque contributeur, qui n'a pas à formuler de requête pour accéder au spécimen sur lequel il travaillera : en reprenant la typologie des régimes de documentalité (Zacklad, 2015), la plateforme inscrit l'herbier dans le régime de la « circulation documentaire », même si le contributeur continue à devoir faire de lui-même des recherches pour accéder aux documents complémentaires nécessaires à transcrire ceux qui lui sont proposés.

La fenêtre de transcription (figure 10.1) constitue une nouvelle « interface qui accompagne l'utilisateur et facilite les usages » (Liquète *et al.*, 2010) du document scientifique. C'est un élément essentiel de la plateforme considérée comme dispositif de médiation documentaire. En plus de donner un accès recontextualisé au document spécialisé, l'interface de transcription oriente la lecture du document dans un sens utile à l'informatisation : sauf ponctuellement dans certaines missions, aucun renseignement n'est demandé concernant l'analyse morphologique du spécimen en lui-même. Cette fonction de cadrage de la lecture du document spécialisé dans un sens conforme aux intentions scientifiques à l'origine de la plateforme et aux objectifs proposés aux contributeurs montre l'efficacité de l'interface documentaire au niveau symbolique. La plateforme mérite bien en ce sens d'être qualifiée de dispositif de médiation documentaire, la notion permettant d'insister sur le rôle des

1. <http://herbonautes.mnhn.fr/>



FIGURE 10.1. – Capture d'écran de l'interface de transcription de la version 1 de la plateforme « Les Herbonautes ». La zone de gauche regroupe les champs à compléter, ouverts selon le niveau du contributeur. La zone de droite permet de zoomer et de se déplacer sur la planche à examiner pour déterminer les informations à transcrire.

artefacts conçus pour faciliter l'accès aux documents dans la construction de leur signification. Étant donné le caractère contributif de la constitution des collections et de leur documentation, la notion de dispositif de médiation documentaire désigne aussi dans leur cas les effets sociaux et organisationnels de mobilisation et de coordination d'une activité collective, qui se trouvent eux aussi amplifiés dans la plateforme des Herbonautes.

L'activité d'informatisation s'y trouve en effet transformée en une activité collective reconfigurée par le statut des contributeurs et les relations qu'elle organise entre eux. L'interface de transcription régule le « répertoire de participants » (Zaklad, 2015), les droits des contributeurs s'ouvrant à mesure de leur activité. Le droit de participer aux tâches qui demandent une lecture plus approfondie du document (comme la transcription de la localité en toute lettre et la géolocalisation) est octroyé à partir d'un niveau plus important de participation dans le site. Les stratégies de contrôle de données rejoignent un enjeu motivationnel par l'organisation des tâches de transcription comme une succession de niveaux à franchir. Ces niveaux correspondent aussi à un objectif d'encadrement pédagogique de l'activité de transcription, puisque réussir un quiz est la condition pour franchir chacun d'eux. La saisie des informations de l'étiquette en champs standardisés avec menus déroulants se justifie encore par un objectif de coordination de l'activité collective : le contrôle du vocabulaire est essentiel pour éviter les erreurs, dont le nombre est d'ailleurs plus important pour les champs à compléter en toutes lettres. La standardisation à laquelle concourt l'interface de transcription permet l'automatisation du contrôle des données produites, en rendant possible leur agrégation en vue du comptage des transcriptions identiques déclenchant la validation, à partir d'un seuil de transcriptions concordantes préalablement paramétré (trois pour

les pays, une seule pour les champs localité et géolocalisation – réservés aux usagers des niveaux 5 et 6). Ainsi, alors que les transactions induites par l’herbier reposaient sur la simple agrégation des produits d’une activité scientifique achevée, la plateforme met en relation les contributeurs, par des formulaires permettant de valider ou d’infirmier une transcription, et donne lieu à des formes de coopération par la mutualisation d’informations entre contributeurs dans des forums de discussion. Pour être intégrées à la base de données du MNHN, les transcriptions réalisées par plusieurs contributeurs à propos d’une même étiquette doivent faire l’objet d’un consensus. Les transcriptions qui se contredisent sont signalées aux internautes qui en sont les auteurs, et qui doivent alors confirmer qu’ils maintiennent leur interprétation, ou s’alignent sur celle des autres. Des échanges de commentaires à propos des planches concernées s’engagent en cas de désaccord. L’examen critique de la planche se poursuit ainsi avec un retour donné par les autres participants sur leurs propres interprétations.

Les effets de mobilisation et d’organisation d’un collectif produits par la plateforme reposent sur des outils d’écriture associant différents niveaux de standardisation et de mise en visibilité des contributions. L’algorithme organise et régule les interactions entre les contributeurs par la mise en visibilité ou en invisibilité des transcriptions individuelles enregistrées – on voit sur la figure 10.2 les marqueurs qui signalent des contributions réalisées sans en indiquer la valeur. Le voile n’est levé sur la transcription entrée par chacun qu’en cas de non-concordance, par l’envoi d’une alerte signalant un conflit et proposant un formulaire pour choisir la transcription à valider parmi les différentes déjà proposées. Il y a toutefois depuis l’un des onglets de chaque page « mission » une visibilité sur les spécimens qui sont en cours de transcription et les conflits non résolus – visibilité sur les conflits accrue dans la version 2 de la plateforme, et fortement utilisée par les contributeurs qui repèrent les conflits en cours et rejoignent les pages des spécimens correspondants pour y contribuer en priorité. Le dispositif donne non seulement accès aux documents scientifiques et à la documentation enrichie à mesure de l’activité des contributeurs, mais à certaines traces des interactions qui ont lieu sur la plateforme, essentielles à comprendre l’efficacité documentaire du site. Pour la comparer avec les autres dispositifs numériques ou non intervenant dans l’utilisation des collections, on peut caractériser la plateforme à partir de la typologie des dispositifs de médiation documentaire de Manuel Zacklad, en la situant au croisement des dispositifs de médiation documentaire contributifs orientés données (puisqu’elle produit des transcriptions standardisées) et orientés dialogue (à cause des forums de discussion). Cette hybridation des écritures documentaires liée à leur standardisation plus ou moins aboutie est déjà présente dans l’herbier qui autorise, en plus des noms d’espèces normalisés, des notes manuscrites. C’est notamment à partir de l’analyse des contributions libres ou standardisées que nous pouvons questionner à présent les formes de communication médiatisées par ce dispositif, par rapport à celles que génère l’herbier.

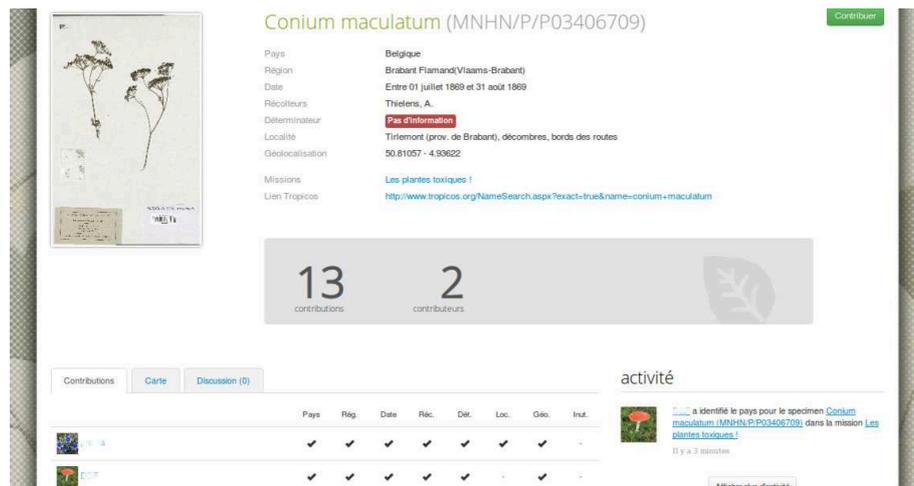


FIGURE 10.2. – « Page spécimen » consacrée à la planche n° P034067091 sur « Les Herbonautes » (capture d'écran). Les treize contributions des deux personnes participant à la transcription du spécimen sont récapitulées par des croix, chacune indiquant le champ rempli mais masquant la valeur qui y a été attribuée.

10.2. Analyse des transactions au cœur de la transcription participative

10.2.1. Un nouveau public pour les herbiers numérisés

L'analyse de l'activité des contributeurs dans sa dimension communicationnelle sous forme de transactions nous conduit à commencer par préciser quels acteurs y prennent part, à partir de leur profil sociologique et de leurs motivations, avant de détailler les interprétations qu'ils font des documents, ainsi que les différentes pratiques d'écriture et interactions auxquelles elles sont associées. Le public des contributeurs est hétérogène dans son niveau d'implication, entre les participants réguliers et les personnes qui découvrent la plateforme au sein d'une exposition ou à l'occasion de la communication institutionnelle qui est faite à son propos, et qui s'inscrivent alors simplement pour la tester le temps de quelques transcriptions. Les participants réguliers constituent une centaine de personnes actives, parmi près de 3000 inscrits depuis l'ouverture en 2012. Il s'agit d'une audience qui correspond, plutôt qu'au recours à la foule, au « community sourcing » (externalisation vers la communauté), organisée autour d'une communauté active, dont les principaux contributeurs ont appris à se connaître à mesure de leurs échanges. C'est la fidélisation de contributeurs réguliers qui est recherchée, plutôt que l'audience maximale. Cela s'est vu dans le fait que l'installation de postes informatiques qui aurait permis de découvrir la plateforme et d'y contribuer directement depuis la Galerie de Botanique au MNHN n'a pas été encouragée : « Les Herbonautes » y sont présents seulement par la présentation vidéo qu'en fait l'un des principaux animateurs de la plateforme, systématicien au MNHN, dans la perspective d'une présentation du document d'herbier, en cohérence avec le projet

de la galerie. L'inscription du site dans un contexte naturaliste est renforcée par le fait qu'il ne fonctionne pas à partir d'un cadre logiciel commun à différents projets (dont la plateforme Zooniverse est l'emblème, et dont on peut aussi rapprocher la fédération des observatoires de la biodiversité du MNHN dans le programme « 65 millions d'observateurs »), mais est inséré à la plateforme de l'infrastructure e-ReColNat. Le choix de communiquer en français implique aussi certaines limites à la participation, même si des missions bilingues ont été réalisées. Le code de la plateforme peut être repris à l'identique afin de créer des sites similaires pour d'autres communautés linguistiques, comme la récente version allemande des Herbonautes l'a montré².

Le questionnaire réalisé avec Tela Botanica, à partir d'un objectif d'enquête de satisfaction et de recueil de besoins en vue de la version 2, permet de préciser le profil et le statut des acteurs impliqués dans les transactions médiatisées par la plateforme. Parmi les contributeurs, on compte un public majoritairement non spécialiste des herbiers, mais déjà intéressé par la botanique (puisque'il a connu le site via Tela Botanica), exerçant parfois une activité en rapport avec la biologie ; le site a attiré des contributeurs impliqués dans d'autres démarches participatives, notamment en généalogie. On trouve aussi des spécialistes, comme des conservateurs d'herbier, qui constituent un public particulier étant donné leurs attentes (pallier l'absence de solution d'intégration des corrections par les usagers dans de nombreuses institutions). Les profils et les motivations des contributeurs laissent apparaître des convergences avec ceux des bénévoles intervenant traditionnellement dans les herbiers au niveau de l'attachage, et parfois de la saisie des données, tels que nous les avons décrits l'une des chercheuses en écologie rencontrées au MNHN : retraités, professions liées à l'enseignement des SVT, partageant un goût pour les collections et la botanique sans en être professionnels. On trouve d'ailleurs dans les principaux contributeurs des Herbonautes une bénévole de l'herbier de Genève. On note encore un recoupement avec la participation à d'autres programmes de sciences participatives, et une proportion importante de membres du réseau Tela Botanica parmi les animateurs de la plateforme.

Le questionnaire donne une idée de ce qui conduit à découvrir le site : la lettre d'information et le site de Tela Botanica ; des séances de formation au MNHN ; les démonstrations de la plateforme dans différentes situations, auxquelles nous avons participé ; mais aussi la communication institutionnelle sur l'herbier. Les communications associées à la rénovation de l'herbier sur des médias nationaux et régionaux ont indéniablement attiré des contributeurs. Des pics de contributions ont été observés lors d'émissions radio et de parution d'articles présentant les travaux de rénovation de l'herbier du MNHN en 2012. L'analyse du nombre de contributions par jour permet de corréliser les deux pics de participation les plus importants avec la parution des articles de deux quotidiens nationaux. Ces journaux sont *Le Monde* du 25 novembre 2013 (5537 contributions ce jour) et *Le Figaro* du 8 avril 2014 (4625 contributions ce jour). Un pic de contributions correspond aussi à la diffusion de l'émission de radio « La tête au carré » (Europe 1) du 23 décembre 2013, dans laquelle Marc Jeanson et Odile Poncy (MNHN) étaient invités.

2. <https://www.herbonautes.de/>

Corrélativement au fait qu'elles impliquent de nouveaux acteurs dans la production de données sur les collections, les transactions à l'origine de la production documentaire de la plateforme obéissent à des motivations particulières. Les entretiens avec des contributeurs plus réguliers, par téléphone ou entrevue, ajoutés aux questionnaires longs (reproduits en annexe D), laissent ressortir un intérêt pris à l'activité proposée, pour la sensation d'utilité qu'elle procure, par une participation à un projet d'intérêt général concourant au travail scientifique, et pour le plaisir d'apprendre et de mener des enquêtes avec rigueur. Le questionnaire en ligne permet de cerner, à partir d'un nombre plus important de répondants (une centaine de contributeurs, et une centaine de visiteurs du site non contributeurs), l'adhésion à l'objectif proposé. Gage de réussite du projet, cette adhésion transparaît dans les scores de satisfaction aux principales questions posées globalement élevés. Dans la question à réponse libre sur les motivations, le registre le plus fréquemment exprimé est celui de la « contribution » à un projet jugé « utile », d'intérêt général ou faisant progresser le travail scientifique. La recherche d'une utilité sociale à travers la contribution est notamment soulignée dans l'un des entretiens comme étant directement liée au départ à la retraite ; les contributeurs comportent effectivement une proportion importante de retraités (avec une moyenne d'âge de 50 ans des contributeurs répondants). Cette motivation importante de participer à un projet utile et scientifique se retrouve dans les principales enquêtes réalisées sur les motivations des contributeurs de projets proposant des tâches à partir de documents scientifiques ou patrimoniaux par l'intermédiaire d'une plateforme ouverte à tous, dans le domaine des sciences de la nature (Raddick *et al.*, 2010) ou des sciences humaines (Dunn et Hedges, 2014).

L'objectif partagé n'empêche pas l'existence d'une motivation par l'émulation, mentionnée par certains, qui cherchent à voir leur pseudonyme figurer sur le podium, à finir les missions, et à obtenir de nouveaux badges. La recherche d'efficacité qui passe par l'émulation est compatible avec des logiques de coopération et de dialogue nécessaires à résoudre les cas les plus complexes. Elle est aussi corrélée avec une préférence pour un rythme de transcription rapide (préférer renseigner un grand nombre de spécimens, alors que ceux qui demandent un travail plus complexe suscitent parfois un certain découragement et la volonté de passer au suivant).

Pour les sources de plaisir venant de l'activité elle-même, en plus de l'adhésion à l'objectif proposé, l'intérêt territorial demeure important, d'autant qu'il va avec une meilleure performance pour la documentarisation des spécimens des régions connues. Comme pour la participation aux collectes, motivée par l'intérêt pour la flore locale sans intention scientifique (Star et Griesemer, 1989), l'intérêt pour la transcription est lié à la connaissance du territoire dans lequel ont été faites les collectes. Ainsi, une contributrice canadienne nous dit avoir arrêté de participer au terme des deux missions consacrées aux Canada ; un des répondants au questionnaire, dans le même cas, exprime son désintérêt pour des missions « franco-françaises » ; le contributeur « Suédonaute » marque son intérêt pour les spécimens provenant de son pays. D'autres sources d'intérêt évoquées attestent d'un effet de la contribution à un projet documentaire pour nourrir un regard patrimonial sur les collections, et de motivations intrinsèquement liées à la nature de

l'herbier comme document d'archive (Moirez, 2012). C'est le cas dans ce témoignage : « J'aime (ré)apprendre la géographie française, d'une France qui souvent n'existe plus, [...] lire qu'on parle des ruines du Château de Pierre-fonds parce que la récolte a été faite avant sa rénovation ». La contributrice a été sensible au lien de l'herbier et d'autres éléments du patrimoine, en l'occurrence un monument. L'examen de l'étiquette qui conduit à faire le lien entre la planche, son auteur, et le lieu de collecte, permet aussi de prendre la mesure du lien de l'herbier avec l'histoire et les territoires dans lesquels il s'est construit. Au-delà de ces traits communs, on trouve des motivations particulières à la participation des professionnels, qui y voient l'intérêt d'avoir accès à des spécimens non encore recensés dans la base de données et concernant les collections dans lesquelles ils sont spécialisés. La démarche de la transcription reste globalement la même dans ces différents cas.

10.2.2. La transcription comme pratique informationnelle

L'analyse de l'activité des contributeurs demande de décrire une démarche de lecture et de saisie standardisée, accompagnée de recherches d'informations plus ou moins collaboratives. Les dimensions d'interprétation, d'annotation et d'interaction que conjugue l'activité de transcription permettent bien de la considérer comme un ensemble de transactions, à condition toutefois de structurer le cadre d'analyse issu de la sémiotique des transactions coopératives à partir des concepts développés pour décrire des pratiques informationnelles dans un contexte de recherche et d'exploitation mobilisant des documents scientifiques, comme nous l'avons proposé au chapitre 2. Nous empruntons ainsi des concepts issus de la modélisation des activités de lecture savante, de recherche d'information, et des techniques intellectuelles mobilisées par les chercheurs pour traiter l'information ; nous structurons avec ces concepts la représentation de l'activité des contributeurs telle qu'elle ressort de la description que certains d'entre eux ont faite dans les entretiens et les questionnaires réalisés pendant notre enquête. Nous avons croisé avec ces descriptions l'analyse des contributions, en particulier les URL citées dans les commentaires laissés sur la plateforme, pour enrichir notre compréhension de l'activité. L'étude des commentaires ouvre aussi sur la dimension sociale de l'activité, que nous aborderons seulement dans la partie suivante.

L'analyse de l'activité des contributeurs réguliers donne à voir des pratiques stabilisées progressivement, après une phase de désorientation initiale qu'ils mentionnent également. Ils décrivent les techniques de transcription acquises à mesure de l'entraînement. Les contributeurs réguliers entretiennent des niveaux de participation très variables ; parmi les herbonautes ayant répondu au questionnaire de satisfaction, deux déclarent avoir passé 50 heures à contribuer pendant les deux semaines précédentes, et deux autres, une vingtaine d'heures — et toujours avec une durée très variable consacrée à chaque document transcrit. L'activité a lieu le plus souvent à domicile, mais peut aussi être réalisée ponctuellement comme une pause au bureau. Nous décrivons ci-après l'activité présentée par des contributeurs néophytes. Nous précisons toutefois préalablement que la participation des spécialistes de taxonomie se distingue par un intérêt plus prononcé pour l'examen du spécimen en lui-même, outre celui des étiquettes, qui permet de se familiariser en

passant avec des flores encore inconnues. Nos entretiens avec les conservateurs et gestionnaires de collection ont mis en évidence la participation de certains d'entre eux à la plateforme, laissant aussi apparaître des témoignages de « confrontation de cultures et d'identités professionnelles » (Heaton *et al.*, 2011) : une frustration liée à l'absence de reconnaissance de leur expertise sur cette plateforme peut apparaître, étant donné qu'elle vise avant tout la participation de néophytes, alors que les gestionnaires ne disposent pas d'un réel espace de travail contributif dont ils auraient pourtant besoin.

Les contributeurs identifient leur démarche comme une enquête, conformément à la connotation du badge « Sherlock », obtenu en rétribution symbolique dès que le contributeur coche la case pour indiquer une déduction de l'information transcrite (par opposition à sa lecture sur l'étiquette). Un contributeur confirme la possibilité de comparer la recherche des éléments de l'étiquette avec l'identification taxonomique, qui procède par l'examen des caractères morphologiques. L'activité associe des étapes de recherche d'information et de résolution de problème (comme déchiffrer un caractère ou faire appliquer la juste règle de transcription). L'étiquette comporte autant d'objectifs de recherche que de champs à transcrire, mais ceux-ci se confirment entre eux. Selon les résultats du questionnaire, c'est davantage la recherche de la localité qui intéresse. Les recherches peuvent s'étaler dans le temps, relancées par les avancées d'autres contributeurs. Elles présentent en cela une dimension collaborative, pour les cas les plus complexes : elles engagent un « dialogue épistémique » (Tricot et Comtat, 2012) entre contributeurs dans l'espace de discussion du spécimen. Les contributeurs mentionnent les émotions qui accompagnent la recherche, notamment un découragement qui conduit à l'abandon dans les cas trop difficiles. Mais si certains privilégient des spécimens plus faciles qui permettent d'avancer rapidement, d'autres recherchent la satisfaction du « challenge » pour les cas plus difficiles. L'analyse des émotions accompagnant la recherche, conformément à l'approche ouverte par Carol C. Kuhlthau (2003), est importante pour comprendre la dynamique de contribution, entre abandon et satisfaction renouvelant l'engagement. Ce sont les difficultés d'application des consignes de transcription qui génèrent des insatisfactions (qui se lisent dans un score moins élevé obtenu par la question de l'intérêt pour l'identification du botaniste). Anticiper « la présence d'incertitude et d'anxiété dans les premières étapes du processus de recherche d'information » (Kuhlthau, 2003) que requiert la transcription des étiquettes anciennes, et le risque de frustration pour les contributeurs débutants qui en résulte, paraît important pour mettre en œuvre des solutions adaptées par l'animation et par l'assistance à la transcription dans l'interface de saisie. Sans avoir pu donner lieu à de nouveaux tutoriels dans la plateforme, ce souci a suscité des missions catégorisées « faciles », plus adaptées pour commencer. La dynamique émotionnelle associée à la recherche d'information progressant vers une plus grande confiance semble être aussi un facteur d'explication pertinent à l'engagement dans la contribution des herbonautes réguliers, qui ressentent d'autant moins les freins à la contribution qu'ils maîtrisent les outils de la plateforme et les ressources en ligne utiles à consulter.

Les questionnaires évoquent un même protocole induit par le dispositif de transcription et les routines qui ont émergé au fil des contributions, centré sur la des-

cription d'une activité solitaire, même si elle fait droit aux interactions à certaines étapes. Ces manières de faire sont justifiées par une recherche d'efficacité documentaire, jusque dans le choix de l'ordre de transcription. La recherche d'efficacité conduit aussi à construire un environnement de « favoris » utiles aux recherches, en plus des ressources de la plateforme principalement mobilisées : consignes, forums, outil de géolocalisation donnant les coordonnées. La description que les herbonautes interrogés font de la routine d'identification des étiquettes, ou de l'utilisation de la carte pour saisir les coordonnées plus rapidement, atteste de compétences de recherche d'information dans le contexte des herbiers, acquises par la fréquentation de la plateforme, qui passent par la capacité à « pré-visualiser le type de documents ainsi que sa structure » (Kembellec, 2012, p.126), et à naviguer dans l'environnement proposé (habileté à « manipuler »), dont Andrew Dillon (1996) souligne l'importance. La description faite par les contributeurs évoque une activité qui part de la visualisation des étiquettes sans considération particulière pour le spécimen — sauf pour les contributeurs qui s'intéressent aussi à la plateforme avec une préoccupation scientifique de découverte de flores inconnues, dans la perspective de recherches futures. L'identification des étiquettes est suivie de la détermination des informations géographiques et des données sur l'auteur du document. Des inférences simples peuvent suffire à identifier l'ensemble des éléments, si les étiquettes sont facilement lisibles, et demandent seulement une recherche simple pour déterminer le département et le pays correspondant à une localité précise. La localité peut même fournir l'ensemble des informations, d'où l'intérêt des contributeurs pour la consultation de bases de données internationales de localités et entités géographiques comme Geonames³ ; la carte de la plateforme fournit les coordonnées des localités, qu'il n'y a alors plus qu'à recopier. De même, quand il peut être déchiffré facilement, le nom du botaniste est trouvé dans un menu déroulant (toutefois critiqué pour ne pas toujours suffisamment être purgé des termes erronés). Même dans ces cas simples, l'identification des termes est suivie de la saisie, qui demande encore d'être réalisée conformément aux règles de la plateforme (préciser le déterminateur, uniquement de la première appellation, seulement si des éléments suffisants permettent de le reconnaître), qui diffèrent parfois des standards professionnels.

Dans les cas les plus problématiques, la démarche adoptée pour déchiffrer une étiquette partage avec la recherche d'information la caractéristique d'être orientée par un but encore indéterminé qui ne se précise qu'à partir de l'identification d'une information pertinente dans les documents trouvés, faisant consensus avec les autres transcriptions. L'activité nécessaire pour déterminer les informations à transcrire mobilise des techniques et sources d'informations variées et adaptées à chaque enquête. Certains évoquent en début de recherche la saisie d'une suite de caractères déchiffrés dans un moteur généraliste. La progression dans l'identification des noms présents sur les étiquettes se fait par la formulation d'une hypothèse suivie d'une recherche documentaire pour la confirmer, passant par exemple par la consultation d'un index de noms (de lieux géographiques ou de botanistes). La vérification est prolongée par la justification de l'identification auprès d'autres contributeurs en cas d'avis divergents. La fin de la recherche est constituée de la

3. <http://www.geonames.org/>

découverte d'un document attestant d'un nom de localité où le botaniste a pu se trouver à la date de la récolte. Elle conduit à réitérer la démarche sur un autre spécimen, avec un caractère addictif souligné par certains.

Les cas complexes donnent lieu à une vérification de la correspondance entre la date de la récolte et la biographie ou les voyages des botanistes. Les recherches d'information réalisées précédemment sont aussi utiles, l'un des contributeurs évoquant la reconstruction dans un tableau de la liste des spécimens collectés par un récolteur à une même période, pouvant faciliter la détermination des stations de récolte encore manquantes. Certains commentaires attestent du rôle de la mémorisation des transcriptions déjà effectuées dans la pratique : « J'ai enfin retrouvé toute la discussion sur les récoltes de Courtois au Anhui/Jiangxi (P04468210) », « Je l'ai retrouvée ! ». Les techniques de mémorisation évoquées par les contributeurs rappellent aussi celles que nous avons rencontrées dans les herbiers, comme la constitution d'un cahier personnel rassemblant des étiquettes et signatures caractéristiques de certains botanistes chez une gestionnaire de collections au MNHN. La répétition de ces tâches fait acquérir des connaissances sur les régions et leur passé. Elle accroît l'expertise des usagers, leurs connaissances acquises par la participation s'ajoutant à celles que détiennent certains usagers « experts » (Casemajor Loustau, 2011) sur des collections ou régions. L'acquisition d'une efficacité croissante sur la base de ces connaissances mémorisées est une source de satisfaction prise à la participation au site, permettant de surcroît d'accélérer le « travail ». La participation à la transcription engendre ainsi un apprentissage informel des spécificités des collections, qui nous a paru se rapprocher de l'habileté à reconnaître des écritures et à connaître la localisation des spécimens les plus utiles décrite par les professionnels des collections. Cet effet d'apprentissage permet de caractériser la plateforme comme un dispositif de médiation dépassant la simple action de diffusion d'information spécialisée ou de communication avec les usagers observées dans la plupart des dispositifs de médiation numérique utilisés par les musées qu'a analysés Patrick Fraysse (2015). La plateforme donne un exemple de médiation numérique qui dépasse la mise à disposition de l'information pour accompagner « la construction des connaissances », conformément aux actions de médiations documentaires conçues dans la perspective d'une médiation des savoirs qui envisagent ensemble le « processus de mise à disposition, de traitement, d'accès (et donc de recherche) et d'appropriation de l'information » (Gardiès et Fabre, 2012).

À part les rares mais notables cas où un botaniste contribue à titre bénévole à la plateforme et découvre un spécimen qu'il a lui-même déposé, les « usagers-témoins » (Casemajor Loustau, 2011), envisagés dans le cas de photos historiques comme des contributeurs pouvant s'appuyer sur leur propre mémoire pour apporter des identifications aux photographies, sont dans notre cas ceux qui connaissent particulièrement bien une région. Les usagers experts seraient ceux qui connaissent la langue de l'étiquette. Le fait de connaître la région rend l'activité à la fois plus facile et intéressante : les bénévoles rencontrés à Montréal ont ainsi expliqué leur motivation pour des récoltes d'une flore qu'ils connaissent, dans des localités qu'ils peuvent facilement identifier ; ils n'ont pas poursuivi la participation une fois achevées les missions consacrées au Canada. Ces préférences conduisent à une forme de spécialisation des contributeurs dans les zones qu'ils connaissent : nous avons

notamment observé un conservateur d'herbier utiliser la plateforme à partir de la carte géographique de la mission, pour préciser les informations manquantes dans les spécimens attribués aux départements qu'il connaît particulièrement bien, et qui sont autant d'occurrences qu'il peut ajouter à son inventaire.

Au-delà de la variété des cas rencontrés, un ensemble de ressources sont utiles, que nous avons pu analyser à partir de l'analyse des pages web citées dans les commentaires. Parmi 7000 URL analysées dans le corpus des commentaires, triées par ordre alphabétique, on trouve des ressources récurrentes : outre les ressources de la plateforme (les pages consignes), Wikipédia (en première place des liens), des bibliothèques d'ouvrages numérisés (bibliothèque numérique du jardin botanique de Madrid⁴), Google Books, des sites d'information géographique (« Get a Map »⁵, Wikimapia), ainsi que des bases de données d'herbiers (de Strasbourg, de Nouméa, du MNHN, du jardin botanique de Kew) et de botaniste (le répertoire de l'université de Harvard⁶). La numérisation des herbiers ne profite pas seulement des autres herbiers en ligne, mais aussi de l'ensemble des ressources numérisées, des projets contributifs, notamment dans le domaine de l'information géographique. Les ressources professionnelles citées dans les commentaires (bases de données des grands herbiers internationaux, et en premier lieu du MNHN) attestent que la plateforme exerce une fonction de formation à l'utilisation des ressources scientifiques, avec un rôle de médiation documentaire comme médiation des savoirs.

Les entretiens traduisent une importance différente accordée aux forums de la plateforme selon les contributeurs : certains cherchent davantage à trouver par eux-mêmes, d'autres y demandent plus spontanément de l'aide. Quant à la répartition des commentaires laissés selon les contributeurs, nous constatons, comme dans les forums de Zooniverse (Luczak-Roesch *et al.*, 2014), qu'elle ne correspond pas à la répartition des contributions, même si on retrouve dans les usagers les plus présents dans les espaces de discussions des contributeurs réguliers. Si les profils plus « sociaux » apparaissent aussi dans les sources de satisfaction des contributeurs, il reste à préciser le rôle des messages échangés sur la plateforme.

10.3. Rôle des transactions dialogiques dans la production de données standardisées

Préciser l'activité des herbonautes demande de comprendre leur investissement dans les échanges entre contributeurs et avec les animateurs, en croisant l'analyse des messages et celle des données de contribution, pour en comprendre le rôle dans la production de données valides. Par rapport à la description de leur activité faite par les contributeurs, l'étude des commentaires permet de préciser la méthode adoptée, y compris dans la dimension sociale des interactions qu'elle suppose, plus ou moins automatisées et standardisées. Nous détaillons maintenant l'analyse des

4. <http://bibdigital.rjb.csic.es/>

5. <http://fr.getamap.net>

6. http://kiki.huh.harvard.edu/databases/botanist_index.html

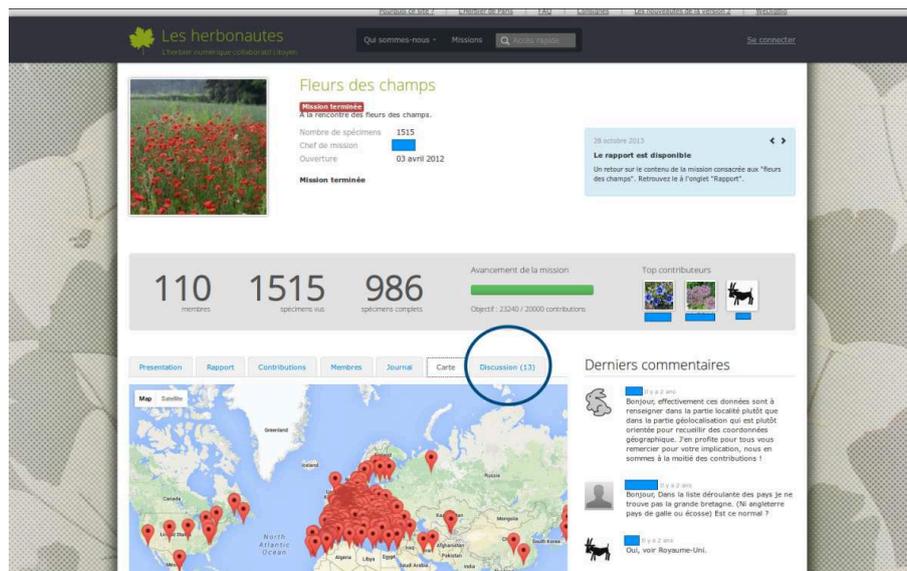


FIGURE 10.4. – Page de la mission « Fleurs des champs » sur « Les Herbonautes » (capture d’écran). Dans la moitié supérieure de la page sont indiqués le nombre de contributeurs et les résultats de la mission (nombre de spécimens transcrits). La partie inférieure de l’écran montre la carte de répartition des spécimens géolocalisés ainsi que, sur la droite, les derniers messages. L’onglet discussion (entouré) donne accès au forum de la mission.

rédigés par les animateurs, plus importante dans les forums de mission. Cette proportion diminue toutefois dans le temps, montrant une prise en charge croissante de la fonction de réponse aux questions par les contributeurs réguliers, qui constitue pour eux une modalité d’implication plus forte, conforme à la progressivité des tâches décrites par Mia Ridge (2014) dans d’autres plateformes.

10.3.1. Coordination de l’activité et apprentissages communautaires dans les forums de mission

La lecture manuelle et l’annotation de messages de forums de mission a permis de les caractériser plus précisément. Ils donnent lieu à des échanges de questions et réponses entre contributeurs et avec l’animateur, dit « chef de mission ». Ces échanges s’articulent en des fils de discussions qui se suivent et s’imbriquent parfois, et abordent principalement, outre le signalement de bugs informatiques, ou les demandes d’amélioration, les règles de transcription dans les différents cas de figure. Une fonction importante est de rappeler aux novices les règles de transcription les plus efficaces pour la communauté – sachant qu’elles peuvent se différencier des conventions adoptées par les professionnels. On le voit par exemple dans le cas du déterminateur, puisque la page de consigne des Herbonautes pour ce champ de saisie précise : « nous indiquons uniquement le déterminateur du nom mentionné entre parenthèses dans la fiche de saisie »⁷. Le forum sert à mutualiser

7. <http://lesherbonautes.mnhn.fr/contents/consignes>

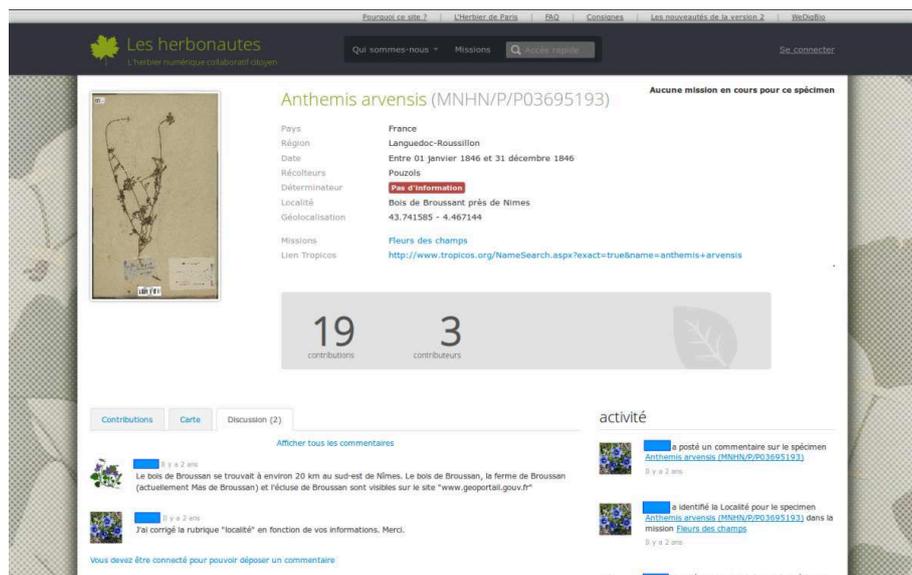


FIGURE 10.5. – « Page spécimen » consacrée à la planche n° P02115067 sur « Les Herbonautes » (capture d'écran). La moitié supérieure de l'écran récapitule les valeurs validées pour chaque champ, et indique le total des 19 contributions (une contribution = un champ de saisie complété) réalisées par 3 contributeurs sur ce spécimen. La partie inférieure de l'écran montre l'espace de discussion consacré à cette planche, avec un échange entre deux herbonautes (pseudos effacés) permettant de déterminer la localité correcte de la collecte.

Répartition des contributions annulées sur l'ensemble des spécimens complets

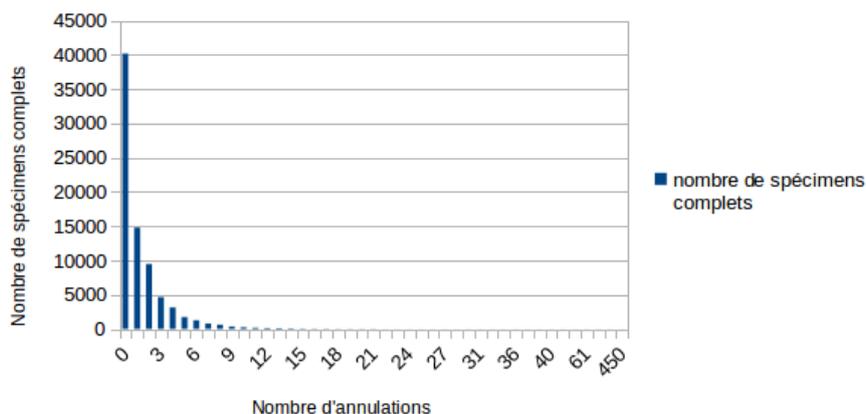


FIGURE 10.6. – Répartition des contributions annulées entre les spécimens complétés

les difficultés rencontrées, prises en charge par l'animateur qui répond à la plupart des questions, sans empêcher les initiatives de certains contributeurs réguliers venant directement en aide à d'autres. Dans cet espace, l'animateur (qui peut être un chercheur, un conservateur d'herbier, ou encore un membre de Tela Botanica) accompagne la pratique des contributeurs sur la base de son expérience des collections, en explicitant par exemple pour certaines règles le travail de documentation du musée qui justifie les consignes de transcription.

On retrouve dans ces espaces de discussion les fonctions mises en évidence pour les forums de la plateforme Zooniverse (Luczak-Roesch *et al.*, 2014) qui, comme les forums de mission des Herbonautes, sont visibles de tous les contributeurs : les interactions entre contributeurs et avec les animateurs scientifiques sont centrales pour les apprentissages nécessaires à la transcription. Les forums sont le lieu du retour des animateurs sur les contributions et de l'expression de la reconnaissance à l'égard des contributeurs. Leur rôle dans la coordination des interventions par le partage de règles communes génère un apprentissage communautaire des connaissances nécessaires pour participer à l'informatisation.

10.3.2. Les messages à l'origine des données valides

Le second ensemble de commentaires est associé à chaque spécimen, ce qui permet de questionner leur rôle dans la production de transcriptions valides, par une approche quantitative du nombre de commentaires, contributions et annulations liés à un même spécimen. La production documentaire de la plateforme passe par une utilisation inégale des fonctions de commentaire, selon les spécimens et les contributeurs. Avec une moyenne de 19 transcriptions nécessaires à valider les 6 champs d'un spécimen (ou le qualifier d'inutilisable), la proportion de spécimens transcrits qui n'ont suscité aucune annulation domine (voir figure 10.6). Les spécimens qui ont généré le plus d'annulations – et donc de contributions –, sont aussi

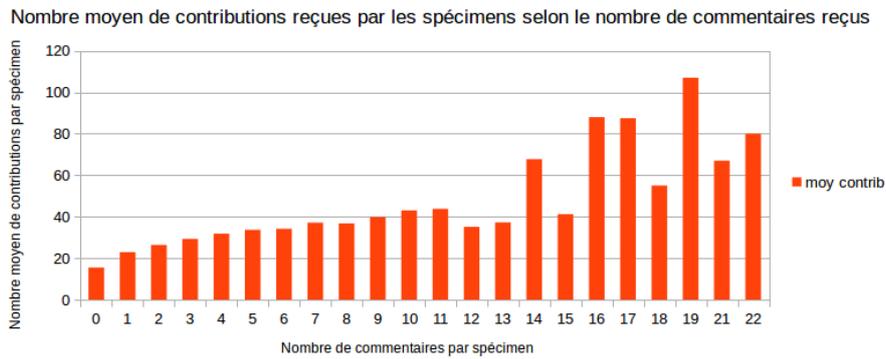


FIGURE 10.7. – Répartition du nombre de contributions selon la répartition des commentaires

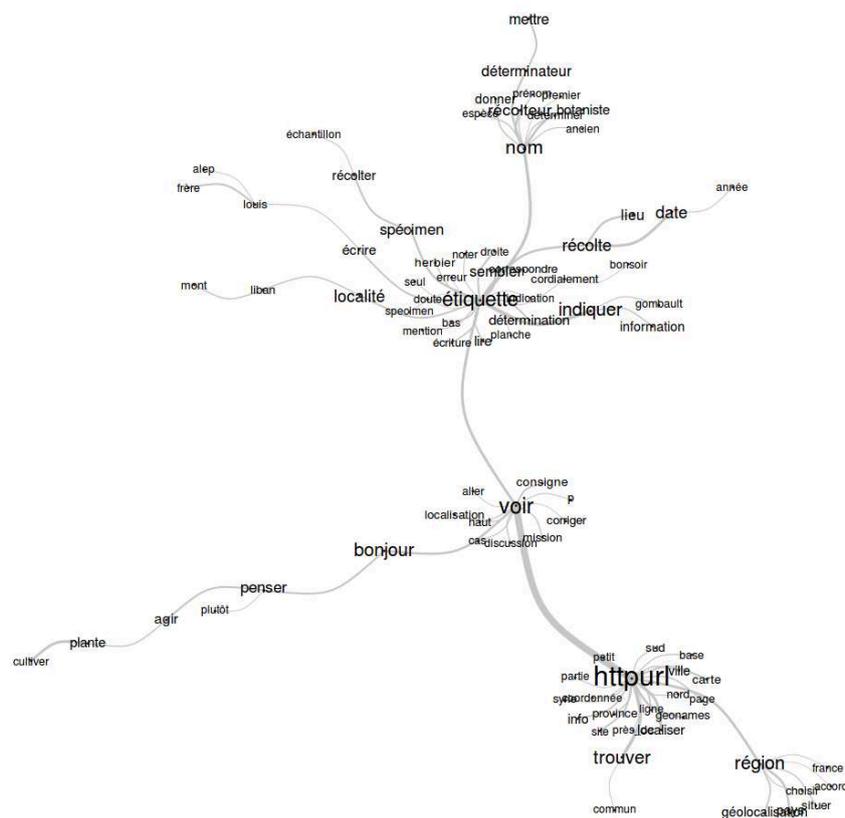


FIGURE 10.8. – Analyse des co-occurrences dans les messages associés aux spécimens, réalisée avec Iramuteq (Ratinaud, 2005)

les plus commentés (voir figure 10.7). Le rôle de la concertation entre contributeurs ressort de la corrélation entre le nombre de contributions et de commentaires de chaque spécimen (voir figure 10.7), qui signifie que des modifications ont été apportées aux contributions à mesure des messages échangés.

L'analyse du lexique permet de préciser le rôle de ces commentaires. Loin d'un fonctionnement en aveugle, on « s'aligne », on « suit » l'argumentation d'un autre, avec plus ou moins de « conviction », on remercie celui qui a débloqué une situation en trouvant l'information précieuse ou en effectuant le changement demandé. Certains relaient l'application des « consignes » (966 occurrences associées à des spécimens) et leur explication, manifestant une « vigilance participative » (Cardon et Levrel, 2009) concernant l'application des règles de transcription. Certains commentaires conservent la forme brève du formulaire de l'« alerte », message automatique envoyé en cas transcription discordante, qu'ils anticipent ou prolongent, par une réponse à une question implicite qu'elle suggère (« Je maintiens 1909 » ; « Pourquoi pas Welwitsch ? »), ou par la justification d'un choix de transcription en vue d'éviter un désaccord futur. À l'opposé de ces échanges brefs (qui constituent une part importante du corpus, puisque 6018 sur les 33 817 commentaires des spécimens font moins de 50 caractères), des discussions plus longues conjuguent le partage d'informations et la réflexion sur les règles de transcription à appliquer.

Outre l'analyse lexicale des formes spécifiques qui fait ressortir les entités géographiques comme les plus caractéristiques des messages associés aux spécimens par rapport à ceux des missions, l'analyse des co-occurrences conduit à préciser la thématique des messages (voir figure 10.8). On distingue des ensembles de termes co-occurents correspondant aux différents champs qui font l'objet des recherches (on voit notamment en haut les formes associées aux messages visant à identifier déterminateur et récolteur), ainsi que des commentaires utilisés pour signaler les « plantes cultivées » pour lesquelles des consignes spécifiques sont en vigueur. C'est surtout la mise en relation des spécimens avec d'autres documents du web qui ressort de la forte corrélation entre « voir » et « httpurl » (codage adopté pour toutes les URL du corpus pour éviter les erreurs de lemmatisation), et avec des termes géographiques génériques. Cela s'explique par le fait que, même en l'absence d'obligation de citer des références, de nombreux herbonautes ont pris l'habitude de justifier leur choix de transcription d'une entité de lieu ou de botaniste.

L'analyse du corpus des messages permet de dégager une fonction qui excède leur rôle dans la détermination des transactions correctes d'un unique spécimen. Outre le fait qu'il peuvent être réutilisés par les contributeurs dans leurs recherches ultérieures, par exemple en permettant de comparer des spécimens d'un même récolteur, leur intérêt apparaît aussi pour l'enrichissement de la documentation des collections. Quand ils mentionnent des cartes anciennes ou des récits d'expédition consultés dans d'autres fonds numérisés comme « Gallica » (73 occurrences), ou encore les publications qui citent les spécimens, les contributeurs tissent des liens entre les images et les documents web utiles à leur compréhension. Ces relations peuvent aussi mettre en lumière les liens internes aux collections, lorsque sont signalés des rapprochements permettant de reconnaître des signatures ou écritures identiques, ou des renvois entre des planches (« feuille 1 » ou « feuille 2 » : 50 occurrences). Quand un contributeur signale un lien fort entre des spécimens dont

la consultation concomitante lui a permis de reconstituer l'ensemble des informations de l'étiquette, le lien peut intéresser le scientifique non seulement pour qu'il puisse vérifier la transcription, mais aussi parce que les spécimens rapprochés ont des chances de provenir d'une même récolte, information qui peut s'avérer déterminante. Des spécimens « types » sont aussi signalés comme tels par les contributeurs qui les ont reconnus (on compte 20 occurrences du terme « isotype » dans le corpus). La question se pose alors de l'intégration de ces commentaires apportant des précisions de lecture : seules les transcriptions standardisées sont importées dans la base institutionnelle d'e-ReColNat ; les contributions non standardisées subissent encore une logique « d'effacement » (Heaton et Proulx, 2012), alors même qu'elles pourraient éclairer la lecture critique de l'étiquette à laquelle se prêtent les scientifiques pour vérifier la cohérence des informations du spécimen. La réflexion concernant l'intégration des commentaires dans l'espace du collaborateur a été l'un des axes de réflexion de la version 2 de la plateforme sur laquelle nous revenons à présent.

10.4. Vers un dispositif de médiation documentaire contributif orienté dialogue valorisant les contributions non standardisées

Nous commençons par caractériser au niveau de l'architecture documentaire développée pour les collections de l'infrastructure e-ReColNat les différences de visibilité des contributions produites depuis les applications développées par chaque « workpackage », afin de comprendre les enjeux de la réflexion de la version 2 et les perspectives qui se dessinent encore suite à cette expérience. Nous avons vu que, par rapport à l'herbier qui ne recense que la localisation des liasses d'un compactus, la base de données accroît la référentialité de la collection numérisée, c'est-à-dire la « permanence » et la « facilité avec laquelle le document ou les fragments qui le composent peuvent être rendus accessibles dans un réseau interne ou sur le Web » (Zacklad, 2015). La plateforme des Herbonautes complexifie encore la référentialité de l'herbier numérisé. En effet, elle associe à chaque spécimen une seconde page web (en plus de celle de la base de données e-ReColNat), qui rassemble les informations issues de l'activité de la plateforme. Deux URL coexistent donc pour un même spécimen — et c'est même bien davantage dans l'ensemble des modules du portail e-ReColNat⁸, étant donné qu'existe aussi une page ouverte aux commentaires pour les spécimens testés dans le logiciel de contrôle de la qualité de l'image à la sortie de la numérisation, en plus d'une notice dans la base locale de l'institution partenaire de e-ReColNat qui détient la planche. Il n'y a pas de mise en relation facilitée par hyperlien entre ces pages. Cette difficulté de coordination est propre au régime de documentalité caractérisé de « transmédia documentaire » par Manuel Zacklad (2015), dans le sens où les données concernant le spécimen d'herbier peuvent être produites par différents contributeurs depuis des terminaux

8. <https://www.recolnat.org/fr/>



FIGURE 10.9. – Photomontage d'étiquettes d'un même botaniste réalisé par des contributeurs

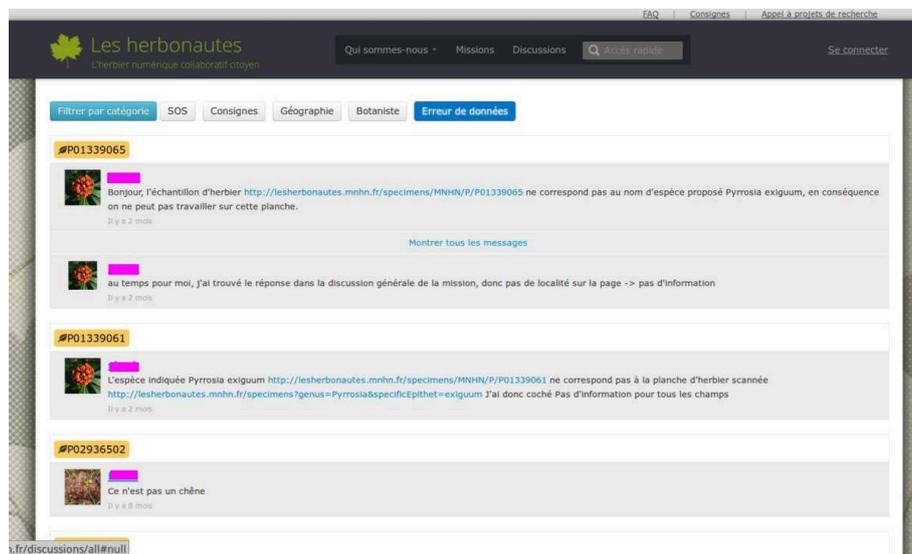


FIGURE 10.10. – Capture d'écran de la page de synthèse des discussions dans la version 2 des Herbonautes

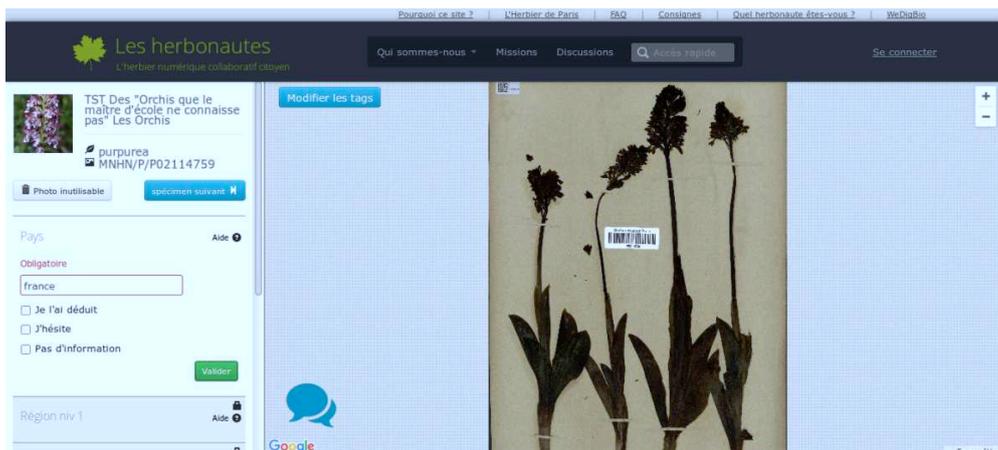


FIGURE 10.11. – Capture d'écran de l'interface de transcription de la version 2 des Herbonautes

variés : en l'occurrence, elles viennent de la base de données régionale et de tous les contributeurs ayant travaillé sur le spécimen numérisé dans « Les Herbonautes ». Ce régime de documentalité interroge la coordination des données produites dans ces différents contextes à propos d'un même objet. Un exemple de la résolution de ces difficultés dans la plateforme se trouve dans la règle d'intégration des données à la base e-ReColNat, qui n'a lieu que si les conservateurs n'ont pas déjà informatisé le spécimen, pour éviter tout conflit. L'absence de lien hypertexte entre les différents modules dans lesquels le spécimen peut être documentarisé, pourtant accessibles depuis le même portail, montre les difficultés restantes.

La mise en évidence du rôle des commentaires a conduit à réfléchir à des solutions pour faciliter leur utilisation et réexploitation par les contributeurs ou d'autres utilisateurs ultérieurs. Ce travail a donné lieu à une version 2 de la plateforme, comportant également d'autres évolutions issues des besoins des animateurs, notamment pour faciliter l'intégration des données à la base e-ReColNat à la fin d'une mission. Des contributeurs avaient souligné le manque de forum (voir annexe D, première répondante). Ils ont aussi besoin de réutiliser facilement leurs précédentes transcriptions. Certains commentaires expriment la volonté de rendre plus accessibles les messages en vue de faciliter leur réutilisation : il s'agit selon l'un d'eux de « constituer une base de connaissances évitant à chacun de chercher dans son coin des références identiques », et de pouvoir « bookmarker » des recherches en cours ou déjà validées. Si les contributeurs réalisent de leur côté des bricolages (par exemple des photomontages d'étiquettes, voir figure 10.9), de nouveaux développements de la plateforme pris en charge par les institutions responsables étaient nécessaires pour la faire évoluer dans le sens d'une meilleure instrumentation des pratiques. C'est un élément de différence important entre le modèle du recours à la foule, dans lequel la plateforme reste entre les mains de l'institution qui la contrôle, et celui des communautés en ligne dont les contributeurs font eux-mêmes évoluer les outils selon leurs besoins – l'évolution des règles de contribution dans Wikipédia et le développement d'interventions de robots, ou encore le choix de la mise en place d'un wiki après une crise de la documentation

dans la communauté WordPress, en sont des exemples (Ruzé, 2013). De ce point de vue, la situation des contributeurs dans la plateforme correspond à celle des bénévoles qui participent aux chantiers d'attachage au MNHN, qui n'ont pas la main sur les règles de fonctionnement ou l'organisation des connaissances de la collection; c'est aussi la situation des participants aux programmes de collecte de données de Tela Botanica (Heaton *et al.*, 2011), du moins pour la majorité d'entre eux, non impliquée dans la définition et la gouvernance des projets. Dans le cas de la plateforme des Herbonautes, la version 2 a été conçue à partir des besoins exprimés par les animateurs et responsables du projet, mais en intégrant le recueil des demandes des usagers, et les conclusions de notre analyse du rôle des contributions. Cela a conduit à une forme de « contribution des usagers » indirecte au « design » du dispositif (Latzko-Toth, 2014), alors que leur rôle relève plus souvent de la contribution aux contenus, sans pouvoir sur l'évolution des plateformes utilisées.

Sans encore intégrer les commentaires des Herbonautes à l'espace de travail en ligne des scientifiques, la version 2 du site a permis de rendre paramétrables les questions posées aux contributeurs. Il est ainsi possible de collecter des données standardisées selon différents modèles, en vue d'utiliser la plateforme pour d'autres programmes participatifs — l'enquête en vue de la conception du laboratoire en ligne d'e-ReColNat avait fait ressortir notamment le projet de faire catégoriser des images selon des caractères repérés, pour faciliter ensuite des recherches de spécimens dans la base de données par caractère morphologique. La version 2 a en outre répondu au projet de mettre en visibilité les espaces de discussion associés aux spécimens, qui n'étaient accessibles que pour les contributeurs, au moment de la transcription ou à partir de la recherche sur la plateforme par numéro de spécimen. À cet effet, une page récapitulant les dernières discussions a été créée (voir figure 10.10); elle était aussi le moyen d'offrir une nouvelle forme de contribution, par la réponse aux demandes d'aide formulées dans les messages catégorisés « SOS ». La mise en visibilité des discussions a renforcé le caractère coopératif du dispositif de médiation documentaire : l'interface de transcription de la version 2 fait apparaître un bouton permettant de commenter le spécimen, qui indique le nombre de messages lui ayant déjà été attribués (voir annexe C.2 pour la maquette, et pour la réalisation la figure 10.11). La logique de contribution indépendante en parallèle reste celle qui régit la participation tant que les usagers ne commentent pas le spécimen ou ne tiennent pas compte des messages déjà postés, mais la nouvelle interface de transcription incite à entrer plus rapidement dans une démarche de recherche d'information mutualisée. Par ces différents outils de communication plus développés, la version 2 vise à renforcer le registre de la contribution orientée dialogue au service de la contribution orientée données. Elle met en visibilité les contributions non standardisées, au service de la communauté en premier lieu, mais aussi dans la perspective d'autres programmes de sciences participatives à partir des documents d'herbiers. Pour faciliter la réutilisation des contributions, des fonctions de recherche plein texte, de catégorisation des messages et de « tags » choisis par l'utilisateur ont été développées : les rubriques des messages, qui structurent l'espace de synthèse des discussions (figure 10.10), sont configurables par les administrateurs de la plateforme ; le contributeur peut

en revanche catégoriser librement ses messages par l'ajout de « tags ».

La plus grande visibilité des commentaires laissés sur un spécimen n'a pas provoqué de hausse significative du nombre de messages échangés rapporté au nombre de contributions réalisées. Les possibilités d'organisation des espaces de discussion par catégorie sont mobilisés ponctuellement, en particulier pour signaler des erreurs de données (comme un nom d'espèce de la notice de la base incohérent avec les étiquettes). Les tags restent utilisés marginalement par rapport au nombre de spécimens traités. De la mise en ligne de la version 2, en janvier 2016, à juin 2017, les 562 tags signalent essentiellement des localités et des botanistes, associant au maximum une dizaine de planches aux tags les plus utilisés. Les discussions ne sont pas titrées ni fermées, ce qui montre que le dispositif dialogal prévu était *a priori* trop complexe pour les besoins de la plateforme.

L'utilisation des nouvelles possibilités de classement libre des images par tag donnaient l'occasion d'observer si les contributeurs les utiliseraient pour conserver la trace de points de vue et interprétations des spécimens autres que ceux requis par les tâches de la contribution. La version 2 donnait ainsi l'occasion d'observer un potentiel effet de valorisation historique ou culturelle des collections opéré par la plateforme à partir du projet d'informatisation participatif. Ce test était un nouveau moyen de mettre à l'épreuve notre hypothèse de la possibilité d'un dispositif numérique de médiation documentaire augmentée de fonctions de découverte des collections, pour des publics non spécialistes participant à son enrichissement. Il s'appuyait sur l'utilisation de l'opération de catégorisation pour témoigner de la forme d'interprétation faite du document, dans une démarche qu'on peut rapprocher de celle qui avait été mise en place dans le cadre du projet LIBRIS de valorisation du patrimoine photographique du Nord, qui visait à permettre l'expression des points de vue de publics non spécialistes sur un fonds d'images historiques mis en ligne dans une base de données, en proposant aux usagers testeurs de classer les images à leur convenance (Després-Lonnet, 2000). Dans notre cas, l'utilisation de ces mots-clés librement associés à des planches s'inscrit surtout dans une problématique d'efficacité documentaire, en cohérence avec le projet proposé aux internautes. Les nombreux termes géographiques taggés attestent d'une fonction de mémorisation en vue des contributions futures. Les tags sont aussi utilisés comme une autre possibilité de mettre en valeur un message. Certains d'entre eux témoignent en outre d'un déplacement du point de vue sur les documents par rapport à la lecture cadrée par l'objectif d'informatisation : des herbiers historiques sont soulignés (« Herbarium Vaillant », « herbarium propriétaire »), mais aussi une caractéristique concernant l'usage du spécimen (« empoisonné »), outre les signalements de type. Les tags sont aussi utilisés pour repérer les signatures d'un botaniste, pratique que nous avons observé chez les gestionnaires : on compte une dizaine de tags « signature » associés au nom d'un botaniste.

Toutefois, sans animation allant en ce sens, l'utilisation de ces marqueurs est restée une initiative personnelle au service de sa propre pratique, sans qu'une convention n'émerge à l'échelle de la communauté pour favoriser le partage d'informations au-delà d'usages de mémorisation individuels, et pour « passer de la mise en mémoire à la transmission de connaissances » (Le Deuff, 2010). Si le dispositif

technique est opérationnel pour permettre une collecte d'informations attestant de différents regards sur les spécimens (par exemple selon les usages des espèces), il n'est pas spontanément utilisé en ce sens — nous notons que la démarche de classement a été laissée à l'appropriation libre de la communauté des Herbonautes, alors qu'elle avait été explicitement proposée, dans le cadre d'une expérimentation, lors du projet LIBRIS (Després-Lonnet, 2000). Il ressort aussi de cette expérience des premiers mois de fonctionnement de la version 2 que les contributeurs ne participent que marginalement à l'organisation des connaissances régissant l'activité de transcription — puisque les catégories nouvelles qu'ils créent restent peu nombreuses et proches des concepts déjà impliqués par la production de données standardisées. Les formes de participation qui rencontrent le plus de succès restent celles qui sont encadrées par des normes déjà bien stabilisées garantissant l'efficacité documentaire : la contribution s'inscrit dans le registre de participation propre à « l'intelligence distribuée » (Haklay, 2015), puisqu'elle repose sur l'interprétation des documents, à la différence des logiques de contribution automatisée passant par l'utilisation de capteurs de surveillance environnementale. Il n'y a pas en revanche de participation aux hypothèses, ou à la fixation d'objectifs de recherche, qui pourrait passer par des règles de tagging élaborées en commun, comme ce serait le cas dans une démarche de « sciences participatives » au plein sens du terme.

De ces limitations de l'appropriation que font les contributeurs des outils de catégorisation libre, ressort un constat pour la démarche de recherche-conception et le rôle qu'y prend le chercheur en communication : l'implication dans le développement de médiations instrumentales pour l'activité de communautés en ligne gagnerait à se prolonger par une participation du chercheur à l'animation de la plateforme, et en particulier au moment de la mise en production des nouveaux outils de communication issus de la démarche de recherche. En l'occurrence, ce sont uniquement les limitations du temps disponible pour suivre l'animation de la plateforme à ce moment critique qui ont conduit à réduire notre participation à l'étude des traces et des usages, à la proposition de spécifications et au recettage. Le travail de conception s'est d'ailleurs prolongé suite à l'étude des données issues de la version 2 de la plateforme. Si les catégories produites confirment un intérêt patrimonial ou historique de certains contributeurs pour les collections, qui se traduisait déjà dans les motivations et sources de satisfaction exprimées dans le questionnaire, les contributions non standardisées dans la plateforme restent surtout mises au service du projet d'informatisation : tags et commentaires font référence aux catégories et critères institutionnels de la conservation. Mais ces commentaires sont porteurs d'un réel enrichissement de la documentation des collections : les signalements de spécimens types et les corrections de données qui ont eu lieu dans la version 2 confirment l'apport des commentaires laissés sur la plateforme pour l'étude scientifique du document. Mettre en valeur ces enrichissements demande de trouver les moyens d'une meilleure visibilité de ces contributions, qui pourrait passer par leur intégration dans d'autres logiciels de l'infrastructure. L'une des maquettes réalisées pour le laboratoire en ligne, qui prévoyait d'intégrer les données du module de contrôle qualité et des Herbonautes, ouvrait cette direction (voir annexe C.3). L'étude des formes de communication à l'origine des données standardisées produites sur la plateforme, en particulier l'analyse des commentaires,

s'est donc ici directement prolongée dans un geste de conception dessinant des pistes de développements ultérieurs susceptibles de favoriser ces enrichissements contributifs par leur mise en visibilité accrue. Le cercle entre recherche et conception a été toutefois interrompu par la priorisation d'autres fonctionnalités moins coûteuses en temps de développement, compte tenu de l'ensemble du calendrier du projet.

10.5. Médiation documentaire augmentée et médiation numérique du patrimoine à l'aune de la plateforme des Herbonautes

La plateforme des Herbonautes témoigne des potentialités de l'ouverture du travail documentaire à des contributeurs non spécialistes pour amplifier, par un dispositif numérique, les caractéristiques de la médiation documentaire complexe à partir de laquelle nous avons caractérisé l'herbier, qui associe accès aux objets des collections, coordination d'une activité collective et échanges de savoirs entre contributeurs de différents statuts, et qui peut s'intégrer à d'autres dispositifs muséaux ou universitaires de transmission des savoirs scientifiques. Outre le fait qu'elle accroît massivement le nombre de contributeurs pouvant prendre part à la documentation des collections, la plateforme rend possible un effet d'amplification dans la description standardisée de chaque planche perfectionnant les possibilités de recherche dans les collections, alors que l'inventaire n'était que rarement effectué avec ce degré de précision avant la numérisation. Le site donne aussi un exemple de dispositif de médiation numérique permettant de coordonner l'action d'un collectif intégrant des amateurs qui participent la constitution de la documentation des collections, en associant aux interfaces de saisie une confrontation algorithmique des données issues des interprétations des étiquettes, et des échanges dialogiques permettant de justifier les choix. Les motivations de certains contributeurs attestent également d'un regard patrimonial porté sur les collections, démontrant un effet de la plateforme qui dépasse la perspective de faciliter l'accès aux documents qui correspondrait à une acception plus restreinte de la médiation documentaire, pour favoriser la découverte de l'herbier dans ses rapports au « monde des hommes qui l'ont produit, utilisé, codifié, embelli, voire au contraire saccagé ou détruit » (Davallon, 2006, p.123). Enfin, en permettant des apprentissages informels et communautaires autour des herbiers et des techniques documentaires qui s'y appliquent, et en accompagnant des amateurs dans l'utilisation de bases de données scientifiques de spécimens numérisés, la réécriture de l'interface de saisie professionnelle dans un contexte scénarisé produit un effet de médiation des savoirs : elle part d'une formation au traitement documentaire pour participer à « transformer l'information » – repérée sur les spécimens – « en connaissance » – des collections et de leur valeur (Liquète *et al.*, 2010).

On retrouve les trois catégories de savoir dégagés par Patrick Fraysse (2015) de ses études sur les dispositifs de médiation numérique muséaux : la plateforme permet d'acquérir des connaissances spécialisées sur les collections, mais cela en

partant des connaissances des publics, à savoir les centres d'intérêts pour les collections à partir desquels sont construites les missions, ou encore les difficultés rencontrées dans la réalisation des transcriptions. L'acquisition de savoirs sur les collections se fonde sur « des habiletés techniques et numériques », et des connaissances documentaires, qu'acquière aussi les contributeurs, comme le remarque Patrick Fraysse dans d'autres expériences de médiation numérique qui reposent sur la maîtrise de connaissances techniques et documentaires. Les situations de démonstration et de test de la plateforme des Herbonautes, destinées à un public non spécialiste dans des contextes de manifestations scientifiques festives, suscitaient d'ailleurs de nombreuses questions sur son fonctionnement technique, en particulier sur les procédures de validation. À travers le processus d'acquisition de connaissances sur les collections observées chez les contributeurs réguliers qu'elle rend possible, la plateforme mérite d'être qualifiée de « dispositif de médiation numérique » au sens où elle permet bien « d'accéder au savoir du musée », conformément à la définition élaborée par Patrick Fraysse à partir d'études de terrains, et contrairement à la plupart des dispositifs qu'il a observés, et qui ne sont de fait utilisés qu'à des fins de communication du discours du musée, ou d'information sur ses collections (Fraysse, 2015).

Le développement de la plateforme au sein d'une infrastructure e-ReColNat intégrant une démarche de recherche-conception a aussi été à l'origine d'une approche originale de la médiation, à la fois dans sa mise en œuvre pratique et dans son étude. La participation à la conception de nouvelles fonctionnalités pour la plateforme, sur la base de retours des usagers et de l'analyse des écrits à l'origine de la production de données valides, inscrit la recherche sur les médiations dans le processus de conception d'outils de transcription adaptés à intégrer des participants de différents niveaux. Cette démarche a permis à la fois de faire évoluer le dispositif de médiation sur la base des conclusions issues de l'analyse des usages et des contributions. Elle a permis de dégager les spécificités de l'entrée en relation des scientifiques avec les contributeurs, en particulier dans les espaces de discussion. La démarche de recherche-conception s'est inscrite facilement dans le développement de la plateforme du fait que les outils de la contribution ne sont pas développés par les usagers eux-mêmes, et qu'une place apparaît donc pour une recherche contribuant à proposer des outils au service des différents profils d'usagers et techniques de contribution. La participation au travail de conception a à la fois donné un but à la recherche orientée à l'efficacité de la communauté, et une problématique guidant la conception des questionnaires et entretiens, autour des usages et besoins des contributeurs. La recherche-conception, par sa démarche itérative (Roy et Prévost, 2013, p.136 ; Parrini-Alemanno, 2004, p.58) conduisant à travailler sur des données issues de deux versions principales de la plateforme, a aussi permis de comparer l'efficacité de la communauté compte tenu de différents outils proposés qui faisaient plus ou moins droit à l'initiative des contributeurs dans la production de catégories. La participation à la conception gagnerait à se prolonger par la suite par une implication dans la définition de missions susceptibles de s'appuyer sur une catégorisation proposée à la communauté, une fois le diagnostic posé d'un besoin d'accompagnement de l'utilisation coordonnée des tags. L'implication du chercheur dans le développement des médiations de l'ac-

tivité de la communauté qui est aussi l'objet de son étude concernerait ainsi la conception d'outils adaptés aux objectifs scientifiques définis, et l'investissement dans l'animation de la plateforme, deux dimensions essentielles des actions de médiation nécessaires à mobiliser la communauté.

11. Conclusion

La démarche de recherche-conception développée pour favoriser l'exploitation scientifique et la valorisation patrimoniale des collections d'histoire naturelle de l'infrastructure e-ReColNat a permis d'appréhender corrélativement les mutations de la communication mobilisant les spécimens dans le contexte de leur numérisation, ainsi que ses nouvelles médiations. Notre questionnement a pris appui sur une conception des enjeux communicationnels des collections d'histoire naturelle : elles constituent un dispositif de médiation documentaire permettant l'archivage, la mise à disposition et la réappropriation par différents publics des documents essentiels à l'élaboration des connaissances. La prise en compte de la matérialité des dispositifs documentaires dans les processus propres à l'élaboration des connaissances dans les disciplines qui mobilisent les collections conduit à s'interroger sur les enjeux de leur numérisation. Notre travail s'est articulé autour d'une hypothèse de recherche qui avait aussi vocation à orienter le travail de conception. Nous avons postulé en effet la possibilité de répondre aux objectifs de valorisation scientifique et patrimoniale des collections de l'infrastructure e-ReColNat à partir de bases de données de spécimens dotées de fonctionnalités susceptibles de reconstituer et d'enrichir celles de l'herbier. Nous avons postulé également que ces bases de données de spécimens numérisés étaient capables d'amplifier le rôle des collections dans la communication scientifique. Cette hypothèse avait pour corrélat la définition d'un modèle de médiation numérique capable de reprendre les fonctions de l'herbier dans la communication scientifique, de l'élaboration au partage des connaissances avec un public non-spécialiste.

Les spécificités de l'herbier constituent déjà un modèle pour penser un dispositif de médiation documentaire qui réponde aussi à des objectifs de production multi-acteurs de données de recherche, et de valorisation du patrimoine scientifique. À partir de ce modèle, nous avons cherché à concevoir un dispositif numérique qui reconstitue des fonctions de médiation documentaire d'accès à l'information scientifique et soit aussi capable de mobiliser des publics de compétences et statuts divers dans l'enrichissement de la documentation des collections, et d'assurer ainsi des fonctions de médiation de la culture et du patrimoine scientifiques. Nous l'avons qualifié de dispositif de médiation documentaire augmentée, pour souligner le rôle central que jouent toujours les systèmes d'organisation des connaissances et plus généralement la documentation des collections dans les nouveaux dispositifs numériques de valorisation des données issues des collections : ces derniers s'appuient en effet toujours sur les informations permettant de relier les spécimens aux territoires des collectes et aux expertises dont ils ont fait l'objet. Pour reconstituer les conditions instrumentales d'exploitation scientifique du spécimen dans un laboratoire en ligne, nous avons envisagé les opérations d'interprétation et d'écriture à l'œuvre dans les activités des usagers – appréhendées notamment

dans leur dimension sensori-motrice, par la prise en compte des gestes de classement à l'origine des catégorisations de spécimens par les systématiciens. Nous avons été conduits par les besoins exprimés à enrichir la réflexion sur la médiation documentaire de l'herbier numérisé à la prise en compte des médiations de la production et de l'archivage des données des recherches issues de l'étude des documents. Dans la perspective d'encourager les collaborations dans le travail d'informatisation des collections, qui progresse grâce à des efforts locaux se nourrissant les uns des autres et constituant un travail en réseau encore morcelé, nous avons réfléchi aux conditions de la mise en place d'un nouveau projet mobilisateur fédérant des dynamiques contributives, chez les professionnels et chez les amateurs. L'enjeu de ce travail de conception est aussi la conceptualisation de modèles de médiation numérique articulant des enjeux liés à plusieurs traditions professionnelles de médiation, dont les convergences sont réinterrogées par l'accès à des ressources numériques.

L'observation de l'utilisation par les usagers des logiciels issus du travail de recherche-conception, jointe aux réactions des usagers testeurs et des responsables de l'infrastructure relatives aux spécifications et aux prototypes présentés par le laboratoire Dicen-IDF, nous ont permis de mettre à l'épreuve le modèle de médiation documentaire augmentée envisagé.

Nous avons ainsi pu comparer le modèle issu de l'analyse des médiations assurées par l'herbier avec les applications effectivement développées compte-tenu des priorités retenues par la politique scientifique de l'infrastructure, des limites des ressources de développement, et des retours d'expériences réalisés avec les prototypes. Les caractéristiques des collections d'herbier ouvraient de multiples directions possibles pour la conception. L'étape de collecte des besoins auprès d'usagers de profils variés a permis de faire ressortir la diversité des attentes, laissant augurer de multiples possibilités d'exploitation des collections numérisées à mesure de l'émergence de nouvelles méthodologies. L'enquête a fait ressortir la pluralité des points de vue adoptés sur les spécimens — systématique, ethnobotanique, historique, ou guidé par un projet nomenclatural de recension des noms d'espèces, ou encore par un travail sur l'œuvre d'un auteur ayant participé à des collectes. Finalement, les fonctions d'exploitation scientifique de l'image ont été privilégiées, en cohérence avec la politique scientifique de l'infrastructure. Les choix des acteurs décisionnaires sont autant d'observations complémentaires nourrissant l'enquête. En proposant des outils adaptés aux attentes des usagers, mais en leur laissant la responsabilité de la hiérarchisation des besoins (dans les limites des prescriptions liées à l'efficacité des développements informatiques), la participation du chercheur ne compromet pas la mise à l'épreuve des hypothèses par les observations sur le terrain. L'étude des dispositifs effectivement développés est enrichie par la connaissance de leur genèse, qui permet d'expliquer certains choix de conception, et parfois leur contingence. On ne saurait donc réduire la définition d'un modèle de médiation numérique à l'existant. On ne peut pas non plus conclure, du fait de la non-réalisation de certaines propositions de spécifications en raison de retards contingents pris par le projet, à l'impossibilité de leur réalisation ultérieure. De multiples perspectives restent donc ouvertes pour valoriser ces points de vue multiples sur les collections. Si elle n'a pas été un objectif pour le laboratoire en ligne,

la valorisation du document patrimonial est effective dans la plateforme participative « Les Herbonautes », qui utilise la diversité des approches des herbiers pour entretenir la motivation des contributeurs. De plus, les convergences en termes de besoins logiciels sont réelles entre les démarches centrées sur la biologie ou le patrimoine, qui profitent toutes deux des opérations d'inventaire et d'informatisation. À ce titre, le catalogue issu du recensement piloté par Tela Botanica, intégré à l'infrastructure e-Recolnat, constitue un module du portail répondant à cette double intention.

La démarche de recherche-conception était adaptée à un terrain caractérisé par le fort investissement des chercheurs dans l'élaboration de méthodologies qui supposent de maîtriser la matérialité des collections pour en extraire l'information, et se prolongent aujourd'hui par une appétence pour l'utilisation d'outils numériques dans le but de renouveler les recherches. L'objectif de collecte des besoins en vue de la réalisation de nouveaux logiciels instaurait une adhésion à l'enquête de la part d'usagers déjà investis dans des réflexions méthodologiques sur l'utilisation des bases de données d'images. La démarche a permis une collecte de données dans des situations variées selon les différentes étapes d'analyse des besoins, de développement et de test ; les mutations des activités mobilisant les collections selon les fonctionnalités développées ont ainsi pu être questionnées. Nous avons ainsi pu évaluer la capacité de plusieurs dispositifs numériques de médiation documentaire à reconstituer les fonctions de l'herbier à différentes étapes de l'élaboration des connaissances et de la diffusion des savoirs. Pour appréhender les formes de communication médiatisées par les collections, dans le contexte de leur inventaire facilité par les efforts simultanés d'informatisation, nous avons analysé les transactions mobilisant les spécimens, numérisés ou non, à partir des types de dispositifs documentaires utilisés, allant des *compactus* des bâtiments d'herbier aux bases de données plus ou moins ouvertes à la contribution des usagers. Les concepts de la sémiotique des transactions coopératives (Zacklad, 2013) ont permis d'analyser, dans leur dimension communicationnelle, et selon les bases de données utilisées, les activités des conservateurs et des usagers des collections. La notion de transaction articule les opérations d'interprétation et de transformation des artefacts, que nous avons aussi appréhendées comme des opérations de lecture et d'écriture qui sont intrinsèquement associées quand on produit de nouvelles connaissances. Cette notion a constitué une forme générale identifiable dans l'activité de tous les acteurs s'appropriant les documents d'herbier. Nous avons convoqué cette notion pour questionner les formes de communication entre acteurs utilisant les collections, en suivant en cela une approche historique et socialement située des phénomènes communicationnels (Jeanneret, 2008). Cela nous a permis de décrire la circulation sociale du document scientifique et l'évolution des significations qui lui étaient données, dans différentes « communautés d'interprétation » (Zacklad, 2007) composées de chercheurs de disciplines variées ou d'amateurs participant à la gestion des collections.

Les catégories de transaction constitutives de la communication médiatisée par les collections, que nous avons dégagées de nos observations, s'inscrivent dans différents registres temporels qui permettent de les structurer. L'auteur d'un document d'herbier et le chercheur qui le réutilise sont séparés par une longue distance ;

les inventaires permettent d'exhumer des collections et d'identifier des taxons méconnus, favorisant la possibilité de leur exploitation future, et ce d'autant plus quand ils se traduisent par la mise en ligne des données. Les transactions entre contributeurs, que nous avons appréhendées à partir des traces qu'elles génèrent, s'inscrivent dans une temporalité de plus court terme. Elles constituent de nouvelles modalités de production de données sur les collections, qui s'ajoutent aux champs standardisés des étiquettes, et posent la question de leur archivage au service de la communauté. L'intégration de corrections faites par les usagers affecte l'organisation de la gestion de la collection, le personnel n'ayant pas toujours le temps d'y répondre. Les solutions d'ajout d'annotations signées et standardisées, associées à une demande d'autorisation à annoter fondée sur les compétences scientifiques de l'usager, sont une réponse émergente adaptée à cette perturbation. Le travail de conservation progresse d'autant plus qu'il capitalise le bénéfice des inventaires déjà réalisés, qui accélèrent certaines recherches. L'exploitation scientifique du spécimen offre des possibilités de traitement accéléré de corpus de tailles accrues, grâce à une automatisation de l'analyse de l'image, qui ne va cependant pas sans un travail manuel et une lecture critique patiente pour vérifier les données.

Les types de transactions distingués correspondent aux caractéristiques des dispositifs qui les encadrent. La diversité des dispositifs intervenant dans les activités, selon leurs objectifs et les étapes où ils sont utilisés, est le corollaire de la présence du spécimen numérisé à différentes étapes de la communication scientifique, dans les fonctions de conservation, de production de connaissances, et de médiation du patrimoine scientifique pour d'autres publics. Au-delà de leur structure commune propre aux bases de données, qui permettent de composer des corpus de spécimens selon les besoins, grâce à la standardisation et à la fragmentation de leur description, nous avons distingué ces dispositifs selon leur manière d'intégrer des fonctions d'écriture et de coordination de l'activité collective ; nous avons ensuite évalué, sur la base de cette grille d'analyse, le degré de reconstitution des spécificités documentaires et fonctions de l'herbier. En mobilisant des typologies développées pour d'autres dispositifs du web soutenant des productions documentaires collectives, nous avons caractérisé les nouvelles formes de documentalité, c'est-à-dire les « manières de faire document » (Zacklad, 2014), propres à chaque dispositif, que nous avons pu comparer grâce à l'utilisation d'un même cadre d'analyse.

Quel que soit le public cible et les modalités d'écriture proposées par les différents types de bases de données, celles-ci exercent une fonction de dispositif de médiation documentaire en facilitant l'accès au document scientifique. Elles associent en effet à la planche d'herbier, par un numéro d'inventaire et un code-barres, un nouveau support d'enregistrement ; ce dernier restitue sur l'écran, qui devient le nouveau support d'appropriation consulté en substitution de la planche d'herbier, la possibilité d'observer le spécimen et de contrôler la fidélité de l'image à l'original (grâce au logiciel de contrôle qualité utilisé en amont de la constitution de la base de données, et à l'outil de calibrage des mesures dans le cas du laboratoire en ligne de e-ReColNat). Faire du spécimen numérisé un document suppose un travail de redocumentarisation, qui passe par une informatisation plus ou moins approfondie des données de l'étiquette permettant de localiser un spécimen

dans la collection. Au-delà de ces points communs, les bases de données conçues pour dépasser les fonctions de catalogue, comme c'est le cas des différents modules accessibles par le portail de e-ReColNat, permettent d'exploiter les données issues des collections dans des contextes spécialisés, et se différencient selon les outils qu'elles offrent pour la visualisation, l'annotation libre ou standardisée, et pour l'export des données en vue d'autres traitements. Elles reconstituent la coordination de la production documentaire collective, en particulier par les dispositifs contributifs de production de données standardisées qui réservent la libre annotation des documents à des espaces bien délimités – à savoir certains champs dont la visibilité ne dépassera pas l'échelle du contexte local de leur production (c'est le cas des commentaires laissés dans le logiciel de contrôle qualité, visibles dans ce seul module, ou dans « Les Herbonautes », même si la deuxième version de la plateforme permet la recherche plein texte dans les commentaires). La production de données standardisées passe par ces fonctions d'annotation libre. Les outils d'écriture non standardisée sont une originalité des herbiers numérisés contributifs, qui rendent exploitables certaines des données relatives au processus d'informatisation, la plupart du temps « effacées » des bases de données d'herbiers. La « mise en visibilité » du travail des « petites mains » (Millerand et Heaton, 2013) est constituée en stratégie de motivation dans la plateforme des Herbonautes (qui joue sur la valorisation des profils des contributeurs, et permet de catégoriser leurs discussions), mais il s'agit d'une visibilité locale limitée à la plateforme, qui n'atteint pas pour le moment les notices de la base de données institutionnelle.

Si les bases de données facilitent la mise en relation des planches qui partagent en commun le collecteur, l'espèce, ou les publications les ayant citées, les dispositifs de médiation documentaire augmentée par différentes fonctionnalités d'annotation, d'organisation dans l'espace de travail des spécimens ou de visualisation des données des collections introduisent une nouveauté dans l'origine multiple des données produites sur un même spécimen. Alors même qu'elles visent à centraliser les données pour faciliter l'accès à l'information scientifique, l'informatisation et la consultation des collections au sein du réseau e-ReColNat ont lieu par plusieurs interfaces, qui recréent chacune un contexte d'usage adapté à certaines activités, correspondant aussi à un projet communicationnel qui permet de rapprocher chaque dispositif d'un nouveau « lieu » documentaire (Després-Lonnet, 2014a). Augmenté de ces nouveaux dispositifs numériques, l'herbier s'inscrit dans le régime de documentalité des dispositifs « transmédiés documentaires » (Zacklad, 2015), conçus pour articuler des données provenant d'une multiplicité de terminaux, rendant essentielle la question de leur coordination, et dont les dispositifs d'écriture collaborative agrégeant des réactions de publics exprimées depuis différents réseaux sociaux ou plateformes constituent d'autres exemples. Par les divers modules donnant accès aux collections, créant autant de contextes différents pour leur consultation, le dispositif de médiation documentaire d'e-ReColNat complexifie celui des compactus et des liasses de l'espace de l'herbier. La relation entre le spécimen papier et la notice dans la base de données n'est pas encore fixée ; elle hésite entre une représentation de la base qui augmente la documentation de chaque spécimen ou qui reproduit le plus fidèlement possible les informations de la planche. L'intégration des différents contextes de production de données reste une perspective

de développement futur, puisqu'à ce jour, la plateforme laisse à l'utilisateur le soin de consulter les annotations dans les modules où elles ont été produites, sans faciliter le passage de l'un à l'autre. Si certaines passerelles pourraient être développées à l'avenir, la tension entre exploitation des collections dans un projet personnel et archivage de données standardisées reste constitutive du travail avec les collections ; elle correspond à la dialectique entre l'objet considéré sous de multiples interprétations spécialisées selon le projet individuel et l'édifice documentaire au service de tous, caractéristique des « objets-frontières » (Star et Griesemer, 1989). Les activités qui se déploient dans les différents dispositifs donnant accès aux spécimens contribuent toutefois à la connaissance globale des collections. À la fragmentation des contextes d'interprétation s'ajoute celle de la multiplicité des bases de données de spécimens en ligne, qui induit des recherches documentaires complexes avant de réunir un corpus ou d'identifier un spécimen. C'est par la diversité des activités mobilisant les collections, dans des dispositifs choisis selon les besoins, que s'élabore une connaissance progressive des collections. La centralisation des données ne sera acquise qu'au terme d'un patient processus d'informatisation et de mise en relation des planches issues de différentes institutions, processus qui suppose des règles et des dispositifs spécifiques selon le statut des acteurs qui y participent.

De l'analyse des pratiques mobilisant les collections dans le contexte de leur numérisation, et plus particulièrement de la mise en réseau des bases de données des partenaires de l'infrastructure et de la conception de nouvelles applications en facilitant l'exploitation, il ressort que la complémentarité des modules de l'infrastructure e-ReColNat répond à des enjeux de médiation documentaire, culturelle, instrumentale, et organisationnelle : les modules participatifs permettent bien de nouvelles possibilités d'implication de chercheurs et d'amateurs à l'enrichissement des collections ; les images documentées précisément alimentent de nouveaux travaux qui se fondent sur des volumes de jeux de données plus importants ; elles suscitent une réflexion sur les nouvelles possibilités d'archivage des données qui en sont issues. Leurs usages permettent de vérifier l'hypothèse d'une amplification du rôle des herbiers dans la communication scientifique : on peut constater cette amplification au niveau du nombre de contributeurs participant à la production de données, qui inclut des internautes nouvellement recrutés aux côtés des personnels impliqués dans les opérations de numérisation ; l'amplification tient aussi au volume de données mises à disposition des usagers, ainsi qu'aux nouvelles modalités d'exploitation qui s'offrent à eux, dans la continuité des méthodologies traditionnelles : tri, comparaison, recherche de proximités entre taxons sur la base de l'identification par méthodes statistiques de caractères similaires à un grand nombre de spécimens. Elle concerne aussi les échanges qu'elle rend possibles entre médiateurs scientifiques, chercheurs en position d'animer la plateforme, et amateurs.

Cette amplification est bornée toutefois par les limites rencontrées dans l'appropriation des nouvelles fonctionnalités développées : quand ces limites ne tiennent pas aux performances techniques du système informatique, elles montrent l'importance de l'accompagnement de l'utilisation des nouvelles fonctionnalités proposées. Selon notre expérience, l'investissement du chercheur-concepteur gagnerait à être prolongé dans le suivi des usagers et la communication sur les outils

développés, puisque leur appropriation par les usagers est une condition de l'efficacité que peut avoir le travail de conception sur les processus de communication qu'il cherche à mieux instrumenter. Les dispositifs numériques développés dans le cadre d'e-ReColNat ont suscité de nouvelles attentes et perspectives, en termes d'archivage des contributions scientifiques standardisées, ou de publication et de partage des données non standardisées. Ces attentes s'inscrivent dans la dynamique de la double utopie scientifique et documentaire constitutive des collections, qui saisit avec le numérique de nouveaux moyens pour se réaliser davantage. Si les réseaux informatiques ont nourri les utopies documentaires naturalistes, réciproquement, la mise en ligne des collections colore aussi les promesses dont le web est porteur. Souvent pensé comme un moyen de tracer les existences, et de générer un « pan-catalogue des individualités humaines » (Ertzscheid, 2009), le web peut aussi être le lieu d'une « vigilance participative » (Cardon et Levrel, 2009) sur l'environnement, et ses plateformes collaboratives, un moyen de documenter la biodiversité.

Bibliographie

- ALBERT-LLORCA, Marlène et GARRETA, Raphaële (2013). « L'association Pierrine Gaston-Sacaze et la patrimonialisation de la flore en vallée d'Ossau (Pyrénées Occidentales – France) ». In : *VertigO, hors-série* 16. DOI : [10.4000/vertigo.13609](https://doi.org/10.4000/vertigo.13609). URL : <http://vertigo.revues.org/13609>.
- ANDRO, Mathieu (2016). « Bibliothèques numériques et crowdsourcing : expérimentations autour de Numalire, projet de numérisation à la demande par crowdfunding ». Thèse de doctorat en sciences de l'information et de la communication. Université Paris VIII.
- APPEL, Violaine et HELLER, Thomas (2010). « Dispositif et recherche en communication des organisations ». In : *Les dispositifs d'information et de communication. Concepts, usages et objets*. Sous la dir. de Violaine APPEL ; Hélène BOULANGER et Luc MASSOU. Bruxelles : De Boeck, p. 39-57. DOI : [10.3917/dbu.massou.2010.01.0039](https://doi.org/10.3917/dbu.massou.2010.01.0039).
- BACHIMONT, Bruno (2001). « L'archive numérique : entre authenticité et interprétabilité ». In : *Archives* 32.1, p. 3-15. URL : http://www.archivistes.qc.ca/revuearchives/vol32_1/32-1-bachimont.pdf.
- BACHIMONT, Bruno (2004). « Arts et Sciences du numérique : ingénierie des connaissances et critique de la raison computationnelle ». Mémoire d'habilitation à diriger des recherches. Université de Technologie de Compiègne. URL : http://www.utc.fr/~bachimon/Livresettheses_attachments/HabilitationBB.pdf.
- BEAUDRY, Guylaine (2011). *La communication scientifique et le numérique*. Paris : Lavoisier.
- BNF (2015). *Réalisation d'une étude d'usages des utilisateurs de la plateforme expérimentale Correct*. URL : http://www.bnf.fr/documents/correct_etude_usages.pdf.
- BOS, Nathan ; ZIMMERMAN, Ann ; OLSON, Judith ; YEW, Jude ; YERKIE, Jason ; DAHL, Erik et OLSON, Gary M. (2007). « From Shared Databases to Communities of Practice : A Taxonomy of Collaboratories ». In : *J. Computer-Mediated Communication* 12, p. 652-672.
- BOTTINI, Thomas (2010). « Instrumenter la lecture critique personnelle multimédia ». Thèse de doctorat en informatique. Université technologique de Compiègne.
- BOTTINI, Thomas (2017). « “Les espaces de la critique” : une étude des conditions de possibilité d'une lecture savante et multimédia ». In : *Écrilecture augmentée dans les communautés scientifiques*. Sous la dir. de Évelyne BROUDOUX et Gérard KEMBELLEC. ISTE éditions, p. 57-70. ISBN : 978-1-78405-220-1.
- BOULANGEAT, Louise (2014). « Le recensement des Herbiers de France : un nouvel enjeu pour la connaissance ». In : *La Lettre de l'OCIM* 156, p. 12-16.

- BOULLIER, Dominique (2013). « Plates-formes de réseaux sociaux et répertoires d'action collective ». In : *Mouvements sociaux en ligne face aux mutations socio-politiques et au processus de transition démocratique*. Sous la dir. de Sihem NAJAR. Paris : Karthala, p. 37-50. ISBN : 9782811109769.
- BOULOGNE, Arlette (coord.) (2004). *Vocabulaire de la documentation*. Paris : ADBS Éditions.
- BOWKER, Geoffrey C. (2000). « Biodiversity Datadiversity ». In : *Social Studies of Science* 30.5, p. 643-683.
- BRIET, Suzanne (1951). *Qu'est-ce que la documentation*. Paris : EDIT.
- BROCHU, Danielle et DAVALLON, Jean (1998). « L'organisation du musée confrontée à l'arrivée des nouvelles technologies ». In : *Médiations sociales, systèmes d'information et réseaux de communication. Onzième congrès national des sciences de l'information et de la communication, 3-4-5 décembre 1998*. Société française des sciences de l'information et de la communication.
- BROUDOUX, Évelyne et KEMBELLEC, Gérald (2017). « Introduction à l'écrilecture scientifique et aux modalités techniques de son augmentation ». In : *Écrilecture augmentée dans les communautés scientifiques*. Sous la dir. de Évelyne BROUDOUX et Gérald KEMBELLEC. ISTE éditions, p. 19-38. ISBN : 978-1-78405-220-1.
- BUCKLAND, Michael (1997). « What is a "document" ». In : *Journal of the American Society of Information Science* 48.9, p. 804-809.
- CARDON, Dominique et LEVREL, Julien (2009). « La vigilance participative. Une interprétation de la gouvernance de Wikipédia ». In : *Réseaux* 154.2, p. 51-89.
- CASEMAJOR LOUSTAU, Nathalie (2011). « La contribution triviale des amateurs sur le web : quelle efficacité documentaire ». In : *Études de communication* 36, p. 39-52.
- CHABIN, Marie-Anne (2012). « L'Opposition millénaire archives/bibliothèques a-t-elle toujours un sens à l'ère du numérique ? » In : *Bulletin des bibliothèques de France* 5, p. 26-30. URL : <http://bbf.enssib.fr/consulter/bbf-2012-05-0026-006>.
- CHARTRON, Ghislaine (2013). « Réseaux et documentation, un lien originel à revisiter ». In : *Documentaliste-Sciences de l'Information* 50.2, p. 24-25. DOI : [10.3917/docsi.502.0024](https://doi.org/10.3917/docsi.502.0024).
- CHARVOLIN, Florian (2004). « Le programme Feederwatch et la politique des grands nombres ». In : *Développement durable et territoires [En ligne], Varia (2004-2010)*. URL : <http://developpementdurable.revues.org/687>.
- CHUPIN, Lisa (2014). « Les dispositifs de transcription comme cadre de lecture. Le cas des herbiers collaboratifs ». In : *Le Livre post-numérique, Actes du 17e Colloque international sur le document électronique, Fés, novembre 2014*. Paris : Europa Productions, p. 69-82.
- CHUPIN, Lisa (2015). « Enjeux de la numérisation des herbiers pour l'information et la communication scientifiques : de la transformation des matières documentaires à l'évolution des pratiques ». In : *Enjeux de l'information et de la communication* 16/2.
- CHUPIN, Lisa (2016). « Documentarisation participative et médiation du patrimoine scientifique numérisé. Le cas des herbiers ». In : *Études de communication* 46, p. 33-49.

- CHUPIN, Lisa (2017). « La construction de normes d'écriture pour la transcription collaborative du patrimoine numérisé : entre algorithme, transmission et élaboration communautaire ». In : *Écriture augmentée dans les communautés scientifiques*. Sous la dir. de Gérald KEMBELLEC et Évelyne BROUDOUX. Londres : ISTE Éditions, p. 89-106.
- CHUPIN, Lisa (2018a). « De la collection à la plateforme de transcription participative : enjeux des bases de données d'herbier de l'infrastructure e-Recolnat pour la communication scientifique ». In : *La reconfiguration du travail scientifique en biodiversité : pratiques amateurs et technologies numériques*. Sous la dir. de Florence MILLERAND ; Lorna HEATON et Patricia DA SILVA. Montréal : Presses Universitaires de Montréal.
- CHUPIN, Lisa (2018b). « Encadrer et animer l'enrichissement des collections scientifiques numérisées par les internautes : de nouveaux enjeux pour penser la médiation et l'action du médiateur ». In : *Tréma* 48, p. 119-131.
- CLAVIER, Viviane et PAGANELLI, Céline (2015). « Patrimoine et collections numériques : politiques, pratiques professionnelles, usages et dispositifs - Présentation du dossier 2015 ». In : *Les Enjeux de l'Information et de la Communication* 16.2, p. 5-13. URL : <http://lesenjeux.u-grenoble3.fr/2015-dossier/00-Clavier-Paganelli/>.
- CORNEY, David P. A. ; TANG, H. Lilian ; CLARK, Jonathan Y. ; HU, Yin et JIN, Jing (2012). « Automating digital leaf measurement : The tooth, the whole tooth, and nothing but the tooth ». In : *PLoS ONE* 7.8. DOI : [10.1371/journal.pone.0042112](https://doi.org/10.1371/journal.pone.0042112).
- DAVALLON, Jean (2004). « La médiation : la communication en procès ? » In : *Médiation et Information* 19, p. 37-59.
- DAVALLON, Jean (2006). *Le don du patrimoine : une approche communicationnelle de la patrimonialisation*. Communication, médiation & construits sociaux. Paris : Hermes science publications - Lavoisier.
- DESJARDINS, Julie et JACOBI, Daniel (1992). « Les étiquettes dans les musées et les expositions scientifiques et techniques ». In : *Publics et Musées* 1, p. 13-32.
- DESPRÉS-LONNET, Marie (2000). « Contribution à la conception d'interfaces de consultation de bases de données iconographiques ». Thèse de doctorat en sciences de l'information et de la communication. Université Charles de Gaulle – Lille 3. URL : <https://tel.archives-ouvertes.fr/tel-01155546v2>.
- DESPRÉS-LONNET, Marie (2014a). « L'écriture du patrimoine, de la documentation à la médiation ». In : *Documentaliste-Sciences de l'Information* 51.2, p. 61-62.
- DESPRÉS-LONNET, Marie (2014b). « Temps et lieux de la documentation : transformation des contextes interprétatifs à l'ère d'internet ». Mémoire d'habilitation à diriger des recherches. Université de Lille. URL : <https://tel.archives-ouvertes.fr/tel-01094789>.
- DILLON, Andrew (1996). « TIMS : A framework for the design of usable electronic text ». In : *Cognitive Aspects of Electronic Text Processing*. Sous la dir. d'Herre van OOSTENDORP et Sjaak de MUL. Norwood NJ : Ablex, p. 99-120. URL : https://www.ischool.utexas.edu/~adillon/BookChapters/Elec%20Text%20Proc-test_files/TIMS.htm.
- DUFRENE, Bernadette et IHADJADENE, Madjid (2010). « La confrontation des formes d'autorité : institutions culturelles et communautés d'amateurs. Médiations do-

- cumentaires : entre réalités et imaginaires ». In : *Colloque international du réseau Mussi, Avignon*, p. 79-92.
- DUNN, Stuart et HEDGES, Mark (2014). « How the Crowd Can Surprise Us : Humanities Crowdsourcing and the Creation of Knowledge ». In : *Crowdsourcing our Cultural Heritage*. Sous la dir. de Mia RIDGE. Surrey : Ashgate, p. 231-247.
- EISENSTEIN, Elisabeth (1991). *La révolution de l'imprimé dans l'Europe des premiers temps modernes*. Paris : La Découverte.
- ERTZSCHEID, Olivier (2009). « L'homme, un document comme les autres ». In : *Hermès, La Revue* 53.1, p. 33-40.
- FABRE, Isabelle (2012). « Médiation documentaire et culturelle dans le musée ». In : *Communication & langages* 173, p. 83-99.
- FLICOTEAUX, Muriel et KINOSHITA, Harumi (2014). « Innovation et réseaux à la Cité Nationale de l'Histoire et de l'Immigration et au Louvre – DNP Museum Lab : vers de nouveaux modèles de coopération et de construction des identités? » In : *Heritage and Digital Humanities : How should training practices evolve?* Sous la dir. de Bernadette SAOU-DUFRÊNE. LIT Verlag, p. 325-336.
- FRAYSSE, Patrick (2015). « La médiation numérique du patrimoine : quels savoirs au musée? » In : *Distances et médiations des savoirs* 12. URL : <http://journals.openedition.org/dms/1219>.
- FUNK, Vicky (2003). « 100 uses for a herbarium : Well, at least 72 ». In : *American Society of Plant Taxonomists Newsletter* 17, p. 17-19.
- GARDIÈS, Cécile (2014). « Enseigner la notion de document par attribution : le cas de « l'herbier » en lycée agricole ». In : *Les transformations du document dans l'espace-temps de la connaissance, actes du 3e colloque International du réseau MUSSI, 10-12 novembre 2014, Salvador de Bahia*, p. 167-186.
- GLASNER, Peter (1996). « From community to "collaboratory"? The Human Genome Mapping Project and the changing culture of science ». In : *Science and Public Policy* 23.2, p. 100-110.
- HAKLAY, Miku (2013). « Citizen Science and Volunteered Geographic Information : Overview and Typology of Participation ». In : *Crowdsourcing Geographic Knowledge*. Sous la dir. de Daniel Z. SUI ; Sarah ELWOOD et Michael GOODCHILD. Dordrecht : Springer.
- HAKLAY, Miku (2015). *Citizen science and policy : A European perspective*. Washington D.C. : The Woodrow Wilson Center/Commons Lab.
- HARRIS, Kari M. et MARSICO, Travis D. (2017). « Digitizing specimens in a small herbarium : a viable workflow for collections working with limited resources ». In : *Applications in Plant Sciences* 5.4. ISSN : 2168-0450. DOI : [10.3732/apps.1600125](https://doi.org/10.3732/apps.1600125).
- HAYTHORNTHWAITE, Caroline (2009). « Crowds and Communities : Light and Heavyweight Models of Peer Production ». In : *Proceedings of the Hawaii International Conference On System Sciences, Big Island, Hawaii*. URL : <http://hdl.handle.net/2142/9457>.
- HEATON, Lorna et MILLERAND, Florence (2013). « La mise en base de données de matériaux de recherche en botanique et en écologie. Spécimens, données et métadonnées ». In : *Revue d'anthropologie des connaissances* 7.4, p. 885-913.

- HEATON, Lorna ; MILLERAND, Florence ; CRESPEL, Élodie et PROULX, Serge (2011). « La réactualisation de la contribution des amateurs à la botanique. Le collectif en ligne Tela Botanica ». In : *Terrains & travaux* 18.1, p. 155-173.
- HEATON, Lorna et PROULX, Serge (2012). « La construction locale d'une base transnationale de données en botanique ». In : *Revue d'anthropologie des connaissances* 6.1, p. 141-162.
- HILL, Andrew ; ROBERT, Guralnick ; SMITH, Arfon ; SALLANS, Andrew ; GILLESPIE, Rosemary ; DENSLow, Michael ; GROSS, Joyce ; MURRELL, Zack ; CONYERS, Tim ; OBOYSKI, Peter ; BALL, Joan ; THOMER, Andrea ; PRYS-JONES, Robert ; TORRE, Javier de la ; KOCIOLEK, Patrick et FORTSON, Lucy (2012). « The Notes from Nature tool for unlocking biodiversity records from museum records through citizen science ». In : *ZooKeys* 209, p. 219-233. DOI : [10.3897/zookeys.209.3472](https://doi.org/10.3897/zookeys.209.3472).
- HINE, Christine (2008). *Systematics as Cyberscience*. Cambridge (Massachusetts), London (England) : The MIT Press.
- HINE, Christine (2013). « The emergent qualities of digital specimen images in biology ». In : *Information, Communication & Society* 16.7, p. 1157-1175.
- HOFF, Michel ; LAUB, Catherine ; CREMERS, Georges et DELUZARCHE, Françoise (2009). « L'herbier Casimir Le Gallo de Strasbourg (STR) ». In : *Actes du Colloque en hommage à Claude Jérôme (1937-2008), Strasbourg les 3-4 octobre 2009*, p. 47-55. URL : https://herbier.unistra.fr/uploads/media/article_6_Hoff_1.pdf.
- HOULLIER, François et MÉRILHOU, Jean-Baptiste (2016). *Sciences participatives en France. État des lieux, bonnes pratiques & recommandations*. Rapport élaboré à la demande des ministres en charge de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche.
- HYAM, Roger ; DRINKWATER, Robyn et HARRIS, David J. (2012). « Stable citations for herbarium specimens on the internet : an illustration from a taxonomic revision of *Duboscia* (Malvaceae) ». In : *Phytotaxa* 73, p. 17-30. DOI : [10.13140/2.1.2267.2008](https://doi.org/10.13140/2.1.2267.2008).
- JACOBI, Daniel (2007). « Communiquer par l'écrit dans les musées ». In : *La révolution de la muséologie des sciences ; des musées pour le 21ème siècle*. Sous la dir. de Bernard SCHIELE et Emlyn KOSTER. Lyon : Éditions multimondes.
- JEANNERET, Yves (2000). *Y a-t-il vraiment des technologies de l'information*. Ville-neuve d'Ascq : Presses universitaires du Septentrion.
- JEANNERET, Yves (2008). *Penser la trivialité I. La vie triviale des êtres culturels*. Paris : Hermes science publications - Lavoisier.
- JEANNERET, Yves (2009). « Faire avec le faire communicationnel ». In : *Actes Sémiotiques [en ligne]*. URL : <http://epublications.unilim.fr/revues/as/3336>.
- JEANNERET, Yves (2014). *Critique de la trivialité : les médiations de la communication, enjeu de pouvoir*. Paris : Éditions Non Standard.
- JUANALS, Brigitte (2002). « L'encyclopédie, des Lumières au numérique : migration d'une utopie ». In : *Communication & langages*, p. 53-65. URL : http://www.persee.fr/doc/colan_0336-1500_2002_num_131_1_3125.
- KEENEY, Elizabeth (1992). *The Botanizers : Amateur Scientists in Nineteenth-Century America*. Chapel Hill : University of North Carolina Press.
- KEMBELLEC, Gérald (2012). « Bibliographies scientifiques : de la recherche d'informations à la production de documents normés ». Thèse de doctorat en sciences

- de l'information et de la communication. Université Paris VIII Vincennes-Saint Denis.
- KIRCHOFF, Bruce K. ; LEGGETT, Roxanne ; HER, Va ; MOUA, Chue ; MORRISON, Jessica et POOLE, Chamika (2011). « Principles of visual key construction – with a visual identification key to the Fagaceae of the southeastern United States ». In : *AOB Plants*. ISSN : 2041-2851. DOI : [10.1093/aobpla/plr005](https://doi.org/10.1093/aobpla/plr005).
- KUHLTHAU, Carol Collier (2003). *Seeking meaning : a process approach to library and information services*. 2^e éd. Westport, Connecticut : Libraries Unlimited.
- LANDRAGIN, Frédéric (2011). « Une procédure d'analyse et d'annotation des chaînes de coréférence dans des textes écrits ». In : *Corpus* 10, p. 61-80. URL : <http://corpus.revues.org/2010>.
- LATOUR, Bruno (1996). « Ces réseaux que la raison ignore – laboratoires, bibliothèques, collections ». In : *Le pouvoir des bibliothèques. La mémoire des livres dans la culture occidentale*. Sous la dir. de Christian JACOB et Marc BARATIN. Paris : Albin Michel, p. 23-46.
- LATOUR, Bruno (2001). « Sol amazonien et circulation de la référence ». In : *L'Espoir de Pandore*. coll. : "L'Armillaire". Paris : La Découverte, p. 33-82.
- LATOUR, Bruno (2011). « Plus elles se répandent, plus les bibliothèques deviennent centrales ». In : *Bulletin des bibliothèques de France (BBF)* 1, p. 34-36. URL : <http://bbf.enssib.fr/consulter/bbf-2011-01-0034-007>.
- LATZKO-TOTH, Guillaume (2014). « La contribution des usagers au développement des médias numériques ». In : *La contribution en ligne. Pratiques participatives à l'ère du capitalisme informationnel*. Sous la dir. de Serge PROULX ; José Luis GARCIA et Lorna HEATON. Québec : Presses de l'Université du Québec.
- LAUTISSIER, Fanny (2009). « Les archives photographiques face aux enjeux de la transition numérique ». Mémoire de master II en histoire visuelle. EHESS. URL : <https://issuu.com/lhivic/docs/lautissier09/138>.
- LE BRAS, Gwenaël ; PIGNAL, Marc ; JEANSON, Marc L. ; MULLER, Serge ; AUPIC, Cécile ; CARRÉ, Benoît ; FLAMENT, Grégoire ; GAUDEUL, Myriam ; GONÇALVES, Claudia ; INVERNÓN, Vanessa R. ; JABBOUR, Florian ; LERAT, Elodie ; LOWRY II, Porter ; OFFROY, Bélangère ; PÉREZ PIMPARÉ, Eva ; PONCY, Odile ; ROUHAN, Germinal et HAEVERMANS, Thomas. (2017). « The French Muséum national d'histoire naturelle vascular plant herbarium collection dataset ». In : *Scientific Data* 4, p. 170016. DOI : [10.1038/sdata.2017.16](https://doi.org/10.1038/sdata.2017.16).
- LE DEUFF, Olivier (2010). « Quelles mnémotechniques pour l'Internet ? » In : *Médiation et information* 32, p. 41-51. URL : http://www.mei-info.com/wp-content/uploads/2014/02/MEI_32_03.pdf.
- LE MAREC, Joëlle (2002). « Situations de communication dans la pratique de recherche : du terrain aux composites ». In : *Études de communication [En ligne]* 25. URL : <http://edc.revues.org/831>.
- LE MAREC, Joëlle (2007). *Publics et musées, la confiance éprouvée*. Paris : L'Harmattan.
- LE MAREC, Joëlle et BABOU, Igor (2003). « De l'étude des usages à une théorie des composites : objets, relations et normes en bibliothèque ». In : *Lire, écrire, réécrire : objets, signes et pratiques des médias informatisés*. Sous la dir. d'Emmanuel SOUCHIER ; Joëlle LE MAREC et Yves JEANNERET. BPI, p. 233-299.

- LEE, Tae Kyoung ; CROWSTON, Kevin ; ØSTERLUND, Carsten et MILLER, Grant (2017). « Recruiting Messages Matter : Message Strategies to Attract Citizen Scientists ». In : *Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. Association for Computing Machinery, Inc., p. 227-230. DOI : [10.1145/3022198.3026335](https://doi.org/10.1145/3022198.3026335).
- LIQUÈTE, Vincent ; FABRE, Isabelle et GARDIÈS, Cécile (2010). « Faut-il reconsidérer la médiation documentaire ». In : *Les Enjeux de l'information et de la communication* Dossier 2010.2, p. 43-57.
- LUCZAK-ROESCH, Markus ; TINATI, Ramine ; SIMPERL, Elena ; VAN KLEEK, Max ; SHAD-BOL, Nigel et SIMPSON, Robert (2014). « Why Won't Aliens Talk to Us? Content and Community Dynamics in Online Citizen Science ». In : *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media*, p. 315-324.
- MATHET, Yann et WIDLÖCHER, Antoine (2011). *Glozz User's Manual*. URL : <http://www.glozz.org>.
- MATHIEU, Daniel et MATHEZ, Joël (2004). « La démarche et le projet Tela Botanica en faveur des herbiers ». In : *Les herbiers : un outil d'avenir. Tradition et modernité. Actes du colloque de Lyon*. AFCEV, p. 143-152.
- MAYA-LASTRA, Carlos Alonso (2016). « ColectoR, a Digital Field Notebook for Voucher Specimen Collection for Smartphones ». In : *Applications in Plant Sciences* 4.7, p. 1600035. DOI : [10.3732/apps.1600035](https://doi.org/10.3732/apps.1600035).
- McKINLEY, Donelle (2013). « Functionality and usability requirements for a crowdsourcing task interface that supports rich data collection and volunteer participation. A case study : The New Zealand Reading Experience Database ». Master of Information Studies. School of Information Management, Victoria University of Wellington. URL : <http://researcharchive.vuw.ac.nz/bitstream/handle/10063/2674/thesis.pdf?sequence=4>.
- MEYER, Morgan (2009). « Objet-frontière ou Projet-frontière ? Construction, (non-)utilisation et politique d'une banque de données ». In : *Revue d'anthropologie des connaissances* 3.1, p. 127-148. DOI : [10.3917/rac.006.0127](https://doi.org/10.3917/rac.006.0127).
- MEYER, Vincent (2006). « De l'utilité des recherches-actions en SIC ». In : *Communication et organisation [En ligne]* 30. URL : <http://communicationorganisation.revues.org/3455>.
- MEYRIAT, Jean (1981). « Document, documentation, documentologie ». In : *Schéma et Schématisation* 14, p. 51-63.
- MILLERAND, Florence (2012). « La science en réseau. Les questionnaires d'information "invisibles" dans la production d'une base de données scientifiques ». In : *Revue d'anthropologie des connaissances* 6.1, p. 163-190. DOI : [10.3917/rac.015.0201](https://doi.org/10.3917/rac.015.0201).
- MOIREZ, Pauline (2012). *Archives participatives, Bibliothèques 2.0 à l'heure des médias sociaux*. Éditions du Cercle de la librairie, p. 87-197.
- MUGAR, Gabriel ; ØSTERLUND, Carsten ; DEVRIES HASSMAN, Katie ; CROWSTON, Kevin et JACKSON, Corey (2014). « Planet Hunters and Seafloor Explorers : Legitimate Peripheral Participation Through Practice Proxies in Online Citizen Science ». In : *17th ACM Conference on Computer Supported Cooperative Work and Social Computing*. DOI : [10.1145/2531602.2531721](https://doi.org/10.1145/2531602.2531721). URL : http://crowston.syr.edu/sites/crowston.syr.edu/files/paper_revised%20copy%20to%20post.pdf.

- MÜLLER, Bertrand (2006). « À la recherche des archives de la recherche. Problèmes de sens et enjeux scientifiques ». In : *Genèses* 63.2, p. 4-24. DOI : [10.3917/gen.063.04](https://doi.org/10.3917/gen.063.04).
- NELSON, Gil ; SWEENEY, Patrick ; WALLACE, Lisa E. ; RABELER, Richard K. ; ALLARD, Dorothy ; BROWN, Herrick ; CARTER, J. Richard ; DENSLow, Michael W. ; ELLWOOD, Elizabeth R. ; GERMAIN-AUBREY, Charlotte C. ; GILBERT, Ed ; GILLESPIE, Emily ; GOERTZEN, Leslie R. ; LEGLER, Ben ; MARCHANT, D. Blaine ; MARSICO, Travis D. ; MORRIS, Ashley B. ; MURRELL, Zack ; NAZAIRE, Mare ; NEEFUS, Chris ; OBERREITER, Shanna ; PAUL, Deborah ; RUHFEL, Brad R. ; SASEK, Thomas ; SHAW, Joey ; SOLTIS, Pamela S. ; WATSON, Kimberly ; WEEKS, Andrea et MAST, Austin R. (2015). « Digitization Workflows for Flat Sheets and Packets of Plants, Algae, and Fungi ». In : *Applications in Plant Sciences* 3.9, p. 1500065. DOI : [10.3732/apps.1500065](https://doi.org/10.3732/apps.1500065).
- OLSON, Gary M. ; ZIMMERMAN, Ann et BOS, Nathan (2008). *Scientific Collaboration on the Internet*. Cambridge, Massachusetts : The MIT Press. DOI : [10.7551/mitpress/9780262151207.001.0001](https://doi.org/10.7551/mitpress/9780262151207.001.0001).
- OWENS, Trevor (2014). « Making crowdsourcing compatible with the missions and values of cultural heritage organisations ». In : *Crowdsourcing our cultural heritage*. Sous la dir. de Mia RIDGE. Surrey : Ashgate, p. 269-280.
- PARRINI-ALEMANN0, Sylvie (2004). « La recherche-action en communication des organisations est-elle une méthode qualitative constructiviste ? » In : *Recherches qualitatives, hors-série 1*. Actes du colloque Recherche qualitative et production de savoirs, UQAM, 12 mai 2004. ISSN : 1203-3839. URL : http://www.recherche-qualitative.qc.ca/documents/files/revue/hors_serie/hors_serie/texte%20P-Alemanno%20actes.pdf.
- PASSERON, Jean-Claude (2006). *Le raisonnement sociologique : un espace non popperien de l'argumentation*. Paris : Albin Michel.
- PÉREZ, Eva et PIGNAL, Marc (2013). « Numériser et promouvoir les collections d'histoire naturelle ». In : *Bulletin des bibliothèques de France* 5, p. 27-31.
- PROULX, Serge (2014). « Enjeux et paradoxes d'une économie de la contribution ». In : *La contribution en ligne. Pratiques participatives à l'ère du capitalisme informationnel*. Sous la dir. de Serge PROULX ; José Luis GARCIA et Lorna HEATON. Québec : Presses de l'Université du Québec, p. 15-29.
- PUIG, Vincent (2013). « Contribuer n'est pas collaborer : un focus sur les dispositifs d'annotation de documents audiovisuels ». In : *Numérisation du patrimoine. Quelles médiations ? Quels accès ? Quelles cultures ?* Sous la dir. de Bernadette DUFRENE ; Madjid IHADJADÈNE et David BRUCKMANN. Paris : Hermann, p. 177-190.
- RADDICK, M. Jordan ; BRACEY, Georgia ; GAY, Pamela L. ; LINTOTT, Chris J. ; MURRAY, Phil ; SCHAWINSKI, Kevin ; SZALAY, Alexander S. et VANDENBERG, Jan (2010). « Galaxy Zoo : Exploring the Motivations of Citizen Science Volunteers ». In : *Astronomy Education Review* 9.1. DOI : [10.3847/AER2009036](https://doi.org/10.3847/AER2009036).
- RASSE, Paul (2013). « Systématique et systémique, la leçon des muséums ». In : *Hermès, La Revue* 66.2, p. 66-72.
- RATINAUD, Pierre (2005). *Iramuteq [en ligne]*. Consultation en septembre 2015. URL : <http://www.iramuteq.org/>.

- RIDGE, Mia (2013). « From tagging to theorizing : deepening engagement with cultural heritage through crowdsourcing. » In : *Curator : The Museum Journal* 56.4, p. 435-450.
- ROMANUC NETO, Serge; THÉBAUD, Gilles; PIGNAL, Marc et CANHOS, Vanderlei (2013). *Herbier virtuel Auguste de Saint Hilaire*. Instituto de Botânica, Brésil; Institut des Herbiers Universitaires, France; Muséum National d'Histoire Naturelle, France; Centro de Referência em Informação Ambiental, Brésil. URL : <http://hvsh.cria.org.br/>.
- ROY, Mario et PRÉVOST, Paul (2013). « La recherche-action : origines, caractéristiques et implications de son utilisation dans les sciences de la gestion ». In : *Recherches qualitatives* 32.1, p. 129-151. ISSN : 1715-8702. URL : [http://www.recherche-qualitative.qc.ca/documents/files/revue/edition_reguliere/numero32\(2\)/32-2-roy-prevost.pdf](http://www.recherche-qualitative.qc.ca/documents/files/revue/edition_reguliere/numero32(2)/32-2-roy-prevost.pdf).
- RUSSELL, Jane M. (2001). « La communication scientifique à l'aube du XXIe siècle ». In : *Revue internationale des sciences sociales* 168.2, p. 297-309. DOI : [10.3917/riss.168.0297](https://doi.org/10.3917/riss.168.0297).
- RUZÉ, Emmanuel (2013). « La constitution et la gouvernance des biens communs numériques ancillaires dans les communautés de l'Internet. Le cas du wiki de la communauté open-source WordPress ». In : *Management & Avenir* 65.7, p. 189-205. DOI : [10.3917/mav.065.0189](https://doi.org/10.3917/mav.065.0189).
- SAHUT, Gilles (2014). « "Citez vos sources" : archéologie d'une règle au cœur du savoir wikipédien (2002-2008) ». In : *Études de communication* 42, p. 97-110.
- SALAÜN, Jean-Michel (2007). « La redocumentarisation, un défi pour les sciences de l'information ». In : *Études de communication [En ligne]* 30. URL : <http://edc.revues.org/428>.
- SARROUY, Olivier (2014). « Faire foule : organisation, communication et (dé)subjectivation à l'ère hyperindustrielle ». Thèse de doctorat en sciences de l'information et de la communication. Université de Rennes 2. URL : <https://tel.archives-ouvertes.fr/tel-01424841/document>.
- SCHÄFER, Peter A. (2004). « Identification de spécimens types dans les herbiers. Coopération européenne et internationale via Internet ». In : *Les herbiers : un outil d'avenir. Tradition et modernité. Actes du colloque de Lyon*. AFCEV, p. 111-123.
- SCHENK, Éric et GUITTARD, Claude (2012). « Une typologie des pratiques de Crowdsourcing : l'externalisation vers la foule, au-delà du processus d'innovation ». In : *Management international* 16, p. 89-100.
- SOLTIS, Pamela S. (2017). « Digitization of herbaria enables novel research ». In : *American Journal of Botany* 104.9, p. 1-4. DOI : [10.3732/ajb.1700281](https://doi.org/10.3732/ajb.1700281).
- SREEKUMAR, V. B.; HUSSAIN, K. H. et RENUKA, C. (2017). « Virtual herbarium of Kerala Forest Research Institute, Peechi, Kerala, India ». In : *Current Science* 112.3, p. 466-470.
- STAR, Susan Leigh (2010). « Ceci n'est pas un objet-frontière ! Réflexions sur l'origine d'un concept ». In : *Revue d'anthropologie des connaissances* 4.1, p. 18-35.
- STAR, Susan Leigh et GRIESEMER, James R. (1989). « Institutional Ecology, Translations, and Boundary Objects ». In : *Social Studies of Science* 19.3, p. 387-420.

- STAR, Susan Leigh et RUHLER, Karen (2010). « Vers une écologie de l'infrastructure : Conception et accès aux grands espaces d'information ». In : *Revue d'anthropologie des connaissances* 4.1, p. 114-161. DOI : [10.3917/rac.009.0114](https://doi.org/10.3917/rac.009.0114).
- STRAUSS, Anselm (1988). « The Articulation of Project Work : An Organizational Process ». In : *The Sociological Quarterly* 29.2, p. 163-178.
- SUHRBIER, Lutz ; KUSBER, Wolf-Henning ; TSCHÖPE, Okka ; GÜNTSCH, Anton et BERENDSOHN, Walter G. (2017). « AnnoSys - Implementation of a generic annotation system for schema-based data using the example of biodiversity collection data ». In : *Database* 2017, bax018. DOI : [10.1093/database/bax018](https://doi.org/10.1093/database/bax018).
- TRANT, Jennifer (2008). « Tagging, Folksonomy and Art Museums : Early Experiments and Ongoing Research ». In : *Journal of Digital Information* 10.1. URL : <http://journals.tdl.org/jodi/article/view/270/277>.
- TRICOT, André et COMTAT, Pauline (2012). « La recherche d'information comme dialogue ». In : *Les Cahiers du numérique* 8.1, p. 15-38.
- UNGER, Jakob ; MERHOF, Dorit et RENNER, Susanne (nov. 2016). « Computer vision applied to herbarium specimens of German trees : testing the future utility of the millions of herbarium specimen images for automated identification ». In : *BMC Evolutionary Biology* 16. ISSN : 1471-2148. DOI : [10.1186/s12862-016-0827-5](https://doi.org/10.1186/s12862-016-0827-5).
- VIDAL, Geneviève (2010). « Les musées et le Web 2.0 : approche méthodologique pour l'analyse des usages ». In : *Web Social, mutations de la communication*. Sous la dir. de Florence MILLERAND ; Serge PROULX et Jacques RUEFF. Québec : Presses universitaires du Québec, p. 145-158.
- VIDAL, Geneviève (2012). « La médiation numérique et les musées : entre autonomie et prescription ». In : *Développer la médiation documentaire numérique*. Sous la dir. de Xavier GALAUP. Villeurbanne : Presses de l'ENSSIB. URL : <http://books.openedition.org/pressesenssib/694>.
- VINCK, Dominique (2009). « De l'objet intermédiaire à l'objet-frontière. Vers la prise en compte du travail d'équipement ». In : *Revue d'anthropologie des connaissances* 3.1, p. 51-72.
- VISSERS, Jorick ; VAN DEN BOSCH, Frederik ; BOGAERTS, Ann ; COCQUYT, Christine ; DEGREEF, Jérôme ; DIAGRE, Denis ; DE HAAN, Myriam ; DE SMEDT, Sofie ; ENGLDOW, Henry ; ERTZ, Damien ; FABRI, Régine ; GODEFROID, Sandrine ; HANQUART, Nicole ; MERGEN, Patricia ; RONSE, Anne ; SOSEF, Marc ; STÉVART, Tariq ; STOFFELEN, Piet ; VANDERHOEVEN, Sonia et GROOM, Quentin John (2017). « Scientific user requirements for a herbarium data portal ». In : *PhytoKeys* 78, p. 37-57. DOI : [10.3897/phytokeys.78.10936](https://doi.org/10.3897/phytokeys.78.10936).
- WELGER-BARBOZA, Corinne (2013). « Les catalogues de collection des musées en ligne, au carrefour des points de vue ». In : *Numérisation du patrimoine. Quelles médiations ? Quels accès ? Quelles cultures ?* Sous la dir. de Bernadette DUFRÈNE ; Madjid IHADJADÈNE et David BRUCKMANN. Paris : Hermann.
- WIGGINS, Andrea et CROWSTON, Kevin (2011). « From Conservation to Crowdsourcing : A Typology of Citizen Science ». In : *Proceedings of the 44th Hawaii International Conference on System Science (HICSS-44)*.
- ZACKLAD, Manuel (2004). « Processus de documentation dans les Documents pour l'Action (DopA) ». In : *Le numérique : impact sur le cycle de vie du document, Uni-*

- versité de Montréal, 13-15 octobre 2004. URL : <http://www.enssib.fr/bibliotheque-numerique/document-1209,%202004>.
- ZACKLAD, Manuel (2007). « Réseaux et communautés d'imaginaire documédiatisées ». In : *A document (re)turn*. Sous la dir. de Roswitha SKARE ; Windfeld Niels LUND et Andreas VARHEIM. Franckfurt am Main : Peter Lang, p. 279-297.
- ZACKLAD, Manuel (2013). « Le travail de management en tant qu'activité de cadrage et de recadrage du contexte des transactions coopératives ». In : *Activités [En ligne]* 10.1. URL : <http://activites.revues.org/650>.
- ZACKLAD, Manuel (2014). « Humanités numériques et digitalisation de la science ». In : *Actes du XIXe congrès de la SFSIC/, Toulon, 4-6 juin 2014*. URL : <http://sfsic2014.sciencesconf.org/3185>.
- ZACKLAD, Manuel (2015). « Genre de dispositifs de médiation numérique et régimes de documentalité ». In : *Les genres de documents dans les organisations, Analyse théorique et pratique*. Sous la dir. de Louise GAGNON-ARGUIN ; Sabine MAS et Dominique MAUREL. Québec : PUC, p. 145-183.
- ZACKLAD, Manuel et CHUPIN, Lisa (2015). « Le crowdsourcing scientifique et patrimonial à la croisée de modèles de coordination et de coopération hétérogènes : le cas des herbiers numérisés ». In : *Revue canadienne des sciences de l'information et de bibliothéconomie* 39.3-4, p. 308-328.
- ZACKLAD, Manuel ; CHUPIN, Lisa ; BERTIN, Gilles et PAYEUR, Cécile (2014). « La documentarisation participative au service de la patrimonialisation des collections naturalistes ». In : *Les cahiers de la SFSIC* 9, p. 147-154.

Quatrième partie

Annexes

A. Documents présentant l'infrastructure e-ReColNat

A.1. Présentation synthétique du projet de recherche financé par l'ANR

Document mis en ligne par le ministère de l'enseignement et de la recherche pour la communication des projets lauréats pour la seconde vague des appels à projets de l'action Santé et biotechnologies, le 7 février 2012. <http://www.enseignementsup-recherche.gouv.fr/cid59286/20-laureats-pour-la-seconde-vague-des-appels-a-projets-de-l-action-sante-et-biotechnologies.html>



Infrastructures nationales en biologie et santé

PROJET E-RECOLNAT



MINISTÈRE
DE L'ENSEIGNEMENT SUPÉRIEUR
ET DE LA RECHERCHE
COMMISSARIAT GÉNÉRAL
À L'INVESTISSEMENT

INTITULE		E-RECOLNAT
COORDINATEUR		Muséum National d'Histoire naturelle (MNHN)
FINANCEMENT TOTAL		16 000 000 €
DISCIPLINE		Systématique, Botanique, Zoologie, Ecologie
DESCRIPTION		Depuis trois siècles, les collections publiques constituées en France, qui conservent une centaine de millions de spécimens au moins dont plus de 60 millions au Muséum, sont parmi les plus importantes du monde quantitativement et scientifiquement, mais sont restées négligées pendant des décennies. E-RECOLNAT propose la valorisation de 350 ans de collections d'histoire naturelle.
APPORTS POUR	LA SCIENCE	Les collections taxinomiques, herbiers et collections d'animaux et de fossiles, constituent à la fois un patrimoine scientifique, une archive de la biodiversité et une grande infrastructure de recherche. Leur informatisation dans des bases de données interopérables est indispensable pour la recherche en taxinomie, en donnant accès aux spécimens sur lesquels sont fondés les concepts des espèces actuelles et fossiles pour les actualiser. Elle est également indispensable pour comprendre et modéliser l'évolution de la biodiversité dans l'espace et dans le temps.
	LE CITOYEN	L'information gérée par E-RECOLNAT est accessible librement et gratuitement. La participation des citoyens sera en partie active : en effet, la plupart des informations procurées par les spécimens anciens demandent à être actualisées, et notamment les noms scientifiques des espèces. E-RECOLNAT propose de faire intervenir les amateurs naturalistes, qui constituent, en France et dans les autres pays européens, des réseaux de compétences et d'expertise précieux. Grâce au web, ils pourront saisir et, lorsque c'est nécessaire, corriger les données écrites accessibles par les images des étiquettes et des spécimens, qui seront ensuite validées collectivement.
	LE SYSTEME DE RECHERCHE	Les collections taxinomiques françaises sont, avec les collections britanniques, les premières en Europe, mais beaucoup sont sous-utilisées par la recherche parce que peu accessibles. L'ensemble de ces données pourra être agrégé à l'ensemble mondial et contribuer ainsi aux modélisations de la biodiversité à toutes les échelles. Avec l'infrastructure E-RECOLNAT l'ensemble de la communauté mondiale des systématiciens aura accès à cette infrastructure unique, qui leur permettra de répondre à certaines questions à distance, par l'observation des images, ou bien d'emprunter ou de venir consulter directement les spécimens qu'ils auront pu repérer sur le web.
	L'ECONOMIE	L'accès aux données de E-RECOLNAT sera public et gratuit, comme le préconisent les conventions internationales qui imposent le libre accès aux données acquises à partir de fonds publics. Au-delà des emplois créés, E-RECOLNAT contribuera à l'expertise environnementale pour laquelle les collections permettent de définir et de vérifier la répartition des espèces, tant géographiquement qu'au cours du temps. Une partie des programmes informatiques de gestion des collections est libre et gratuite, mais certains des nouveaux outils informatiques créés, en particulier pour le web 2.0, feront l'objet de licences commerciales. Les développements informatiques commerciaux peuvent permettre la création de nombreux emplois d'informaticiens hautement qualifiés.
LOCALISATION	REGION(S)	Auvergne; Bourgogne; Languedoc-Roussillon; Guadeloupe
	VILLE(S)	Clermont-Ferrand; Dijon; Montpellier; Petit-Bourg

A.2. Page de présentation du réseau botanique e-ReColNat sur le site du MNHN

Le réseau botanique e-ReColNat a été constitué en 2014 à partir du réseau des herbiers de France, ensemble d'institutions qui avaient déjà mutualisé leurs bases de données de collections avec celle du MNHN, Sonnerat.

RÉSEAU BOTANIQUE E-RECOLNAT

LE RÉSEAU DES HERBIERS DE FRANCE REGROUPE DES INSTITUTIONS NATIONALES, UNIVERSITAIRES, MUNICIPALES OU ASSOCIATIVES QUI TRAVAILLENT À L'INFORMATISATION DE LEURS COLLECTIONS SUR UN OUTIL COMMUN, LA BASE DE DONNÉES SONNERAT/BRYOMYCO. ELLE RÉSULTE DE LA FUSION DE SONNERAT CONSACRÉE AUX PLANTES VASCULAIRES (ET FUT NOMMÉE EN L'HONNEUR DU BOTANISTE PIERRE SONNERAT) ET DE BRYOMYCO CONSACRÉ AUX CRYPTOGAMES NON VASCULAIRES ET AUX CHAMPIGNONS.

GESTION DES COLLECTIONS VÉGÉTALES SÈCHES

La gestion des collections végétales sèches obéit à des pratiques internationales non écrites qui se sont forgées à la fin du XVIII^e siècle. Elle s'appuie sur la répartition des spécimens dans les herbiers. Ainsi, tous les systématiciens ont envoyé et envoient encore leurs plantes dans de nombreux établissements. Le terme de "double", consacré par l'usage, est peut-être mal choisi. Il conviendrait de parler plutôt de matériel complémentaire, issu d'un même individu (ou d'une même station dans le cas des plantes herbacées).

La comparaison avec le système en cours dans les bibliothèques est pertinente. En effet, les notices bibliographiques sont enrichies par le pool des bibliothèques qui travaillent à l'informatisation de leurs catalogues. Une notice peut être utilisée pour "exemplariser" un ouvrage présent dans plusieurs bibliothèques. Ainsi, les spécimens d'une même "récolte" partagent comme les livres d'une même édition, de nombreuses données en commun : récolteur, date de récolte, localité, descriptions écologiques etc.

Il résulte de cette pratique une histoire commune de l'ensemble des échantillons. Informatiser l'herbier de Paris ou de Nancy par exemple, c'est aussi mettre à la disposition de la communauté scientifique des données de récoltes qui intéressent de nombreuses collections.

Nombre de spécimens

L'*Index herbariorum* qui recense les herbiers du monde, avoue 23 millions de spécimens dans les collections françaises. Ce nombre est certainement largement inférieur à la réalité, car l'index est loin d'être complet. L'inventaire attentif des herbiers de la région Rhône-Alpes réalisé par Andrine Faure montre en effet que seuls 4 % des herbiers figurent dans cet index. Il est vrai cependant que "les grosses collections" y sont répertoriées.

Par ailleurs, plusieurs projets d'inventaires des collections en régions sont en cours. Citons en particulier ceux coordonnés par l'association *Tela Botanica* pour la région Languedoc-Roussillon et par le Muséum d'Aix-en-Provence pour la région PACA.

Informatiser est une entreprise de longue haleine

Un technicien bien formé peut entrer dans la base environ 80 spécimens par jour. On comprendra donc l'importance de mettre en place le plus rapidement possible les procédures qui vont permettre d'augmenter la vitesse de saisie. La mise en commun des moyens en fait naturellement partie.

Mise en évidence du matériel commun entre les institutions

Exemple des Campanulaceae

La famille des *Campanulaceae* a été choisie pour tester les récoltes communes entre Paris et Nancy. 12 % des récoltes sont communes aux deux collections. Ceci est loin d'être négligeable.

Avantages de la mise en commun des moyens

La base de données SONNERAT sous ORACLE et l'interface JACIM qui permet de l'utiliser sont les fruits d'un travail et d'échanges réguliers entre les informaticiens et les botanistes. Ceci nous a permis d'avoir un outil, qui, s'il n'est pas parfait, se révèle adapté (expérience des botanistes) et évolutif ("veille" technologique des informaticiens). Par ailleurs, les procédures de sauvegarde mises en place par le DSI (Direction des services informatiques du Muséum national d'Histoire naturelle) (y compris un stockage hors site) permettent au Muséum national d'Histoire naturelle d'assurer la sauvegarde des données. Si le Muséum national d'Histoire naturelle met ses moyens informatiques à la disposition du réseau, les moyens en développement des autres partenaires sont naturellement bienvenus.

Fonctionnement et évolution du réseau

Le réseau fonctionne sur la base d'une série de conventions bipartites entre le Muséum et chacune des institutions gestionnaires d'un herbier. Les données entrées restent la propriété de chacun, à l'exception de celles qui seront partagées entre plusieurs institutions, comme des récoltes communes, des noms de plantes etc.

L'organe qui définit les orientations est un comité de pilotage comprenant un représentant de chaque partenaire. Les demandes d'adhésion sont à envoyer à Marc Pignal (pignal@mnhn.fr).

Le comité est un lieu de discussions ouvert où les participants font état de l'avancement de leurs projets et de leurs souhaits d'évolution de la base. En 2008, le réseau s'est élargi à nos collègues du Maroc : l'institut scientifique de l'université Mohamed V de Rabat.

[Accédez à SUDOC](#)

Pour mieux comprendre SUDOC : [article Wikipédia](#)

Notes bibliographiques

Faure A., Bange C., Barale G., Danet E., Dutarte G., Fayard A., Guignard G., Pautz F., Poncet V. & Ronot P. 2006a - *Herbiers de la région Rhône-Alpes* - 1ère partie - Bignon, J. (éd.), *Herbiers de la région Rhône-Alpes*, 89-99.

**B. Tableau récapitulatif des tâches
entreprises pour l'infrastructure
e-Recolnat et des situations
d'observations de l'enquête**

Récapitulatif des activités et observations réalisées dans le contexte de l'infrastructure e-Recolnat au sein du WP5 (Lisa Chupin, Dicen-IDF, CNAM).

Les activités d'enquête d'usages et d'assistance à la conception pour l'infrastructure e-Recolnat décrites ci-après ont été réalisées dans le cadre d'un doctorat financé par un contrat d'ATER à mi-temps au Cnam, du 1er septembre 2013 au 31 août 2017, après l'obtention du Capes de documentation en juin 2013 et d'un master "Documentation et médiation des savoirs" obtenu à l'Université Lille III en septembre 2013.

Période	Principales activités dans le projet	Principales situations d'observations
2 ^e semestre 2013	<ul style="list-style-type: none"> ▶ Bibliographie orientée sites de sciences participatives et herbiers numériques ▶ Conception d'une méthodologie pour la recherche-action ▶ Conception d'un questionnaire en ligne sur le site Les Herbonautes 	<ul style="list-style-type: none"> ▶ Septembre 2013 : Réunion WP1,WP2,WP5 (MNHN) <p>Ajout d'un objectif dans les livrables du laboratoire Dicen-IDF : logiciel de contrôle qualité des images issues de la numérisation</p> <ul style="list-style-type: none"> ▶ Réunions de conception au sein du WP5, avec le partenaire Agoralogie ▶ Décembre 2013 : Réunion WP5, Tela Botanica <p>Initiation du questionnaire en ligne pour les Herbonautes avec Tela Botanica</p>
1 ^{er} semestre 2014	<ul style="list-style-type: none"> ▶ Enquête sur les usages des collections d'histoire naturelle et des bases de données sur la biodiversité (poursuivie jusqu'en 2015) ▶ Entretiens et observations filmées en situation, auprès de chercheurs et de conservateurs de collections partenaires de l'infrastructure Recolnat 	<ul style="list-style-type: none"> ▶ Entretien responsable d'un projet de sciences participative du MNHN sur la flore, janvier 2014 ▶ Entretien coordinateur du projet, janvier 2014 ▶ Participation à un colloque sur la valorisation scientifique des herbiers (Montpellier, Cirad) ; participation à la rencontre nationale des animateurs de Sauvages de ma rue, Tela Botanica (Montpellier) ; ▶ Entretiens à Tela Botanica : mise en commun des idées sur les attentes du laboratoire en ligne et de la V2 des Herbonautes ▶ Participation à la journée de lancement de l'infrastructure e-Recolnat (Montpellier, janvier 2014) ▶ Entretien avec la responsable des collections historiques et de leur valorisation au MNHN ▶ Entretien avec une chercheuse écologue ayant dirigé des recherches sur l'histoire d'une collection toulousaine ▶ Entretiens à la société linnéenne, au jardin botanique et à l'herbier universitaire de Lyon ▶ Entretiens avec des chercheurs écologues (1) et systématistes (3) au MNHN ▶ Entretien avec un mycologue à l'herbier de l'université de Lille

Période	Principales activités dans le projet	Principales situations d'observations
2 ^e semestre 2014	<ul style="list-style-type: none"> ▶ Analyse des retours du questionnaire en ligne sur Les Herbonautes et des entretiens et questionnaires complémentaires avec des contributeurs réguliers ▶ Synthèse des retours d'usage des contributeurs des Herbonautes et des situations de démonstration ▶ Rédaction de spécifications fonctionnelles pour la version 2 des Herbonautes avec les ingénieurs du WP5 	<ul style="list-style-type: none"> ▶ Octobre 2014 : démonstration, test et retour des usagers découvrant la plateforme lors de manifestations scientifiques (220 ans du CNAM, fête de la science au muséum) ▶ Entretiens complémentaires à Tela Botanica en vue de la version 2 des Herbonautes. Entretiens à l'herbier de Rabat (à l'occasion du colloque CIDE, novembre 2015). ▶ Décembre 2014 : réunion avec le WP1 (MNHN) pour prioriser les fonctions proposées pour le cahier des charges de la version 2 des Herbonautes
1 ^{er} semestre 2015	<ul style="list-style-type: none"> ▶ Travail de conception du « collaboratoire » avec les ingénieurs du WP5 ; élaboration des spécifications fonctionnelles à partir du retour de l'enquête et de l'analyse des différentes activités et gestes mobilisant les herbiers 	<ul style="list-style-type: none"> ▶ Juin 2015 : Réunion de travail au CNAM en vue de l'intégration du logiciel Xper et du laboratoire en ligne ; réunion avec le prestataire de développement de la version 2 des Herbonautes ▶ Juin 2015 : Démonstration et test de la plateforme des Herbonautes lors du festival Futur en Seine au CNAM
2 ^e semestre 2015	<ul style="list-style-type: none"> ▶ Test et recettage pour la V2 de la plateforme des Herbonautes. Rédaction d'une vingtaine de « tickets » pour les développeurs. 	<ul style="list-style-type: none"> ▶ Participation à la journée de formation sur la recherche de type organisée par Tela Botanica (octobre 2015, Montpellier). ▶ Entretiens complémentaires avec le personnel de l'herbier de l'université et un usager de l'herbier responsable d'un conservatoire botanique national. ▶ Entretien avec un contributeur régulier des Herbonautes ▶ Entretien avec un conservateur de cryptogamie (MNHN) ▶ Entretien avec le conservateur de l'herbier de Strasbourg
1 ^{er} semestre 2016	<ul style="list-style-type: none"> ▶ Fin du travail de recettage (version 2 des Herbonautes) ▶ Suivi des retours d'usagers testeurs sur le laboratoire en ligne à mesure des développements 	<ul style="list-style-type: none"> ▶ Participation à l'un des ateliers de test (CNAM, mars 2016) ▶ Entretien avec l'un des responsables de la gestion des données de l'infrastructure e-Recolnat
2 ^e semestre 2016		<ul style="list-style-type: none"> ▶ Entretien complémentaire avec un chercheur (MNHN) sur l'utilisation du logiciel Xper en vue de son intégration aux outils e-Recolnat ▶ Participation aux réunions au MNHN en vue de l'adaptation du laboratoire en ligne aux besoins d'une méthodologie de thèse en écologie
1 ^{er} semestre 2017	<ul style="list-style-type: none"> ▶ Réalisation de maquettes pour des développements ultérieurs du laboratoire en ligne ▶ Réunion WP5 sur la priorisation des demandes de corrections et d'amélioration du laboratoire en ligne 	<ul style="list-style-type: none"> ▶ Entretien avec la doctorante en écologie ayant testé le laboratoire en ligne ▶ Présentation du portail web e-Recolnat et des différents modules à l'équipe de l'herbier Marie Victorin à Montréal (à l'occasion d'un colloque de l'ACFAS organisé par Florence Millerand et Lorna Heaton).

C. Documents de conception internes du WP5

C.1. Synthèse de l'enquête sur les pratiques mobilisant les collections en préalable autravail de spécification avec les ingénieurs du WP5

Instrumenter l'exploitation des collections naturalistes numérisées. Bilan sur les pratiques et propositions en vue d'un herbier numérique collaboratif.

Les collections d'histoire naturelle sont le fruit du travail de multiples acteurs agissant avec des objectifs et pratiques diverses depuis plusieurs siècles. La numérisation offre de nouvelles perspectives pour la valorisation et l'exploitation de ces collections et leur ouverture pour un public plus large que celui des spécialistes ayant accès aux herbiers physiques. Concevoir une interface de consultation et de travail sur les collections numérisées suppose de connaître les pratiques et besoins des différents acteurs amenés à utiliser les collections (partie 1), pour outiller le travail avec les images numérisées et fédérer une communauté de pratiques autour de l'utilisation des collections (partie 2).

Partie 1 : Acteurs, documents et usages des collections

Les acteurs de la gestion, de l'exploitation et de la valorisation des collections

Les collections sont entretenues et utilisées par une grande diversité d'acteurs qui interviennent à différents niveaux dans les opérations présidant à la constitution des collections et à la production des savoirs. La division entre amateurs et professionnels est difficilement applicable dans le contexte botanique où les savoirs sont détenus par des non-professionnels. Ce serait plutôt la division personnel de l'institution responsable des collections / usager qui serait pertinente, les deux catégories intervenant en complémentarité. En effet, les usagers participent à la valorisation des collections par l'actualisation des spécimens, la recherche de type, l'envoi de publications aux institutions, la collecte de spécimens. Le personnel rattaché aux institutions contrôle la participation des usagers en leur autorisant l'accès aux collections, mais sans mettre en doute l'expertise exercée par les usagers autorisés. Les conservateurs peuvent être des chercheurs chargés de la conservation des collections. Ce sont les premiers usagers des collections dont ils ont la charge et d'autres collections numérisées.

La division des domaines de recherche mobilisant les collections

On distingue parmi les acteurs de l'exploitation scientifique des collections :

- systématiciens (étude des spécimens des collections pour décrire une nouvelle espèce ou proposer un réagencement de la classification de certaines espèces)
- taxonomistes (recherche des noms anciens donnés aux espèces *et leurs relations* dans l'histoire de la discipline – exemple de relation : synonymie)
- écologues (suivi de l'évolution d'une espèce dans l'espace et le temps de sa répartition)
- généticiens, autres chercheurs venant utiliser ponctuellement l'herbier comme « matériel » pour prélèvement
- histoire des collections
- chercheurs dans d'autres disciplines : l'herbier peut être consulté pour des recherches non directement liées à la botanique (par exemple, les collections de Jean-Jacques Rousseau).

Ces domaines de recherches peuvent faire intervenir des amateurs. Nous avons rencontré dans notre enquête davantage de chercheurs professionnels – à cause de la priorité donnée au collaboratoire pour le CNAM – mais aussi quelques amateurs (un instituteur taxonomiste ; une spécialiste amateur d'un collecteur de sa région).

La diversité des acteurs participant à la gestion et à la valorisation des herbiers

Participent à la gestion et à la valorisation des herbiers :

- des professionnels (ingénieurs, techniciens) participant à la constitution des collections : aide à la détermination des spécimens récoltés sur le terrain pour les insérer dans l'herbier
- des chercheurs chargés de conservation au MNHN (activités de recherche + répond aux besoins des collègues pour une partie des collections sous leur responsabilité)
- conservateurs curateurs sans activité de recherche liées aux collections (mais impliqués toutefois dans la rédaction de bilans d'activité retraçant l'histoire des collections)
- amateurs, retraités restaurant les collections (restauration matérielle + recherche taxonomique ou historique)
- techniciens participant à des programmes de numérisation et recherche de types
- chargés de communication scientifique (médiation patrimoniale)
- personnels d'institutions détenant des herbiers dans leurs collections : conservateurs de bibliothèques, archivistes.
- étudiants (participent à l'inventaire d'une collection, à la mise à jour des noms attribués aux spécimens, à la vérification de leur détermination, à la recherche de types...)

Usagers des collections physiques et numériques

On compte :

- des personnes réalisant des activités de recherche : chercheurs, professionnels ou non, dans différents domaines.
- des conservateurs des collections cherchant dans d'autres collections des informations utiles pour les leurs
- des enseignants utilisant les collections comme support de cours (enseigner les différentes espèces et déterminations)
- des professionnels travaillant à partir de l'exploitation de collections. Par ex, botanistes des conservatoires nationaux, qui utilisent l'herbier comme une aide à la détermination des espèces dans des cadres d'inventaires

commandés par les bureaux d'étude, ou comme source de données d'inventaires utiles pour la réalisation de politiques de protection d'espaces fragiles.

- le grand public lors des visites des herbiers

Usages et non-usages des collections en ligne

Les usages des portails donnant accès à des collections numérisées sont notamment :

- la recherche d'informations utiles à d'autres collections (types, collecteurs)
- la simple préparation de la visite ou du prêt de matériel
- la recherche des spécimens types détenus par d'autres institutions
- la réalisation d'observations des images en ligne : mesures de spécimen (2 chercheurs observés) ; recherche de planches ayant appartenu à un botaniste et des noms qu'elles comportent.

Il y a aussi des non-usages : crainte chez les chercheurs qui travaillent sur les collections physiques d'une perte de l'autonomie de l'herbier physique et du changement de son statut documentaire + mise en doute de la possibilité de faire des déterminations ayant valeur scientifique à partir des seules images.

Usagers des sites de transcription collaborative

À partir du questionnaire en ligne sur les Herbonautes on peut dire que les participants sont intéressés par « la nature ». Une partie n'a pas de connaissance botanique, mais d'autres pratiquent la botanique, comme loisir ou de manière plus ou moins directe dans leur profession. Retraités, notamment professeurs de SVT à la retraite, sont bien représentés parmi les herbonautes, comme c'est le cas parmi les bénévoles participant à la gestion des herbiers physiques. On note l'implication de la communauté professionnelle sur la plateforme, notamment pour avoir accès à des planches pas encore transcrites sur les sites des institutions, mais aussi pour participer à la documentation des collections sans passer par le personnel de l'herbier, de manière plus autonome.

Caractéristiques du système documentaire des herbiers dans le contexte numérique

Les collections d'une institution sont souvent divisées entre un herbier général classé en genres et espèces, et des herbiers historiques. Ces derniers sont « figés », cela signifie qu'on ne va pas classer les spécimens qu'ils contiennent avec les autres de la même espèce. Ils peuvent contenir des types, des spécimens appartenant à la collection d'un autre botaniste, qui sont « prisonniers » dans l'herbier historique. Le fait de numériser les collections doit permettre d'exploiter les collections par différentes entrées pour faire tomber la contradiction entre différentes logiques de classements : un classement historique, un classement botanique, un classement par région ou par récolte. Ces différentes logiques de classements correspondent à différentes activités scientifiques. L'évolution des différents classements des herbiers reflète en partie l'évolution des rapports de force entre ces différentes approches scientifiques.

Consulter les collections par des entrées comme les usages, les récolteurs ou les récoltes, différentes des entrées que propose un herbier classé selon une classification systématique, prend tout son sens en ethnobotanique. La recherche en ethnobotanique s'intéresse à l'unité de la collecte d'un récolteur, qui va permettre de rassembler des informations botaniques sur les spécimens récoltés, mais aussi sur les savoirs locaux et usages des plantes observés dans cette collecte. C'est l'unité de la récolte ou de la région qui fait sens. Les collections d'ethnobotanique physiques restent pour le moment séparées de l'herbier général car ces collections sont rangées par « boîtes » conservant des objets divers.

Si la base de données permet différentes entrées dans la collection par différentes requêtes possibles, cela suppose que les différentes informations associées à une planche soient entrées, et que la structure de la base permette par différentes tables la conservation des différentes dimensions informatives d'une planche d'herbier. Dans Sonnerat, l'interopérabilité de la collection avec des bases bibliographiques, ou d'autres collections, a été avant tout recherchée. Pour cela, on a besoin d'associer chaque spécimen de la collection aux noms scientifiques pris pour référence de manière consensuelle. Ce qui fait référence est un référentiel, mais ce n'est pas forcément la validité scientifique. De plus, si l'on associe seulement la planche au nom valide dans le référentiel utile, on perd toutes les informations historiques sur la planche (les dénominations antérieures), et on efface la diversité des expertises. L'enjeu de l'interface e-Recolnat est de réconcilier les approches des collections des scientifiques qui privilégient la consultation des collections par référentiel pour l'interopérabilité et l'approche plus historique. Il faudra pour cela associer à chaque planche d'herbier :

- un nom valide pour un référentiel choisi, (voire plusieurs référentiels choisis ? Il reste à préciser quels sont les référentiels choisis)
- les différents noms donnés au spécimen de la planche historiquement.
- il serait également très intéressant d'intégrer des informations permettant la mise en correspondance des informations historiques et actuelles.

Une précision s'impose. Il faut distinguer les informations associées au spécimen lui-même de celles associées à l'espèce à laquelle il est rattaché. Pour permettre l'interopérabilité des collections, il faut associer les spécimens à des noms correspondant à un consensus actualisé matérialisé par un référentiel. Les relations (synonymies / nom valide - nom ancien) entre différents noms sont gérées au niveau des noms associés à des spécimens (taxons), et non au niveau des spécimens. Cette distinction est nécessaire aussi pour voir comment réaliser le projet d'associer des informations aux usages des plantes : les usages ne portent pas sur un spécimen, mais sur son espèce. *Si on veut permettre aux internautes d'ajouter des informations relatives à l'espèce, il faut distinguer les informations sur le spécimen de celles qui portent sur le lieu ou le récolteur (communs à différents spécimens).*

Pour les usages et les noms d'espèces, il peut y avoir recoupement entre un territoire et une information : par exemple il faut pouvoir signaler un usage d'une plante propre à telle région, ou une dénomination vernaculaire d'une espèce, propre à tel pays (langue du pays) ou région (langue régionale). *Ci-après, on expliquera aussi l'intérêt de pouvoir associer des informations morphologiques (caractères), tantôt à une espèce, tantôt à un spécimen, pour faciliter la recherche dans la base et son utilisation pour la détermination de l'espèce de nouveaux spécimens. Il faut aussi sans doute (c'est à préciser) une entité (table?) récolte* : il faut pouvoir signaler que ces trois spécimens-là ont été collectés par le même botaniste, à la même date, au même endroit : c'est la même récolte. Cela permet de mutualiser des informations de récoltes entre différentes planches.

Autres outils documentaires associés à la consultation des herbiers

L'herbier est utilisé avec d'autres documents et systèmes documentaires.

Carnets de récolte

L'herbier en lui-même perd de sa valeur si l'institution n'est pas en possession du carnet de récolte qui contient des informations plus précises sur le contexte de la collecte (lieu, écologie) que l'étiquette de la planche. *Il est à prévoir dans e-Recolnat de numériser les carnets de récoltes détenus par les institutions et d'associer spécimens et carnets de récoltes*, comme c'est le cas sur l'herbier virtuel Saint-Hilaire fait par Marc, Simon et quelques autres : <http://hvsh.cria.org.br/caderno?a1>. C'est particulièrement important pour les institutions qui détiennent des collections mais pas le carnet de récolte (l'herbier de Rabat lorgne sur les carnets de récoltes vraisemblablement laissés à Montpellier).

Outils de recherche des spécimens pertinents.

Notamment :

- sites référençant des collections (COEL, JSTOR, Missouri Botanical Garden)
- portails en ligne de collections numérisées utilisés pour préparer la visite (rechercher les spécimens du récolteur, de l'espèce, de la région qui intéressent l'utilisateur). Enrichir la base de donnée de la collection peut aussi être le but de l'activité de consultation de l'herbier.
- mails : échanges avec les gestionnaires pour avoir accès aux collections (autorisation + détermination de l'offre de l'institution correspondant aux besoins de l'utilisateur)
- bases de données taxinomiques (pour chercher les correspondances de l'appellation ancienne ou nouvelle et accéder aux spécimens des collections dont les noms n'ont pas été mis à jour, et pour les mettre à jour). La consultation de l'herbier peut aussi avoir pour objectif de compléter ces bases taxinomiques (activité des taxonomistes).

Outils de recherche dans la bibliographie scientifique :

- les bases de données bibliographiques : pour consulter les publications scientifiques qui mentionnent certains spécimens des collections. La diagnose est la description d'une espèce. Le protologue mentionne les spécimens de collections ayant servi à la description (holotypes), avec photos et schémas associés.
- les ressources mises en ligne par les bibliothèques spécialisées, bulletins des sociétés savantes en partie numérisés
- les bibliothèques spécialisées pour consulter la bibliographie non numérisée (ex : bibliothèque de la société linnéenne de Lyon, bibliothèque du MNHN au premier étage du bâtiment de l'herbier)

Outils de détermination :

- clés de détermination, autres spécimens d'herbiers, physiques ou en ligne
- flores interactives électronique, ou flores papier. La flore concerne une région déterminée sur laquelle on travaille et dont on consulte les spécimens. Elle contient elle-même des photos, des dessins et schémas, des cartes de répartition, des clés de détermination – outil permettant de déterminer l'espèce d'un spécimen à partir d'une succession d'observations systématisées. Réaliser la flore d'une région peut être aussi l'aboutissement du travail d'exploitation de l'herbier.

Autres

- sites référençant des données d'observations (sites des conservatoires) associées à des données de collections
- les référentiels floristiques : l'herbier est utilisé pour noter les occurrences de spécimens collectés pour connaître la répartition des espèces, les noms qui leur ont été donnés. La consultation de l'herbier sert à constituer/compléter un référentiel floristique, dans lequel le chercheur consigne les occurrences correspondant aux planches qu'il consulte.
- bases de données de prélèvements génétiques (bases de mesures morphométriques réalisées sur les plantes ? À vérifier)

Documents produits par l'utilisateur

Base de données personnelles réalisées par l'utilisateur au cours de la consultation. Elle peut rassembler selon les disciplines occurrences de spécimen à un certain endroit, noms mentionnés dans l'herbier...

Les activités botaniques entre coopération et collaboration

Les collections naturalistes se développent par agrégation des travaux de collecte et d'identification des botanistes dans l'histoire. Elles ont une structure intrinsèquement coopérative (chacun réalise une tâche concourant à une production commune sans interagir avec les autres). Les classifications et référentiels communs ont pour fonction de permettre de travailler à partir des collections sans collaborer (interagir pour travailler ensemble) avec les chercheurs qui en sont à l'origine. Un systématique, qui va étudier un genre botanique, utilise la bibliographie et les spécimens collectés et décrits par les botanistes qui l'ont précédé. Les apports de ses travaux se traduisent par des annotations (étiquettes) sur des spécimens existants ou par le dépôt de nouveaux spécimens. En cela, l'herbier est le support d'une coopération scientifique distribuée à très grande échelle dans l'espace et le temps, sans qu'il n'y ait de travail

collaboratif (interactions en vue de la réalisation d'un même projet) entre chercheurs. La structure des collections reflète ces coopérations puisque les institutions possèdent des spécimens ayant été récoltés par les mêmes botanistes et répartis entre différentes collections au fur et à mesure des échanges scientifiques et des legs de collections. L'herbier physique est un document médiateur d'interactions distribuées dans l'espace et le temps à l'origine de la production des savoirs.

La numérisation demande de renforcer les efforts communs en vue du traitement des images numérisées. Mes observations attestent là encore d'une logique davantage coopérative : chacun travaille sur sa collection, ou une partie d'une collection qui lui a été attribuée. La mise en ligne des planches déjà traitées constitue une ressource commune qui va permettre au travail en cours de progresser plus facilement. Les interactions collaboratives se limitent à l'entraide entre techniciens. La collaboration concernerait davantage les personnes impliquées dans le montage des projets de numérisation qui vont réfléchir à des règles communes pour permettre l'interopérabilité entre différentes institutions. De mes observations, le travail du chercheur en lui-même est solitaire, un peu décevant pour le « collaboratoire » donc. J'ai observé un cas de collaboration de deux chercheurs du MNHN. Ils ont réalisé en commun l'ensemble des opérations de la recherche, jusqu'à la publication de l'article. Quelques interactions existent aussi sur JSTOR pour discuter le caractère de type attribué à un spécimen, mais elles sont peu nombreuses par rapport au nombre de commentaires apportant simplement une indication à l'institution détentrice des collections.

De ces constats résultent les propositions suivantes pour le collaboratoire : il s'agit de faciliter la gestion coopérative des collections à distance, ce qui suppose de repenser les droits accordés aux usagers pour leur permettre d'intervenir sur les bases de données des collections en ligne. Actuellement, les collections en ligne ne permettent que de formuler des commentaires/signalement d'erreurs qui sont intégrés par le personnel de l'herbier à la base institutionnelle. Il faut par exemple permettre à des usagers accrédités de transcrire à distance les informations de toutes les étiquettes des spécimens, et de procéder à des précisions et corrections des informations déjà transcrites. Pour permettre des interactions en direct à propos des collections il faut rendre publiques les interventions des usagers (ce qui est rarement le cas sur les sites existants), pour permettre des discussions scientifiques à propos des collections. Ces dernières pourraient être facilitées aussi par des *systèmes de communications synchrones (permettant de voir qui est en ligne)*. Il s'agit aussi de favoriser le développement de ressources communes gérées collaborativement *comme des pages ressources sur les collecteurs (comprenant des étiquettes représentatives de leur pratiques et de leur écriture)*. Il s'agit aussi de faciliter les recherches dans les ressources communes (*moteurs de recherche adaptés aux besoins de la transcription – banque d'écritures de botanistes – ou aux activités de systématique – recherche par caractère morphologique*).

La participation aux sites de transcription collaborative laisse envisager des formes de travail réellement collaboratif à partir des images des collections concernant l'identification des lieux, dates, récolteurs. Il y a bien dans ces sites de transcription collaborative des interactions pour résoudre les problèmes posés par les opérations de transcription (difficulté de déchiffrement, recherche d'information, questions relatives à l'application des règles de transcription) : on pourrait envisager ces mêmes questions à propos de la reconnaissance d'un type. Les discussions qui ont lieu sur les mailing list et forum de Tela Botanica concernant l'identification d'espèces à partir de photos en pièce jointe pourraient avoir lieu également dans l'herbier numérique. Cela pose toute de même des problèmes : *il faudrait sans doute distinguer dans la base les identifications faites à partir de la seule consultation des images de celles fondées sur un réexamen du spécimen physique.*

Licences et droits d'exploitation

Les sciences citoyennes sont souvent (mais pas toujours) associées à la production de données libres de droit. Tela Botanica fonctionne sur ce principe et milite en ce sens. Plusieurs usagers des Herbonautes exerçant par ailleurs une profession liée à la conservation de collection ou à la réalisation d'inventaires (Conservatoire Botaniques Nationaux) demandent à ce qu'il soit possible de récupérer les données produites sur les collections (export de données en vue de l'intégration à leurs inventaires – *rendant nécessaire une réflexion sur l'interopérabilité des données géographiques et botaniques*).

On observe dans les différents herbiers numérisés en ligne que la consultation des spécimens en haute définition est gratuite, mais chaque image porte un copyright de l'institution qui les détient.

Marc rappelle qu'il faut réfléchir à un financement du projet pour répondre aux contraintes de rentabilité fixées par l'ANR. Il demande une réflexion sur les apports des partenaires d'e-Recolnat à l'infrastructure sur le long terme (apports financiers, ou en matériel, ou en personnel). Il envisage que le téléchargement d'un nombre important d'images soit soumis à paiement. L'idée est en tous les cas d'empêcher la réutilisation commerciale des données e-Recolnat sans contrepartie à l'infrastructure.

+ *On pourrait envisager une page « crowdfunding » dans le portail : dons pour les collections, je peux parrainer x plantes collectées dans telle région... (possibilité de versement en ligne).*

+ *question : peut-on utiliser des données « CC BY SA » (par exemple intégrer les infos sur les usages associés aux différentes espèces de plantes recensées dans le wiki Plantuse) si le site e-Recolnat comprend des formes d'exploitations payantes ?*

Analyse des principales activités en jeu dans la consultation et l'exploitation des collections

L'étude des usages des bases de données d'images d'herbiers existantes nous conduit à décrire plus particulièrement les activités dont les collections numériques sont le support et qu'il s'agit d'instrumenter au mieux.

Recherche de spécimens répondant à différents critères

Recherche par espèce, par entité géographique, par catégorie de spécimen (recherche dans les types). Ce sont des recherches par les chercheurs, à des fins d'étude (révision d'un genre par exemple), par les conservateurs investis dans le récolement de leurs propres collections, par les conservateurs et gestionnaires de collections pour répondre aux demandes de prêts et aux besoins des visiteurs. Il faut aussi envisager les recherches des usagers des plateformes de transcription, et permettre des requêtes par un public néophyte, même si l'entrée publique passe plutôt par la navigation. Les recherches dans les collections numériques sont couplées à des recherches dans des bases taxinomiques et bibliographiques.

Signalement des spécimens « types »

La valorisation d'une collection passe en partie par la connaissance des spécimens de référence qu'elle contient, car ce sont ceux qui sont le plus demandés. La plupart des institutions ayant numérisé leurs collections ont procédé à une recherche de leurs « types » avant la numérisation. Cela ne se passe pas ainsi pour toutes les institutions partenaires d'e-Recolnat. Il faut donc prévoir que la recherche de types aura lieu à partir des images.

La recherche de types faite par les membres d'une institution à des fins de valorisation de la collection diffère de la recherche de types par les chercheurs : le personnel des institutions ne cherche pas les types d'une espèce particulier, ou un auteur particulier (qui est l'objet de recherche du chercheur), mais cherche les types de toute la collection. Les chercheurs, quand ils consultent et cherchent les types nécessaires à leurs travaux, trouvent aussi des types non encore signalés et peuvent concourir ainsi à la valorisation des collections en contribuant au recensement des types qu'elles contiennent, qui est plutôt la mission des conservateurs et du personnel engagé à cet effet. La recherche des types est l'occasion de coopération entre différents acteurs, mais aussi d'interactions comme on le voit dans les commentaires de JSTOR où des spécialistes viennent critiquer le statut de type qui a été associé par un technicien, ou un conservateur, à un spécimen.

Quel que soit leur objectif, les opérations de recherche de type se rejoignent. C'est leur point de départ qui diffère le plus. On peut soit partir des publications et rechercher dans les liasses correspondant aux espèces décrites si une planche pourrait être le type utilisé pour la publication, ou un spécimen collecté lors de la même récolte que celle ayant servi à la description (isotype). C'est plutôt ainsi que procèdent les chercheurs. Quand il s'agit d'examiner toute une collection, on peut aussi parcourir les spécimens liasse par liasse et se demander systématiquement s'il peut s'agir d'un type.

Les fins de recherche et de valorisation des collections se rejoignent dans l'opération consistant à examiner une planche, ou son image, pour déterminer son rapport avec le spécimen utilisé pour décrire l'espèce. Différents documents sont consultés dans cette opération :

- la planche physique ou son image dont on questionne le statut de type
- les autres types déposés pour l'espèce en question.

- le protologue : la publication dans laquelle le nom scientifique du spécimen de la planche en question a été publié.

L'opération consiste à comparer la date de la planche que l'on examine avec la date de la publication du nom scientifique qu'elle comporte, et celle des autres types déjà connus. Elle a pour préalable le déchiffrement de l'étiquette des différentes planches en question, une compréhension minimale des publications (parfois en latin).

Nombreuses de ces publications sont numérisées. Certaines sont détenues par les bibliothèques des herbiers. Certaines planches associent une copie du protologue (description de l'espèce) au spécimen. *Ce lien pourrait être systématisé dans l'interface e-Recolnat, avec le lien du spécimen type vers la publication.* Ce lien existe déjà dans certaines bases. Trouver la publication et vérifier qu'elle mentionne bien un spécimen détenu par l'institution (ou ayant un rapport avec lui) requiert une recherche d'information et une lecture du document. Ce n'est pas une opération automatisable. En revanche, elle peut être prise en charge par des personnes formées à ces opérations. L'examen du contexte de la collecte du spécimen à partir de son image, auquel sont rompus les herbonautes réguliers, est une étape essentielle de la recherche de types. *On pourrait alors ouvrir la recherche de types comme une étape supplémentaire (ultime) des opérations participatives de transcription proposées aux internautes. On pourrait en tous les cas proposer cette recherche de type aux personnes ayant suivi une formation adéquate ou accréditées.*

Il serait souhaitable de réaliser un tutoriel sur la recherche de types, pourquoi pas même des sessions de formation dans les institutions partenaires, et pourquoi pas un MOOC. La session de formation à la recherche de type pourrait même avoir lieu lors d'un rassemblement physique des herbonautes et être associé à la visite de l'herbier du MNHN qu'envisage Marc.

Transcription d'étiquettes

Étant donné le volume des collections, le personnel des institutions n'a pas retranscrit toutes les informations utiles sur une planche. Les usagers des collections sont amenés à déchiffrer des étiquettes ; il pourrait leur être proposé de transcrire au fur et à mesure qu'ils parcourent des collections et lisent des spécimens. Les transcriptions sont faites par les usagers en fonction de différents objectifs. Il s'agit donc de proposer une *interface de transcription modulable*, pour permettre à la fois la mutualisation des informations, ainsi que des annotations personnalisées qui n'obéissent pas forcément à des champs transcrits habituellement et ne suivent pas de vocabulaire contrôlé.

La principale question posée par la transcription est celle des droits : une étiquette mal transcrite peut rendre introuvable un spécimen car le lieu de sa collecte aura été mal renseigné, ce qui l'exclurait des résultats d'une recherche portant sur les spécimens de la région où il a été collecté. Notons que même quand elle est réalisée par des professionnels, la transcription peut susciter des erreurs. Il faut donc rendre les corrections de transcription possibles.

Opérations de catégorisations de spécimens

C'est une opération préalable à la révision d'un genre qui va conduire à le diviser en espèces différentes de celles qui existent pour le moment. Le chercheur va rassembler un grand nombre de spécimens appartenant aux différentes espèces du genre qu'il révisé. Il va regarder les images des spécimens du genre qui l'intéresse dans les différentes bases de données d'herbier, et se procurera sans doute une partie des spécimens physiques par prêt. Si, pour la majeure partie des chercheurs rencontrés, la révision taxonomique d'un genre nécessite l'examen des spécimens physiques, notamment parce que les critères distinctifs entre différentes espèces vont nécessiter analyse microscopique, prélèvement, ou dissection, certains pensent possible de réaliser en ligne certaines opérations en vue de la révision. C'est la demande principale de Marc, qui souligne l'analogie qu'il y a entre le travail qu'il peut faire en classant des images d'herbiers, et celui qu'il ferait avec des planches physiques. La description suivante résulte de la façon dont il envisage les opérations de systématique en ligne telles qu'il les a déjà réalisées.

Le chercheur va partir de la consultation des images des spécimens appartenant au genre auquel il s'intéresse. Il va les trier en fonction de différentes caractéristiques. Il va notamment procéder en regardant certains caractères morphologiques, qui vont varier ou être identiques selon les spécimens. Une espèce se définit par un ensemble de caractères distinctifs. Une même espèce connaît aussi une variation morphologique qui fait sa diversité. Le systématicien a donc besoin d'annoter chaque image, par tags, ou par champs prenant différentes valeurs (les états d'un caractère selon les spécimens consultés). Il va ensuite chercher à construire des catégories à partir des images annotées : il va regrouper celles qui contiennent tel et tel tag, ou celles auxquelles il a attribué telle ou telle valeur de caractère. Ces nouveaux regroupements sont la base pour constituer de nouvelles espèces, ou sous-espèces potentielles. Dans la base de données, le systématicien a donc besoin d'un espace où il puisse consigner ses observations – sur telle image, il y a tel état de caractère – en associant à l'image un champ (le caractère), qui pourra comprendre différentes valeurs (les états du caractère).

Il faut distinguer l'association d'un caractère à un spécimen ou à une espèce. Dans la base de données, il est aussi possible d'associer des caractères non pas à chaque image mais à une espèce : quand tout spécimen d'une espèce donnée se caractérise par un état de caractère particulier (par exemple toutes les marguerites ont des pétales (caractère) blancs (état de caractère)), on peut associer l'état blanc pour le caractère couleur des pétales à tous les spécimens portant le nom marguerite. Associer ces caractères aux espèces demande que des « référentiels » (pris au sens particulier ici de liste associant à chaque espèce d'un genre une liste de caractères) aient été constitués. Xper est un outil permettant de constituer de tels référentiels. Le fait d'associer ce type de référentiel aux espèces recensées dans la base de données, au moins pour partie d'entre elles, facilite l'utilisation de la base pour déterminer à quelle espèce appartient un spécimen que je cherche à déterminer. Elle prépare l'utilisation de la base de données comme clé de détermination : il suffirait de sélectionner par requêtes successives des états de caractères dans l'ordre prescrit par la clé pour arriver aux images qui peuvent correspondre au spécimen que je souhaite identifier.

Renseigner des caractères de manière collaborative

Il s'agit de faciliter l'utilisation des images d'herbier dans le cadre de l'aide à la détermination de l'espèce d'un nouveau spécimen collecté. Le botaniste parvient à déterminer la famille de la plante, mais il ne sait pas quel est son genre, encore moins son espèce. S'il faut une recherche par famille, il peut tomber sur des centaines d'images. Pour faciliter la recherche d'images, on voudrait pouvoir faire des recherches par genre (botanique), croisées avec un caractère : par exemple, je vais chercher tous les spécimens de « rosaceae », dont la planche comporte des fleurs, et dont les fleurs comportent des pétales de couleur bleu.

Cela suppose que d'autres chercheurs, conservateurs, voire de gentils internautes, aient préalablement annoté les planches pour une liste de caractères définis par des chercheurs comme pertinents pour faciliter leurs recherches dans les collections. Ce sont bien ici des caractères associés aux images et non aux espèces dont il est question. Mais des collaborations sont aussi possibles pour constituer des référentiels (une réalisation a été faite en partenariat avec Tela). Mais cela sortirait du champ d'e-Recolnat tant que Xper n'y est pas officiellement intégré.

Autres usages des collections, entre constat et perspectives

Sans les décrire de manière approfondie, nous rappelons les différents usages motivant la consultation d'herbiers que nous avons observés :

- *Gestion des prêts : demande de prêt en ligne, gestion des retours (lettre de rappels). L'intégration de ces fonctions au site est demandée.*
- Suivi des usagers des collections : il faut trouver comment remplacer le suivi que les institutions ont sur l'activité de consultation des collections grâce au suivi des visites physiques.
- Inventorier les collections détenues dans une région, un pays (d'où COEL intégré à e-Recolnat)
- Inventorier les occurrences d'observation d'un spécimen dans une région
- Mise à jour de la collection : les noms d'espèce associés aux spécimens doivent correspondre à l'état actuel de la nomenclature
- ... liste non limitative.

Voir Vicky Funk, « 100 Uses for an herbarium » <http://www.virtualherbarium.org/vh/100UsesASPT.html>

Aux usages existants s'ajoutent les fonctions à envisager pour ouvrir les collections à différents publics. Il faut donc faire de la base d'images numérisées un outil utile à la communauté investie dans la transcription des herbiers sans être compétente concernant les noms scientifiques. Il s'agit donc de permettre de visiter les collections autrement que par des requêtes par noms scientifiques dans la base de données. La navigation dans des collections par regroupement

thématique d'images de spécimens serait privilégiée. Il s'agit aussi de faire droit à la dimension patrimoniale des collections. Une éditorialisation participative permettrait à la fois de présenter les collections sous des angles multiples susceptibles d'intéresser un public non spécialiste et éventuellement de collecter des connaissances sur les collections entrant ou non dans la perspective académique sur les herbiers. Par exemple : ajouter des photographies, raconter une anecdote propre à un lieu de collecte – les collections entre objet de mémoire et patrimoine scientifique.

Partie 2 : Propositions pour la conception de l'interface des collections

Pages objets, et pages collections d'objets

Page « collection d'images »

Elle regroupe des informations relatives à un ensemble de planches et donne accès à ces planches présentées en miniature en bas de la page.

Tout usager (simplement inscrit – pour qu'il puisse retrouver sa collection) peut créer une page collection. Il faut donc prévoir une interface de création de la page collection. Elle peut être accessible à partir de la consultation d'une image (par navigation ou recherche) : une icône permet de créer une collection à partir de l'image consultée, ou d'ajouter l'image à une collection existante. Cette collection d'images peut être rendue publique ou gardée privée. Elle se verrait attribuer différents statuts en fonction de l'accréditation ou non de l'utilisateur à l'origine.

La page « collection d'images » donne accès à un ensemble d'images sélectionnées correspondant au résultat d'une requête dans la base de données, définie par le créateur de la page « collection d'images ». La collection peut aussi être constituée à partir d'un panier de spécimens regroupés dans un dossier de l'espace personnel de l'utilisateur, avec une fonction « créer une collection à partir de ce dossier ».

On peut choisir d'exposer sur la page « collection d'images » une carte affichant les localisations des collectes des spécimens de la collection. On peut accéder directement aux spécimens à partir de cette carte.

Sur la page « collection d'images », on peut choisir d'afficher une galerie d'images significatives de la collection. Cette galerie est un sous-ensemble de la collection. Elle peut comporter :

- une sélection par le créateur de la collection de quelques spécimens particulièrement intéressants
- ensuite par exemple on peut présenter les spécimens appartenant à la collection d'images qui ont été récemment tagués par d'autres usagers d'e-Recolnat, ou tagués « beaux spécimens ».

La galerie peut n'afficher qu'une sélection d'images ou toutes les images de la collection.

En plus de la galerie et de la carte, il y a des espaces d'édition comme ceux des blogs (facile, texte + images, liens vers des sites ressources, possibilité d'écrire en HTML pour ceux qui savent...)

On peut aussi afficher sur cette page des statistiques sur la « collection d'images » : proportion de spécimens correspondant à différents critères que l'utilisateur peut définir.

Il y a aussi possibilité de présenter les images de la collection d'images par une arborescence correspondant à une classification des images présentées. Peut-être faudrait-il une interface indépendante pour la construction de la classification, type outil de carte conceptuelle Freemind, si possible le plus simple d'utilisation possible. Par exemple, si l'exposition présente les spécimens de rosaceae de l'herbier de Paris, la page présente l'arborescence correspondant aux différentes espèces et sous-espèces de rosaceae. Mais si la collection d'images a lieu sur une thématique non directement botanique, ce pourrait être un principe de classification tout autre : éducatif, ludique, onirique, esthétique, mystique... Quand on clique sur le nœud représentant une sous-espèce, on a accès aux spécimens correspondant à la sous-espèce et à une explication des caractères distinctifs de cette sous-espèce ou sous-catégorie.

NB : Une page collection d'images peut comporter un seul élément

Page récolteur/botaniste

C'est une page « collection d'images » particulière en ce que la sélection d'images correspond à la requête : récolteur = un tel. Elle est générée automatiquement (à partir de la table des récolteurs de la base de données e-Recolnat). Elle est accessible par un onglet spécifique, en tous cas il y a un accès à l'ensemble des pages récolteurs, qui pourrait comprendre le moteur de recherche dans les pages récolteurs (par chaîne de caractère) décrit plus loin.

Elle pourrait importer des contenus à partir de bases de données de récolteurs existantes, dont certaines comportent déjà des collections d'étiquettes.

En plus des contenus produits automatiquement, elle comporte des possibilités d'édition pour compléter la vie du récolteur. Au fur et à mesure des transcriptions, les internautes associeraient des étiquettes d'herbier réalisées par le récolteur en question.

Elle comprend une carte présentant la géolocalisation des récoltes du récolteur recensées dans les collections.

Il faudrait aussi un outil de construction d'une arborescence type Freemind qui permettrait de représenter la classification botanique utilisée par le récolteur en question. Cela demande :

- d'identifier les planches déterminées par le botaniste en question
- d'identifier les noms qu'il utilise, et les éléments de classification : quels espèces il distingue dans tel genre.
- éventuellement de déterminer la relation entre le nom employé par le botaniste et le nom actuel

Le principe serait de pouvoir naviguer dans les planches d'herbiers du botaniste en question en suivant sa classification : cliquer sur une espèce de l'arborescence donnerait accès aux spécimens correspondants.

Page institution

C'est aussi une page « collection d'images » particulière. Elle est créée pour chaque institution.

Elle comprend des statistiques sur la composition des collections de l'institution : le nombre de spécimens, la proportion des spécimens selon les genres, selon les récolteurs, les régions de collecte...

Elle comprend une galerie d'images des principaux collecteurs représentés dans les collections de l'institution, chaque image renvoyant à la page collecteur.

Elle comprend une galerie d'images sélectionnées par le personnel de l'institution.

Elle donne accès aux images de l'institution par navigation dans un arbre de classification.

Spécificités de la page institution

Elle comprend des contenus produits automatiquement pour chaque institution partenaire à partir de métadonnées la concernant (recensées dans COEL ???)

Les possibilités d'édition de contenus sur cette page sont réservées au personnel de l'institution.

Elle regroupe les fonctions pour entrer en communication avec les conservateurs d'herbiers, pour faire des demandes de prêt. + statistiques sur les prêts dans l'institution.

Page collections e-Recolnat

Ce serait une page rassemblant les informations et statistiques sur l'ensemble des collections e-Recolnat (proportion de spécimens par genres, récolteurs...) Agrégation du nombre de prêts dans toutes les collections e-Recolnat...

On pourrait aussi parcourir l'ensemble des collections à partir d'une arborescence représentant une classification prise pour référence (APG 3 *a priori*).

Page spécimen

Elle comprend les éléments suivants :

- Miniature cliquable du spécimen. Le clic donne accès à l'image telle qu'elle apparaît pour la consultation, en haute définition et avec des icônes (appel du collaborateur, appel de l'interface de transcription, création d'une collection d'images).

- Carte localisant le lieu de collecte du spécimen

- Liste des objets associés au spécimen avec des liens cliquables : parties de l'image, publications, liens dans les bases de données, liens vers les types relatifs aux espèces du spécimen en question.

- Les informations transcrites sur ce spécimen, notamment les différents noms d'espèce qui lui ont été associés, avec nom du déterminateur et date de la détermination. Il faudra distinguer (par un code couleur ?) les informations validées, intégrées dans R, ou seulement dans R+. Les informations venant des Herbonautes sont signalées comme telles, et il y a un lien vers la page Herbonautes du spécimen.

- les discussions des différents modules relatives à ce spécimen, classées par catégorie de module.

Pages objets

Je sais que dans le modèle Reclonat il y a des pages objets, mais je ne visualise pas tellement. Quel accès à ces pages ? Une visibilité particulière sur le site ? Un annuaire ? Ou l'intérêt est juste d'avoir une URL associée à l'objet et regrouper les discussions qui y sont associées ?

Instrumenter le travail sur les images des collections

Comparer des spécimens

Visualisation d'une sélection d'images (résultat d'une requête ou sélection manuelle par une icône « ajouter à la galerie ») en mode galerie pour faciliter la comparaison de spécimens. Nombre variable d'images dans ma galerie.

Choix entre différents modes de visualisation. Notamment :

- une image seule,

- une image + son interface de transcription,

- 4 (ou 8, ou 24) images seules, 4 (ou 8, ou 24) images + leur interface de transcription minimale.

« Punaïser » un spécimen (ou plusieurs) : il s'ouvre dans une autre fenêtre le temps que je travaille ailleurs sur le site : cela me permet d'avoir mon ou mes spécimens toujours visibles dans une partie de l'écran pendant que j'ouvre d'autres images ou que je consulte d'autres sites.

Superposer des images (une idée que j'ai eue à partir de ce que m'a dit Marc, mais je ne lui ai pas soumise) : je visualise une première image sur la partie gauche de mon écran, une deuxième sur la partie droite. Je fais glisser celle de droite sur celle de gauche : par transparence, je vois encore celle de gauche, et je peux comparer ainsi les 2.

Zoom et ciseaux

Quand je zoome, il y a une miniature de la totalité de l'image qui s'affiche et on voit la partie du spécimen sur laquelle je zoome, matérialisée par une zone grisée. Quand je déplace la zone grisée sur la planche en miniature, je circule sur la planche avec le zoom.

Je peux sélectionner une partie de mon image, la faire tourner (utile quand une étiquette n'est pas collée dans le bon sens). Faire tourner le spécimen entier de manière continue, à 360 degrés.

Découper un fragment d'image, remplir des métadonnées associées, l'agréger à une page objet (par exemple la page spécimen, ou la page récolteur si c'est une étiquette).

Découper (et détourner ?) une image (plutôt pour les fossiles ? À préciser)

Mesures

L'outil de mesure doit pouvoir s'étalonner automatiquement à partir de la règle présente sur le spécimen. Les mesures que je prends sont enregistrées automatiquement dans un tableau, associées au numéro du spécimen, et éventuellement aux coordonnées de la zone que j'ai mesurée par rapport à l'ensemble de l'image. Ce tableau comporte

des fonctions classiques de calcul et permet l'export des données dans les formats habituels.

Faire une mesure sur ligne courbe : règle déplaçable et incurvable.

Annotations graphiques

Outils d'annotations graphiques

On définit une zone sur l'image. Cette zone devient un objet (URL, métadonnées), elle peut être associée à d'autres pages. Site intéressant pour les outils graphiques, adaptable même si c'est pour de l'image 3D : Smithsonian X3D <http://3d.si.edu/explorer?modelid=44>

Étiquettes virtuelles

Il a été proposé que : quand un spécialiste fait une redétermination sur le spécimen physique, ou une redétermination en ligne – qui se traduira ou non, selon la politique de l'institution, par l'ajout d'une étiquette sur le spécimen physique – vienne automatiquement s'ajouter sur l'image de l'herbier une étiquette virtuelle : la reproduction d'une étiquette « collée » sur l'image. Cela permettrait de maintenir au maximum le parallélisme herbier papier/image de l'herbier. Ce à quoi les inconditionnels de la base de données pourront répondre que c'est ce parallélisme qu'il faut faire voler en éclats, mais cela suscite une forte opposition.

Organisation de documents

Pour organiser des documents images à retenir, à classer dans ses dossiers, on a plusieurs possibilités :

- des tags, plutôt destinés à des classements personnels assez simples, ou à des classements définis par la communauté.
- faire glisser des images dans des dossiers personnels, mais que l'on peut aussi partager à d'autres usagers choisis, ou à tous (comme les dossiers Google Drive). Les dossiers peuvent contenir des images, fragments d'images, et autres documents (texte, tableur...).
- associer à des images des champs de transcription comprenant différentes valeurs. Il y a ensuite une fonction créer un dossier regroupant les images comportant telle valeur pour tel champ.
- il faut pouvoir exporter sous forme de tableau les identifiants des images avec les valeurs des différents caractères qui leur sont associés.
- il faut pouvoir importer une liste de caractères qui vont constituer des champs de transcription (constituer une interface de transcription à partir d'une liste de caractères).

Il faut pouvoir importer des classifications, ou les créer sur le site à partir d'un logiciel type Freemind. Par exemple il faut pouvoir importer un tableau classant telles espèces sous tel genre. Ce tableau va me créer des dossiers (des dossiers de genre contenant des dossiers d'espèces), et je vais pouvoir mettre des images dedans. Je veux aussi pouvoir représenter mon arborescence de dossier sous la forme d'un arbre.

Recherche d'images

Recherche de types

Case à cocher « rechercher les types » dans l'interface de recherche dans les collections.

Menu de sélection de la catégorie de type recherchée (holotype, isotype, lectotype...).

Aide à la recherche

Aide à la recherche par menus déroulants de noms scientifiques et proposition de synonymes de recherches

Recherche par des traits morphologiques : menu déroulant des caractères renseignés dans la base.

Historique de recherche permettant de chercher dans mes recherches antérieures.

Recherche de récolteurs

Une fonction destinée à faciliter la transcription. Il faudrait pouvoir faire des recherches par nom incomplet ou approximatif dans l'ensemble des récolteurs, et croiser ces recherches par caractères avec d'autres critères (région et époque où il a collecté). Taper quelques caractères permettrait d'arriver à plusieurs récolteurs possibles pour la chaîne de caractères entrée.

Recherches dans R+

Il faut pouvoir rechercher dans les informations validées par les institutions, ou l'ensemble des annotations (valeurs des champs), avec possibilité d'inclure une recherche dans le texte des commentaires.

Il faut prévoir une fonction de recherche dans les noms actuels / tous les noms ayant été attribués aux spécimens.

Il faut définir des règles de sélection des résultats en cas d'informations contradictoires. Par exemple, si un usager dit que ce spécimen est un type, mais que quelqu'un conteste cette information, est-ce qu'on inclut ou non le spécimen dans la recherche qui demande exclusivement des types ? *A priori* oui pour avoir plutôt du bruit que du silence.

Autorités

Il y aura des contenus qui sont soumis ou non à validation institutionnelle. La classification que fait un chercheur dans son espace, même s'il décide de la partager, peut être ouverte aux commentaires, mais elle n'est pas intégrée aux données des collections et n'a pas à être validée.

Le problème de la validation des données produites se pose pour les données qui intéressent les collections : les transcriptions des étiquettes, le statut de type à accorder à un spécimen, et les déterminations faites à partir des images. La transcription ouverte à tous a lieu dans l'espace de science participative (Herbonautes) avec un système d'auto/rétrocontrôle. *Une institution peut autoriser des personnes à transcrire ses collections. Pour une personne non accréditée, la transcription est possible, mais elle serait placée « en attente de validation ». La validation pourrait*

être le fait du personnel de l'institution, ou d'une personne accréditée et accréditée comme validateur des transcriptions réalisées par des personnes non accréditées.

On peut laisser les institutions accréditer les personnes qui en font la demande auprès d'elles (prévoir un formulaire/fonction de contact pour l'accréditation). L'autorisation peut être gérée au niveau du réseau e-Recolnat ou d'une institution en particulier (si certaines veulent valider elles-mêmes les accréditations).

L'accréditation est modulable : elle donne le droit de réaliser un ensemble défini d'opérations pour un ensemble défini de spécimens. Par exemple un chercheur peut être accrédité à remplir les champs de lieu et date de récolte pour toutes les collections e-Recolnat, mais à ne faire de déterminations en ligne que sur les rosaceae, parce que c'est le seul genre qu'il connaît suffisamment pour cela.

On peut aussi envisager qu'une personne ne soit pas accréditée, mais qu'elle soumette des données à validation : on ne lui donne pas l'accréditation, mais on valide ponctuellement ses données (= on accepte l'intégration de ses données à la base R). En tous les cas, il faut que la validation se fasse en un clic, pour celui qui valide, mais que les données soient intégrées automatiquement, sans besoin de rentrer à nouveau les données saisies par la personne non accréditée.

Il reste un problème : si des usagers tous les deux accrédités ne sont pas d'accord sur les données à rentrer, comment fait-on ?

Par ailleurs, on peut paramétrer l'affichage pour voir les images associées aux annotations et commentaires laissés seulement par des usagers accrédités.

Le collaboratoire en termes de navigations et interfaces

Je verrais bien qu'on entre plutôt dans le site par la consultation des images, par navigation (à partir des différentes pages objets), ou par requête. On arrive donc à la consultation d'un spécimen.

Quand on consulte le spécimen, on aurait par exemple sur la gauche la possibilité d'afficher un onglet regroupant les fonctions d'organisation documentaires : les différents dossiers personnels et partagés que j'ai créés, les différentes classifications que j'ai éventuellement importées ou créées (par la transformation d'une arborescence de dossiers en classification), les tags, historiques de recherches et de consultation d'image.

À droite on aurait des icônes appelant différents outils (annotations, mesures) et les différentes façons de visualiser l'image :

- mode transcription diviserait l'écran entre l'image et une interface de transcription à champs multiples que je pourrais configurer.

- mode galerie où je peux visualiser plusieurs images. Je pourrais faire glisser une image de ma galerie dans un dossier partagé qui reste affiché dans mon menu organisation documentaire de la partie gauche de l'écran.

En haut il faudrait qu'il y ait les fonctions de recherche qui restent toujours visibles.

En bas par exemple, il faudrait qu'il y ait toujours les liens vers les fonctions de partage de mes contenus dans le réseau e-Recolnat, une barre présentant mes « amis Recolnat » présents, l'activité sur le site...

Fonctions collaboratives

Elles porteraient sur l'attribution d'états de caractères à certaines images, à partir d'une sélection de caractères et d'un ensemble d'images pertinentes. Cela peut se faire seulement au niveau du collaboratoire : les chercheurs renseignent certains caractères pour des ensembles d'images. Ils peuvent déclarer leur intérêt pour tels caractères et tels lots d'images. (Il faut un espace pour cette déclaration et cette publication d'intérêts). Ils vont alors regarder eux-mêmes les images et signaler les valeurs de certains caractères.

Si cela doit se faire avec des internautes novices, plusieurs solutions sont possibles.

- Il faut soit des annotations « trans-missions », ce qui a été demandé pour la V2 des herbonautes, mais n'a pas été retenu pour les prochains développements. Sinon, cela peut passer par des missions spécifiques qui demandent de renseigner des caractères supplémentaires par rapport à la seule transcription de l'étiquette, pour des ensembles d'images qui constituent des missions. Cela rejoint notre demande de champs variables en fonction des missions.

- Ou alors il faut permettre à tout internaute de parcourir un lot de planches défini par un chercheur et de renseigner les caractères que les chercheurs ont demandé. Cela veut dire que c'est un autre espace participatif que celui des herbonautes. Ça peut être une forme d'exposition virtuelle participative. Ça veut dire que le chercheur qui serait intéressé par renseigner tel et tel caractère pour un groupe d'image va réaliser une page « collection d'image » *ad hoc*. Dans cette page, il va expliquer ses objectifs scientifiques, les caractéristiques des images qu'il propose aux internautes de parcourir, donner des détails susceptibles d'intéresser le public sur les espèces qu'il étudie. Cette page va alors donner accès à la série d'images qu'il a sélectionnées et présentées. En cliquant sur chaque image, un menu s'affiche et pose à l'internaute une série de questions concernant l'image qui lui est présentée. L'internaute remplit alors les états des caractères proposés.

Discussions : voir spécifications de Thomas

Il faut le même système de demande d'aide, qui portera sur une hésitation sur le statut de type, sur une transcription... Les discussions sur un spécimen/une partie/un champ doivent être agrégées pour « remonter » sur une page non pas mission mais discussions du collaboratoire, visible de tous.

C.2. Maquette de l'interface de transcription réalisée dans le cadre du travail de conception de la version 2 des Herbonautes

Lisa Chupin, décembre 2014

Maquette interface de transcription

Icône donnant accès À la rédaction de commentaire Concernant le champ concerné

Que dois-je faire ?

Laisser ou voir des commentaires.
(Il y a 5 commentaire sur ce spécimen.)

Travailler sur le spécimen dans le collaboratoire

Ressources pour faciliter la transcription

C.3. Maquette d'une version du laboratoire en ligne intégrant les contributions des Herbonautes et du logiciel de contrôle qualité des images

Slide de présentation de la maquette, Lisa Chupin, juin 2017

Intégration des retours des usagers

The screenshot shows a web interface for plant specimen management. The main content area displays two images of plant specimens, both labeled 'Asteraceae Centaurea sp.'. The interface includes a navigation menu on the left with 'Organisation' and 'Labbranch' sections. A right-hand sidebar contains 'Reference Data' (Specimen, Determination 1), 'Données sociales' (Herbonautes (2)), 'Éditer les données', and 'Tags and Measures' (a table with columns Name, Value, Sheet). Below this is a 'Properties' section with 'Creation date' and 'Angle'.

Name	Value	Sheet
étiquette	50.18 mm	MPU333617
apex	52.78 °	MPU333617

Properties
apex
Creation date : Monday, 5/15/2017, 10:49 PM
Angle = 52° 47' 0"

Callout boxes on the right highlight the following features:

- Numéro de spécimen visible (points to the specimen ID in the sidebar)
- Signaler les commentaires réalisés dans les autres modules (points to the 'Données sociales' section)
- Faciliter le lien vers l'interface d'édition de données (points to the 'Éditer les données' button)
- Ajouter moyenne, écart type des mesures d'un même type du set (points to the 'Tags and Measures' table)

C.4. Échéancier prévisionnel des développements du Cnam concernant le collaboratoire

Document envoyé le 11 septembre 2015 par Manuel Zacklad au coordinateur de l'infrastructure.

Echéancier prévisionnel des développements du CNAM concernant le Collaboratoire

Première livraison été 2016 : package minimal de fonctions pour pouvoir faire un travail de recherche sur les planches

Composant « Visionneuse » permettant, entre autres :

- de définir des zones graphiques sur des images (formes simples, sélections polygonales)
- de poser des marqueurs graphiques sur les planches (répertoire d'icônes significatives)
- d'effectuer des mesures
- d'associer du contenu métier (détermination, caractère, etc.) ou libre (commentaire libre) aux zones et repères graphiques posés sur la planche

Composant « Bureaux Virtuels », permettant :

- d'importer des planches à partir de l'outil de recherche développé par le WP2
- de créer librement des collections & sous-collections et d'y organiser les planches en les spatialisant (notion de bureau virtuel)
- superposer des planches (avec niveau de transparence)

Les aspects techniques (sécurité, gestion des comptes utilisateurs, communication avec la base du WP2), de même que les fonctions basiques (visualisation des métadonnées de base, par exemple) ne sont pas détaillées ici mais constituent bien entendu un chantier préalable (déjà très amorcé).

Seconde livraison printemps 2017 : fonctions scientifiques plus spécifiques et premières fonctions sociales/contributives/coopératives

- outils d'aide à la classification (constitution de catégories émergentes, extraction de fragments d'images pour, par exemple, « mettre en série tous les bourgeons d'un certain type », etc.)
- travail coopératif en temps-réel entre chercheurs sur les planches (notion de « bureau virtuel partagé »)
- modèle de « caractères » et intégration des principes d'XPer3
- prémisses du système de contribution généralisée (tout le monde peut contribuer, mais les contributions de chacun sont associées à un rang d'autorité)
- exports Excel
- système de Tags collaboratif

Troisième livraison fin du projet : fonctions de collaboration/contribution avancée réseau social botanique savant complet.

Les fonctions afférentes seront ajustées en fonction de l'expérience accumulée suite aux livraisons 1 et 2 (détails sur demande)

**D. Réponses aux questionnaires détaillés
envoyés à des contributeurs réguliers
de la plateforme « Les Herbonautes »**



Les herbonautes
L'herbier numérique collaboratif citoyen



RETOUR D'EXPERIENCE DES HERBONAUTES

Voici un questionnaire destiné à mieux appréhender votre expérience d'herbonaute. Vous pouvez y apporter directement vos réponses après chaque question, et me renvoyer par mail le document complété. Je vous remercie de votre participation très précieuse pour mon étude qui servira elle-même aux développements ultérieurs des Herbonautes.

Lisa Chupin, Conservatoire National des Arts et Métiers

Votre utilisation du site

1) Comment avez-vous connu le site *Les Herbonautes* ?

En écoutant France-Info, en décembre 2013, j'ai entendu la fin de la chronique info-sciences qui parlait d'un herbier participatif sur internet. J'ai vaguement compris et je suis allée réécouter la chronique sur le site de France-Info, c'est comme ça que j'ai découvert le site des Herbonautes et je m'y suis inscrite de suite.

2) Dans quelles circonstances participez-vous aux *Herbonautes* ? (Vous pouvez préciser dans quels lieux, à quels moments, pour combien de temps vous avez déjà contribué au site, seul(e) ou entouré(e)...)
Chez moi, quand je surfe et que j'ai un moment, je vais sur le site. Je le fais seule.

3) Comment procédez-vous pour remplir les différents champs (pays, région, date, déterminateur, localisation...) associés aux planches d'herbier qui vous sont proposées ? (Vous pouvez expliquer notamment dans quel ordre vous examinez les différentes zones de la planche scannée, les raisonnements auxquels vous avez recours, les sources d'information que vous consultez, ou encore les personnes et avis que vous sollicitez pour vous aider...)

Il est souhaitable de remplir les champs dans l'ordre, compte-tenu de la configuration du site. Si on le fait dans le désordre, on est obligé à plus de manipulations, il faut cliquer sur la petite flèche pour avoir accès au détail du champ et remplir les informations. Si on le fait dans l'ordre, dès qu'un champ est rempli et validé, le suivant est directement accessible sans manipulation.

Je visualise globalement d'abord pour voir où sont les étiquettes à lire, il y en a de toutes petites dans les coins parfois et il y a également des annotations au crayon pas très visibles. Je les parcours rapidement pour voir si j'arrive à tout lire.

C'est le lieu le plus important pour moi. Ayant déchiffré la ville, je vais sur google pour déterminer la région administrative (pour la France et les pays étrangers). Si je ne trouve pas une ville en France, j'ai recours aux différents wiki « liste des anciennes communes + département concerné ». Si je n'ai pas le département, je copie le lieu indiqué par le botaniste et je fais une recherche sur google ou plus souvent sur google-livres, quelquefois directement sur google-maps qui peut aussi donner des résultats.

La date pose généralement peu de problèmes. Il m'est arrivé en cas de doute de vérifier les dates (naissance, décès) du botaniste sur wikipedia (wiki espagnol mieux doté pour les botanistes français que le wiki français et les autres wikis en italien ou anglais) ou la base Harvard pour confirmation qu'il avait pu ou n'avait pas pu récolter à la date que je croyais lire.

Les champs botanistes sont les plus difficiles car il y a 2 infos à valider : le récolteur et le déterminateur. Si le nom est difficile à lire, j'ai recours à la base Harvard d'abord, très précieuse même si non exhaustive ; je recherche également sur <http://science.mnhn.fr/institution/mnhn/search/form> dans le premier champ « recherche globale ». Il m'arrive de lire des notices biographiques de botanistes sur Gallica pour vérifier s'il ont participé à des expéditions lointaines.

Ensuite, il faut être très attentif pour remplir ces deux champs, vérifier les abréviations leg, ded, com, ipse legit, det, coll, etc... vérifier que le déterminateur -s'il est noté clairement- a bien déterminé la bonne espèce.

Ce champ est source de multiples alertes entre les herbonautes.

Le champ localité est assez facile à condition de bien déchiffrer les écritures anciennes, connaître un peu de latin. Il suffit de recopier entièrement ce que le botaniste a indiqué.

Enfin, la géolocalisation que je remplis uniquement en utilisant la carte sur la page ; je ne maîtrise pas assez pour recopier des coordonnées avec degrés, minutes etc. Je remplis le champ précision en recopiant ce qui est indiqué dans la petite fenêtre de la carte. Pour les lieux en France, j'indique toujours dans la recherche sur la carte le département dont il s'agit.

4) Vous souvenez-vous avoir eu beaucoup de difficultés à déterminer en particulier une localisation, ou un récolteur ? Si oui, pourriez-vous expliquer ce qui posait problème et la façon dont vous êtes parvenu(e) à trouver tout ou partie de l'information cherchée ?

Je me souviens de certaines recherches car elles ont été très longues. Cela pouvait concerner un lieu ou bien un nom de botaniste ou bien un paraphe à déchiffrer. Je vous propose de vous donner des liens concernant les spécimens où vous pourrez lire mes commentaires :

<http://lesherbonautes.mnhn.fr/specimens/MNHN/P/P00699656> : discussion sur Héribaud-Joseph nom ecclésiastique d'un botaniste.

<http://lesherbonautes.mnhn.fr/specimens/MNHN/P/P01584265> : discussion sur le nom du botaniste et le nom du lieu.

<http://lesherbonautes.mnhn.fr/specimens/MNHN/P/P02163439> : discussion (monologue) sur les botanistes et le lieu.

<http://lesherbonautes.mnhn.fr/specimens/MNHN/P/P02101459> : discussion botaniste et lieu.

<http://lesherbonautes.mnhn.fr/specimens/MNHN/P/P04006937> : recherche d'un botaniste.

Votre intérêt pour le site

5) Que trouvez-vous de particulièrement plaisant dans *Les Herbonautes* ?

Lire les étiquettes manuscrites anciennes du 19e siècle, concernant des lieux en France, surtout. Et participer à la conservation de ces données.

6) Y a-t-il des informations contenues dans certaines planches d'herbier ou certaines étiquettes qui vous intéressent en particulier ?

J'aime beaucoup tomber sur (ou rechercher sur la carte de géolocalisation globale de la mission) des lieux que je connais bien (dans le Var et en Bretagne). J'aime (ré)apprendre la géographie française, d'une France qui souvent n'existe plus, où on localise un bois, une petite forêt sur une zone commerciale, lire qu'on parle des ruines du Château de Pierrefonds parce que la récolte a été faite avant sa « rénovation », lire « prairies maritimes » où il y a maintenant des quartiers d'habitation, des immeubles. J'aime bien lire les notes ou copies de courrier qui figurent sur les planches également.

7) Les échanges à propos du site, que ce soit avec des animateurs de missions, ou d'autres personnes, sont-ils une source de motivation importante pour votre participation au site ?

Concernant les responsables de mission : une source d'aide pas une source de motivation. Une aide pour des réponses aux questions qu'on peut se poser sur la manière de renseigner les fiches. Une aide également avec les autres herbonautes, lorsqu'on discute sur comment renseigner tel ou tel champ.

8) Avez-vous le sentiment d'avoir acquis des connaissances ou des savoir-faire grâce à votre contribution au site ? Si oui, qu'avez-vous appris ?

J'ai acquis quelques connaissances générales sur la géographie, la biographie de quelques botanistes, quelques noms d'espèces et sous-espèces de plantes que j'ignorais.

9) Contribuer au site présente-t-il aussi des désagréments ? Merci de bien vouloir les décrire.

Les désagréments sont d'ordre technique.

Je n'aime pas le fait de devoir me connecter à chaque fois que je viens sur le site (et je passe plusieurs fois par jour) alors que je ne clique jamais sur déconnexion. Et pour se connecter il faut non seulement indiquer pseudo et mot de passe mais cliquer avec la souris sur « se connecter » car les infos ne sont pas prises en

compte avec la touche « entrée ».

Je sais par expérience qu'on ne reçoit pas toutes les alertes car j'ai l'habitude de lire tous les commentaires dans la rubrique « activité des herbonautes ». Certains auraient dû m'être envoyés car j'ai participé sur le specimen mais je ne les reçois pas tous. Je suppose que c'est le cas de tous et c'est un peu frustrant car des fois on pose des questions dans les commentaires des specimens et on reçoit rarement des réponses.

10) Quelles sources de motivation vous amènent à continuer à participer aux *Herbonautes* malgré les difficultés ou désagréments que vous avez pu rencontrer ?

Ce ne sont que des désagréments techniques qui pourront être réglés, j'espère. Ca n'entame pas le plaisir que j'ai à contribuer aux herbonautes.

11) Diriez-vous que votre participation au site a changé votre regard sur les herbiers ? Si oui, en quoi consiste ce changement ?

Je n'avais pas de regard particulier sur les herbiers. Avec le site, j'ai appris à connaître un peu les herbiers. Je vois que c'est une activité qui doit être faite avec rigueur, tout est dans la précision aussi bien en remplissant les étiquettes qu'en disposant le specimen. J'ai remarqué qu'il y avait des specimens mieux disposés que d'autres.

Vos autres pratiques naturalistes

12) Hormis la participation aux *Herbonautes*, avez-vous d'autres activités concernant la botanique ou les sciences naturelles, que ce soit à titre personnel ou professionnel ? Si oui, merci de bien vouloir décrire en quoi elles consistent, et préciser si vous les pratiquez au sein d'une association ou organisation en particulier. (N'hésitez pas à mentionner si vous vous êtes déjà investi(e) dans la constitution d'un herbier ou autre collection naturaliste.)

Non, pas d'activité botanique ou de sciences naturelles et je n'ai jamais fait d'herbier. J'aime les fleurs et les plantes et les prendre en photo également.

13) Dans les deux dernières années, avez-vous visité une exposition présentant des herbiers ou d'autres collections d'histoire naturelle ? Si oui, laquelle ou lesquelles ?

Non.

14) Hormis les *Herbonautes*, consultez-vous d'autres sites Web consacrés à la botanique, aux sciences naturelles, à la biodiversité, ou en lien avec des collections naturalistes ? Si oui, quels sont ceux qui apportent le plus à vos activités ? Et qu'apportent-ils effectivement ?

Je reçois la newsletter de Noé-Conservation, j'avais été tentée de participer au recensement des papillons et escargots de jardins mais j'avais trouvé ça trop fastidieux, je ne l'ai pas fait.

Informations complémentaires

15) Hormis votre participation aux *Herbonautes*, contribuez-vous à d'autres programmes de science participative, ou à des sites Web collaboratifs (blog, encyclopédie en ligne, forum, journalisme participatif, espace collaboratif de travail...) ?

Aucun autre programme de science participative. J'ai contribué longtemps sur un site de paroles de chansons où je rédigeais et/ou corrigeais des textes ; sur un wiki de chansons également où je cherchais beaucoup d'illustrations, photos, partitions, vidéos ; sur un site de chansons anciennes où j'ai fait des recherches de généalogie, iconographie, notices biographiques, détails sur articles de journaux sur Gallica entre autres.

16) Avez-vous des compétences en programmation informatique ? Si oui, pouvez-vous préciser les réalisations ou projets dans lesquels vous êtes investis ?

Non.

~~17) Pour nous permettre de mieux appréhender les caractéristiques des usagers du site, quelques renseignements sur votre situation socio-professionnelle seraient également utiles. Si vous n'y voyez pas d'inconvénient, vous pouvez nous laisser pour finir quelques indications concernant votre quotidien (Où habitez vous ? Quel est votre âge ? Quelle est votre situation professionnelle et familiale ?~~

N'hésitez pas à ajouter ci-dessous d'autres informations sur votre expérience d'herbonaute qui n'auraient pas trouvé leur place dans le questionnaire et à indiquer si certaines questions vous ont posé problème.

Merci encore !

J'aimerais bien qu'un forum soit créé en parallèle au site des herbonautes. Un forum pas vraiment officiel si ce n'est pas envisageable, mais verouillé, accessible aux seuls herbonautes, où nous pourrions échanger sur les specimens, les botanistes, les lieux posant problème avec un côté discussion de tout de rien, entre nous. Et/ou il serait bien de pouvoir s'envoyer des messages privés sur le site des herbonautes.



Voici un questionnaire destiné à mieux appréhender votre expérience d'herbonaute. Vous pouvez y apporter directement vos réponses après chaque question, et me renvoyer par mail le document complété. Je vous remercie de votre participation très précieuse pour mon étude qui servira elle-même aux développements ultérieurs des Herbonautes.

Lisa Chupin, Conservatoire National des Arts et Métiers

Votre utilisation du site

1) Comment avez-vous connu le site *Les Herbonautes* ?

Lecture d'un article dans la revue *Sciences et Avenir* de novembre 2013

2) Dans quelles circonstances participez-vous aux *Herbonautes* ? (Vous pouvez préciser dans quels lieux, à quels moments, pour combien de temps vous avez déjà contribué au site, seul(e) ou entouré(e)...))

Je suis en retraite et j'ai beaucoup de temps libre. Je participe depuis novembre 2013, chez moi, surtout le matin, à raison d'une heure ou plus, mais pas tous les jours... je travaille seule et j'évalue à une centaine de « contributions » par heure ce qui donne environ 200 heures depuis novembre.

3) Comment procédez-vous pour remplir les différents champs (pays, région, date, déterminateur, localisation...) associés aux planches d'herbier qui vous sont proposées ? (Vous pouvez expliquer notamment dans quel ordre vous examinez les différentes zones de la planche scannée, les raisonnements auxquels vous avez recours, les sources d'information que vous consultez, ou encore les personnes et avis que vous sollicitez pour vous aider...)

Examen de la planche sans grossir l'image pour repérer les étiquettes, code barre pour éventuellement sélectionner les spécimens non utilisables.

Grossir l'image pour une lecture facile de l'étiquette en cherchant les éléments dans l'ordre du questionnaire à compléter.

Pour les cas qui ne sont pas évidents,

- chercher les indications utiles dans les consignes du site des Herbonautes, dans Internet, dans un atlas, sur des cartes papier...

- comparer les réponses des autres participants quand il y en a,

- pour les dates ne comportant que 2 chiffres (sans 1800 ou 1900) revoir la biographie du botaniste pour s'assurer de la période concernée,

4) Vous souvenez-vous avoir eu beaucoup de difficultés à déterminer en particulier une localisation, ou un récolteur ? Si oui, pourriez-vous expliquer ce qui posait problème et la façon dont vous êtes parvenu(e) à trouver tout ou partie de l'information cherchée ?

Pour les localisations : imprécisions de l'indication menant à plusieurs possibilités – par exemples nom de lieu-dit uniquement, ou communes multiples avec le même nom, anciens noms de communes modifiés depuis ou regroupement de communes, noms d'anciens pays ayant été modifiés ou divisés aujourd'hui.

En cherchant dans tous les sens, compte tenu de l'activité connue du botaniste, des fois en relation avec

d'autres planches du même récolteur et parfois avec une sorte de « pari » sur la bonne réponse, toujours en cochant la case « J'hésite ».

C'est aussi pourquoi je choisis souvent d'appeler les planches en utilisant le n° (par exemple) <http://lesherbonautes.mnhn.fr/missions/1511211/specimens/1544842#> et en augmentant 1 par 1, ce qui permet de rester dans des séries du même botaniste, sur un même lieu ou une même campagne de récolte.

Pour les récolteurs, écriture manuscrite difficile à déchiffrer, manque d'un prénom.

Votre intérêt pour le site

5) Que trouvez-vous de particulièrement plaisant dans *Les Herbonautes* ?

Je m'intéresse depuis longtemps à la botanique à titre privé, et comme membre de l'association des amis du Conservatoire et Jardin botanique de Genève (CJB), j'ai souvent parlé avec les conservateurs du CJB de la rénovation des locaux de l'herbier de Genève, du travail de numérisation en cours de cet herbier, de la constitution des collections et des travaux des chercheurs sur ces collections.

Du coup, la proposition des *Herbonautes* m'a paru très intéressante, comme démarche proposant aux amateurs curieux de visiter en quelque sorte toutes ces collections anciennes.

En plus je suis du genre à bien aimer les collections, les classements, les répertoires...

Il me plaît bien de participer modestement à une action qui peut être utile aux chercheurs et qui permet de compléter ce travail énorme de numérisation.

Sur le fonctionnement, cela m'apparaît presque comme un jeu de décrypter les étiquettes et de finir par « connaître » un peu ces anciens botanistes, de repérer les lieux de collecte.

6) Y a-t-il des informations contenues dans certaines planches d'herbier ou certaines étiquettes qui vous intéressent en particulier ?

Je me suis aperçue que je participe plus volontiers aux collections qui concernent des plantes ou des régions que je connais, par exemple L'Herbier Fenoul, les gentianes, dans lesquelles je retrouve des lieux ou des plantes que j'ai moi-même eu l'occasion de visiter ou de rencontrer.

7) Les échanges à propos du site, que ce soit avec des animateurs de missions, ou d'autres personnes, sont-ils une source de motivation importante pour votre participation au site ?

Les échanges sont intéressants et permettent d'aider dans certains cas. Je les vois un peu comme une conversation où je retrouverais des personnes connues, surtout avec les animateurs et les herbonautes qui contribuent beaucoup.

Je découvre avec étonnement et plaisir comme sont nombreux ceux qui s'intéressent dans l'anonymat à ces actions.

8) Avez-vous le sentiment d'avoir acquis des connaissances ou des savoir-faire grâce à votre contribution au site ? Si oui, qu'avez-vous appris ?

C'est peut être dans la méthode de travail que j'ai fait des acquisitions en examinant avec soin les étiquettes et sur le travail de recherche. Découverte de sites internet.

Les connaissances sont très intéressantes aussi mais c'est une avalanche de renseignements qu'il m'est difficile de mémoriser dans son entier.

9) Contribuer au site présente-t-il aussi des désagréments ? Merci de bien vouloir les décrire.

Pour moi, les désagréments se situeraient plutôt au niveau de la lecture sur écran qui est plutôt fatigante.

Mais en fait je ne considère pas cela comme un handicap puisque c'est un travail volontaire, réalisé quand j'en ai envie, sans obligation de délai ou de rendement.

De plus, la consultation de ces collections, à distance et ouverte à tous n'est pas possible autrement.

En fait l'inconvénient principal serait que j'y passe un temps fou, puisque on a toujours envie de renseigner une planche supplémentaire.

10) Quelles sources de motivation vous amènent à continuer à participer aux *Herbonautes* malgré les difficultés ou désagréments que vous avez pu rencontrer ?

[Voir question 5](#)

Mais j'avoue quand même avoir abandonné quelques planches trop complexes à décrypter (langue, écriture, renseignements trop incomplets...)

11) Diriez-vous que votre participation au site a changé votre regard sur les herbiers ? Si oui, en quoi consiste ce changement ?

[J'en conçois mieux la valeur patrimoniale et l'utilité pour les recherches actuelles.](#)

Vos autres pratiques naturalistes

12) Hormis la participation aux *Herbonautes*, avez-vous d'autres activités concernant la botanique ou les sciences naturelles, que ce soit à titre personnel ou professionnel ? Si oui, merci de bien vouloir décrire en quoi elles consistent, et préciser si vous les pratiquez au sein d'une association ou organisation en particulier. (N'hésitez pas à mentionner si vous vous êtes déjà investi(e) dans la constitution d'un herbier ou autre collection naturaliste.)

Ma participation à des activités concernant la botanique se résume à la constitution, il y a fort longtemps d'un herbier « amateur », au départ pour intéresser les enfants, mais abandonné depuis longtemps et qui s'est détruit au fur et à mesure faute de bons moyens de conservation, au profit d'un herbier photographique avec l'apparition de la photo numérique.

La participation aux activités de l'Association des amis de jardin Botanique de Genève (AAJB) – sorties sur le terrain avec des spécialistes et conférences ou cours, depuis une dizaine d'années est venue enrichir mes connaissances.

Une autre activité qui me prend beaucoup de temps actuellement depuis 4 ans est la participation aux collectes de photos d'insectes pour le SPIOLL.

13) Dans les deux dernières années, avez-vous visité une exposition présentant des herbiers ou d'autres collections d'histoire naturelle ? Si oui, laquelle ou lesquelles ?

[Herbier du CJB Genève après sa rénovation en 2013](#)

[Museum de Genève](#)

[Musée d'histoire naturelle de Montauban](#)

14) Hormis les *Herbonautes*, consultez-vous d'autres sites Web consacrés à la botanique, aux sciences naturelles, à la biodiversité, ou en lien avec des collections naturalistes ? Si oui, quels sont ceux qui apportent le plus à vos activités ? Et qu'apportent-ils effectivement ?

Je « clique » un peu de façon anarchique jusqu'à trouver les infos et les confirmer, sites généralistes ou plus spécialisés par pays (selon le pays concerné par mon voyage – j'aime bien regarder la flore des pays que je visite) ou régions.

Mais certains sites restent souvent présents : [Tela botanica](#), [info flora.ch](#), pour les insectes [aramel.free](#) ou le site [SPIOLL](#), [papillons Poitou-Charente...](#)

Ces sites me permettent les identifications ou donnent les précisions pour compléter les connaissances,

Informations complémentaires

15) Hormis votre participation aux *Herbonautes*, contribuez-vous à d'autres programmes de science participative, ou à des sites Web collaboratifs (blog, encyclopédie en ligne, forum, journalisme participatif, espace collaboratif de travail...) ?

[SPIOLL - insectes pollinisateurs, collectes de photos](#)

[Globeatnight - astronomie, collecte d'observations tout à fait élémentaires sur la magnitude d'étoiles](#)

16) Avez-vous des compétences en programmation informatique ? Si oui, pouvez-vous préciser les réalisations ou projets dans lesquels vous êtes investis ?

non

17) Pour nous permettre de mieux appréhender les caractéristiques des usagers du site, quelques renseignements sur votre situation socio-professionnelle seraient également utiles. Si vous n'y voyez pas d'inconvénient, vous pouvez nous laisser pour finir quelques indications concernant votre quotidien (Où habitez vous ? ~~Quel est votre âge ? Quelle est votre situation professionnelle et familiale ?~~).

Genève,

Souvent dehors dans la campagne dès qu'il fait beau, toujours le nez par terre dans les randos ici ou ailleurs pour voir les fleurs et les insectes, et parfois avec les petits-enfants pour essayer de les intéresser.

D'ailleurs l'intérêt pour la botanique a commencé à cause d'une de mes filles (12 ans à l'époque) qui voulait récolter des plantes médicinales et nous faire goûter des tisanes – prudence oblige.

N'hésitez pas à ajouter ci-dessous d'autres informations sur votre expérience d'herbonaute qui n'auraient pas trouvé leur place dans le questionnaire et à indiquer si certaines questions vous ont posé problème.

Merci encore !



Voici un questionnaire destiné à mieux appréhender votre expérience d'herbonaute. Vous pouvez y apporter directement vos réponses après chaque question, et me renvoyer par mail le document complété. Je vous remercie de votre participation très précieuse pour mon étude qui servira elle-même aux développements ultérieurs des Herbonautes.

Lisa Chupin, Conservatoire National des Arts et Métiers

Votre utilisation du site

1) Dans quelles circonstances participez-vous aux *Herbonautes* ? (Vous pouvez préciser dans quels lieux, à quels moments, pour combien de temps vous avez déjà contribué au site, seul(e) ou entouré(e)...))

Je participe de façon non régulière, chez-moi, parfois aussi lors de pauses de mon travail, pendant dix ou vingt minutes, parfois plus si la fiche est difficile à analyser. Je suis toujours seul.

2) Comment procédez-vous pour remplir les différents champs (pays, région, date, déterminateur, récolteur localisation, géolocalisation) associés aux planches d'herbier qui vous sont proposées ? (Vous pouvez expliquer notamment s'il y a un ordre dans lequel vous examinez les différentes zones de la planche scannée, les raisonnements auxquels vous avez recours, les sources d'information que vous consultez, ou encore les personnes et avis que vous sollicitez pour vous aider...)

Je suis l'ordre indiqué dans la planche. Dans le cas les plus complexes, j'utilise Internet pour exploiter les indices que j'ai repérés: cela va du nom de l'espèce (pour en vérifier par exemple la zone de diffusion), jusqu'au nom du botaniste (parfois, si la graphie est incertaine, essayant plusieurs écriture approchées). Les outils plus performants sont (pas forcément dans l'ordre), la base Sonnerat, Google, Wikipedia, Plant List et Tropicos, archives.org, la Biodiversity Heritage Library, Gallica, Telabotanica...

2bis) Vous arrive-t-il d'hésiter à propos des informations d'une étiquette ? Comment procédez-vous quand vous n'êtes pas sûr de la valeur d'un champ à remplir ?

Plusieurs types d'hésitation : celles liées à la graphie (difficulté à déchiffrer une écriture), celles liées à la cohérence des informations (type « ours blancs dans le désert »), celles liées aux spécimens de la planche (cela est plus rare, mais parfois les étiquettes ne semblent pas correspondre à la plante présentée par la planche). En cas de doute, je laisse un message expliquant mes perplexités.

2ter) Pensez-vous avoir déjà entré des informations inexactes dans certains champs ? Savez-vous si elles ont été corrigées ?

J'ai en effet commis parfois des erreurs d'interprétation, la plus part du temps corrigés par d'autres herbonautes qui m'ont signalé leur point de vue.

3) Quand vous avez commencé à participer aux Herbonautes, vous souvenez-vous avoir éprouvé des difficultés particulières que vous avez ensuite surmontées ?

Je dû me construire ma « boîte à outils » (voir question précédente), mais cela s'est réglé assez vite... Et puis les autres herbonautes donnent de bonnes indications qui mettent rapidement un « bleu » dans le droit chemin...

4) Vous souvenez-vous avoir eu beaucoup de difficultés à déterminer en particulier une localisation, ou un récolteur ? Si oui, pourriez-vous expliquer ce qui posait problème et la façon dont vous êtes parvenu(e) à trouver tout ou partie de l'information cherchée ?

Les difficultés plus importantes sont liées à l'orthographe (certains noms restent illisibles) et – en ce qui concerne les localités, surtout étrangères – au changement de dénomination des lieux. Je me souviens que certains spécimens d'Iris d'Asie centrale portaient des indications de localité en allemand qui ont donné pas mal de problèmes, car les noms des lieux actuels ne correspondent plus. De vieilles flores du XIXe siècle et des comptes rendu d'expéditions scientifiques (repérés grâce à archive.org) ont permis de trouver une solution satisfaisante dans plusieurs cas. De façon générale, il faut savoir que la connaissance, même approximative, d'une ou deux langues étrangères (l'anglais et l'allemand, par exemple) et du latin est TRES utile dans ce travail.

5) Avez-vous constaté des améliorations dans le fonctionnement du site depuis que vous y contribuez ?

Je ne sais pas, depuis un moment je ne participe plus au travail... (problèmes de temps)

Votre intérêt pour le site

6) Que trouvez-vous de particulièrement plaisant dans *Les Herbonautes* ?

Etant un inconditionnel de la botanique et de l'histoire, je trouve les Herbonautes un fantastique voyage dans le temps et l'espace. Il y a aussi le défi intellectuel et la possibilité de visionner des planches d'herbier autrement inaccessibles pour le simple amateur que je suis...

7) Y a-t-il des informations contenues dans certaines planches d'herbier ou certaines étiquettes qui vous intéressent en particulier ?

Personnellement, j'ai passé des heures et des heures sur le Gentianes, regardant les échantillons, faisant des recherches bibliographiques en parallèle, notant parfois les sites de récolte. La flore européenne m'intéresse particulièrement.

8) Les échanges à propos du site, que ce soit avec des animateurs de missions, ou d'autres personnes, sont-ils une source de motivation importante pour votre participation au site ?

C'est une source fondamentale ! J'ai l'impression par contre qu'il n'y ait pas de véritables botanistes qui suivent le travail, car les quelques remarques « techniques » faites à propos de la détermination de certains spécimens sont restées sans réponse (parfois on m'a répondu par exemple que si le MNHN avait nommé de cette façon la planche cela indiquait que le nom saisi était le bon...)

9) Avez-vous le sentiment d'avoir acquis des connaissances ou des savoir-faire grâce à votre contribution au site ? Si oui, qu'avez-vous appris ?

Une technique de recherche intégrée « image-texte-outils internet-réflexion personnelle-partage collaboratif » qui est très enrichissante. Et puis, énormément d'informations spécifiques dans mes domaines d'intérêt (prolongées aussi grâce au site Sonnerat).

10) Contribuer au site présente-t-il aussi des désagréments ? Merci de bien vouloir les décrire.

Seulement si vous restez collés au site pendant des heures (cela m'est arrivé rarement), car à un moment

donné on se remplit de courbatures et de problèmes ophtalmologiques☺. Un autre désagrément est dû à la frustration de ne pas arriver à une détermination. On peut aussi avoir de petites blessures narcissiques quand un ou plusieurs herbonautes vous informent que vous vous êtes trompés (surtout si vous restez de l'avis que cela n'est pas le cas).

11) Quelles sources de motivation vous amènent à continuer à participer aux *Herbonautes* malgré les difficultés ou désagréments que vous avez pu rencontrer ?

Malheureusement je n'ai plus vraiment le temps d'y participer sérieusement.... Mais de façon générale, on voit bien l'importance de ce travail. Des bases de données consolidées comme le site Sonnerat ou – dans d'autres domaines – Gallica ou archive.org vont changer notre façon de se rapporter à la connaissance. C'est un travail historique, une espèce de « révolution » qui est en marche. On pourra dire : j'y étais (je me contente de peu, j'en conviens... ☺)

12) Diriez-vous que votre participation au site a changé votre regard sur les herbiers ? Si oui, en quoi consiste ce changement ?

Vous m'imaginez, avant Sonnerat ou Herbonautes, mettre le nez dans l'herbier du Muséum ? N'étant pas un professionnel, on m'aurait prié plus ou moins gentiment d'aller voir ailleurs... Les seuls spécimens abordables étaient derrière des vitrines. Aujourd'hui je peux comparer, en quelques minutes, des planches situées aux quatre coins du monde et réalisées par les plus célèbres botanistes de tous les temps... c'est tout simplement fantastique ! Enfin, je savais déjà qu'un herbier n'est pas seulement la collection farfelue d'un excentrique (moi-même, lors de mes études universitaires, j'ai réalisé plusieurs centaines de planches), mais ce travail en montre la valeur à la puissance mille !

Vos autres pratiques naturalistes

13) Hormis la participation aux Herbonautes, avez-vous d'autres activités concernant la botanique ou les sciences naturelles, que ce soit à titre personnel ou professionnel ? Si oui, merci de bien vouloir décrire en quoi elles consistent, et préciser si vous les pratiquez au sein d'une association ou organisation en particulier. (N'hésitez pas à mentionner si vous vous êtes déjà investi(e) dans la constitution d'un herbier ou autre collection naturaliste.)

J'ai déjà un peu répondu. Je m'intéresse à la flore française et européenne (et aussi aux sciences naturelles en générale, mais la botanique reste mon intérêt prépondérant) depuis de longues années et de façon plus approfondie depuis l'université (mon mémoire a consisté dans l'exploration d'une petite vallée des Alpes du sud). J'ai aussi travaillé un peu pour un parc naturel et dans le domaine de l'animation scientifique pour les écoles. Je ne fais pas partie d'une association, même si j'en connais plusieurs (notamment les réseaux participatifs de Telabotanica).

12) Dans les deux dernières années, avez-vous visité une exposition présentant des herbiers ou d'autres collections d'histoire naturelle ? Si oui, laquelle ou lesquelles ?

Je passe régulièrement aux MNHN (mais je dois avouer que je m'intéresse surtout aux jardins...)

14) Hormis les *Herbonautes*, consultez-vous d'autres sites Web consacrés à la botanique, aux sciences naturelles, à la biodiversité, ou en lien avec des collections naturalistes ? Si oui, quels sont ceux qui apportent le plus à vos activités ? Et qu'apportent-ils effectivement ?

Vois ci-dessus ! A partir de Telabotanica (« must » pour ceux qui s'intéressent à la flore française) jusqu'aux différents blogs de passionnés, je trouve que l'information sur le net ouvre des possibilités fantastiques. Un exemple : la possibilité de repérer rapidement de centaines de publications scientifiques sur un sujet donné, ce qui était impensable à l'époque des mes études...

Informations complémentaires

14) Hormis votre participation aux Herbonautes, contribuez-vous à d'autres programmes de science

participative, ou à des sites Web collaboratifs (blog, encyclopédie en ligne, forum, journalisme participatif, espace collaboratif de travail...) ?

Pas pour l'instant.

15) Avez-vous des compétences en informatique ? Si oui, pouvez-vous précisez les réalisations ou projets dans lesquels vous êtes investis ?

Non...

16) Pour nous permettre de mieux appréhender les caractéristiques des usagers du site, quelques renseignements sur votre situation socio-professionnelle seraient également utiles. Si vous n'y voyez pas d'inconvénient, vous pouvez nous laisser pour finir quelques indications concernant votre quotidien (Où habitez-vous ? ~~Quel est votre âge ? Quelle est votre situation professionnelle et familiale ?~~).

Paris

N'hésitez pas à ajouter ci-dessous d'autres informations sur votre expérience d'herbonaute qui n'auraient pas trouvé leur place dans le questionnaire et à indiquer si certaines questions vous ont posé problème.

Merci encore !



Les herbonautes
L'herbier numérique collaboratif citoyen



RETOUR D'EXPERIENCE DES HERBONAUTES

Voici un questionnaire destiné à mieux appréhender votre expérience d'herbonaute. Vous pouvez y apporter directement vos réponses après chaque question, et me renvoyer par mail le document complété. Je vous remercie de votre participation très précieuse pour mon étude qui servira elle-même aux développements ultérieurs des Herbonautes.

Lisa Chupin, Conservatoire National des Arts et Métiers

Votre utilisation du site

- 1) Comment avez-vous connu le site *Les Herbonautes* ?

Par un article dans la presse (Le Monde ou le Nouvel Obs ou Libé, je ne me souviens plus)

- 2) Dans quelles circonstances participez-vous aux *Herbonautes* ? (Vous pouvez préciser dans quels lieux, à quels moments, pour combien de temps vous avez déjà contribué au site, seul(e) ou entouré(e)...))

Chez moi, pas d'horaires particuliers, de 10 mn à 1 heure, voire plus. J'ai démarré ma contribution avec la Mission Ophrys, Le printemps me voit moins assidu, néanmoins je m'efforce de répondre à toutes les alertes 1 à 2 fois par semaine. Je suis seul devant mon ordinateur.

- 3) Comment procédez-vous pour remplir les différents champs (pays, région, date, déterminateur, localisation...) associés aux planches d'herbier qui vous sont proposées ? (Vous pouvez expliquer notamment dans quel ordre vous examinez les différentes zones de la planche scannée, les raisonnements auxquels vous avez recours, les sources d'information que vous consultez, ou encore les personnes et avis que vous sollicitez pour vous aider...)

Je ne regarde que les étiquettes, les plantes séchées ne me sont d'aucun secours, sauf pour déterminer s'il y a des spécimens différents. Je commence par la localisation en m'appuyant sur Google-Maps proposé, et si je ne trouve pas, je cherche sur Géoportail pour les spécimens récoltés en France, sur Wikipedia, sur Géonames, enfin tout ce qu'on peut trouver de sérieux sur Internet, afin de repérer des indices qui nous permettent de découvrir le lieu de récolte. Les récits de voyage, trouvés sur Google nous permettent de progresser et d'éliminer des hypothèses et d'en confirmer d'autres. Ceci est valable aussi pour les dates et les noms des botanistes.

Concernant ceux-ci, j'ai souvent quelque difficulté à ne pas indiquer comme déterminateur le botaniste qui a créé son herbier. La consigne ne m'apparaît pas suffisamment claire à ce sujet.

Le débat entre les contributeurs nous permet d'affiner les propositions et bien souvent d'arriver à un consensus.

- 4) Vous souvenez-vous avoir eu beaucoup de difficultés à déterminer en particulier une localisation, ou un récolteur ? Si oui, pourriez-vous expliquer ce qui posait problème et la façon dont vous êtes parvenu(e) à trouver tout ou partie de l'information cherchée ?

La principale difficulté réside dans le déchiffrement de l'écriture, vient ensuite la différence entre les dénominations du passé et les actuelles. Lors des colonisations les colons ont adapté à leur langage d'origine la toponymie des lieux. Depuis il y a eu décolonisation et redénomination dans la langue indigène, viennent

s'ajouter des erreurs d'orthographe. Si en plus l'écriture est arabe, cyrillique, chinoise,.. !!! cela se complique. A ce titre, les logiciels de traduction automatique proposés sur Internet me sont bien utiles. Et finalement, c'est par la discussion avec les autres contributeurs que nous approchons de la vérité.

Votre intérêt pour le site

- 5) Que trouvez-vous de particulièrement plaisant dans *Les Herbonautes* ?

Ce qui est gratifiant, c'est de participer à cette aventure en apportant ma contribution. Cela nous force à se remuer les méninges et à ne pas se contenter de l'à peu près. Je retrouve la rigueur de la démarche scientifique à ce modeste niveau. Et puis on visite de nombreuses régions et pays, on réapprend sa géographie et un peu d'histoire, en constatant les bouleversements de ces 100 dernières années.

- 6) Y a-t-il des informations contenues dans certaines planches d'herbier ou certaines étiquettes qui vous intéressent en particulier ?

Non

- 7) Les échanges à propos du site, que ce soit avec des animateurs de missions, ou d'autres personnes, sont-ils une source de motivation importante pour votre participation au site ?

Ces échanges sont indispensables pour affiner les propositions (cf les commentaires ci-dessus)

- 8) Avez-vous le sentiment d'avoir acquis des connaissances ou des savoir-faire grâce à votre contribution au site ? Si oui, qu'avez-vous appris ?

J'ai acquis une certaine méthodologie qui m'a permis de procéder de manière plus rapide et efficace

- 9) Contribuer au site présente-t-il aussi des désagréments ? Merci de bien vouloir les décrire.

Non, car nous sommes libres d'arrêter à tout moment et d'intervenir n'importe quand.

- 10) Quelles sources de motivation vous amènent à continuer à participer aux *Herbonautes* malgré les difficultés ou désagréments que vous avez pu rencontrer ?

Cf réponse n° 5

- 11) Diriez-vous que votre participation au site a changé votre regard sur les herbiers ? Si oui, en quoi consiste ce changement ?

J'avais, lors de mes études, constitué un herbier, mais j'étais loin d'imaginer l'importance de certains herbiers et l'acharnement de certains botanistes à les réaliser, et les moyens que cela supposait.

Vos autres pratiques naturalistes

- 12) Hormis la participation aux *Herbonautes*, avez-vous d'autres activités concernant la botanique ou les sciences naturelles, que ce soit à titre personnel ou professionnel ? Si oui, merci de bien vouloir décrire en quoi elles consistent, et préciser si vous les pratiquez au sein d'une association ou organisation en particulier. (N'hésitez pas à mentionner si vous vous êtes déjà investi(e) dans la constitution d'un herbier ou autre collection naturaliste.)

Je suis ingénieur agro de formation et ai toujours été intéressé par les sciences naturelles et plus particulièrement la biologie animale et végétale. Je n'ai jamais besoin d'avoir recours à cet aspect de ma formation lors de ma vie professionnelle et ce n'est donc que ma curiosité personnelle sur l'observation de la nature qui font que je cherche systématiquement à connaître le nom des plantes et des animaux que je croise.

- 13) Dans les deux dernières années, avez-vous visité une exposition présentant des herbiers ou d'autres collections d'histoire naturelle ? Si oui, laquelle ou lesquelles ?

Non

- 14) Hormis les *Herbonautes*, consultez-vous d'autres sites Web consacrés à la botanique, aux sciences naturelles, à la biodiversité, ou en lien avec des collections naturalistes ? Si oui, quels sont ceux qui apportent le plus à vos activités ? Et qu'apportent-ils effectivement ?

Je dispose bien entendu de la Flore Bonnier, et j'ai mis en favoris sur mon ordinateur de nombreux sites de reconnaissance des plantes et des animaux. L'intérêt est de pouvoir croiser les recherches de détermination et de disposer de multiples photos qui limitent l'incertitude.

Ci-après quelques uns des sites que je consulte :

<http://www.jpdugene.com/botanique.htm>

<http://www.flora-electronica.com/content/flore-bonnier>

<http://mycorance.free.fr/index.htm>

<http://www.oiseaux.net/>

<http://www.luontoportti.com/suomi/fr/>

Informations complémentaires

15) Hormis votre participation aux Herbonautes, contribuez-vous à d'autres programmes de science participative, ou à des sites Web collaboratifs (blog, encyclopédie en ligne, forum, journalisme participatif, espace collaboratif de travail...) ?

Non

16) Avez-vous des compétences en programmation informatique ? Si oui, pouvez-vous précisez les réalisations ou projets dans lesquels vous êtes investis ?

Non

17) Pour nous permettre de mieux appréhender les caractéristiques des usagers du site, quelques renseignements sur votre situation socio-professionnelle seraient également utiles. Si vous n'y voyez pas d'inconvénient, vous pouvez nous laisser pour finir quelques indications concernant votre quotidien (Où habitez vous ? ~~Quel est votre âge ? Quelle est votre situation professionnelle et familiale ?~~).

ville mitoyenne de Bayonne., au carrefour de la vallée de l'Adour, de la forêt des Landes, du Pays Basque et de l'océan : un magnifique terrain de jeux.

N'hésitez pas à ajouter ci-dessous d'autres informations sur votre expérience d'herbonaute qui n'auraient pas trouvé leur place dans le questionnaire et à indiquer si certaines questions vous ont posé problème.

Merci encore !

E. Archivage des documents de l'enquête de terrain

Les notes, transcriptions et documents complémentaires issus des entretiens et des différentes situations d'observation de l'enquête font l'objet d'un dépôt pérenne sur la plateforme d'archivage et valorisation des données de la recherche en sciences humaines et sociales de l'infrastructure Huma-Num.

URL de consultation du corpus des documents issus de l'enquête :

<https://herbier-num.nakalona.fr/>

Les données de l'enquête sont déposées et documentées sur l'entrepôt Nakala et accessibles à l'URL : <https://www.nakala.fr/page/collection/11280/58f69d2a>

le cnam

Lisa CHUPIN

**Enjeux communicationnels de la
conception de dispositifs de
médiation documentaire
augmentée pour les herbiers
numérisés**

Dicen *IDF*
Dispositifs d'Information et de Communication à l'Ère Numérique - Paris 8 de France

Résumé : Nous décrivons les mutations des formes de communication médiatisées par les collections de plantes séchées dans le contexte de leur numérisation. Les herbiers ont servi de paradigme pour penser le rôle des artefacts matériels dans la production de données scientifiques standardisées mobilisant des acteurs aux compétences hétérogènes, y compris des amateurs. Nous abordons leur numérisation en faisant l'hypothèse que des dispositifs numériques de médiation documentaire peuvent reconstituer leurs spécificités. Cette hypothèse ouvre la perspective d'une médiation des collections numérisées capable d'intégrer des objectifs d'exploitation scientifique et de valorisation patrimoniale qui pourrait être utile à d'autres corpus.

L'augmentation des fonctionnalités des bases de données au-delà de la seule consultation des collections était l'un des objectifs de l'Infrastructure Nationale en Biologie Santé e-Recolnat, en plus de la numérisation des collections des institutions partenaires. Ce projet a conduit à la participation d'un laboratoire de sciences de l'information et de la communication, Dicen-IDF, à l'infrastructure, avec une mission de conception et de développement d'applications destinées à la valorisation des collections numérisées. Elle a donné lieu à une démarche de recherche-action articulant des phases d'observation, de conception et de test des dispositifs développés.

À partir d'un cadre d'analyse de l'activité des usagers comme flux de transactions, notre enquête a mis en évidence différentes formes de communication médiatisées par les collections numérisées, selon les types de dispositifs utilisés et les spécialités des usagers. Les collections numérisées facilitent la progression concomitante d'inventaires locaux. Les spécimens numérisés jouent un rôle qui va du simple intermédiaire facilitant l'accès à la planche à celui de document de référence. La conception du logiciel a montré la nécessité d'articuler médiation documentaire et médiation de la donnée. Dans une plateforme proposant la transcription des étiquettes des spécimens à des internautes non spécialistes, se déploient de nouvelles formes de médiation scientifique, entre animateurs impliqués dans des activités de conservation et de recherche et contributeurs réguliers qui s'en font les relais.

Mots-clés : Médiation documentaire, Numérisation, Collection scientifique, Recherche-action, Crowdsourcing, Sciences participatives

Abstract : We describe evolving patterns of communication mediated by naturalistic collections in the context of their digitisation. Natural history collections have been a reference for theorising the role played by material artifacts in the process of standardised data production lead by multiple and heterogenous actors, including amateurs. We study the digitisation of naturalistic collections with the aim of validating the hypothesis that the characteristics of their documents and the principles of their organisation can be reproduced with a digital device. If so, digitised specimens could be used at each step of the scientific communication involving uses of naturalistic collections: sometimes collected during leisure botany practices, specimens are a source of information used during the scientific study, and in the end they are the medium for publishing the results of the study, using a new label ; over time, they become a scientific heritage documenting the history of knowledge creation. The digital reproduction of these functions of herbaria creates a model for digital mediation enabling scientific collections to be considered as both scientific information and scientific heritage. This model could be used for other kinds of digitised collections.

Keywords : Document and mediation, Digitization, Scientific collection, Action research, Crowdsourcing, Citizen science