



HAL
open science

Intégration de données guidée par une ontologie : Application à l'assistance personnalisée aux personnes âgées

Rahma Dandan

► **To cite this version:**

Rahma Dandan. Intégration de données guidée par une ontologie : Application à l'assistance personnalisée aux personnes âgées. Base de données [cs.DB]. Université Paris-Nord - Paris XIII, 2022. Français. NNT : 2022PA131024 . tel-03945528

HAL Id: tel-03945528

<https://theses.hal.science/tel-03945528>

Submitted on 18 Jan 2023

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.

Intégration de données guidée par une ontologie

Application à l'assistance personnalisée aux personnes âgées

THÈSE DE DOCTORAT

présentée par

Rahma DANDAN

pour l'obtention du grade de
DOCTEUR EN INFORMATIQUE

soutenue le 21 janvier 2022 devant le jury d'examen composé de :

PERNELLE Nathalie, Université Sorbonne Paris NordPrésidente du jury
LAMOLLE Myriam, IUT de Montreuil Rapporteure
BRINGAY Sandra, Université de Montpellier Rapporteure
AZZI Rabia, Université de Bordeaux Examineur
ABEL Marie-Hélène, Université de Technologie de Compiègne Examinatrice
DESPRES Sylvie, Université Sorbonne Paris Nord Directrice de thèse
NOBECOURT Jérôme, Université Sorbonne Paris Nord Co-encadrant de thèse

A Gilles Guezennec,
A ma famille,

Remerciements

Enfin, cette partie qui me tenait à coeur ! De manière conventionnelle, elle apparaît au début de la thèse, mais je ne l'écris qu'à la fin comme pour marquer le coup d'une fin de thèse pleine de rebondissements et de leçons de vie. Je l'écris surtout pour vous remercier d'avoir été présents à chaque étape et challenge que j'ai dû relever tout au long de ma thèse. Toujours avec bienveillance et sympathie, vous m'avez guidé pour trouver et suivre mon chemin. Je vous en suis sincèrement reconnaissante.

Je souhaite remercier en premier lieu Sylvie Despres, ma directrice de thèse, pour m'avoir proposé de réaliser une thèse auprès d'elle et de travailler sur un sujet qui mêle des domaines qui me passionnent tout autant, l'informatique et la santé. Je lui suis reconnaissante pour le temps conséquent qu'elle m'a accordé, pour sa patience, sa franchise, sa rigueur scientifique et pour toutes les opportunités d'enseignements et d'élargissement de mes champs de compétences. J'ai beaucoup appris à ses côtés et je lui adresse ma gratitude pour tout cela. Je ne pourrais pas citer tout ce que nos échanges m'ont apporté, je n'oublie rien.

Je tiens également à remercier Jérôme Nobécourt, mon co-encadrant, pour son apport scientifique, ses conseils avisés et son écoute durant ma thèse. Je lui suis reconnaissante pour toute l'attention apportée lorsque je le demandais. Je n'oublie pas chaque conseil, chaque échange et tout ce que cela a pu avoir comme conséquence sur mon travail.

J'adresse également de sincères remerciements à Mme Myriam Lamolle et Mme Sandra Bringay, rapporteuses de mon travail, pour l'évaluation, les suggestions et les commentaires qu'elles ont apportés à l'amélioration de ce travail.

Je remercie également Mme Marie-Hélène Abel, Mme Nathalie Pernelle et M. Rabia Azzi, membres de mon jury, d'examiner mon travail et tout ce que vous m'apportez pour l'améliorer.

J'associe à ces remerciements Mme Marie-Christine Jaulent, Directrice du LIMICS, pour son accueil au sein du laboratoire et ses retours pertinents sur mon travail lors des différentes présentations réalisées.

Merci à tous les membres du LIMICS pour tous les moments de partage que nous avons eu et tous les échanges durant ma thèse, merci Karima Sedki, Pascal Vaillant, Chan Le Duc, Jean-Batiste Lamy, Rim Djedidi, Mélanie Courtine. Merci Corinne Fabbro pour ta bonne humeur et ces moments de détente que nous avons partagée.

Merci à toi Bernard Virginie, j'ai tellement apprécié nos moments de rire, de détente, tes conseils et tout le soutien que tu m'as apporté pendant cette période.

A toi, Gilles Guezennec, qui m'a laissé une empreinte indélébile sur ma thèse. Je n'oublierai jamais tous les enseignements que tu m'as apportés. Tu fais partie à jamais de ceux qui m'ont inspiré pour mener et venir à bout de ce projet.

Merci à toi Rosy Tsopra pour avoir animé mes deux premières années de thèse et de n'avoir cessé d'être là.

Merci à toi Gaoussou Camara qui n'a cessé d'être un soutien. J'ai tellement apprécié tous ces moments partagés, tes retours d'expériences, tes enseignements et ta bienveillance.

Merci à toi Rabia Azzi pour avoir toujours été là et pour tout ce que tu m'as apporté. Merci pour ton partage d'expériences et de compétences tout au long de ma thèse, merci pour ton écoute même lorsque je n'avais pas grand-chose d'intéressant à raconter.

Je remercie également tous les stagiaires, Kamar, Dimitri, Abdelmalek, Leonard, Huihui, Yanis, Florent, Mickaël, Steeven, Massilia, Chaofu, Jose, Massinissa que j'ai eu le plaisir d'encadrer durant ma thèse et qui ont contribué au développement de certains modules du Serious Game.

Merci à mes parents, mon frère, ma soeur, mes petits-neveux, mes "mitochondries" pour votre amour inconditionnel qui me donne des ailes et sans qui ce travail n'aurait jamais vu le jour.

Enfin, un grand merci à tous ceux qui ont été dans l'ombre aux yeux de tous, mais pas aux miens. Je ne vous citerais pas un par un, mais vous vous reconnâtriez sans l'ombre d'un doute !

Sommaire

Préambule	ix
Synopsis de la thèse	ix
Guide de lecture du manuscrit de thèse	1
Introduction	1
Objectifs	5
Contributions de la thèse	5
Liste des publications scientifiques internationales	6
Structuration de la thèse	7
Chapitre 0 Préliminaires	9
0.1 Ontologies et Web sémantique	10
0.2 RDF et SPARQL	11
0.3 OWL2	13
0.4 Méthodologies d'ingénierie d'ontologie	16
0.5 Ontologies et principes FAIR	19
0.6 Graphes de connaissances	21
I Etat de l'art et positionnement	23
Chapitre 1 Personnalisation de la recommandation	24
1.1 Qu'est-ce que la personnalisation ?	25
1.2 Aperçu des approches de personnalisation	26
1.3 Les phases de la personnalisation	28
1.4 Positionnement	31
Chapitre 2 Application : prise en charge personnalisée pour les per- sonnes âgées	33
2.1 Introduction	34
2.2 Bien vieillir : un objectif de santé publique	34
2.3 Outils de collecte de données en santé	42

II Contributions	48
Chapitre 3 Évolution des ressources sémantiques modulaires	49
3.1 Introduction	51
3.2 Synthèse des approches d'évolution existantes	52
3.3 Évolution de l'ontologie MIAM	55
3.4 Conclusion	80
Chapitre 4 Support pour la collecte de données et la présentation de la recommandation	82
4.1 Introduction	83
4.2 Conception des questionnaires en ligne	83
4.3 Apport des <i>Serious Games</i>	84
4.4 Approches de conception du <i>Serious Games</i>	86
4.5 Conception du <i>Serious Games</i>	89
4.6 Construction d'un jeu de données fictif	95
4.7 Conclusion	96
Chapitre 5 Intégration de données guidée par des ressources sémantiques	98
5.1 Introduction	99
5.2 Méthodologie proposée : l'approche DIKG2	101
5.3 Interrogation des ressources sémantiques	106
5.4 Implémentation de l'approche DIKG2	109
5.5 Conclusion	109
Conclusion et perspectives	112
Bibliographie	129
Glossaire	130
Liste des abréviations, des sigles et des symboles	133
Table des matières	136
Index des figures	139
Index des tableaux	140
Résumé	141
Abstract	142

Préambule

Synopsis de la thèse <input type="checkbox"/>			
	Problématiques	Objectifs	Enjeux/Résultats
Chapitre 1	Comment améliorer la personnalisation des recommandations ?	Exploiter les technologies du Web sémantique et les approches d'ingénierie des connaissances pour personnaliser les recommandations	Déterminer une stratégie pour améliorer la personnalisation s'appuyant sur des ressources sémantiques
Chapitre 2	Comment participer à la prévention de la dénutrition et de la sédentarité des personnes âgées fragilisées et dépendantes ?	Exposer les stratégies actuelles de prévention et de prise en charge des personnes âgées dénutries et sédentaires	Identifier les outils de collecte pour répondre à un besoin de personnalisation de prise en charge des personnes âgées
Chapitre 3	Quelle approche adopter pour faire évoluer une ressource sémantique modulaire de domaine (MIAM) ?	Exploiter les ressources textuelles de recommandation pour faire évoluer MIAM en tenant compte des besoins nutritionnels et d'activité des personnes dénutries et des pathologies spécifiques liées au vieillissement	Conception de l'ontologie modulaire ONAFE en réutilisant certains modules de MIAM et en développant de nouveaux modules nécessaires pour personnaliser les recommandations nutritionnelles et d'activités physiques
Chapitre 4	Quel support permet d'améliorer la collecte des données et la personnalisation des recommandations ?	Explorer les questionnaires en ligne puis les Serious Game comme support de collecte de données	Conception d'un Serious pour l'aide à la collecte de données, au coaching personnalisé, et à l'acceptation de la recommandation
Chapitre 5	Comment structurer, intégrer et rendre interopérable des données hétérogènes et semi-structurées ? Comment exploiter les ontologies pour intégrer des données et construire un graphe de connaissances personnel de santé (PHKG) ?	S'appuyer sur le schéma de la base de données pour faciliter l'intégration de données sémantiques quelque soient leurs sources et leurs caractéristiques Alignement entre les PHKG et ONAFE pour construire le profil de santé Exploiter les PHKG pour l'élaboration de recommandations personnalisées	Approche permettant l'intégration de données semi-structurées issues d'une base de données relationnelle Création et enrichissement de graphe de connaissances à partir des entités de ONAFE Inférence de nouvelles connaissances pour la personnalisation de la recommandation

TABLEAU 1 – Tableau synoptique du manuscrit de thèse

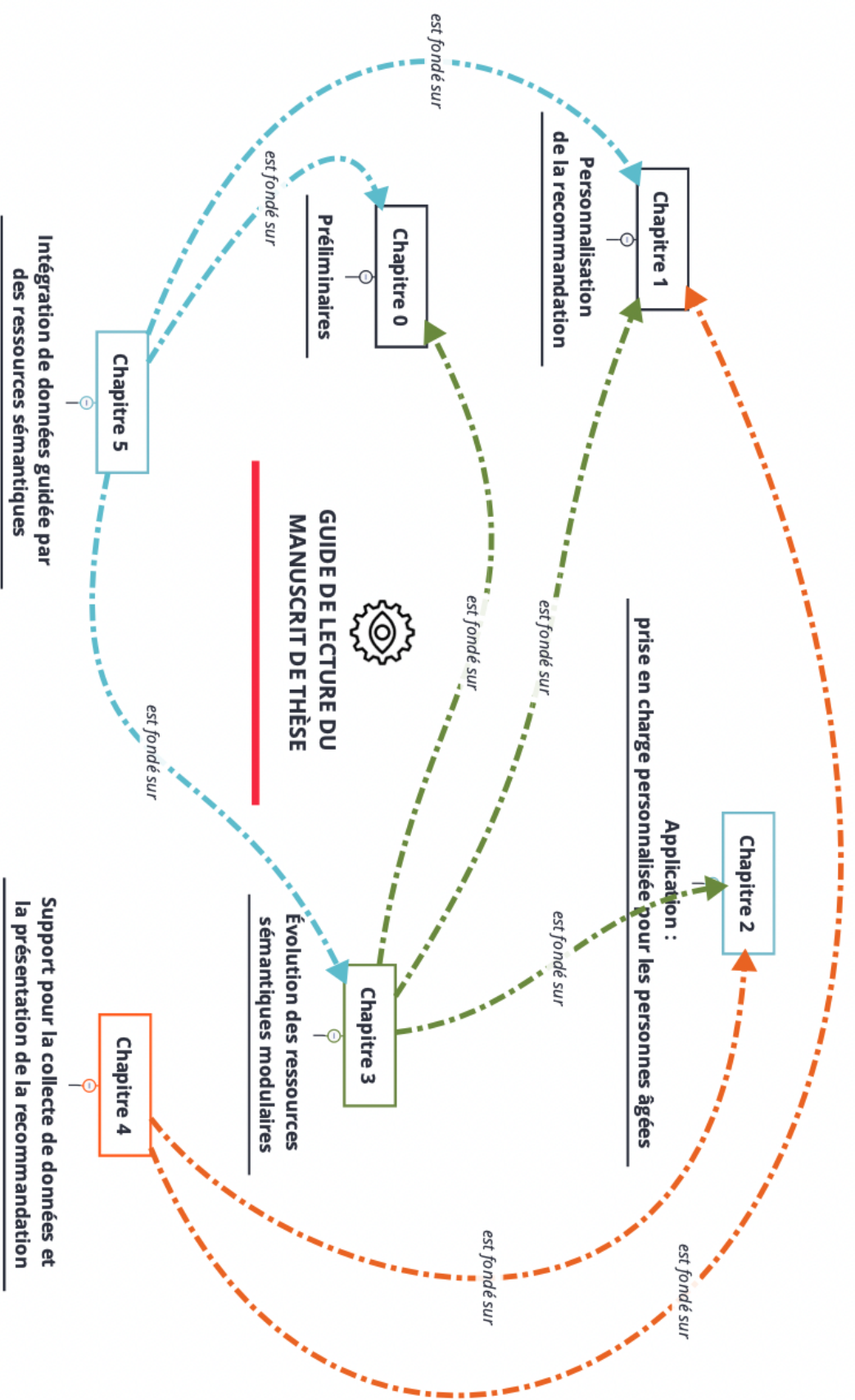


FIGURE 1 – Guide de lecture du manuscrit de thèse

Introduction

Depuis quelques années, les technologies de l'information sur la santé (TIS) participent aux défis de l'informatisation de la santé dans l'objectif ultime de faire évoluer les pratiques et le développement d'applications et de dispositifs intelligents en e-santé. Les TIS sont définies comme « l'application du traitement de l'information impliquant à la fois du matériel informatique et des logiciels qui traitent du stockage, de la récupération, du partage et de l'utilisation d'informations, de données et de connaissances sur les soins de santé pour la communication et la prise de décision » (YEN et al., 2017).

L'évolution des TIS et de l'intelligence artificielle (IA) a permis l'émergence des systèmes dits à intelligence ambiante. Ces systèmes fournissent des solutions centrées sur l'utilisateur. Les programmes d'assistance à l'autonomie à domicile (AAL, *Active Assisted Living*) font partie des TIS qui suscitent un intérêt croissant dans le domaine de la recherche en IA. Ils ont l'avantage d'intégrer le comportement de l'utilisateur, les besoins, le contexte et les exigences des utilisateurs pour leur apporter une assistance appropriée en les aidant à vivre une vie pleinement indépendante. Ils trouvent toute leur utilité dans le cas d'assistance, d'aide à domicile, de réadaptation et de soutien physique des personnes âgées dépendantes.

Le contexte applicatif sur lequel s'appuie le travail présenté est celui de la prévention de la dénutrition et de la sédentarité chez les personnes âgées. Ils représentent les principaux facteurs responsables d'un vieillissement dit symptomatique et constituent actuellement un problème de santé et de santé publique majeur. Nous nous intéressons à cette population car elle présente des besoins et des spécificités qu'il est possible d'identifier et qui sont intéressants à traiter dans le cadre de la thèse. En effet, le vieillissement démographique est un phénomène qui ne cesse d'augmenter. Quelle que soit l'étude statistique, toutes les projections démographiques prédisent la même tendance : à l'horizon 2050-2070, une personne sur trois sera âgée de plus de 60 ans contre une sur cinq au début du siècle (INSEE, 2020). La population française sera par conséquent plus âgée. Dans ce contexte, l'un des enjeux majeurs de notre société actuelle est de permettre à tout un chacun de « bien vieillir ». Toutefois, quel que soit l'âge ou le niveau de dépendance de la personne, une mauvaise alimentation et une activité physique inférieure aux recommandations augmentent le niveau de dépendance et diminuent la qualité de vie.

La dénutrition représente un problème particulièrement important chez les personnes âgées notamment en institution (EHPAD, centre hospitalier, etc.). La dénutrition touche 25% des plus de 70 ans vivant seules et une personne sur deux des personnes âgées

hospitalisées (SANTÉ (HAS), 2020). Un certain nombre de pathologies se révèlent être de puissants facteurs de risque de dénutrition chez la personne âgée tels que la dépendance, la dépression et les troubles cognitifs, de même que les maladies chroniques telles que les cancers et insuffisances d'organes sévères. Face à ces pathologies, d'autres symptômes peuvent apparaître altérant ainsi la perception sensorielle (goût, odorat, etc.) et la qualité sensorielle (saveur, structure, consistance, etc.) des aliments chez ces personnes âgées (MINGIONI, 2016). Ainsi, de nombreuses questions se posent quant à l'optimisation de la prise en charge nutritionnelle de la personne âgée. La sédentarité, quant à elle, a été identifiée comme étant le quatrième facteur de risque des maladies non transmissibles (ORGANIZATION, 2009). Actuellement, une personne âgée sur trois ne suit pas les recommandations en termes d'activité physique (SANTÉ PUBLIQUE FRANCE, 2020). En raison des changements physiologiques, économiques ou sociaux, il n'est pas toujours évident pour une personne âgée de maintenir son activité.

Bien que chaque individu cherche à rester en bonne santé tout au long de son existence, la détection d'une dégradation de son propre état de santé n'est pas si évidente pour réagir avant que son état de santé ne devienne irréversible (DREVET et GAVAZZI, 2019). Par exemple, suite à la récente crise sanitaire due à la COVID-19 et des mesures restrictives liées au confinement, l'état de santé des personnes âgées s'est fortement aggravé. Les derniers chiffres publiés montrent que 4 à 10 % des personnes âgées à domicile sont dénutries, 40% des personnes âgées sont hospitalisées pour cause de dénutrition, 100% des malades âgés positifs à la COVID-19 sont à risque de dénutrition et 6 personnes sur 10 ont été sédentaires durant cette période. Ces chiffres témoignent de l'impact direct du confinement sur la nutrition, l'activité physique et l'état de santé de la population de tout âge et plus spécifiquement des personnes âgées¹ (LAËTTIA, 2020).

Au cours des dernières années, la prise de conscience sur la nécessité d'améliorer la nutrition et l'activité physique des personnes âgées s'est accrue parmi les politiques et les professionnels de la santé (SOLIDARITÉS ET DE LA SANTÉ, 2019). Les mesures mises en place par les politiques de santé et adressées aux aidants et aux personnes concernées pour lutter contre la dénutrition et la sédentarité sont globales et disponibles sous forme de brochures (généralement sous format PDF). Cependant, ce type de publication ne permet pas de rendre les recommandations incitatives et faciles à intégrer par le public visé et leur entourage. En milieu institutionnel, malgré les efforts fournis, les mesures actuelles visant à suggérer des produits et des exercices aux besoins des personnes âgées ne permettent pas de satisfaire et de répondre aux besoins et aux préférences sensorielles des personnes âgées (POUYET et al., 2015). Ceci peut s'expliquer par des mesures non adaptées aux besoins individuels permettant une alimentation et une activité physique personnalisées. Ces éléments ne sont pas toujours évidents à prendre en compte par tous les acteurs de la santé en raison notamment de la diversité des pathologies et des besoins individuels. Ainsi, la considération des facteurs déclenchant une meilleure prise en compte et une personnalisation des recommandations est nécessaire pour prévenir la dénutrition et la sédentarité.

Au-delà des campagnes gouvernementales sur l'alimentation équilibrée et la pratique de l'activité physique, qui incitent par exemple à consommer 5 fruits et légumes par

1. www.has-sante.fr/jcms/p_3165944/fr/diagnostic-de-la-denu-trition-de-la-personne-agee-note-de-cad-rage

jours ou à exercer une activité physique régulière, d'autres solutions doivent être proposées pour répondre aux problèmes spécifiques des personnes âgées. Parce que les besoins de tout un chacun varient et évoluent, les solutions proposées pour prévenir la dénutrition et la sédentarité doivent s'adapter à l'évolution des nouvelles technologies pour répondre aux problèmes spécifiques de chaque individu. La personnalisation exploitant ces nouvelles technologies devient ainsi une solution majeure en santé.

L'enjeu de la personnalisation de l'alimentation et des activités est de fournir aux personnes âgées, les clés d'une nutrition et d'une activité physique répondant à leurs besoins spécifiques et les amenant vers un « mieux vieillir ». La majorité des travaux sur les recommandations se sont principalement concentrés sur le développement des algorithmes pour améliorer la précision des recommandations. Jusqu'à présent, très peu d'études ont été effectuées pour permettre aux utilisateurs d'interagir avec le système notamment pour améliorer la collecte de données et la visualisation des recommandations. Or, la personnalisation des recommandations ne peut aboutir efficacement si l'utilisateur est exclu du processus de conception. Pour approcher au mieux de cette personnalisation, notre stratégie consiste à faire participer la personne âgée, au moyen d'interface Web (questionnaire en ligne, etc.), pour élaborer son propre plan alimentaire et d'activité suivant les recommandations nutritionnelles actuelles. Faire interagir l'utilisateur avec le système permet d'améliorer l'intégration de ses besoins et de l'impliquer dans la collecte de ses données.

La personnalisation dépend en grande partie de la capacité des systèmes à intégrer, partager et à relier les données spécifiques sur l'utilisateur tout en tenant compte des connaissances du domaine. La gestion et l'intégration de données représentent jusqu'à ce jour un réel défi en e-santé. Ce défi se complexifie lorsque les données sont de sources et de formats hétérogènes et que le domaine d'intérêt est d'une complexité multifactorielle. La personnalisation représente également un véritable challenge lié au traitement, à la représentation et au raisonnement sur des quantités de données dynamiques et hétérogènes, diffusées sur le Web, sur les réseaux sociaux et dans l'Internet des objets (Internet of things).

Au cours des deux dernières décennies, l'intégration de données s'appuyant sur la sémantique et plus précisément sur les ontologies ont largement contribué à fournir des solutions efficaces en termes de représentation, de partage et de réutilisation des connaissances du domaine. Les ontologies sont fondamentales pour donner du sens aux données et assurer la cohérence entre les données à intégrer et les connaissances du domaine. Elles permettent de représenter et de définir formellement les connaissances (concepts et les relations entre ces concepts) du domaine étudié. En plus d'améliorer l'exploitation et la représentation des informations provenant de différentes sources, elles permettent de fournir un vocabulaire partagé et formel des différents concepts du domaine d'intérêt et de favoriser la découverte de nouvelles connaissances implicites et favorise l'interopérabilité des données entre différentes applications et systèmes à base de connaissances.

Dans le domaine de la représentation des connaissances, les graphes de connaissances ou "*Knowledge Graphs*" (KG) sont de plus en plus utilisés notamment auprès de la communauté du Web sémantique et dans les projets de « Web des données » et des « données ouvertes liées » (*Linked Open Data* (LOD)). S'appuyant sur les caractéristiques des techniques du Web sémantique et ils représentent également un choix intéressant

pour répondre aux exigences de représentation, de structuration et d'intégration de données. La combinaison des ontologies avec les graphes de connaissances permet de lier et d'améliorer les résultats attendus par les systèmes. Ils représentent également un moyen pour agréger les connaissances de différents domaines pour déduire rapidement et efficacement de nouvelles informations sur les entités.

Objectifs

Pour relever ces défis, nous proposons dans cette présente thèse d'exploiter les technologies du Web sémantique et de l'ingénierie des connaissances pour traiter 4 principaux axes. Ces axes concernent : la construction et l'évolution d'ontologies modulaires, la collecte de données *via* des questionnaires en ligne, l'intégration sémantique de données *via* des graphes de connaissances et la construction d'un profil utilisateur *via* l'exploitation des ressources sémantiques construites (Figure 2). Ces 4 axes sont respectivement abordés dans les chapitres 3, 4, 5.

Avant d'aborder le travail réalisé sur ces 4 axes, nous présentons en préliminaire les principales ressources et technologies sémantiques, et ce, dans les deux premiers chapitres un état de l'art sur les approches existantes pour la personnalisation des recommandation et les stratégies existantes pour répondre aux besoins nutritionnels et d'activités des personnes âgées.

Pour lever les différents verrous identifiés dans cette thèse (Figure 2), le travail réalisé consiste à proposer des solutions et des moyens s'appuyant sur les technologies du Web sémantique et d'ingénierie des connaissances pour permettre de construire un profil santé de la personne âgée après l'intégration de ses données. A partir du profil de santé construit, l'idée consiste à induire par raisonnement de nouvelles connaissances et d'associer à chaque personne la recommandation qui lui correspond selon son profil de santé.

L'objectif du premier axe consiste à faire évoluer une ressource ontologique modulaire de domaine (MIAM) pour répondre aux problèmes spécifiques nutritionnels et d'activités physiques rencontrés chez les personnes âgées. Dans cette thèse, l'ontologie conçue sert de support pour :

- la description du domaine de la nutrition et de l'activité physique ;
- la représentation du profil de la personne âgée à partir des données collectées ;
- l'intégration des données collectées ;
- la construction d'un graphe de connaissances personnel de santé (PHKG) permettant la construction du profil de la personne âgée à partir des nouvelles inférences obtenues par raisonnement ;
- l'élaboration de règles exploitant les ressources sémantiques créées afin de suggérer des régimes nutritionnels et des programmes d'activité selon le profil de santé construit.

L'objectif du second axe est d'explorer les questionnaires en ligne comme stratégie de collecte de données *via* une interface utilisateur exploitant les questionnaires de dépistage et d'auto-évaluation.

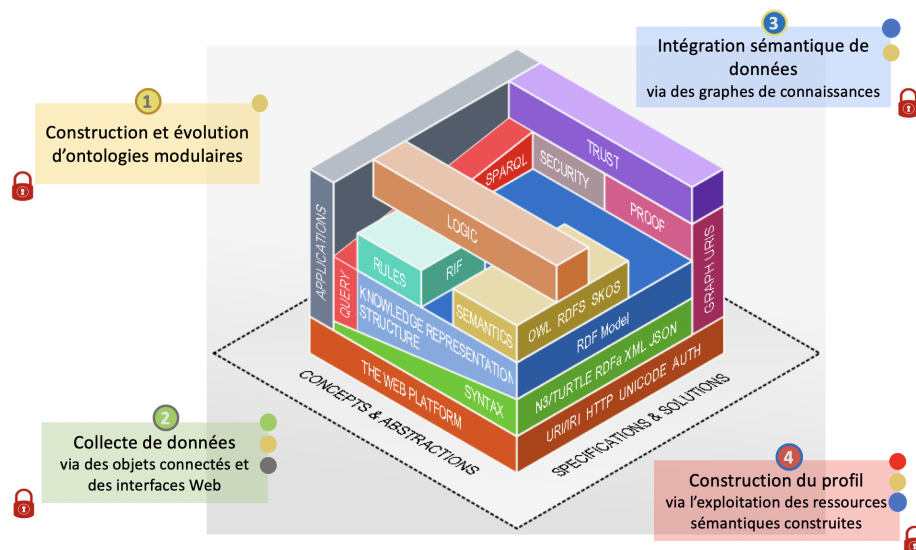


FIGURE 2 – Les axes sur lesquels s’appuient les objectifs de la thèse (d’après la pile des technologies et des langages du Web sémantique de NOWACK, 2009)

L’objectif du troisième axe est d’élaborer une approche d’intégration sémantique de données (DIKG2) qui exploite les ontologies modulaires précédemment conçues et les données collectées. En s’appuyant sur le vocabulaire commun entre les questionnaires et celui des entités de l’ontologie évoluée, l’approche DIKG2 permet de construire un graphe de connaissances personnel de santé (PHKG) correspondant au profil santé de la personne âgée. Ce profil est utilisé pour l’inférence de nouvelles connaissances et la personnalisation de la recommandation.

Contributions de la thèse

La présente thèse a permis d’apporter les contributions suivantes :

- Évolution d’ontologies modulaires pour la construction de recommandations nutritionnelles et d’activités physiques personnalisées et adaptées aux besoins des personnes âgées (verrou 1, chapitre 3).
- Exploration des outils et des supports technologiques pour répondre à un besoin de collecte de données et de personnalisation de la prise en charge des personnes âgées (verrou 2, chapitre 2 et 4).
- Intégration de données semi-structurées issues d’une base de données relationnelle corrélée à l’exploitation des ressources sémantiques (verrou 3, chapitre 5).
- Construction de PHKG correspondant au profil de santé à partir de ressources sémantiques (verrou 4, chapitre 5).

Liste des publications scientifiques internationales

Le travail présenté s'appuie sur des publications scientifiques internationales produites durant la période de la thèse et sont citées au début de chaque chapitre :

1. **Dandan R.**, Despres S. : "*Evolution of Modular Ontology : Application to Personalization*", Terminology & Ontology : Theories and applications, TOTh 2021, Chambéry, France, 2021. (A paraître)
2. **Dandan, R.**, Despres, S. : "*DIKG2 : A Semantic Data Integration Approach for Knowledge Graphs Generation from Web Forms*", 34th International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems, IEA/AIE 2021, Lecture Notes in Computer Science 12798, Springer, p. 255-260, Kuala Lumpur, Malaysia, 2021. (www.springerprofessional.de/en/dikg2-a-semantic-data-integration-approach-for-knowledge-graphs-/19370986)
3. **Dandan R.**, Despres S., Nobecourt J. : "*OAFE : an Ontology for the Description of Elderly Activities*". 14th International Conference on Signal-Image Technology & Internet-Based Systems, SITIS 2018 (IEEE), pp. 396–403, Gran Canaria, Italy, 2018. (<https://ieeexplore.ieee.org/document/8706249>)
4. **Dandan R.**, Despres S., Nobecourt J. : "*OAPA : Une Ontologie pour la description de l'Activité des Personnes Agées*", 7ème Journée Francophones sur les Ontologies, JFO 2018, pp.16-27, 2018.

Enfin, le rapport ci-dessous est issu d'une participation à l'école d'été ISWS 2018.

- Ghor, T. A. ; Agrawal, E. ; Alam, M. ; Alqawasmeh, O. ; D'amato, C. ; Annane, A. ; Azzam, A. ; Berezovskyi, A. ; Biswas, R. ; Bonduel, M. ; Brabant, Q. ; Bucur, C.-i. ; Camossi, E. ; Carriero, V. A. ; Chari, S. ; Fraga, D. C. ; Ciroku, F. ; Cochez, M. ; Curien, H. ; Cutrona, V. ; **Dandan, R.** ; Dess, D. ; Carlo, V. D. ; Djebri, A. E. A. ; Erp, M. V. ; Falakh, F. M. ; Izquierdo, A. F. ; Futia, G. ; Gangemi, A. ; Gasperoni, S. ; Grall, A. ; Heling, L. ; Henri, P. ; Herradi, N. ; Issa, S. ; Jozashoori, S. ; Juniarta, N. ; Kaffee, L.-a. ; Keles, I. ; Khare, P. ; Kovtun, V. ; Leone, V. ; Li, S. ; Lieber, S. ; Lisena, P. ; Makhalova, T. ; Marinucci, L. ; Minier, T. ; Moreau, B. ; Loustaunau, A. M. ; Nandini, D. ; Ozdowska, S. ; Moura, A. P. D. ; Padhee, S. ; Palma, G. ; Jimnez, P. D. P. ; Presutti, V. ; Reda, R. ; Rizza, E. ; Rosales-mndez, H. ; Rudolph, S. ; Sack, H. ; Sciallo, L. ; Simanjuntak, H. ; Stomeo, C. ; Thanapalasingam, T. ; Tietz, T. ; Varanka, D. ; Vidal, M.-e. ; Wolowyk, M. ; Zocholl, M. : *Linked Open Data Validity – A Technical Report from ISWS 2018*, In : CoRR abs/1903.1255, **2019**. (<https://arxiv.org/abs/1903.12554>)

Structuration du manuscrit

Le manuscrit de cette thèse s'articule autour de six chapitres :

- **Chapitre 0 – Préliminaires**

Dans ce chapitre préliminaire, nous présentons les principales ressources sémantiques exploitées dans le travail réalisé. Nous présentons également leur langage, les principes et les méthodologies sur lesquels nous nous appuyeront au cours de notre étude concernant l'intégration sémantique de données.

- **Chapitre 1 – Personnalisation de la recommandation**

Dans ce chapitre, nous présentons un état de l'art des principales approches de recommandation ainsi que l'intérêt d'exploiter les ressources sémantiques pour répondre aux challenges de la personnalisation.

- **Chapitre 2 – Application : prise en charge personnalisée pour les personnes âgées**

Nous présentons à travers ce chapitre les problématiques et les stratégies préventives liées à la dénutrition et à la sédentarité afin de cibler les principaux challenges liés à la personnalisation et l'intérêt d'exploiter de nouveaux outils pour une prise en charge personnalisée et adaptées aux besoins des personnes âgées.

- **Chapitre 3 – Évolution des ressources sémantiques modulaires**

Dans ce chapitre, nous décrivons l'approche d'évolution de l'ontologie MIAM et la construction de nouveaux modules nécessaires à l'élaboration d'une recommandation personnalisée nutritionnelle et d'activité physique.

- **Chapitre 4 – Support pour la collecte de données et la présentation de la recommandation**

Dans ce chapitre, nous décrivons les différents supports explorés lors de collecte de données, les questionnaires en ligne et les Serious Game. Nous présentons alors des éléments de conception d'un nécessaire à la collecte, à la présentation et à la personnalisation des recommandations.

- **Chapitre 5 – Intégration de données guidée par des ressources sémantiques**

Enfin, dans ce chapitre, nous présentons l'approche d'intégration de données semi-structurées et issues de formulaires Web (DIKG2) et de construction du graphe de connaissances personnel de santé. Ces derniers permettront de générer le profil utilisateur nécessaire à la personnalisation des recommandations.

Le nuage de mots-clés ci-dessous (Figure 3) offre un survol des différentes thématiques et problématiques abordées dans cette thèse.



FIGURE 3 – Nuage de mots reflétant les thématiques et les problématiques abordées dans la thèse

Chapitre 0

Préliminaires

Sommaire

0.1	Ontologies et Web sémantique	10
0.2	RDF et SPARQL	11
0.3	OWL2.....	13
0.4	Méthodologies d'ingénierie d'ontologie	16
0.5	Ontologies et principes FAIR	19
0.6	Graphes de connaissances	21

L'objet de ce chapitre est de présenter les ressources sémantiques exploitées dans le travail réalisé. Nous présentons également leur langage et les principes sur lesquels ils s'appuient dans le cadre de notre étude sur l'intégration sémantique de données.

0.1 Ontologies et Web sémantique

En 1999, Tim Berners-Lee présentait pour la première fois sa vision du Web sémantique : « Le Web sémantique n'est pas un Web séparé, mais une extension du Web actuel, dans lequel l'information reçoit un sens bien défini, permettant aux ordinateurs et aux gens de mieux travailler en coopération » (BERNERS-LEE et al., 2001). Ainsi, le Web sémantique se distingue par une représentation plus significative de l'information pour les humains et les ordinateurs, fournissant une description de son contenu et de ses services sous une forme lisible par machine (PAOLUCCI et al., 2002). Depuis, le Web a évolué d'un Web de documents à un Web de données (BIZER et al., 2009).

Avec l'émergence du Web sémantique, la notion d'ontologie a connu un nouvel essor. Autrefois réservées au domaine de la philosophie, et plus tard au domaine de l'intelligence artificielle (IA), les ontologies représentent désormais l'épine dorsale de la technologie du Web sémantique. Dans la littérature, plusieurs définitions de l'ontologie ont été présentées et ont évolué au fil du temps. GRUBER, 1993 définit une ontologie comme étant « une description (comme une spécification formelle d'un programme) des concepts et relations qui peuvent exister formellement pour un agent ou une communauté d'agents. BERNERS-LEE et al., 2001 la définit comme étant : « une description conceptuelle et formelle des connaissances d'un domaine. Elle offre une description sémantiquement riche des concepts pertinents du domaine d'intérêt de leurs attributs, des relations entre ces concepts et des assertions logiques caractérisant les connaissances du domaine ».

Une ontologie est constituée d'un ensemble de concepts et de relations entre ces concepts qui décrivent formellement un domaine d'intérêt fondé sur un ensemble d'individus (également appelés entités ou objets), des classes d'individus et les relations existant entre ces individus. Les énoncés logiques sur les appartenances des individus à des classes ou les relations entre individus forment une base de faits. Les instructions logiques sont utilisées pour spécifier les connaissances sur les classes et les relations. Ils spécifient les contraintes sur la base de faits et forment la base de connaissances. En fonction de l'usage qui en est fait, une ontologie peut comprendre des faits et des contraintes.

Différents modèles de conceptualisation des connaissances peuvent être distingués selon leur complexité :

- une **taxonomie** est une classification de termes ou de concepts d'un domaine organisé selon une structure hiérarchique.
- un **thésaurus** ajoute des relations sémantiques non hiérarchiques entre les termes et les concepts, telles que les relations d'équivalence et les relations d'association, et d'autres propriétés à chaque concept ou terme, telles que ses étiquettes préférées ou alternatives dans SKOS¹. Les concepts et les termes peuvent également être organisés en collections.
- une **ontologie** est une formalisation plus complexe d'un domaine d'intérêt dans lequel des contraintes sont ajoutées pour affiner la sémantique des entités impliquées et leurs relations.
- une **ontologie peuplée** ajoute la description d'instances (ou d'individus) via des assertions d'appartenance à une classe et des assertions de propriété entre

1. <https://www.w3.org/2004/02/skos/>

instances.

- un **graphe de connaissances** est un moyen de représenter des ontologies peuplées.

Les ontologies s’inscrivent dans la continuité de nombreux travaux sur la représentation des connaissances, réseaux sémantiques, cartes conceptuelles, graphes conceptuels. Ces dernières années, l’ingénierie des systèmes fondés sur les ontologies s’est considérablement développée. Les ontologies sont de plus en plus exploitées pour permettre la représentation, le partage et la réutilisation des informations provenant de différentes sources et de favoriser l’interopérabilité des données entre différentes applications et systèmes à base de connaissances. Elles permettent de fournir un vocabulaire partagé et formel des différents concepts du domaine d’intérêt et de favoriser la découverte de nouvelles connaissances implicites. Le vocabulaire ainsi défini permet d’identifier sans ambiguïté les entités dans les systèmes d’information hétérogènes et de spécifier les relations qui relient ces entités entre elles. L’utilisation de la logique de description au moyen de raisonneurs permet l’inférence de nouvelles connaissances qui n’ont pas été explicitement introduites dans l’ontologie et permettent de vérifier la cohérence de l’ontologie.

0.2 RDF et SPARQL

Le modèle RDF

RDF (*Resource Description Framework*)², développé par le World Wide Web Consortium (W3C), est un modèle standard de données de graphe pour représenter, structurer et échanger des données et des métadonnées en particulier sur le Web.

Une instruction RDF est une structure composée de trois éléments (sujet, prédicat, objet) appelés triplets (Figure 1). Cette structure forme un graphe orienté et étiqueté, où les arêtes représentent le lien nommé entre deux ressources, représenté par les nœuds du graphe. Un triplet est une association entre : sujet, prédicat, objet. Les propriétés (prédicats) relient les ressources entre elles (sujet, valeur). Lorsque le sujet et l’objet sont des ressources, ils peuvent être identifiés par une URI ou être des nœuds anonymes (nœud blanc). L’utilisation de nœuds blancs permet de capturer une certaine forme d’individus inconnus. Le prédicat est nécessairement identifié par un URI. Par exemple, nous représentons l’information « John est dénutri » en RDF comme il suit : (John aPourPathologie Denutrition) où « John » désigne le sujet, « aPourPathologie » désigne le prédicat et « Denutrition » désigne l’objet.

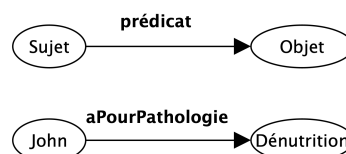


FIGURE 1 – Exemple d’un triplet RDF

2. <https://www.w3.org/RDF/>

L'utilisation d'une syntaxe dite de sérialisation RDF (RDF/XML, Turtle, RDFa, etc.) est nécessaire pour publier des données sur le Web.

RDFa

RDFa (*Resource Description Framework in Attributes*) définit une syntaxe permettant d'ajouter des données structurées dans une page HTML ou n'importe quel document XML. La syntaxe RDFa, conforme au modèle, est constituée d'un ensemble d'éléments et d'attributs conforme au modèle Resource Description Framework (RDF) : des données décrites en RDFa peuvent donc être facilement transformées en données RDF.

L'idée principale derrière la syntaxe de RDFa est de séparer le vocabulaire du code HTML et de faciliter la mise à jour des données existantes sur le Web. Les annotations sémantiques sont intégrées directement au code HTML, ceci en utilisant des attributs existants en HTML (class, id, rel, rev et href) et d'autres attributs propres au langage RDFa (about, property, content, datatype, resource). L'extraction des triplets se fait à la suite de l'analyse de la page Web.

Le langage de requêtes SPARQL

SPARQL³ est destiné à interroger les graphes RDF. C'est un langage de requête et un protocole qui permet de rechercher, d'ajouter, de modifier ou de supprimer des données RDF. Tout comme RDF, SPARQL fait partie de la pile technologique du Web sémantique du W3C. De nombreux langages de requête destinés à interroger les graphes RDF ont été développés ces dernières années. L'utilisation de requêtes SPARQL permet d'interroger les éléments représentés dans l'ontologie et d'inférer des connaissances supplémentaires. La réponse aux requêtes peut être considérée comme étant le processus de calcul des réponses dérivant logiquement de la combinaison de ces connaissances incomplètes et des axiomes d'ontologie. L'utilisation d'une ou de plusieurs ontologies permet de combiner efficacement des données ou des informations provenant de plusieurs sources hétérogènes.

Dans l'exemple qui suit, nous décrivons John et ses pathologies. Les prédicats « foaf » appartiennent au vocabulaire FOAF⁴ et les prédicats « : » correspondent à des espaces de noms non nommés et appartiennent au vocabulaire de notre exemple d'ontologie (ONAFE).

```
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix : <http://www.semanticweb.org/rahmadandan/ONAFE> .

:John foaf:age _:a .
_:a owl:hasValue "69" .
:John :aPourPathologie _:p1 .
:John :aPourPathologie _:p2 .
_:p rdfs:type :Denutrition .
```

3. <https://www.w3.org/TR/2008/REC-rdf-sparql-query-20080115/>

4. <http://xmlns.com/foaf/spec/>

```
_:p      rdf:type :Diabete .  
:Denutrition      rdfs:subClassOf      :Pathologie .
```

Ci-dessous, un exemple de requête SPARQL qui modélise la question : « Quelles sont les pathologies de John ? » Dans le cas des requêtes qui lisent les données de la base de données ou d'un jeu de données (dataset), le langage SPARQL spécifie quatre variantes de requêtes différentes à des fins différentes :

- SELECT : pour extraire les valeurs brutes d'un point de terminaison SPARQL, les résultats sont renvoyés sous forme de tableau.
- CONSTRUCT : pour extraire des informations du point de terminaison SPARQL et transformer les résultats en graphe RDF valide.
- DESCRIBE : pour extraire un graphe RDF du point de terminaison SPARQL.
- ASK : pour fournir un simple résultat "*True*" / "*False*" pour une requête sur un point de terminaison SPARQL.

Les variables sont indiquées par un « ? ». Un triplet se termine par un point. S'il se termine par un point-virgule, le sujet de ce triplet complétera implicitement la paire suivante en un triplet entier. Par exemple, `:aPourPathologie ?pathologie` est l'abréviation de `?x :aPourPathologie ?pathologie`.

```
PREFIX foaf:      <http://xmlns.com/foaf/0.1/>  
PREFIX :         <http://www.semanticweb.org/rahmadandan/ONAFE>
```

```
SELECT ?name ?pathologie  
WHERE  
{ ?x foaf:name ?name .  
  ?x :aPourPathologie ?pathologie }
```

Le processeur de requêtes SPARQL recherche ainsi les ensembles de triplets qui correspondent à ces deux modèles de triplets, liant les variables de la requête aux parties correspondantes de chaque triplet. Le résultat obtenu est obtenu sous forme de liste.

0.3 OWL2

OWL2⁵, une extension de OWL1 est conçu pour faciliter le développement d'ontologies et leur partage sur le Web pour les rendre accessibles aux machines. C'est un langage de représentation des connaissances conçu pour formuler, échanger et raisonner avec des connaissances sur un domaine d'intérêt. Une terminologie fournit un vocabulaire ainsi que des informations d'interrelation. OWL2 est déclaratif, c'est-à-dire qu'il décrit un état de fait de manière logique. Des outils appropriés (appelés raisonneurs) peuvent être utilisés pour déduire des informations supplémentaires sur cet état de fait.

Dans les sections qui suivent, nous proposons quelques exemples pour décrire quelques fonctionnalités de modélisation essentielles qu'offre la syntaxe OWL2.

5. <https://www.w3.org/TR/owl2-syntax/>

Les instances de classes

Un nom de classe est représenté syntaxiquement comme une référence URI. Une description de classe est syntaxiquement représentée comme une instance nommée de owl:Class, une sous-classe de rdfs:Class :

```
<owl:Class rdf:ID="Pathologie"/>
```

Nous commençons par présenter l'individu John. Cette déclaration signifie que l'individu nommé John et déclare que cet individu est une PersonneAgee. Plus techniquement, être une Personne s'exprime en déclarant que John « est une instance de » la classe de toutes les PersonneAgee. En général, les classes sont utilisées pour regrouper des individus qui ont une chose en commun afin de s'y référer.

```
ClassAssertion( :PersonneAgee :John )
```

Hierarchie de classes

Les déclarations de *SubClassOf* sont utilisées pour modéliser des hiérarchies de classes entières en spécifiant les relations de généralisation de toutes les classes dans le domaine d'intérêt. Dans cet exemple, nous pouvons déclarer que les PersonneAgee sont des Personne et que PersonneJeuneSenior est une sous-classe de PersonneAgee.

```
SubClassOf( :PersonneAgee :Personne )
```

En déclarant que John est une PersonneAgee, le raisonneur peut déduire pour chaque individu classé comme PersonneJeuneSenior, qu'il s'agit également d'une PersonneAgee (et par conséquent d'une Personne). Il peut également déduire par classification que PersonneJeuneSenior est une sous-classe de Personne.

Les classes de notre vocabulaire peuvent effectivement faire référence aux mêmes ensembles, et OWL fournit un mécanisme par lequel elles sont considérées comme sémantiquement équivalentes. Par exemple, nous utilisons les termes PersonneAgee et Senior de manière interchangeable, ce qui signifie que chaque instance de la classe PersonneAgee est également une instance de la classe Senior, et vice versa. Deux classes sont considérées comme équivalentes si elles contiennent exactement les mêmes individus. L'exemple suivant indique que la classe PersonneAgee est équivalente à la classe Senior.

```
EquivalentClasses( :PersonneAgee :Senior )
```

Disjonction des classes

Un individu peut être une instance de plusieurs classes. Cependant, dans certains cas, l'appartenance à une classe exclut spécifiquement l'appartenance à une autre. Par exemple, si nous considérons les classes PersonneAgee et PersonneDenutrie, nous savons qu'aucun individu ne peut être une instance des deux classes.

```
DisjointClasses( :PersonneAgee :PersonneDenutrie )
```

Cette « relation d'incompatibilité » entre les classes est appelée « disjonction de classe ». L'information selon laquelle deux classes sont disjointes fait partie des connaissances de base et doit être explicitement indiquée pour qu'un système de raisonnement puisse l'utiliser.

Types de données

Dans de nombreux cas, les individus doivent être décrits par des valeurs de données. Pensez à la date de naissance d'une personne, son âge, son adresse e-mail, etc. À cette fin, OWL fournit un autre type de propriétés, appelées propriétés de type de données. Ces propriétés relient les individus à des valeurs de données (au lieu d'autres individus), et de nombreux types de données XML Schema (XML Schema Datatypes (XSD)⁶) peuvent être utilisés. L'exemple qui suit utilise une propriété de type de données et indique que l'âge de John est de 69 ans.

```
DataPropertyAssertion( :hasAge :John "69"^^xsd:integer )
```

Propriété d'objets

Les entités décrivant de quelle manière les individus sont liés sont appelées propriétés d'objet (*ObjectProperty*). Dans cet exemple, la propriété d'objets permet de déclarer que John a comme pathologie une dénutrition.

```
ObjectPropertyAssertion( :aPourPathologie :John :Denutrition )
```

Restriction de Domain et de Range

Avec ces restrictions, dans le cas où John est déclaré comme ayant une dénutrition, un raisonneur peut déduire que John est une Personne et dénutrition est une pathologie.

```
ObjectPropertyDomain( :aPourPathologie :Personne )  
ObjectPropertyRange( :aPourPathologie :Pathologie )
```

Restriction de propriétés

Une restriction de propriété, appelée « quantification existentielle », définit une classe comme l'ensemble de tous les individus qui sont connectés via une propriété particulière à un autre individu qui est une instance d'une certaine classe. Cela signifie que l'on s'attend à ce que pour chaque instance de *PersonnePeurDeChuter*, il existe au moins une *FreinActivite* et que *PeurDeChuter* soit un membre de la classe *FreinActivite*. Ceci est utile pour capturer des connaissances incomplètes.

```
EquivalentClasses(  
  :PersonnePeurDeChuter  
  ObjectSomeValuesFrom( :aPourFreinActivite :PeurDeChuter )  
)
```

Une autre restriction de propriété, appelée « quantification universelle », est utilisée pour décrire une classe d'individus pour laquelle tous les individus liés doivent être des instances d'une classe donnée. Nous pouvons utiliser la déclaration suivante pour indiquer que quelqu'un est une personne qui a peur de chuter (*PersonnePeurDeChuter*) exactement si tous ses *FreinActivite* sont *FreinActivite*.

6. <https://www.w3.org/TR/2012/REC-xmlschema11-2-20120405/>

```
EquivalentClasses(  
  :PersonnePeurDeChuter  
  ObjectAllValuesFrom( :aPourFreinActivite :FreinActivite )  
)
```

Propriétés de classe avancées

Au moyen des éléments de langage décrits jusqu'ici, des ontologies simples peuvent être modélisées. Afin d'exprimer des connaissances plus complexes, OWL fournit des constructeurs de classes logiques. En particulier, OWL fournit des éléments de langage pour les logiques : et, ou, et non. Les termes OWL correspondants sont empruntés à la théorie des ensembles : (classe) intersection, union et complément. Ces constructeurs combinent des classes atomiques – c'est-à-dire des classes avec des noms – en classes complexes.

L'exemple ci-dessous décrit une `FrequenceCardiaqueNormale` et les valeurs incluses. `FrequenceCardiaqueNormale` fait partie des `SignesVitaux` et ses valeurs sont incluses entre 60 et 70. L'intersection de deux classes se compose exactement des individus qui sont des instances de ces deux classes.

```
EquivalentClasses( :FrequenceCardiaqueNormale  
  ObjectIntersectionOf( :SignesVitaux  
    DataSomeValuesFrom( :aFrequenceCardiaque  
      DataIntersectionOf(  
        DatatypeRestriction(xsd:float xsd:minInclusive "60.0"^^xsd:float)  
        DatatypeRestriction(xsd:float xsd:maxInclusive "70.0"^^xsd:float)  
      )))
```

De nombreux éditeurs d'ontologies ont été conçus pour faciliter la création et la manipulation des ontologies (Protégé⁷, NeOn Toolkit⁸, SWOOP⁹, etc.).

0.4 Méthodologies d'ingénierie d'ontologie

Depuis le début du Web sémantique, de nombreuses études ont été menées pour faire avancer les pratiques liées à l'ingénierie des ontologies (IO) afin de fournir des méthodologies, des bonnes pratiques, des langages et des outils pour soutenir les experts de cette discipline. L'IO est un sous-domaine de l'ingénierie des connaissances (IC) qui étudie les principes de conception, les méthodologies et les techniques pour soutenir la conception efficace d'ontologies. Cette discipline couvre les aspects et les étapes du cycle de vie de l'ontologie (développement, maintien, modification, évolution, etc.). En effet, comme tout artefacts, les ontologies ont un cycle de vie. Elles sont conçues, implémentées, évaluées, exploitées et réutilisées pour un but précis.

Le processus qui conduit à l'élaboration d'une ontologie est long et difficile. De nombreuses méthodologies d'ingénierie d'ontologies (MIO) ont été proposées ces 20

7. <https://protege.stanford.edu/>

8. http://neon-toolkit.org/wiki/Main_Page.html

9. <https://www.w3.org/2001/sw/wiki/SWOOP>

dernières années pour supporter la conception d'ontologies. La mise en oeuvre d'une MIO efficace vise à produire des ontologies utilisables, réutilisables et faciles à maintenir. Une méthodologie est composée de méthodes, de techniques, de processus et d'activités (GOMEZ PEREZ, 1999).

Une MIO doit prendre en charge la participation active de différents rôles au cours du cycle de vie de l'ontologie, en fournissant des outils pour leur participation active à toutes les étapes du cycle de vie de l'ontologie. Les rôles impliqués sont les suivants (K. I. KOTIS et al., 2020) :

- **Experts du domaine** : Ce sont les personnes qui ont la connaissance/l'expertise du domaine et/ou des sources de données. Ce sont généralement des praticiens, non familiarisés avec les langages d'ontologie, les spécifications, etc.
- **Ingénieurs d'ontologie** : ce sont les personnes qui ont les connaissances/l'expertise pour faire des spécifications ontologiques et coordonner une tâche IO.
- **Travailleurs de la connaissance ("Knowledge worker")** : Ce sont les personnes qui exploitent l'ontologie dans des conditions « opérationnelles », afin de résoudre des problèmes, ou d'effectuer des tâches d'analyse axées sur les données. Il peut s'agir d'experts du domaine qui, dans certains cas, peuvent ne pas être impliqués dans le processus d'IO.

Les méthodologies diffèrent largement dans leurs approches de développement, la participation des parties prenantes au processus de développement, les détails dans la définition des tâches et le soutien fourni à différentes étapes. Selon K. I. KOTIS et al., 2020, une MIO doit prendre en charge toutes les parties prenantes impliquées pendant le cycle de vie de l'ontologie, c'est-à-dire depuis le développement des ontologies germes à l'évolution et au maintien continu des ontologies, en les gardant « vivantes » pour façonner les connaissances. Ainsi, une ontologie peut se caractériser tout au long de sa vie selon trois dimensions :

- **Vivacité** : une ontologie est vivante du moment où elle est active, qu'elle soit en cours de développement, d'évolution, de maintenance, de mise à jour, d'évaluation ou d'utilisation.
- **Evolution** : ontologie évoluée signifie que plusieurs versions de l'ontologie existent au cours de son cycle de vie.
- **Réutilisabilité** : une ontologie réutilisée signifie qu'elle a été réutilisée pour spécifier d'autres ontologies ou être exploitée par d'autres applications.

Les MIO peuvent être divisées en trois grandes catégories, selon les types d'acteurs impliqués dans le processus d'IO (K. I. KOTIS et al., 2020) :

- **MIO collaboratives** : cette méthodologie définit clairement et systématiquement (en phases, tâches et flux de travail). Elle met l'accent sur l'implication active et décisive des ingénieurs du savoir, des travailleurs du savoir ainsi que des experts du domaine dans toutes les phases de l'IO (spécification, mise en oeuvre, exploitation et évaluation) *via* leur collaboration étroite et continue, en vue de façonner des connaissances communément admises. Des exemples représentatifs de cette catégorie sont HCOME (K. KOTIS et VOUIROS, 2006), DILIGENT (VRANDEČIĆ et al., 2005), DOGMA-MESS (DE MOOR et al., 2006), NeOn (GOMEZ-PEREZ et SUÁREZ-FIGUEROA, 2009), GOSPL (DEBRUYNE et al., 2013) et SAMOD (NARULA et al., 2018).

- **MIO non collaboratives** : cette méthodologie ne met pas l’accent sur la collaboration des parties prenantes à l’IO communément convenues. Elle définit néanmoins clairement les phases, les tâches et les flux de travail de l’approche IO, de manière systématique et formelle. Des exemples représentatifs de cette catégorie sont METHONTOLOGY (FERNÁNDEZ-LÓPEZ et al., 1997), USCHOLD et KING, 1995, On-To-Knowledge Methodology (OTKM) (SURE et al., 2004) et FMCLGO (GIUNCHIGLIA et al., 2012).
- **MIO personnalisées** : cette catégorie ne définit pas nécessairement les phases, les tâches et les flux de travail de manière formelle et systématique. Cependant, elle engage toujours l’implication active des communautés de pratique et l’utilisation d’outils (collaboratifs ou non) pour développer des ontologies de manière agile et décentralisée et la plupart du temps de manière collaborative. La plupart de ces approches utilisent la technologie Wiki et GitHub comme moyen pour prendre en charge la collaboration distribuée fondée sur des outils dans le développement d’ontologies communément acceptées. Les travaux représentatifs de cette catégorie sont dans REBELE et al., 2016, CALBIMONTE et al., 2017, NARULA et al., 2018, SALATINO et al., 2018, TOMMASINI et al., 2018 et ARNDT et al., 2019.

Parmi ces MIO, nous présentons la méthodologie NeOn qui sera utilisée dans le chapitre 3.

La méthodologie Neon

La méthodologie NeOn (SUÁREZ-FIGUEROA et al., 2012) est constituée d’un ensemble de neuf scénarios flexibles pour la construction collaborative d’ontologies et de réseaux d’ontologies. Elle met un accent particulier sur la réutilisation et la réingénierie des ressources de connaissances (ontologiques et non ontologiques). Les scénarios sont les suivants :

- **Scénario 1 : De la spécification à la mise en œuvre.** Le réseau d’ontologies est développé à partir de zéro, c’est-à-dire sans réutiliser les ressources de connaissances disponibles.
- **Scénario 2 : Réutilisation et réorganisation des ressources non ontologiques.** Ce scénario couvre le cas où les développeurs d’ontologies doivent analyser des ressources non ontologiques et décider, en fonction des exigences que l’ontologie doit remplir, quelles ressources non ontologiques peuvent être réutilisées pour construire le réseau d’ontologies. Le scénario couvre également la tâche de réingénierie des ressources sélectionnées en ontologies.
- **Scénario 3 : Réutilisation des ressources ontologiques.** Dans ce scénario, les développeurs d’ontologies réutilisent les ressources ontologiques (les ontologies dans leur ensemble, les modules d’ontologie et/ou les déclarations d’ontologie).
- **Scénario 4 : Réutilisation et réingénierie des ressources ontologiques.** Dans ce scénario, les développeurs d’ontologies réutilisent et réorganisent les ressources ontologiques.
- **Scénario 5 : Réutilisation et fusion des ressources ontologiques.** Ce scénario ne se déroule que dans les cas où plusieurs ressources ontologiques dans le même domaine sont sélectionnées pour être réutilisées et lorsque les développeurs

d'ontologies souhaitent créer une nouvelle ressource ontologique à partir de deux ressources ontologiques ou plus.

- **Scénario 6 : Réutilisation, fusion et réingénierie des ressources ontologiques.** Ce scénario est similaire au scénario 5. Cependant, dans ce scénario, les développeurs décident de ne pas utiliser l'ensemble de ressources fusionnées tel quel, mais de le reconcevoir.
- **Scénario 7 : Réutilisation des modèles de conception d'ontologies (ODP).** Les développeurs d'ontologies accèdent aux référentiels ODP pour les réutiliser.
- **Scénario 8 : Restructuration des ressources ontologiques.** Les développeurs d'ontologies restructurent (modularisation, élagage, extension et/ou spécialisation) les ressources ontologiques à intégrer dans le réseau d'ontologies en cours de construction.
- **Scénario 9 : Localisation des ressources ontologiques.** Les développeurs d'ontologies adaptent une ontologie à d'autres langues et communautés culturelles, produisant ainsi une ontologie multilingue.

0.5 Ontologies et principes FAIR

En 2016 et grâce à l'initiative du Web des données, un nouveau paradigme est apparu, celui des données FAIR pour définir des principes et des bonnes pratiques à suivre et soutenir la recherche sur l'interopérabilité des données (WILKINSON et al., 2016). La notion de données FAIR est distincte de celle des données ouvertes (*Open Data*). En effet, des données peuvent être FAIR sans pour autant être accessibles.

Depuis leur apparition, les principes FAIR ont largement été adoptés autant dans le cadre de la recherche sur les données que par de nombreuses organisations privées et publiques dans le monde (Programme H2020, l'European Open Science Cloud (EOSC), etc.). L'objectif final des principes FAIR est d'améliorer la réutilisation des données de la recherche. Chaque principe FAIR se décline en un ensemble de caractéristiques que doivent présenter les données et les métadonnées pour favoriser la découverte (*Findable*), l'accès (*Accessible*), l'interopérabilité (*Interoperable*) et la réutilisation (*Reusable*) des données et des métadonnées aussi bien pour les humains que pour les machines. A chaque lettre du sigle FAIR est associé des bonnes pratiques de gestion des données. Ainsi, des données FAIR doivent donc être :

- **Facile à trouver** - Pour faciliter la découverte des données par les humains et les systèmes informatiques, une description et une indexation des données et des métadonnées est nécessaire suivant ces 4 principes :
 - Identifier les données et les métadonnées par un identifiant global unique et pérenne : le PID (*Persistent Identifier*). A chaque type de données est attribué un PID (DOI, URN, etc.). Les métadonnées sont un ensemble de données qui décrivent et donnent des informations sur d'autres données.
 - Décrire finement les données à l'aide de métadonnées. Les métadonnées servent à décrire et à caractériser une ressource (par exemple, le titre d'un livre, ses auteurs, etc.).

- Enregistrer et indexer les données dans un dispositif permettant de les rechercher.
- Spécifier explicitement l’identifiant de la donnée dans une ressource consultable.
- **Accessible** - Rendre les données accessibles consiste à faciliter et à sécuriser l’accès aux données par l’utilisation de protocole de communication standardisé et d’usage universel (HTTP ou HTTPS). Lorsque les données ne sont plus accessibles, les métadonnées doivent le rester.
- **Interopérable** - Rendre les données interopérables consiste à les rendre exploitables quel que soit l’environnement informatique utilisé, ceci par l’utilisation de vocabulaires standards et l’indication de lien vers d’autres ressources. Leur utilisation est conditionnée par la condition de les rendre téléchargeable, utilisable, intelligible, et combinable. Pour cela, l’interopérabilité doit être sémantique et syntaxique. Elle implique l’utilisation des métadonnées contextuelles précises et de formats respectant les standards internationaux et utilisant un vocabulaire respectant lui-même les principes FAIR, ce qui permet d’éviter les confusions entre les systèmes et la mauvaise interprétation des informations.
- **Réutilisable** - Enfin, la réutilisation des métadonnées est facilitée par l’utilisation de vocabulaires standards pour décrire les métadonnées. La description de métadonnées avec une pluralité d’attributs précis et pertinents, incluant des détails sur leur provenance et par l’utilisation de standards communs permet de les rendre utilisable pour de futures recherches ou d’autres finalités.

Récemment, de nombreux projets ont été lancés et ont vu l’intérêt d’étendre l’application de ces principes aux ontologies (FAIRsFAIR Project, etc.). Des études ont également été menées dans ce sens pour implémenter les principes FAIR sur les ontologies (BREWSTER et al., 2019). POVEDA-VILLALÓN et al., 2020 ouvre la discussion sur les conséquences techniques et sociales de l’adoption et l’application des principes FAIR pour la publication et le partage d’ontologies et montre les différents liens qui existent entre les principes FAIR et celles des bonnes pratiques et directives du Web sémantique. Ainsi, l’utilisation des ontologies soutiennent et renforcent l’interopérabilité des données et l’application des principes FAIR par l’application des principes suivants :

- Utilisation d’un langage formel, accessible, partagé et largement applicable pour la représentation des connaissances ;
- Réutilisation des URI et du vocabulaire de l’ontologie pour identifier et décrire les données ;
- Description des données avec des métadonnées riches incluant des références qualifiées à d’autres (méta)données ;
- Réutilisation des ressources sémantiques FAIR par les (méta)données rendus FAIR ;
- Adoption des vocabulaires ouverts (FOAF, PROV-O, etc.) pour associer les données aux ressources sémantiques ;
- Rendre l’ontologie accessible dans le Web des données.

Pour soutenir les chercheurs et garantir l’application des principes FAIR tout au long du cycle de vie des données, des outils d’auto-évaluation (*FAIRifier tool*) ont été

développés dans l’objectif d’évaluer et d’améliorer la qualité des données manipulées (THOMPSON et al., 2020). L’outil FAIRdat¹⁰<https://www.w3.org/2001/sw/wiki/SWOOP>, publié en 2017, a été conçu pour s’adresser aux conservateurs de données. Il permet de créer un schéma de badge appelé « Profil FAIR » qui évalue sur une échelle de 5 étoiles dans quelle mesure un ensemble de données est conforme à chaque principe FAIR. Cet outil notamment s’inspire de l’initiative du schéma 5 étoiles de Tim Burners-lee (Open Data 5¹¹) soutenant la publication de données ouvertes sur le Web.

Pour évaluer le niveau d’application des principes FAIR sur les ontologies, l’expression « *FAIRness assesment* » est apparue. Ainsi, une ontologie FAIRness est une ontologie qui applique, selon un certain degré, les principes FAIR (AMDOUNI et JONQUET, 2021). Les référentiels d’ontologies (BioPortal, OntoBee, etc.) jouent un rôle important pour rendre les ontologies plus FAIR ou FAIRness.

0.6 Graphes de connaissances

Depuis 2012 et depuis leur adoption par Google, la popularité des graphes de connaissances n’a cessé d’augmenter et leur contenu n’a cessé d’évoluer devenant plus modulaires et réutilisables. Le graphe de connaissances de Google (GKG), DBPedia et YAGO en sont des exemples marquants. Ils se trouvent actuellement au coeur de la stratégie du moteur de recherche de Google. Par exemple, lorsque un terme est recherché (ex : « melon »), le moteur de recherche affiche en plus des résultats de recherche, un certain nombre d’informations nutritionnelles (calories, teneur en lipides, etc.) provenant principalement de Wikipedia, Wikidata, et d’autres outils ou partenaires de Google (Maps, Weather.com, etc.). Pour être visible dans le GKG, l’information doit être dans ces sources d’informations ou d’utiliser des balises de données structurées.

Les graphes de connaissances sont des structures mathématiques assimilées à une base de connaissances modélisant les données sous forme de représentation graphique pour lier les concepts et entités (HOGAN et al., 2021). Un graphe est un ensemble d’entités reliées entre elles, composé de nœuds qui représentent les entités et d’arcs (ou d’arêtes) qui représentent les relations. Les relations peuvent être enrichies par des attributs ou encore une valeur quantitative représentant le poids de la relation.

Un graphe de connaissances peut être construit explicitement à partir des données, en suivant une approche dite ETL (BERGAMASCHI et al., 2011), pour extraction-transformation-charge des données, permettant d’effectuer des synchronisations massives d’information d’une source de données vers une autre. Cependant, cette approche présente plusieurs inconvénients liés à la duplication des données, à leur mise à jour et aux conflits rencontrés avec les politiques de gestion des données et les exigences de confidentialité. La pile technologique du Web sémantique fournit les éléments nécessaires pour définir et construire un graphe de connaissances.

De la même manière que pour les ontologies, l’utilisation des graphes de connaissances renforce l’application des principes FAIR. Ils ont récemment été appliqués sur les graphes de connaissances Wikidata pour simplifier l’ajout et la synchronisation de données avec les bases de données sources (WAAGMEESTER et al., 2020).

10. <https://www.surveymonkey.com/r/fairdat>

11. <https://5stardata.info/fr/>

Graphes de connaissances personnels en santé

Les graphes de connaissances ont été appliqués à de nombreux domaines pour rassembler et organiser les connaissances d'un domaine. Dernièrement, ils ont été appliqués aux connaissances liées à la santé (ETTORRE et al., 2020, GENTILE et al., 2019) ou l'alimentation (HAUSSMANN et al., 2019) introduisant ainsi de nouveaux concepts intéressants à explorer.

Le concept de « graphe de connaissances personnelles » (PKG) a été récemment introduit pour exprimer l'idée d'une vision personnalisée ou d'un résumé d'un graphe de connaissances. Une telle interprétation du terme touche un aspect pertinent de la personnalisation. BALOG et KENTER, 2019 définissent un PKG comme « une source de connaissances structurées sur les entités et la relation entre elles, où les entités et les relations entre eux ont une importance personnelle plutôt que générale ».

L'application des graphes de connaissances au domaine de la santé ont fait émerger un nouveau concept de « graphe de connaissances personnel de santé » (PHKG pour *Personal Health Knowledge Graph*). Un PHKG est une représentation des connaissances médicales et des données personnelles pertinentes pour un patient (GYRARD, GAUR et al., 2018). Les PHKG capturent ainsi des informations personnelles sur la santé qui peuvent être utilisées dans le cadre de soins de santé personnalisés axés sur les connaissances. Ils demeurent jusqu'à ce jour un sujet quelque peu sous-exploré. GYRARD, ATEMEZING et al., 2021 appliquent la vision des PHKG au traitement des maladies chroniques et décrivent certains défis impliqués dans la conception et l'utilisation des PHKG. Par exemple, un graphe de connaissances pour la dénutrition peut décrire les signes, les symptômes et les recommandations préventives, et le PHKG peut être le sous-graphe contenant uniquement les signes, les symptômes et les recommandations applicables à un patient donné (Figure 2).

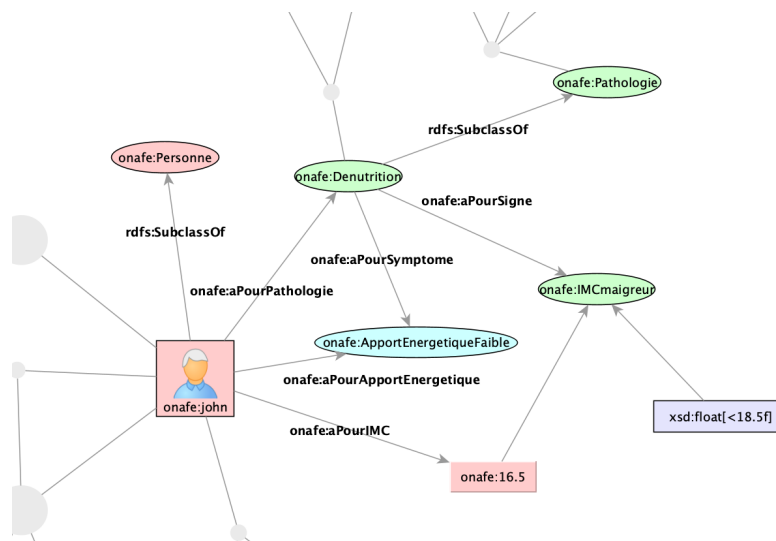


FIGURE 2 – Exemple d'un PHKG

Première partie

Etat de l'art et positionnement

Personnalisation de la recommandation

Sommaire

1.1	Qu'est-ce que la personnalisation ?	25
1.2	Aperçu des approches de personnalisation	26
1.2.1	Les approches de recommandation	26
1.2.2	Impact du Web dans l'évolution des approches de recommandation	26
1.2.3	Les sources de connaissances	27
1.2.4	En résumé	28
1.3	Les phases de la personnalisation	28
1.4	Positionnement	31

L'objet de ce chapitre est d'exposer un état de l'art sur les différentes approches existantes permettant la personnalisation des recommandations. Nous nous positionnons ainsi par rapport à ces approches et montrons l'intérêt d'exploiter les technologies du Web sémantique et de l'ingénierie des connaissances pour améliorer la personnalisation des recommandations.

1.1 Qu'est-ce que la personnalisation ?

Quotidiennement, nous sommes confrontés à des choix dont la prise de décision en dépend grandement. Quel film regarder ? Quel livre acheter ? Quel hôtel choisir ? Quel plat préparer ? Quel aliment consommer ? La personnalisation vient apporter des réponses à ces questions. Elle représente une stratégie efficace autant pour les fournisseurs de services que pour les utilisateurs : pour augmenter le trafic d'un site ou d'une application en créant de la demande et en fidélisant l'utilisateur ; ou pour réduire les possibilités qui s'offrent à lui en l'influençant dans ses prises de décision qui consomment de son temps et de son énergie.

De plus, face à la quantité d'informations sur le Web, il est nécessaire de pouvoir filtrer, hiérarchiser et fournir des informations pertinentes aux utilisateurs qui les protègent de la surcharge d'information. La personnalisation permet d'orienter l'utilisateur dans ses choix face à la surcharge informationnelle (*overload*) disponible sur le Web. La plupart du temps, la surcharge informationnelle pollue le message à transmettre aux utilisateurs et augmente le niveau de difficulté pour les systèmes de recommandation et pour les utilisateurs afin de cibler les informations les plus pertinentes. Par exemple, lorsqu'il s'agit d'améliorer son alimentation ou son activité physique, à moins d'accéder directement à la bonne source, il est difficile de faire le tri parmi toutes les recommandations disponibles sur le Web. La personnalisation résout en partie le problème de surcharge informationnelle en fournissant dynamiquement aux utilisateurs un contenu et des services personnalisés.

Aujourd'hui, la personnalisation en ligne fait partie de notre quotidien, que ce soit par l'intermédiaire des plateformes de diffusion de contenus multimédias (Netflix, YouTube, Spotify, etc.), des réseaux sociaux (Facebook, Twitter, LinkedIn, etc.), des plateformes de vente en ligne (Amazon, Alibaba, etc.), des services de livraison de courses ou de repas (Uber Eats, Deliveroo, etc.) et d'autres encore. Durant la dernière décennie, la personnalisation *via* les systèmes de recommandation a suscité l'intérêt de nombreux chercheurs et industriels.

Bien que la personnalisation ait connu une importante croissance ces dernières années avec le développement des plateformes multimédia (Netflix, etc.) et de commerce (Amazon, etc.), la personnalisation d'une recommandation reste une tâche complexe. Cette complexité augmente lorsque la quantité de données à exploiter et le nombre d'*items* proposé deviennent importants. Dans les faits, la génération d'une recommandation personnalisée n'est pas évidente en raison de plusieurs raisons comme : la prise en compte des besoins implicites de la personne ; de l'hétérogénéité de l'environnement et des objectifs ; de l'évolution d'une recommandation et des outils de visualisation selon les besoins et les exigences de l'utilisateur ; l'évolution continue en temps réel du contexte d'utilisation.

Par définition, la personnalisation est considérée comme étant un moyen pour fidéliser la clientèle en établissant une relation individuelle significative, en comprenant les besoins de chaque individu et en aidant à atteindre un objectif qui répond efficacement, et en connaissance de cause, aux besoins de chaque individu dans un contexte donné (RIECKEN, 2000). Cette définition, qui résume assez efficacement toutes celles retrouvées dans la littérature, montre que la personnalisation dépend principalement de quatre facteurs principaux : l'utilisateur, les besoins, le contexte et les capacités des dispositifs.

Une recommandation personnalisée est donc constituée d'*items* satisfaisant les besoins des utilisateurs tout en tenant compte de ces différents facteurs. Elle est proposée à l'utilisateur par l'intermédiaire d'un support. Ainsi, prendre en considération le support de collecte, ayant la capacité de tenir compte de tous ces facteurs, participe à l'amélioration de la personnalisation. Améliorer la personnalisation consiste donc à améliorer la prise en compte des différents facteurs impliqués dans la personnalisation et des connaissances du domaine de la recommandation.

Pour motiver l'approche de personnalisation adoptée et présentée en fin de ce chapitre, nous présentons dans la section qui suit un état de l'art succinct sur les différentes approches existantes pour personnaliser les recommandations. A partir des approches présentées, nous nous positionnons puis présentons les défis à soulever pour contribuer à améliorer la personnalisation des recommandations.

1.2 Aperçu des approches de personnalisation

Historiquement, la personnalisation a vu le jour dans les années 90 *via* les systèmes de recommandation. Ces systèmes étaient principalement conçus pour être des systèmes sociaux à travers lesquels les utilisateurs partagent leurs préférences sur des *items* ou des produits simples tels que les livres ou les films (MELVILLE et SINDHWANI, 2010). Ces systèmes englobent un ensemble de techniques et d'outils logiciels qui ont pour objectif de produire des recommandations pertinentes sur des *items* ou des produits susceptibles d'intéresser ou d'orienter le choix de l'utilisateur. Le terme "*item*" représente la recommandation fournie par ces systèmes. Ils ont été étudiés dans de nombreux domaines : le e-commerce, la recherche d'informations, le Web, l'exploitation des usages du Web, la modélisation de l'utilisateur, les sciences cognitives, les interactions humain-machine, l'apprentissage artificiel et bien d'autres encore. Chaque compétence (IA, BDD, BigData, Statistiques, Monitoring, Génie logiciel, Développement Web, Systèmes distribués, etc.) apporte sa contribution à l'évolution de ce pan de recherche.

1.2.1 Les approches de recommandation

Classiquement, la classification des approches de recommandation dépend des connaissances exploitées par un algorithme de filtrage (cf. section 1.2.3) pour générer la recommandation, soit :

1. un filtrage fondé sur le contenu ;
2. un filtrage fondé sur un aspect collaboratif (social) (BURKE et al., 2011).

A partir de ces deux types de filtrage, une approche dite hybride est apparue pour combiner les points forts des précédentes approches, répondre à de nouveaux besoins et pallier certains problèmes rencontrés dans l'une ou l'autre approche.

1.2.2 Impact du Web dans l'évolution des approches de recommandation

Les approches de recommandation ont évolué en parallèle avec l'évolution de la vision du Web. L'émergence du Web et l'avènement des nouvelles technologies ont

bouleversés les stratégies impliquant l’exploitation des données de l’utilisateur.

A l’origine, le Web (ou le World Wide Web), inventé par Tim Berners-Lee, était un moyen d’accéder à des informations sur Internet en affichant des pages Web statiques. Ces pages Web sont connectées par des hyperliens et peuvent contenir différents types de ressources (du texte, des graphiques, des vidéos, etc.). Dans cette vision du Web 1.0, l’utilisateur peut uniquement lire les informations mais ne peut pas interagir avec le contenu des pages Web. Cette vision du Web a soutenu l’apparition des approches de recommandation fondées sur le contenu.

Depuis, ce dernier n’a cessé d’évoluer, passant d’une version statique et qualifiée de Web informationnel, à une version dynamique qualifiée de « Web participatif », nommée le « Web 2.0 ». Il s’appuie sur l’intelligence collective pour mettre en réseau les pages Web et permettre aux utilisateurs de communiquer entre eux. Ceci, non pas uniquement pour lire le contenu des pages Web, mais également d’écrire, de modifier et de mettre à jour le contenu en ligne. Cette vision du Web a permis l’émergence des approches de recommandation fondées sur le filtrage collaboratif.

Pour personnaliser les recommandations, les approches de recommandation s’appuient généralement sur des algorithmes de recommandation permettant de trier les différentes sources de connaissances et les exploiter pour fournir une recommandation. Récemment, et grâce à l’émergence du Web sémantique, de nombreux chercheurs ont vu le potentiel et l’intérêt d’exploiter les technologies du Web sémantique notamment pour maintenir à jour les connaissances d’un domaine, pour contextualiser la recommandation ou encore pour conceptualiser l’utilisateur et ses besoins (BEDI et al., 2007). Les approches de recommandation sémantiques sont apparues pour répondre à ces besoins (AMEEN et HEAD, 2019). Toutefois, l’enjeu demeure le même, comprendre les besoins du consommateur pour lui offrir une suggestion de valeur et personnalisée.

A partir des concepts de données de partage ou encore d’accessibilité, certains font référence à la notion du Web 3.0 pour inclure les fonctionnalités prises en charge par la sémantique. L’idée principale est de transformer le Web en un langage plus compréhensible et mieux interprété par les systèmes que les êtres humains. La notion du Web 3.0 est apparue pour désigner un Web sémantique, un Web des données ou encore un Web de personnalisation (NATH et al., 2014).

1.2.3 Les sources de connaissances

Les approches de recommandation diffèrent entre elles selon les sources de connaissances et le rôle qu’elles peuvent remplir. Quatre sources de connaissances ont été identifiées comme étant nécessaires à la génération d’une recommandation : les connaissances sur l’utilisateur ; les connaissances sur la communauté ; les connaissances sur les données et les *items* de recommandation ; les connaissances sur le domaine de la recommandation.

Sur ce principe, BURKE, 2000 propose une classification dans laquelle il distingue les recommandations selon les connaissances sur lesquelles le système de recommandation s’appuie pour fournir la recommandation. Ces connaissances peuvent être : sociales (collaboratives), individuelles ou liées au contenu.

Les connaissances « sociales » sont issues d’évaluation numérique des avis ou des opinions des utilisateurs (notation d’articles achetés sur un site internet, etc.) ou d’une

mesure de similitude entre les données démographiques de l'utilisateur (âge, sexe, lieu d'habitation, profession, etc.). Ces données permettent de mieux appréhender le groupe cible d'un site Web ou pour créer des « *personas* » ou des profils utiles à l'orientation stratégique utilisées dans l'analyse et les rapports de performance des systèmes de recommandation. Les connaissances « utilisateurs » sont issues : de l'opinion de l'utilisateur, de ses données démographiques ou encore de ses propres besoins. Lorsqu'un utilisateur arrive sur une plateforme, il a une idée en tête, le système doit être capable de capturer et de répondre à ses besoins. Pour finir, la connaissance liée au « contenu » concerne l'élément recommandé, le contexte ou le domaine. La connaissance liée au contenu d'un domaine représente un large spectre. En effet, l'identification du besoin et des exigences attendus du système aide à cerner les connaissances à inclure. Ces besoins tout comme les caractéristiques et les contraintes du domaine permettent d'alimenter le système en connaissances.

1.2.4 En résumé

Initialement, la recherche sur la personnalisation des recommandations se concentrait principalement sur la recommandation d'éléments (ou d'*items*) simples (vidéos, films, etc.). La recommandation était générée à partir de l'exploitation des données dans une base de données. Depuis, une attention particulière a été apportée aux recommandations plus complexes nécessitant de prendre en compte un grand nombre de contraintes et d'éléments concernant les utilisateurs et le contexte de recommandation. L'exploitation d'une base de données répertoriant les *items* susceptibles d'intéresser l'utilisateur n'est plus suffisante pour répondre aux attentes des utilisateurs et aux problématiques qui sous-tendent les besoins de la recommandation.

Dans la majorité des études, la personnalisation de la recommandation est abordée à travers des approches de recommandation. Les approches de recommandation s'appuient sur une grande variété de compétences (BDD, BigData, Statistiques, Monitoring, Génie logiciel, Développement Web, Monitoring, Systèmes distribués, etc.) et de techniques d'IA pour fournir des recommandations pertinentes. Chacune d'entre elles apporte sa contribution à l'évolution de ce pan de recherche. La plupart des études réalisées dans ce cadre se concentrent sur les techniques de filtrages liées à l'apprentissage automatique (classification, etc.) pour effectuer des prédictions sur la recommandation (RICCI et al., 2015).

Dans notre étude, nous nous sommes intéressés plus spécifiquement aux phases du processus de la personnalisation et des facteurs intervenants dans l'amélioration de la personnalisation pour adapter notre approche de personnalisation. Pour situer notre approche, nous présentons dans la section qui suit les phases impliquées dans la personnalisation et les approches existantes permettant de les appliquer.

1.3 Les phases de la personnalisation

D'après les différentes études retrouvées dans la littérature (MOBASHER et al., 2002, ADOMAVICIUS et TUZHILIN, 2005, HIRSCH et ABERNETHY, 2012), nous pouvons résumer les phases de personnalisation par les étapes suivantes : la collecte de données,

la construction du profil, la mise en correspondance, la livraison et la présentation puis l'évaluation de la personnalisation (Figure 1.1).

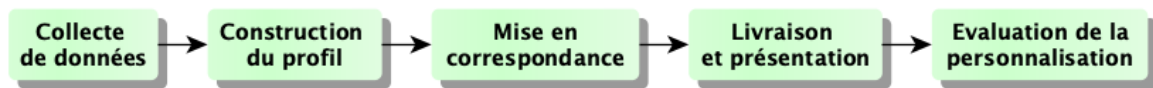


FIGURE 1.1 – Phases de la personnalisation

Collecte de données

Le processus de personnalisation commence par la collecte de données. Cette phase consiste à comprendre l'utilisateur en collectant des informations à son sujet à la suite de son interaction avec le système (ordinateur, tablette, smartphone, etc.) et à partir de diverses autres sources de données hétérogènes (montre connectée, pèse-personne, etc.) (cf. 2.3). Les données collectées décrivant l'utilisateur sont souvent regroupées sous forme de profils. Le contenu du profil d'un utilisateur varie selon les approches et les applications (GAUCH et al., 2007, FARID et al., 2018). Une fois les données collectées, elles subissent une série de pré-traitements (nettoyage, structuration, l'agrégation et le filtrage inutile (manuel ou automatique), complétion, etc.) avant d'être stockées.

L'accès aux données par l'intermédiaire de l'ontologie (Ontology-Based Data Access, OBDA) est un nouveau paradigme. Il s'appuie sur l'utilisation de techniques de représentation des connaissances et de raisonnement, pour contrôler les ressources (données, métadonnées, services, processus, etc.) et les systèmes (CALVANESE et al., 2018). La distinction entre l'ontologie et les sources de données reflète la séparation entre le niveau conceptuel (celui présenté au client) et le niveau logique et physique du système (celui stocké dans les sources).

Construction du profil

La construction de profils utilisateurs précis et complets représente une étape clé dans le processus de personnalisation. Un profil précis doit permettre la description complète de l'utilisateur selon ses données personnelles et son comportement. Il inclut des données simples (données socio-démographiques : sexe, âge, situation familiale; données socio-économiques : niveau d'étude, statut financier; lieu de résidence; etc.) ou de données plus complexes par exemple en lien avec son comportement sur un site (statistiques sur la durée de temps passé sur une page, etc.). FARID et al., 2018 classifient les différents éléments intervenant dans la construction du profil utilisateur.

La majorité des approches de recommandation (cf. section 1.2) s'appuient sur une suite d'algorithmes (filtrage, règle d'association, statistiques sur les attributs produits, etc.) pour construire le profil de l'utilisateur. Une fois construit, le profil est majoritairement stocké dans une base de données. De plus, dans la majorité de ces approches, le profil et les *items* de recommandation sont stockés dans une base de données, ce qui représente un coût élevé. Ce coût est plus ou moins conséquent selon le nombre d'utilisateurs, le nombre de profils uniques stockés et le nombre de recommandations

personnalisées disponibles dans la base de données. L'amélioration de ce processus de personnalisation dépend en grande partie de l'amélioration des approches de profilage et de mise en correspondance.

Récemment, la construction du profil exploitant les technologies du Web sémantique et de l'ingénierie des connaissances sont apparue (ELALLIOUI et EL BEQQALI, 2012, PRATIWI et al., 2018). De nombreux efforts ont été fournis jusqu'à présent dans différents domaines pour exploiter les ontologies lors de la construction de profil (MIDDLETON et al., 2004, OBEID et al., 2018, WERNER et al., 2012, CANTADOR et al., 2008 RIMITHA et al., 2019). La construction de profil fondée sur les graphes de connaissances fait partie de ces nouvelles approches (TALEBI et al., 2019). Certains travaux proposent de peupler l'ontologie pour construire des profils lors du processus d'intégration de données sans évoquer les problèmes qui y sont liés (MIDDLETON et al., 2004). TALEBI et al., 2019 proposent d'exploiter les graphes de connaissances pour construire des profils utilisateurs Twitter. Dans le domaine de la santé, la construction de profil s'appuyant sur les PKG demeure peu explorée.

Mise en correspondance

Cette étape représente la mise en correspondance entre le profil de l'utilisateur et la recommandation générée. Les approches de recommandation (cf. 1.2) permettent d'établir cette correspondance.

Livraison et présentation

La présentation de la recommandation peut se faire sous différentes formes. Elle peut prendre la forme de liste ordonnée par pertinence, ou de liste non ordonnée d'alternatives. La technique principale utilisée pour personnaliser les recommandations consiste à prédire des scores pour les *items* qui n'ont pas été évalués par les utilisateurs pour les lui présenter selon leurs scores de manière décroissante. Cependant, à elle seule, cette technique ne s'avère pas suffisante pour satisfaire les intérêts des utilisateurs.

Evaluation de la personnalisation

L'un des principaux problèmes rencontrés lors de la personnalisation de la recommandation est l'absence de personnalisation. Dans certains cas, les *items* ainsi recommandé ne représentent pas une plus-value significative pour l'utilisateur. Leur redondance et l'absence de nouveauté et de diversité affaiblissent des recommandations. SILVEIRA et al., 2019 présentent un aperçu des principaux concepts liés à cette évaluation.

Ajustement de la stratégie de personnalisation

L'exploitation des métriques d'évaluation permet de mesurer la qualité et l'impact de la personnalisation. Cette étape permet d'identifier les améliorations possibles pour chacune des cinq autres étapes du processus de personnalisation. De ce fait, des améliorations pourront être apportées sur la collecte de données, la création de meilleurs profils utilisateurs, la diffusion et la présentation des recommandations, ou encore utiliser des mesures supplémentaires pour augmenter l'impact de la personnalisation.

L'utilisation efficace de ce *feedback* permet d'améliorer le processus de personnalisation. Dans le cas contraire, un effet de dépersonnalisation pourra se faire ressentir.

1.4 Positionnement

Dans ce travail, nous abordons la personnalisation comme étant un processus à suivre et une finalité à atteindre pour aboutir à des recommandations pertinentes. L'objectif étant d'exploiter les ressources sémantiques (ontologie et PHKG) pour apporter des solutions permettant de soutenir les premières phases du processus de personnalisation et faire le lien entre les différentes couches du système (Figure 1.2).

A travers la vision de OBDA, nous identifions trois couches du système de personnalisation :

- la couche « Applicative » correspond aux composants du système (interfaces Web, objets connectés, etc.) permettant la collecte de données. Pour assurer la phase de collecte, la prise en compte des différentes sources de données et du support de collecte de données sont nécessaires pour améliorer la personnalisation. Nous abordons la stratégie adoptée pour répondre à cet enjeu dans la section 2.3 ;
- la couche « Sémantique » correspond à l'ontologie représentant le domaine d'intérêt, l'utilisateur et les *items* recommandés ;
- la couche « Intelligente » correspond au "*mapping*" entre les deux précédentes couches. Ce "*mapping*" doit fournir des moyens permettant l'exploitation des données de l'utilisateur et des connaissances du domaine. A l'issue de cette étape, l'enjeu consiste à construire le profil et permettre la correspondance entre le profil et les *items* recommandés.

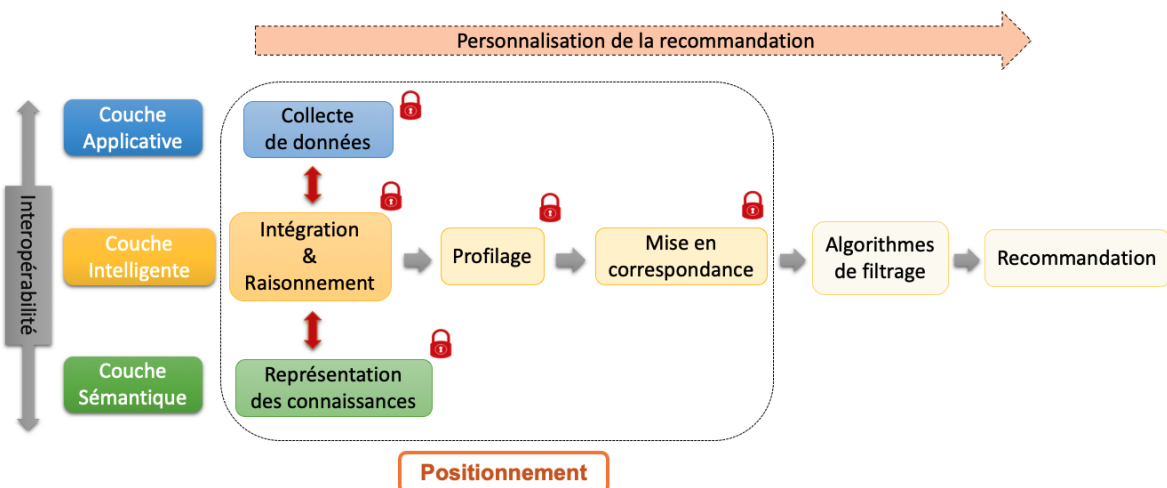


FIGURE 1.2 – Positionnement vis-à-vis des approches de personnalisation

La construction des profils utilisateurs représente une étape essentielle dans la personnalisation. L'objectif étant de transformer les données collectées en connaissances exploitables stockées dans un modèle appelé « profil ».

Par conséquent, nous explorons dans notre étude l'enjeu lié à l'exploitation des ontologies et des PHKG pour améliorer les trois premières phases de personnalisation et fournir des moyens permettant de :

- résoudre le coût lié au stockage du profil et des *items* de recommandation ;
- prendre en compte les connaissances et les facteurs impliqués dans la personnalisation ;
- définir un vocabulaire unifié pour représenter l'utilisateur, les facteurs impliqués dans la personnalisation (ses besoins, son environnement, ses contraintes, etc.), les données collectées et les connaissances du domaine ;
- unifier les données qui partagent une certaine sémantique commune mais qui proviennent de sources différentes ;
- fournir une représentation explicite de la relation entre les données et l'ontologie *via* les PHKG ;
- utiliser l'expressivité de l'ontologie pour construire le profil utilisateur et établir la correspondance entre le profil et la recommandation.

Chapitre 2

Application : prise en charge personnalisée pour les personnes âgées

Sommaire

2.1	Introduction	34
2.2	Bien vieillir : un objectif de santé publique	34
2.2.1	Problématiques liées à la dénutrition	35
2.2.1.1	Stratégies préventives pour la dénutrition	37
2.2.1.2	Recommandations nutritionnelles	38
2.2.2	Problématiques liées à la sédentarité	38
2.2.2.1	Stratégies préventives pour l'activité physique	39
2.2.2.2	Recommandations d'activité physique	40
2.2.3	Outils d'évaluation actuels	41
2.3	Outils de collecte de données en santé	42
2.3.1	Le DMP	42
2.3.2	Les objets connectés	42
2.3.2.1	Apports des objets connectés pour la collecte de données	42
2.3.2.2	Problématiques et challenges liés aux objets connectés	43
2.3.3	Les questionnaires en lignes	46

Les objectifs de ce chapitre sont doubles. Le premier objectif consiste à identifier les connaissances intéressantes à exploiter sur les personnes âgées et sur les recommandations pour répondre au besoin de représentation des connaissances et d'intégration des données. Pour cela, nous présentons dans un premier temps le champ applicatif sur lequel nous nous appuyons pour répondre aux problématiques liées à la dénutrition et à la sédentarité. Le second objectif consiste à déterminer notre stratégie de collecte de données. Pour cela, nous identifions les principaux outils exploitables tout en tenant compte des difficultés associées à chacun d'entre eux.

2.1 Introduction

L'enjeu de la personnalisation de l'alimentation et des activités est de fournir aux personnes âgées les clés d'une nutrition et d'une activité physique répondant à leurs préférences sensorielles et leurs besoins spécifiques, les amenant vers un « mieux vieillir ». De plus, les besoins de tout un chacun varient et évoluent au cours du temps. Les solutions proposées pour prévenir la dénutrition et la sédentarité doivent s'adapter à l'évolution des nouvelles technologies pour répondre aux problèmes spécifiques de chaque individu.

Actuellement, la détection des facteurs de risque se fait principalement par le personnel de santé lors de la consultation. Les outils utilisés pour détecter un signe, un symptôme ou une pathologie, sont des questionnaires d'évaluation sous format papier. Cette approche traditionnelle ne permet pas de détecter de manière précoce la dégradation d'un état de santé. L'évolution des TIS présente actuellement un développement important. Ils constituent une opportunité intéressante et une alternative dans la prise en charge non thérapeutique, à la fois pour le diagnostic et pour le suivi des pathologies, ceci, en complément aux outils diagnostiques habituels. Pour bénéficier d'une telle évolution, toutes les connaissances (recommandations, conseils, données, etc.) potentiellement exploitables par le système doivent être identifiées pour pouvoir agir sur les facteurs de risques modifiables favorisant la dénutrition et la sédentarité.

Avant d'aborder les solutions potentiellement exploitables pour répondre au besoin de collecte de données, nous présentons dans ce chapitre la problématique de dénutrition et de sédentarité et les stratégies actuelles de prises en charge permettant d'orienter la personnalisation de la recommandation.

2.2 Bien vieillir : un objectif de santé publique

La notion de personne âgée ou senior est variable selon les secteurs. Pour les industriels, un individu devient senior au-delà de 50 ans. Pour les autorités publiques, la notion de personne âgée est rattachée aux prestations sociales et représente les personnes âgées de plus de 60-65 ans.

Les personnes âgées d'au moins 65 ans représentent actuellement plus de 20% de la population française¹. L'accélération du vieillissement de la population entraîne, dans les pays industrialisés, un déséquilibre entre le nombre de personnes âgées et retraitées, et le nombre de personnes jeunes et actives modifiant ainsi le fonctionnement de la société. Cette situation conduit les politiques et les personnels de soins à prendre des mesures pour pallier les conséquences et les contraintes dues à cette augmentation. En effet, le vieillissement démographique de la population engendre non seulement des conséquences économiques, sociales et culturelles mais également des effets sur la santé et la prise en charge des personnes âgées. La « silver économie » est apparue pour répondre aux conséquences liées au papy-boom. Elle tente d'apporter des solutions transversales pour répondre aux problèmes spécifiques des personnes âgées. Ces solutions concernent les secteurs d'activité de la santé, du maintien à domicile, de l'aménagement de l'habitat, des services à la personne.

1. <https://www.insee.fr/fr/statistiques/3353488>

Les conséquences dues au vieillissement sont déterminées par plusieurs mécanismes ou facteurs dits intrinsèques et extrinsèques (isolement social, changement des habitudes de vie, etc.) (BELMIN et al., 2016). Les facteurs intrinsèques ou chronologiques se caractérisent principalement par des modifications biologiques, physiologiques et génétiques. Ces modifications font partie des changements inévitables liés à l'âge qui, en soi, est un phénomène naturel. Les avancées de la recherche et des pratiques de soins ont, avec le temps, amélioré la prise en charge des pathologies chroniques ou aiguës plus fréquentes chez les personnes âgées. Ces avancées ont pu ces dernières années augmenter la durée de vie sans déficiences et prolonger la vie en bonne santé des personnes âgées (DREVET et GAVAZZI, 2019). Cependant, le rôle des facteurs extrinsèques (alimentation, activité, environnement, etc.) a longtemps été sous-estimé. Toutefois, il paraît de plus en plus évident qu'une meilleure considération de ces facteurs extrinsèques et une meilleure prise en charge des maladies chroniques, en dehors des interventions médicales, permettent de tendre vers un vieillissement dit « optimal » voire sans pathologie ni handicap.

Au cours de la vie, il suffit d'un incident de santé, familial ou environnemental pour que la personne se retrouve dans un état de fragilité voire de dépendance. Dans la majorité des cas, cet état accentue les symptômes et pathologies présents chez la personne. Actuellement, 58 % des plus de 65 ans sont considérés comme préfragiles ou fragiles² et un quart des personnes âgées se trouvent à un épisode de leur vie dans une situation de fragilité et/ou de dépendance (SANTOS-EGGIMANN et al., 2009). Cette situation ne fait qu'accentuer les conséquences associées au vieillissement, à l'isolement et à la dépendance des personnes âgées. Ainsi, la détection des facteurs favorisant la fragilité des personnes âgées permet également d'améliorer la prévention des facteurs associés à la dénutrition et à la sédentarité.

La dénutrition et la sédentarité font partie des facteurs de risque évitables comportant des risques de maladies et diminuant l'espérance de vie. Les avantages d'une alimentation et d'une activité physique qui suivent les recommandations sont bien nombreux. Elles contribuent en partie à la prévention primaire et secondaire des maladies chroniques (diabète, hypercholestérolémie, hypertension, cancer, etc.), la réduction et le contrôle de certaines maladies (neurologiques, cardiaques, etc.), l'amélioration des défenses immunitaires, la diminution du risque de chute et de dépression, etc. (SOLIDARITÉS ET DE LA SANTÉ, 2019). En somme, elles améliorent la qualité de vie et contribuent au « bien vieillir ».

2.2.1 Problématiques liées à la dénutrition

La nutrition est un élément essentiel pour le maintien de la santé et du bien-être pour tout individu. Un apport adéquat de nutriments est fondamental pour renforcer le système immunitaire et maintenir une santé et un bien-être optimal tout au long de la vie. En France, environ 2 millions de personnes souffrent de dénutrition. La prévalence de la dénutrition est estimée à environ 10 % à domicile, à 38 % en institution, et à 30 à 70 % à l'hôpital (SANTÉ (HAS), 2020).

Souvent utilisée de manière interchangeable avec la malnutrition, la dénutrition est considérée comme étant un type de malnutrition. La malnutrition est un déséquilibre

2. <https://solidarites-sante.gouv.fr/IMG/pdf/strategie-prise-en-charge-personnes-agees-covid-19.pdf>

entre les nutriments dont l'organisme a besoin et les nutriments qu'il obtient. La dénutrition représente quant à elle un état pathologique caractérisé par un déséquilibre de la balance énergétique résultant d'apports nutritionnels insuffisants en regard des dépenses énergétiques de l'organisme. En gériatrie, le terme malnutrition protéino-énergétique (MPE) représente une carence en protéines et en énergie, le plus souvent accompagné d'une carence en vitamines et en oligo-éléments (MEIJ et al., 2017).

La dénutrition représente un problème particulièrement important chez les personnes âgées. Le vieillissement physiologique implique des modifications physiologiques comme la diminution de la masse maigre corporelle (dont la masse musculaire) ou l'augmentation de la pression artérielle au repos et à l'exercice. Ces modifications rendent la personne plus vulnérable et risquent de faire rentrer la personne dans la spirale de la dénutrition (Figure 2.1) augmentant ainsi la survenue d'autres pathologies ou de syndromes gériatriques (chute, etc.). Une fois que la personne est entrée dans la spirale de la dénutrition, il lui est difficile d'en sortir sans aide, qu'elle soit professionnelle ou familiale. Plus les symptômes augmentent, et plus la sortie de cette spirale devient irréversible.

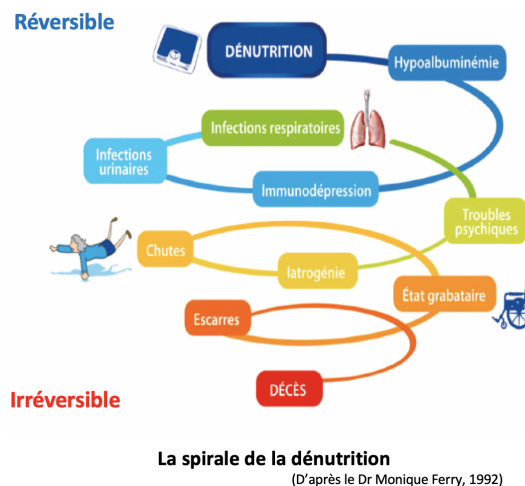


FIGURE 2.1 – Spirale de la dénutrition (FERRY, 2010)

De nombreux facteurs peuvent expliquer l'évolution de la consommation avec l'âge comme les changements physiologiques, sensoriels, psychologiques et sociologiques. Ces changements sont susceptibles d'avoir un impact délétère sur la prise alimentaire et le statut nutritionnel de la personne âgée. Cependant, ils ne peuvent aucunement justifier d'une MPE. Ainsi, de nombreuses questions se posent quant à l'optimisation de la prise en charge nutritionnelle de la personne âgée notamment en institution. En effet, la situation actuelle en institution (EHPAD, établissement de soins, etc.) est particulièrement préoccupante où la grande majorité des personnes âgées ne satisfait pas à ses besoins caloriques et protidiques, notamment en raison d'une consommation insuffisante de repas riche en protides servi au déjeuner et au dîner. En effet, une enquête récemment réalisée a montré que 83 % des résidents en EHPAD ne satisfaisaient pas à leurs besoins caloriques et protidiques et 43 % présentaient un apport inférieur aux 2/3 des apports conseillés et 13 % seulement des personnes âgées vivant en institution satisfaisaient

à leurs besoins caloriques et protidiques (apport calorique et protidique supérieur ou égal à 100 % du besoin) (MEIJ et al., 2017). Chez les personnes âgées autonomes, les principaux motifs de changement de consommations sont liés à l'évolution des préférences et l'état de santé. Chez les personnes âgées dépendantes, le fait d'imposer des menus représente le principal motif lié aux changements de consommations (SULMONT-ROSSÉ et VAN WYMELBEKE, 2019). Le plat principal actuellement proposé en HEPAD comprend une composante protidique (viande, poisson...) et un accompagnement (légumes, féculents...), représente les insatisfactions des personnes âgées dépendantes pour leur alimentation. En général, les personnes âgées présentent des attentes pour un plat « familial » qui n'est malheureusement pas toujours compatible avec les contraintes de la restauration collective des HEPAD. Ce constat montre le besoin de proposer un outil d'aide à la préparation de recettes destinées au personnel de cuisine œuvrant en restauration collective et tenant compte des attentes et des caractéristiques individuelles des résidents. Cet outil devra s'appuyer sur les stratégies et les recommandations actuelles pour lutter cette problématique de santé publique.

2.2.1.1 Stratégies préventives pour la dénutrition

La dénutrition peut être suspectée avec l'apparition d'un certain nombre de signes (constipation, prise de plus de 3 médicaments par jour). Certains signes sont mesurables d'autres ne le sont pas. L'objectif est de détecter au plus tôt ces signes pour adapter au mieux la prise en charge.

Un certain nombre de situations peuvent contribuer à la survenue d'une dénutrition. Certaines d'entre elles sont indépendantes de l'âge (cancer, pathologies à l'origine de maldigestion, etc.), et d'autres y sont liées (troubles bucco-dentaires, régimes restrictifs, troubles de la déglutition etc.) (SANTÉ (HAS), 2007).

Pour augmenter les apports énergétiques et protidiques, la HAS a établi en 2007 une stratégie de prise en charge de la dénutrition. Les recommandations fournies sont destinées aux professionnels de santé pour : les aider à déterminer les situations à risque de dénutrition ; leur fournir les modalités de dépistage ; les critères de diagnostics de dénutrition ; la stratégie de prise en charge nutritionnelle à adopter et les modes de suivi (SANTÉ (HAS), 2007).

La stratégie de prise en charge nutritionnelle prend trois formes : orale en première intention (conseils nutritionnels, aide à la prise alimentaire, alimentation enrichie et compléments nutritionnels oraux (CNO)) ; entérale en cas d'impossibilité ou d'insuffisance de la nutrition orale ; ou parentérale dans les cas extrêmes. Nous citons un exemple sur les conseils nutritionnels diététiques fournis par la HAS³ retrouvés et potentiellement exploitables :

- respecter les repères du Programme national nutrition santé (PNNS)⁴ ;
- augmenter la fréquence des prises alimentaires dans la journée ;
- éviter une période de jeûne nocturne trop longue (> 12 heures) ;

3. https://www.has-sante.fr/jcms/c_546549/fr/strategie-de-prise-en-charge-en-cas-de-denuitration-proteino-energetique-chez-la-personne-agee

4. <https://www.mangerbouger.fr/1-essentiel/les-recommandations-sur-l-alimentation-l-activite-physique-et-la-sedentarite>

- privilégier des produits riches en énergie et/ou en protéines et adaptés aux goûts du patient ;
- organiser une aide au repas (technique et/ou humaine) et favoriser un environnement agréable.

2.2.1.2 Recommandations nutritionnelles

Dans l’optique d’améliorer l’alimentation, le PNNS fournit les principales recommandations nutritionnelles disponibles pour le grand public. Après avoir revu la formulation de ses conseils alimentaires, la dernière version du PNNS (PNNS 4) parue en 2019 fournit à la fois des repères nutritionnels et des recommandations. Les repères nutritionnels sont établis par le HCSP. Ils délivrent pour chaque groupe alimentaire le repère principal qui correspond souvent à une fourchette de quantité à consommer, et des données complémentaires qui renseignent sur le « grammage », le mode de substitution ou d’autres informations utiles. Les recommandations se veulent plus simples à suivre. Elles sont regroupées selon les repères nutritionnels réunissant les groupes alimentaires en trois catégories : augmenter, aller vers, réduire. Lorsqu’elles sont détaillées, elles fournissent dans chaque catégorie pour chacun des groupes alimentaires un repère principal et des données complémentaires.

Le PNNS 4 inclus également dans son programme le Nutri-Score. Ce score aide les consommateurs à distinguer clairement les produits à favoriser et ceux à éviter en leur attribuant un score allant de A à E. Le Nutri-Score est fonction de la teneur en aliment à favoriser et à limiter pour 100 grammes de produits. Il s’appuie sur un calcul complexe tenant compte d’un bon nombre de critères tels que les valeurs nutritionnelles, la proportion d’additifs, les pesticides, etc.

Avec quelques adaptations, les recommandations nutritionnelles et d’activité délivrée pour la population générale sont valables pour les personnes âgées. Le PNNS fournit des recommandations nutritionnelles spécifiques pour les aidants dans lesquelles l’accent est mis sur 4 points essentiels à surveiller : le poids, l’alimentation, l’hydratation, le niveau d’activité.

Récemment, le site d’assurance maladie Ameli propose *via* sa plateforme « Coaching Santé Active »⁵ des conseils pour la prévention et le dépistage des principales maladies chroniques tels que le diabète ou l’hypertension nécessitant un régime particulier. Ces conseils textuels concernent les principales maladies chroniques rencontrées notamment chez les personnes âgées (diabète, hypertension, etc.). En revanche, il ne fournit pas d’outils pour appliquer ces recommandations.

2.2.2 Problématiques liées à la sédentarité

La nutrition et l’activité physique sont étroitement liées. Au-delà de leurs effets bénéfiques sur la santé, l’activité physique permet notamment de stimuler l’appétit et de potentialiser l’effet d’une prise en charge nutritionnelle. L’activité physique doit ainsi être systématiquement envisagée en association avec la prise en charge nutritionnelle. L’Organisation Mondiale de la Santé (OMS) définit l’activité physique comme « tous

5. <https://www.ameli.fr/paris/assure/sante/assurance-maladie/coaching-sante-active>

les mouvements que l'on effectue notamment dans le cadre des loisirs, sur le lieu de travail ou pour se déplacer d'un endroit à l'autre » (SANTÉ (OMS), 2020).

Selon le Programme National de Nutrition et Santé (PNNS), la sédentarité ou « comportement sédentaire » est caractérisée par une dépense énergétique inférieure ou égale à la dépense de repos en position assise ou allongée tels que les déplacements en véhicule automobile ou toutes autres activités réalisées au repos en position allongée (lire, écrire, converser par téléphone, etc.) Actuellement, plus de 50 % des plus de 65 ans sont inactifs (SANTÉ PUBLIQUE FRANCE, 2020). L'inactivité est caractérisée par un niveau insuffisant d'activité physique d'intensité modérée à élevée, ne permettant pas d'atteindre le seuil d'activité physique recommandé.

L'inactivité physique associée à un mode de vie sédentaire sont des facteurs favorisant : un risque de fracture et de maladies chroniques, de maladie cardiaque, d'accident cérébrovasculaire, de dépression, d'ostéoporose, ainsi que l'apparition de syndromes gériatriques (mauvaise tonicité musculaire, peur de tomber, mobilité réduite, isolement, perte d'autonomie, etc.). Il est reconnu que l'activité physique représente une stratégie efficace dans la prise en charge de nombreuses maladies telles que les maladies chroniques (TURNER et al., 2017) ainsi que dans la prévention de nombreux syndromes gériatriques tels que la chute (KARINKANTA et al., 2010). Ce dernier étant un indicateur de fragilité et un facteur prédictif de sédentarité, de dépendance et d'admission en institution. Par exemple, dans le cas d'une personne venant de chuter, son état de santé peut l'empêcher non seulement d'exercer ses activités sportives habituelles, mais également ses activités de la vie quotidienne, comme faire le ménage ou les courses.

2.2.2.1 Stratégies préventives pour l'activité physique

L'augmentation de la sédentarité de toute la population occidentale met en évidence le besoin d'inscrire l'activité physique comme stratégie thérapeutique non médicamenteuse en première intention ou associée à la prise en charge médicale. Depuis 2017, la prescription d'activité physique adaptée (APA) représente une recommandation de la Haute Autorité de Santé (HAS)⁶ et fait partie des stratégies de prises en charge notamment pour les patients qui souffrent d'une affection longue durée (ALD).

Une ordonnance d'activité physique comprend entre autres :

- des conseils sur l'augmentation des activités physiques de la vie quotidienne, en particulier les déplacements actifs ;
- des conseils sur la réduction du temps passé à des comportements sédentaires et son fractionnement ;
- un programme d'APA à la pathologie du patient, en ville ou en institution, en individuel ou en groupe (en fonction des ressources locales) ;
- une activité sportive de loisirs choisie avec le patient, avec les coordonnées du club sportif de proximité, ou une activité de sport-santé (selon l'état de santé du patient), en s'aidant du dictionnaire médicosport-santé⁷.

6. https://www.has-sante.fr/jcms/p_3111587/fr/maladie-chronique-une-ordonnance-d-activite-physique-sur-mesure

7. <https://www.vidal.fr/sante/sport/infos-sport-medicosport-sante/>

Or, les médecins ne sont majoritairement pas formés pour de telles prescriptions. Ceci montre le besoin de trouver d'autres moyens pour soutenir la prévention et la prise en charge des personnes âgées. Détecter au plus tôt les facteurs de risque et les signes de dénutrition et de sédentarité permet de prévenir l'apparition des facteurs associés.

En termes de mesures proposées pour améliorer l'alimentation, PNNS propose une plateforme de planification de menu (la Fabrique à menu⁸) qui jusqu'à présent ne tient pas compte des caractéristiques et des besoins individuels. Pour favoriser l'activité, le PNNS propose également des outils de test du niveau, de sélection et de planification d'activités physiques dans la semaine.

2.2.2.2 Recommandations d'activité physique

Bien que la durée d'activité des personnes de plus de 65 ans diminue avec l'âge (SANTÉ PUBLIQUE FRANCE, 2020), les besoins en termes d'activité sont identiques aussi bien en dessous qu'au-dessus de 65 ans.

Les recommandations en termes d'activité physique disponibles prennent plusieurs formes : Le PNNS 4 recommande d'exercer : 30 minutes par jour d'activité, d'intensité modérée, au moins 5 fois par semaine ; ou de 25 minutes d'activité physique d'intensité élevée, au moins 3 jours par semaine.

Selon l'OMS, le nombre de pas recommandé est fixé à 10000 pas par jour afin de se maintenir en forme et préserver sa santé. Elle préconise également d'exercer : des activités variées à plusieurs composantes qui mettent l'accent sur l'équilibre fonctionnel ; et des exercices d'intensité modérée ou supérieure, 3 fois par semaine ou davantage, afin d'améliorer les capacités fonctionnelles et de prévenir les chutes.

Les recommandations fournies prennent parfois en compte certains aspects d'actualité. Par exemple, en période de confinement, le PNNS avait ajouté à ses recommandations de : - pratiquer 30 minutes d'activité physique dynamique au quotidien ; - privilégier les exercices de renforcement musculaire et de souplesse ; - répartir la pratique en plusieurs fois sur la journée (par exemple par tranche de 10 minutes).

Selon la HAS, une activité physique d'intensité légère (< 3 METs), comme la marche, peut être recommandée sans risque excessif et sans consulter un médecin. L'évaluation du niveau de risque cardio-vasculaire est cependant recommandée chez tous les patients âgés avant de commencer un programme d'activité physique d'intensité au moins modérée (SANTÉ (HAS), 2007).

En résumé

Bien que les recommandations présentées dans ce chapitre soient pertinentes, elles ne permettent pas de diagnostiquer, de suivre et de prendre en charge de manière précoce ou en continu les personnes isolées susceptibles de rentrer dans la spirale de la dénutrition (Figure 2.1). Les recommandations générales sur la nutrition ou l'activité physique largement véhiculées (manger 5 fruits et légumes par jour, exercer 30min d'activité physique modérée par jour, etc.), sont bien connues par le grand public, cependant, d'autres aspects de ces recommandations essentiels ne viennent pas forcément à l'esprit. S'il est demandé à une personne âgée : « Quelles sont les recommandations permettant

8. <https://www.la-fabrique-a-menus.fr/front/>

d'améliorer votre alimentation et votre activité physique ? », elle aura forcément du mal à tous nous les citer ou à suivre des recommandations telles qu'elles sont actuellement (cf. section 2.2.1.1) pour respecter les repères du Programme national nutrition santé (PNNS) ; augmenter la fréquence des prises alimentaires dans la journée, etc.).

De plus, les recommandations générales ne tiennent pas compte des spécificités alimentaires et sensorielles rencontrées chez les personnes âgées. Au-delà des campagnes gouvernementales sur l'alimentation équilibrée et la pratique de l'activité physique - qui incitent par exemple à limiter la quantité de sel en cas d'hypertension artérielle ou à exercer une activité physique régulière - d'autres solutions doivent être proposées pour répondre aux problèmes spécifiques des personnes âgées.

Les conseils fournis par le PNNS comme « Surveiller le poids tous les mois » ou « Augmenter son apport en eau dans certains cas (petit appétit, fièvre, la prise de certains médicaments etc.) » peuvent être automatisés. Dans ce contexte, le développement d'un outil éducatif et de surveillance de l'état de santé de la personne âgée tenant compte des spécificités de chaque individu est intéressant à exploiter. A partir de toutes les connaissances citées, l'objectif est de les exploiter dans le chapitre 3 pour conceptualiser le domaine de la nutrition et de l'activité en mettant au centre de nos intérêts les besoins spécifiques de la personne âgée.

Pour personnaliser au mieux les recommandations élaborées, un certain nombre de données sont collectées. Ces données concernent les pathologies, les habitudes alimentaires, les préférences sensorielles ainsi que celles issues d'objets connectés. De ce fait, la question du choix de l'outil de collecte de données s'est alors posée.

2.2.3 Outils d'évaluation actuels

Les questionnaires (de dépistage, d'auto-évaluation) font partie des outils potentiellement exploitables dans une démarche de prévention ou de détection des signes de dénutrition et de sédentarité. En milieu gériatrique, ils sont utilisés pour déterminer des symptômes gériatriques. Ils permettent également de recueillir des informations sur l'état de santé et nutritionnelles de la personne âgée.

Un certain nombre de questionnaires ont été élaborés et ont démontré leur efficacité pour la confirmation d'une suspicion de maladie, d'un syndrome gériatrique ou encore pour l'évaluation de l'état de santé perçu par une personne, ce dernier étant un indicateur puissant des capacités fonctionnelles et du bien-être (GOLDMAN et al., 2004). De même pour l'analyse des habitudes alimentaires des personnes âgées, de nombreux questionnaires existent pour identifier un lien entre le comportement alimentaire comme l'appétit et la nature des aliments consommés. Par exemple, il a été montré que les personnes avec un faible appétit consomment moins de protéines, de fibres alimentaires, d'aliments solides, d'aliments riches en protéines, de céréales complètes et de fruits et légumes (MEIJ et al., 2017).

Concernant l'activité physique, les questionnaires représentent la méthode d'évaluation la plus répandue. Ils peuvent être autoadministrés ou remplis lors d'un entretien. Les questions portent sur les différents types d'activités (professionnelles, domestiques, loisirs, sport, ou des activités spécifiques) à l'aide de réponses ouvertes ou fermées. Quelques questionnaires comprennent des questions spécifiques sur la sédentarité (par ex., nombre d'heures quotidiens passés devant un écran ou assis) (WASHBURN et

MONTOYE, 1986).

Ces outils d'évaluation représentent actuellement la principale stratégie pour identifier les signes permettant de personnaliser la prise en charge des personnes âgées dénutries et sédentaires. Aujourd'hui, d'autres moyens peuvent être exploités pour soutenir la collecte de données auprès de la personne âgée. Nous les présentons dans la section qui suit.

2.3 Outils de collecte de données en santé

Dans le cadre applicatif de notre étude, nous identifions trois principaux outils de collecte de données : le dossier médical patient (DMP), les objets connectés, les questionnaires en ligne.

2.3.1 Le DMP

Les données de santé issues du dossier médical patient (DMP, *Electronic Health Records* (EHRs)) font partie des TIS (ALOTAIBI et FEDERICO, 2017). Cependant, l'accès et la disponibilité des données en santé présentent une réelle difficulté. Lorsque l'accès aux données de santé est possible (HealthData Hub, etc.), le cadre réglementaire d'utilisation et d'exploitation présente un réel frein pour la recherche sur des données de santé. En effet, la confidentialité des données exploitées rend cette tâche plus ardue. Nous nous sommes tout de même intéressés aux éléments inclus dans le DMP (ou les données issues de la carte vitale) pour une future exploitation des données par les personnels de soins qui y sont inclus.

Actuellement, le DMP⁹ se compose de huit rubriques. La rubrique principale est une synthèse sous forme de fiche standardisée regroupant les principales informations de santé telles que les antécédents médicaux, allergies, groupe sanguin, ou les traitements en cours. Nous y trouvons également les résultats d'examens tels que les radio ou les analyses biologiques, les comptes rendus d'hospitalisation, ou encore les coordonnées des proches à prévenir en cas d'urgence. Il comprend également les dernières volontés du patient (dons d'organe, refus de décès à l'hôpital, etc.). Ces données sont intéressantes à exploiter, cependant, la procédure d'accès est bien complexe. L'accès est rendu possible uniquement lorsque les raisons d'accès et d'exploitation sont justifiables. Pour pallier cette difficulté, nous nous sommes donc intéressés à l'exploitation des objets connectés pour faciliter cette étape.

2.3.2 Les objets connectés

2.3.2.1 Apports des objets connectés pour la collecte de données

L'évolution des nouvelles technologies (dispositifs, réseau, moyens de communication, etc.) offre dans le domaine de la santé de nouvelles promesses en termes de collecte de données : implicites ; explicites ; de la manière la plus discrète possible ; en temps réel ou tout au long du quotidien. Parmi ces nouvelles technologies, les objets connectés ou les données issues de capteurs, présents dans différents périphériques (objets connectés,

9. <https://www.dmp.fr/cnil>

smartphones, etc.), séduisent de plus en plus les personnels de santé pour la prise en charge des personnes âgées et le maintien de leur qualité de vie.

Les objets connectés sont utilisés dans un grand nombre de domaines (e-santé, transport, domotique, etc.). La e-santé est l'un des secteurs qui profite le mieux à l'évolution de l'économie et du marché des objets connectés. Depuis, de nouveaux champs d'application sont apparus pour tirer profit des fonctionnalités liés à :

- l'automesure connectée ("*Quantified-Self*") : De nos jours, un certain nombre d'applications mobiles offrent à l'utilisateur la possibilité de renseigner la nature et la quantité des aliments consommés lors d'un repas. L'analyse apportée par ces applications est sous forme d'un descriptif statistique de la composition nutritionnelle par aliment voire par repas. Ils ont également la possibilité d'exploiter certaines données individuelles pour calculer l'IMC et fournir en retour l'apport calorique recommandé à atteindre ou à ne pas dépasser.
- les observations de la vie quotidienne "Observations of Daily Living" : Les objets connectés font de plus en plus partie du quotidien de beaucoup de personnes, que ce soit pour quantifier le nombre de calories brûlées selon l'activité physique pratiquée par l'intermédiaire notamment de montres ou de bracelets connectés, ou encore pour mesurer et suivre ses variables physiologiques (pression artérielle, poids, etc.) au cours du temps ou d'une activité pratiquée. Lorsqu'ils sont connectés à Internet, ils constituent un réseau appelé l'internet des objets (*Internet of Things* (IoT)).

Ce constat nous a poussé à vouloir exploiter la puissance des objets connectés dans l'optique de faciliter la collecte de données auprès de la personne âgée. Leur intégration dans le quotidien des personnes âgées permettra au personnel de santé, aux proches et même au sujet concerné de surveiller son état de santé et de contribuer au maintien de son autonomie. Cependant, leur exploitation nous a confronté à de nombreux défis que nous présentons dans la section qui suit.

2.3.2.2 Problématiques et challenges liés aux objets connectés

Un déclin d'innovation des objets connectés

Nous distinguons deux catégories d'objet connecté dans le domaine de la e-santé : - les montres/bracelets connectés et les « autres ». On distingue les montres des bracelets connectés par leurs fonctionnalités. La montre a pour caractéristique d'être une extension du support connecté. Elle a donc de multiples fonctionnalités du fait qu'elle contient différents types de capteurs, contrairement au bracelet qui quant à lui a pour fonction de suivi (*tracker*) d'activité. L'avantage qu'ils offrent comparé aux périphériques connectés (smartphones, tablettes) est qu'ils s'intègrent plus facilement au quotidien et se font rapidement oublier. Parmi les montres connectées présentes sur le marché, très peu d'entre elles sont adaptées aux personnes âgées. Celles qui le sont, au-delà de donner l'heure, ils présentent trois types de fonctionnalités ciblées pour les personnes âgées : une fonction de géolocalisation utile au cas où la personne âgée se perd, des fonctions SOS pour alerter en cas de besoin, des fonctions de santé pour rappeler la prise de médicaments ou mesurer en temps réel l'activité, le rythme cardiaque, etc. Ces produits présentent plusieurs dissemblances : nécessité d'avoir un forfait pour apporter

une assistance (techniques, interventionnelle, etc.) (ex : Linkoo seniors), ceux qui ont une utilisation indépendante d'un support connecté (UnaliWear) et ceux qui ont des fonctionnalités incomplètes (Zembro). En ce qui concerne les « autres » objets connectés, il en existe une multitude liée à la santé et à l'alimentation, parmi eux le tensiomètre connecté, le pèse-personne connecté, le thermomètre connecté, le pilulier connecté, etc. Concernant les objets connectés liés à l'alimentation (verre connecté, assiette connectée, fourchette connectée, balance alimentaire connectée, boîtes alimentaires de conservation connectée, scanner connecté, etc.), la majorité d'entre eux sont depuis leur buzz restés en phase de prototypage. Ceux qui se sont retrouvés sur le marché n'ont pas réussi à attirer suffisamment les consommateurs (bug, gadget, application non-ergonome, etc.).

En général, pour qu'un consommateur puisse adopter une technologie, la notion du besoin associé au plaisir d'utiliser la produit technologique doit être créée. Pour que les consommateurs puissent intégrer les objets connectés dans leur quotidien, ces objets doivent avoir une utilisation simple et un design qui fait oublier les établissements de soins. Pour répondre aux besoins du secteur de la silver économie, l'association française de normalisation et de certification (Afnor) a mis en place un label pour identifier les objets du quotidien réellement adaptés en termes de confort, d'ergonomie, de fonctionnalité et de simplicité d'utilisation. Le but consiste à encourager les fabricants à améliorer leurs produits de différents secteurs pour mieux répondre aux besoins des personnes de plus de 60 ans. Cette organisation est la seule à représenter la France auprès de l'organisation internationale de normalisation (ISO) et du comité européen de normalisation (CEN). Le manque de rapport détaillé pour les produits labellisés incite à être plus vigilant quant à cette certification fournie par certains industriels. Actuellement, seules deux tablettes ont reçu le label « testé et approuvé par les seniors » (Ardoiz de la Poste et Facilotab) et aucun objet connecté n'en a obtenu jusqu'à ce jour. Le manque de concurrence freine les fournisseurs dans l'avancée de ces technologies (version d'Android non mis à jour, etc.). Ce déclin technologique contraint leur exploitation et les rend non compatibles avec les objets connectés technologiquement plus avancés.

Les objets connectés : essor attendu non observé

En 2020, les estimations considéraient que le marché des objets connectés risquait d'exploser dans les prochaines années. D'après les derniers chiffres publiés par Gartner et Idate (2017), entre 20 et 80 milliards d'objets connectés seront sur le marché en 2020 dépassant largement celui des smartphones et des tablettes et même le nombre d'humain sur terre. La consommation également est en pleine expansion. La même étude montre qu'en 2020, près de 13 Milliard de consommateurs auront adopté les objets connectés dans leur quotidien contre 8,4 milliards de personnes en 2017, soit une augmentation de 65 %.

Au départ, une partie de la problématique était centrée sur la collecte de données issues des objets connectés. En effet en 2017, les objets connectés avaient une place centrale dans l'innovation et la recherche. Les études spéculatives sur l'évolution des objets connectés étaient similaires à l'évolution qu'a connu le monde de la téléphonie mobile. Or depuis, les objets connectés ont technologiquement parlant reculé. De nombreuses options disparaissent peu à peu pour permettre à ces objets d'être fonctionnels tout en gardant une autonomie conséquente. Si nous prenons l'un des plus grands acteurs sur le marché comme exemple, la montre connectée Samsung Gear S3 (novembre 2016) et la

Samsung Galaxy Watch (3 août 2020) nous n’observons aucune évolution technologique majeure. De plus, le risque dans les projets d’objets connectés est, qu’avec le temps, ces produits soient de plus en plus délaissés (objets gadgets et non utiles) ou ne trouvent pas suffisamment de consommateurs cibles. Prenons l’exemple des « frigos connectés » qui peut plaire à un grand nombre de personnes, combien de personnes seraient prêtes à dépenser dix fois plus pour un réfrigérateur qui renseigne sur les courses à envisager.

Manque d’interopérabilité des objets connectés

La problématique liée au manque d’interopérabilité des objets connectés entraîne un coût important dans la recherche, la sélection et l’intégration de composants multiples IoT (appareils connectés, plateformes IoT, outils d’analyse, tableaux de bord de visualisation et applications logicielles). Le terme IoT, Internet of Things, est employé pour représenter les échanges d’informations et de données provenant de dispositifs retrouvés dans le monde réel (physique) et transmis vers le monde virtuel (réseau Internet). Ces dispositifs peuvent être soit des capteurs soit des actionneurs. Pour interagir entre eux et atteindre un but commun, ils utilisent une multitude de réseaux de communication sans fil (Bluetooth, WIFI, etc.). Par abus de langage, le terme IoT est parfois employé pour désigner l’interaction de plusieurs objets connectés entre eux ayant des usages diversifiés et dans différents domaines (la e-santé, la domotique ou encore le *quantified self* (mesure, analyse et partage de données personnelles)).

Le manque d’homogénéité dans les systèmes d’exploitation des objets connectés contraint à multiplier le développement d’interfaces compatibles pour la collecte des données sur lesdits objets connectés. De plus, l’hétérogénéité des formats des données pour chaque objet connecté complique également la tâche de la collecte des données. A l’heure actuelle, la récupération et l’exploitation des données se fait de façon propriétaire, ce qui demande un développement d’API personnalisé pour chaque objet connecté. Ajouté à cela, il suffit d’une mise à jour du système des objets connectés pour que toute l’API développée soit obsolète.

Pour répondre à ce problème d’interopérabilité des IoT, jusqu’à ce jour, il n’existe pas de normes définies par des réglementations officielles (telles que les normes W3C, IEEE, etc.). De nombreux projets de recherche internationaux (IEEE¹⁰, OneM2M¹¹) oeuvrent pour cet objectif et fournissent des solutions de normalisation, de développement et de déploiement intéressantes à exploiter dans de futurs projets. Parmi ces projets, citons le projet OneM2M qui inclut une centaine de membres (universitaires et industriels) travaillant actuellement sur la définition et le développement de normes pour répondre aux besoins de systèmes IoT interopérables et évolutifs. Le projet vise à fournir des standards ouverts, accessibles et reconnus mondialement pour permettre la communication et le transfert de données entre les systèmes (GEZER et TASKIN, 2016). Les solutions fournies par oneM2M pourraient être exploitées dans le futur pour rassembler notamment tous les composants de la pile de solutions IoT. Son architecture, qui définit une technologie middleware commune dans une couche horizontale entre les appareils et les réseaux de communication et les applications IoT, permet de standardiser

10. <https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/iot.pdf>

11. <https://www.onem2m.org/>

les liens entre les appareils connectés, les passerelles, les réseaux de communication et l’infrastructure cloud de stockage des données. Ainsi, il est nécessaire d’appréhender les systèmes incluant des objets connectés de manière distribuée et coopératives pour éviter de réinventer pour chaque objet connecté de nouvelles solutions. Néanmoins, jusqu’à présent, aucune obligation n’est décrite pour l’ensemble des acteurs commerciaux qui développent ces objets, ce qui rend impossible la pérennité des données à travers ces outils.

2.3.3 Les questionnaires en lignes

Les questionnaires en lignes présentent de nombreux avantages sur le plan opérationnel. Ils permettent de collecter rapidement les données de manière simple, accessible (REGMI et al., 2016). Les questionnaires en ligne ont également l’avantage, contrairement aux autres outils précédemment cités, d’être facilement modulaires et adaptatifs à l’évolution des connaissances. Ils peuvent également être utilisés par le personnel médical pour compléter les données du patient. Ce processus de collecte de données permet le suivi à long terme sans nécessiter une évolution majeure du support technologique. La majorité des questionnaires d’évaluation sont trop longs. Il n’est donc pas envisageable de poser toutes les questions en une seule fois. Les questionnaires sélectionnés et inclus dans les questionnaires en ligne ont été choisis selon leur contenu exploitable pour l’étude et le nombre de questions qu’ils contiennent. En effet, leur contenu doit nous permettre d’obtenir des informations pertinentes pour personnaliser les recommandations.

A partir des différents questionnaires analysés, nous dressons ci-dessous une liste de questionnaires courts qu’il est envisageable d’exploiter et d’intégrer dans des questionnaires en ligne :

- VQ11 : pour mesurer la qualité de vie en général (NINOT et al., 2010) ;
- Geriatric Quality of Life Questionnaire (GQLQ) : pour la mesure qualité de vie spécifiquement pour les personnes âgées (GUYATT et al., 1993) ;
- Mini Nutritional Assessment - MNA (KAISER et al., 2009) : pour identifier les personnes âgées dénutries ou présentant un risque de dénutrition ;
- Q-AAP+ : sur l’aptitude à l’activité physique pour tous ;
- International Physical Activity Questionnaire (IPAQ) (CLELAND et al., 2018) : pour interroger sur la pratique d’activités physiques et le temps passé assis au cours des sept derniers jours ;
- Starting The Conversation (STC) : pour évaluer des habitudes alimentaires (AE et al., 2011) ;
- Grille Fragilité SEGA (Short Emergency Geriatric Assessment)¹² : pour évaluer le profil gériatrique à l’admission ;
- Questionnaire sur les habitudes alimentaires¹³ : pour connaître certaines habitudes alimentaires.

12. <http://www.la-cosse.org/>

13. www.hss.gov.nt.ca/sites/hss/files/resources/healthy-eating-assessment-fr.pdf

En résumé

La plupart du temps, la personne ne connaît pas ses propres besoins, de ce fait, il lui est plus difficile de personnaliser et de suivre une recommandation. Ainsi, l'identification des besoins de la personne âgée à partir de la collecte de ses données, comme les habitudes alimentaires, (GOUVERNEMENT DES TERRITOIRES DU NORD-OUEST, 2017) représente un bon moyen pour suggérer des recommandations menant à une meilleure qualité de la prise alimentaire. A partir des outils de collecte potentiellement exploitables, notre stratégie consiste à exploiter les connaissances citées permettant de représenter la personne âgée et de personnaliser les recommandations. Dans la Figure 2.2, nous présentons un exemple d'exploitation des données au moyen de plusieurs outils de collecte (montre connectée, DMP, questionnaires en ligne) pour déduire de nouvelles connaissances.




Source acquisition	Relation	Type acquisition donnée	Donnée brute	Nouvelles relations inférées	Nouvelles connaissances déduites
	Nb Pas Quotidien	mesurée	3800	Mode de vie	Sédentaire
	Temps inactivité continue	mesurée	4,3		
	Opération	collectée	Prothèse hanche	Activité autorisée Activité interdite	Golf, natation, vélo Hockey, Karaté, baseball
	Frein Activité	déclarée	Peur chuter	Qualité physique à développer	Equilibre
	Localisation Activité Préférée	déclarée	Intérieur	Activité suggérée	?
	Contexte Activité Préférée	déclarée	Loisir		

FIGURE 2.2 – Exemple de déduction de nouvelles connaissances et de personnalisation des recommandations à partir des données collectées

Au début du projet de thèse, nous prévoyons d'exploiter les données issues d'objets connectés pour faciliter la collecte de données auprès de la personne âgée. Cependant, nous avons été confrontés à de nombreuses problématiques intrinsèques au manque d'interopérabilité entre les systèmes, les langages et les applications cités dans la section 2.3.2.2. En effet, lors du développement d'une API spécifique à une montre connectée, il a suffi d'une mise à jour de la montre connectée pour que l'API ne soit plus fonctionnelle. Ceci montre la nécessité de répondre à ce problème lié au manque d'interopérabilité en raison notamment de l'hétérogénéité des OS, des périphériques, des protocoles réseau, des applications et des données fournies, transmises et exploitées par ces objets connectés.

En raison des difficultés présentées pour la collecte de données en santé, nous avons décidé d'explorer les questionnaires en ligne pour la conception du support de collecte de données exposée dans le chapitre 4.

Deuxième partie
Contributions

Évolution de ressources sémantiques modulaires

Sommaire

3.1	Introduction	51
3.2	Synthèse des approches d'évolution existantes.....	52
3.3	Évolution de l'ontologie MIAM	55
3.3.1	Etape 1 : Détection des besoins pour l'évolution	55
3.3.2	Etape 2 : Suggestion des changements	61
3.3.2.1	Définition des questions de compétences et des scénarios d'usage.....	61
3.3.3	Etape 3 : Suggestions des changements pour l'évolution de MIAM.....	63
3.3.3.1	Présentation de MIAM	63
3.3.3.2	Modules concernés par l'évolution.....	63
3.3.3.3	Importation des modules.....	63
3.3.3.4	Evolution du module Aliment et Nutrition	64
3.3.3.5	Evolution du module Personne	66
3.3.4	Développement de nouveaux modules	67
3.3.4.1	Les modèles conceptuels existants d'activité	67
3.3.4.2	Ontologies existantes	70
3.3.4.3	Scénarios et questions de compétences	72
3.3.4.4	Module Activité.....	72
3.3.4.5	Module ConditionPhysiologique	75
3.3.4.6	Modules ObjetConnecté et Mesure.....	77
3.3.5	Etape 4 : Évaluation de l'impact de l'évolution	78
3.3.6	Etape 5 : Gestion des versions	79
3.4	Conclusion	80

Ce chapitre a pour objectif de présenter les étapes menant à l'évolution de l'ontologie MIAM vers ONAFE dans le but de proposer des services contribuant à la construction de recommandations nutritionnelles et d'activités physiques personnalisées et adaptées aux besoins des personnes âgées. Nous présentons à travers cette évolution des exemples concrets impliquant l'évolution d'ontologie

modulaire de domaine.

Textes de références

- Dandan R., Despres S. : "Evolution of Modular Ontology : Application to Personalization", Terminology & Ontology : Theories and applications, TOTh 2021, Chambéry, France, 2021. (A paraître)
- Dandan R., Despres S., Nobecourt J. : "OAFE : an Ontology for the Description of Elderly Activities". 14th International Conference on Signal-Image Technology & Internet-Based Systems, SITIS 2018 (IEEE), pp. 396–403, Gran Canaria, Italy, 2018.
- Dandan R., Despres S., Nobecourt J. : "OAPA : Une Ontologie pour la description de l'Activité des Personnes Agées", 7ème Journée Francophones sur les Ontologies, JFO 2018, pp.16-27, 2018.

3.1 Introduction

L'évolution d'ontologie fait partie des activités du cycle de vie d'une ontologie. En effet, une fois qu'une ontologie a été conçue, les changements des exigences, la modification des besoins, l'évolution des connaissances modélisées et la correction des défauts de conception nécessitent d'appliquer un certain nombre de changements altérant la hiérarchie de l'ontologie. En raison de la variété des conséquences causées par ces changements, l'évolution d'ontologie n'est pas un problème trivial et représente un réel challenge dans le domaine de l'ingénierie des ontologies.

De plus en plus d'études soutiennent l'intérêt et le besoin de développer des ontologies modulaires, accessibles et compréhensibles par les experts du domaine (KHAN et KEET, 2021). Une ontologie modulaire contient des composants réutilisables, autonomes qui entretiennent des relations définies avec d'autres ontologies modulaires. La modularité représente une solution efficace pour simplifier le développement d'une ontologie ayant un domaine large et complexe (D'AQUIN et al., 2009). De plus, le fait que chaque module ait été conçu de sorte à être auto-suffisant permet de maximiser à la fois la réutilisabilité ainsi que la maintenabilité de l'ontologie globale (JARRAR, 2005). De ce fait, l'ajout de nouveaux modules et l'évolution d'un ou plusieurs modules pourra se faire de manière indépendante tout en générant le minimum d'effets de bord.

Précédemment, l'ontologie modulaire MIAM (DESPRES, 2016) a été conçue pour la suggestion de recettes « bien-être » pour la population générale tenant en compte des préférences nutritionnelles et sensorielles des utilisateurs et permettant aux utilisateurs de s'alimenter de manière équilibrée (DESPRES, 2016). Les recommandations nutritionnelles s'appuient sur les recommandations du PNNS 3 et sur l'expertise en nutrition de l'Unité de Recherche en Épidémiologie Nutritionnelle (UREN). Dans ce projet, nous nous sommes intéressés aux recommandations nutritionnelles et d'activités physiques selon les données collectées sur personnes âgées (besoins, pathologies, habitudes alimentaires, préférences sensorielles, etc.). Pour être en mesure de tenir compte des données et des connaissances nécessaires à la personnalisation des recommandations nutritionnelles et d'activité disponibles, l'évolution de MIAM doit être étudiée.

L'objectif de ce travail consiste à présenter les étapes menant à l'évolution de l'ontologie modulaire MIAM nécessitant le développement d'autres modules spécifiques au contexte. Tout comme l'évolution d'ontologie, le développement de ressources terminologiques (RTO) impliquant la réutilisation de ressources déjà existantes représente un domaine de recherche actif dans l'ingénierie d'ontologie. A travers cet objectif, nous aborderons les questions de recherche suivantes :

- quelle approche d'évolution d'ontologie adopter pour faire évoluer une ontologie modulaire nécessitant la construction de nouveaux modules ?
- comment exploiter les connaissances existantes du domaine pour personnaliser les recommandations ?
- les ontologies et les connaissances du domaine sont-elles fiables pour être réutilisées (par exemple, maintenues, représentatives) ?

Avant de présenter le processus d'évolution de MIAM, nous présentons dans la section qui suit une synthèse des différentes approches pour la conception et l'évolution d'ontologie.

3.2 Synthèse des approches d'évolution existantes

Il n'existe pas de consensus quant à la définition de l'évolution d'ontologie. De nombreuses définitions ont été attribuées à l'évolution d'ontologie de par la complexité et la diversité des étapes incluses dans ce processus. Parmi les définitions sélectionnées, l'évolution d'ontologie est définie comme :

- « une adaptation adéquate afin de la maintenir à jour tout en respectant les changements survenus dans le domaine et les nouvelles exigences du système d'information utilisé par l'ontologie » (HAASE et STOJANOVIC, 2005).
- « un processus de modification d'une ontologie en réponse à un certain nombre de changements du domaine ou de sa conceptualisation » (FLOURIS, PLEXOUSAKIS et al., 2006).
- « l'adaptation opportune d'une ontologie aux changements survenus et la propagation cohérente de ces changements aux artefacts dépendants. Elle doit être considérée comme un processus qui englobe l'ensemble des activités, à la fois techniques et de gestion, qui garantissent que l'ontologie continue de répondre aux objectifs organisationnels et aux besoins des utilisateurs de manière efficace et efficace » (STOJANOVIC, 2004).

Ces définitions montrent la complexité de cette activité en raison de la variété des activités qui y sont liées et des conséquences liées aux changements des exigences et du domaine. En effet, l'évolution d'ontologie implique une dizaine de sous-domaines : la correspondance, l'alignement, l'articulation, la traduction, l'évolution, le débogage, le versionnage, l'intégration, le fusionnement, etc.) (WARDHANA et al., 2018).

Ces 20 dernières années, un intérêt croissant s'est porté sur l'étude du processus d'évolution d'ontologies et une variété d'approches, de techniques et d'outils ont été proposés pour assurer la réussite des changements opérés lors de l'évolution d'ontologie. Nous présentons pour commencer une synthèse des différentes approches d'évolution retrouvées dans la littérature.

Pour réaliser notre comparaison entre les différentes approches existantes concernant l'évolution d'ontologie, nous nous sommes principalement intéressés aux travaux présentant des étapes d'évolution. Pour structurer notre synthèse des différents cadres d'évolution retrouvés dans la littérature, nous nous sommes appuyés sur le cycle d'évolution proposé par ZABLITH, ANTONIOU et al., 2015 qui fait la distinction entre ces différents termes.

ZABLITH, ANTONIOU et al., 2015 synthétisent les différentes étapes retrouvées dans les différentes études au sein d'un cycle d'évolution suivant cinq étapes, où chaque étape s'appuie potentiellement sur différentes entrées et informations contextuelles représentées dans les cercles :

1. **Détection du besoin** : cette étape est le point de départ pour toute évolution. Elle permet de détecter si de nouveaux concepts et relations doivent être ajoutés à l'ontologie, ou si certains éléments de l'ontologie peuvent être supprimés ;
2. **Suggestion des modifications** : cette étape permet de spécifier les opérations concrètes de changement nécessaire pour faire évoluer l'ontologie. Elle s'apparente aux étapes de représentation de STOJANOVIC et MOTIK, 2002, à l'étape de

transformation des données de KLEIN et al., 2003 et à la tâche de découverte de relation d'Evolva ZABLITH, SABOU et al., 2009 ;

3. **Validation des modifications** : cette étape filtre les changements qui ne devraient pas être ajoutés à l'ontologie. Elle permet de s'assurer à la fois que l'ontologie est logiquement cohérente et consistante selon les contraintes spécifiées (validation formelle) et que les changements du domaine sont pertinents ;
4. **Évaluation de l'impact** : L'évaluation est réalisée en fonction de l'impact sur les artefacts externes qui dépendent de l'ontologie (capacité de répondre à des requêtes) ou selon des critères spécifiques (coût, avantages, etc.). Elle correspond à l'étape de propagation de NOY et KLEIN, 2004 ;
5. **Gestion des modifications** : Cette étape représente l'étape d'application, d'enregistrement des modifications et de suivi des différentes versions de l'ontologie.

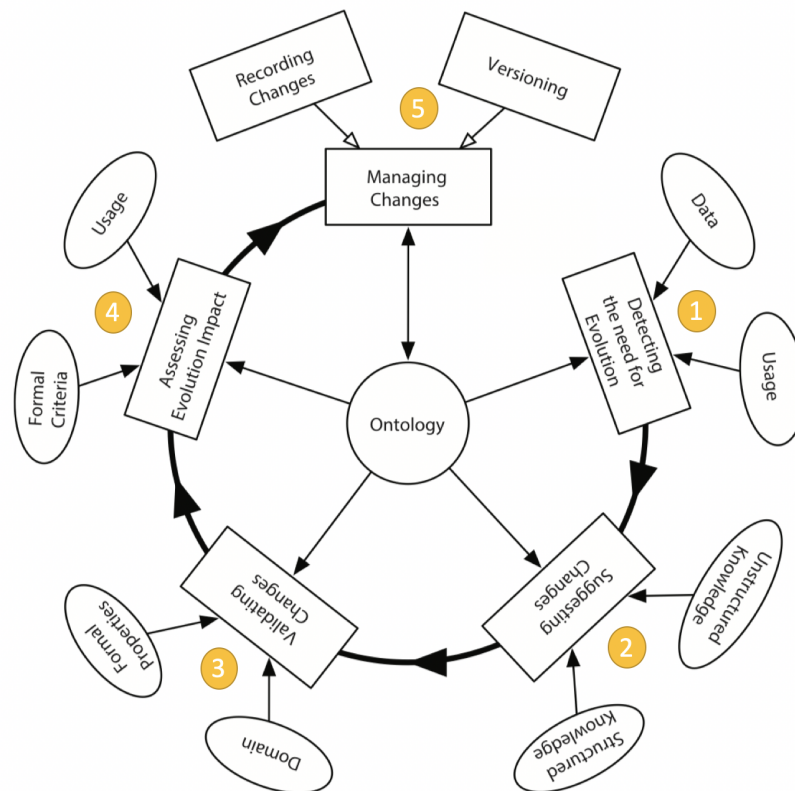


FIGURE 3.1 – D'après le cycle d'évolution proposé par ZABLITH, ANTONIOU et al., 2015

Pour résumer les différentes approches d'évolution retrouvées dans la littérature, nous présentons dans la Figure 3.2 un schéma comparatif des différentes étapes d'évolution pour un panel d'études d'évolution d'ontologie, ceci en s'appuyant sur les étapes du cycle de ZABLITH, ANTONIOU et al., 2015. Lors de cette comparaison, nous faisons la distinction entre les phases du processus d'évolution et les opérations de changement à appliquer sur l'ontologie (ajout, suppression, modification) sur les entités de l'ontologie (concept, propriété, les instances et l'héritage de concepts).

L'étape de la détection du besoin est retrouvée dans la majorité des étapes d'évolution (STOJANOVIC, 2004, A. M. KHATTAK et al., 2009, etc.).

STOJANOVIC, 2004 aborde les différentes étapes du processus d'évolution et propose une vue d'ensemble du problème d'évolution d'ontologie. Par la suite, FLOURIS, MANAKANATAS et al., 2008 fournissent une étude clarifiant ces étapes en se concentrant sur les changements opérés sur l'ontologie lors de son évolution. Il identifie également les domaines de recherche traitant des aspects de changement d'ontologie, et délimite leurs frontières tout en fournissant des clarifications terminologiques pour chaque type de changement. DE LEENHEER et MENS, 2008 explorent les approches de l'évolution d'ontologies conçues dans des environnements distribués et collaboratifs dans lesquels l'aspect dynamique et le niveau de complexité du domaine est plus important que les ontologies mono-utilisateur. Plus récemment, K. I. KOTIS et al., 2020 ont exploré les méthodologies d'ingénierie d'ontologie pour l'évolution et la réutilisation d'ontologies vivantes notamment dans le cadre d'évolution d'ontologies collaboratives. A partir de son étude, il met l'accent sur l'importance de ne pas négliger les recommandations des MIO pour l'évolution d'ontologies vivantes et fournit des recommandations sur la base de l'analyse réalisée. Chacune de ces approches apporte une vision parfois similaire, parfois différente quant à l'activité d'évolution d'ontologie. Bien que ces approches soient motivées par le besoin de construire un cadre d'évolution de l'ontologie, elles proviennent d'une variété de domaines de recherche disjoints rendant la compréhension, l'adoption et l'application des tâches d'évolution plus difficile.

La majorité des cadres d'évolution reconnaissent la nécessité de la tâche de gestion des versions de l'ontologie, cependant, ils présentent cette tâche comme étant une activité distincte, en dehors de la tâche d'évolution. FLOURIS, MANAKANATAS et al., 2008 considèrent d'une part cette tâche en dehors du processus d'évolution d'ontologie et d'autre part comme étant une tâche permettant la propagation des changements. L'avantage du cycle d'évolution proposé par ZABLITH, ANTONIOU et al., 2015 est qu'il inclut la tâche de versionnage dans les tâches d'évolution d'ontologie. De plus, il offre une granularité plus élevée pour faire évoluer les ontologies. En effet, pour chaque changement opéré sur l'ontologie, l'application des étapes de cycle est effectuée de manière itérative. Enfin, leur cycle inclut les fonctionnalités associées à chaque étape (cercles), souvent considérées en dehors du cycle.

Cette comparaison met en évidence l'inclusion des étapes pour chaque cadre et la différence entre le vocabulaire employé. La majorité des études appliquent le processus identifié par STOJANOVIC, 2004. Très peu d'études montrent les conséquences et les répercussions liées à l'application des opérations de changement STOJANOVIC, 2004. SAFYAN et al., 2019 identifie trois opérations de changements possibles : ajouter des sous-concepts spécialisés, ajouter des sous-concepts étendus, modifier des concepts existants. Certaines études identifient les changements à opérer avant l'évolution et considèrent l'étape de gestion des changements comme étant le début de l'évolution d'ontologie (HAASE et STOJANOVIC, 2005). Dans la majorité des cadres d'évolution, la distinction entre la gestion, la modification, l'évolution et le versionnage d'ontologies est, dans certains cas, imprécise (WARDHANA et al., 2018, ZABLITH, ANTONIOU et al., 2015, A. KHATTAK et al., 2013, etc.).

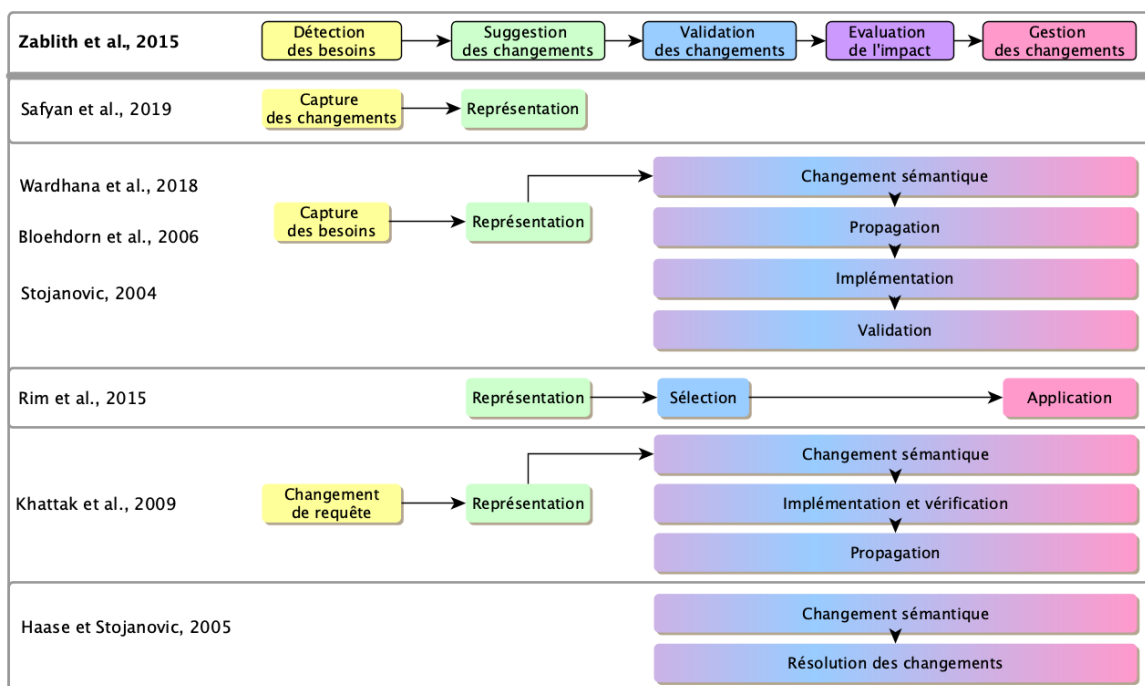


FIGURE 3.2 – Comparaison des étapes d’évolution en référence avec les étapes du cycle proposé par ZABLITH, ANTONIOU et al., 2015

Approche d’évolution adoptée

La synthèse réalisée nous a permis de nous positionner au sein des approches existantes et d’identifier quelle approche répond au mieux à nos besoins d’évolution. En général, la stratégie adoptée pour faire évoluer une ontologie dépend de l’ampleur des modifications à apporter sur l’ontologie à faire évoluer. Pour que l’approche choisie puisse accompagner l’évolution d’ontologie, il est nécessaire d’identifier clairement les tâches et les sous-tâches qui sous-tendent cette évolution. Pour cela, en plus des étapes d’évolution, les opérations de changement possibles selon les entités de l’ontologie doivent être intéressantes à prendre en considération lorsqu’elles sont précisées. Au vu de la granularité du cycle d’évolution proposé par ZABLITH, ANTONIOU et al., 2015 et de l’inclusion du processus de gestion des versions, nous nous sommes appuyés sur les étapes du cycle (Figure 3.1) pour faire évoluer MIAM.

3.3 Évolution de l’ontologie MIAM

3.3.1 Etape 1 : Détection des besoins pour l’évolution

L’étape de détection des besoins représente la première étape d’évolution d’ontologie et nécessite une acquisition de connaissances (cf. section 2). L’identification des changements nécessite une acquisition de connaissances. Elle est réalisée à partir de connaissances non structurées (texte brut, schéma, tableau, etc.) provenant principalement de sources officielles (PNNS, HAS, OMS, etc.) ou de structures de santé

spécialisées (HUG, etc.).

Les questions qui se sont posées sont par exemple : Quelles sont les connaissances exploitables pour spécifier un état de santé ? Quelles sont les outils exploitables pour collecter des données relatives à un état de santé ? Quelles ressources, éléments des questionnaires et données permettent de faire le lien entre la donnée de la personne et la recommandation à fournir ?

L'évolution de l'ontologie MIAM est motivée par les changements concernant :

- les connaissances du domaine : concernant les recommandations nutritionnelles ;
- les utilisateurs cibles : les personnes âgées pour prévenir la dénutrition et sédentarité et répondre à des problématiques spécifiques non concernés lors de la conception de MIAM ;
- les besoins de collecte de données : pertinentes et nécessaires à la spécification des recommandations fournies ;
- l'usage applicatif et des recommandations fournies : l'environnement dans lequel l'ontologie sera utilisée inclue les moyens de collecte de données (OC, questionnaire en ligne, etc.)

Changement des recommandations nutritionnelles du PNNS

MIAM a été conçue en s'appuyant sur les recommandations du PNNS 3 et sur l'expertise en nutrition de l'Unité de Recherche en Epidémiologie Nutritionnelle (UREN). Depuis la conception de MIAM, les recommandations du PNNS ont évolué passant du PNNS 3 au PNNS 4. Le but recherché à partir de cette évolution dépendait de la nécessité de simplifier les messages, d'être plus interactif et d'inclure les aliments délaissés. Cette mise à jour nous a poussés à devoir modifier certains concepts existants dans MIAM.

Dans la nouvelle version du PNNS, trois familles d'aliments supplémentaires ont été créées (les légumes secs, les charcuteries, les fruits à coque), certaines familles existantes ont été modifiées (féculents complets, poisson), et certaines familles à limiter ont été regroupées (produits gras, sucrés, salés et ultra-transformés). Dans certains cas, comme pour celui-ci, il devient plus évident de créer à nouveau le concept décrivant les nouvelles familles de PNNS plutôt que d'appliquer une série d'opérations sur les concepts existants.

Le vocabulaire employé pour désigner les catégories d'aliments n'est pas toujours identique selon les sources. Ainsi, une attention particulière a été accordée à ce point pour éviter d'introduire des incohérences sémantiques. Une incohérence sémantique se produit lorsque la signification d'une entité change au cours du processus d'évolution (TAMMA et BENCH-CAPON, 2001). En plus du changement de vocabulaire employé, les recommandations ont également évolué. Par exemple, la recommandation de produits laitiers passe de 2 par jour pour les adultes à 3 ou 4 par jour pour les personnes âgées. Ces différences de recommandations nécessitent le remaniement d'un certain nombre de concepts décrivant les groupes alimentaires du PNNS, les portions recommandées pour les définir selon le profil de l'individu.

Changement des besoins de collecte de données

Les besoins de collecte de données sont conditionnés par les besoins spécifiques de la personne âgée. Pour répondre à leur besoin, des données sur leur état de santé, leurs préférences (alimentaires, sensorielles, etc.), leurs contraintes, leur contexte, leur environnement (social, etc.), leur niveau de dépendance et autres encore doivent être collectées. Les besoins en termes de collecte de données sont également guidés par le besoin de construire un profil à partir duquel la recommandation pourra être personnalisée. En effet, l'objectif étant de détecter les signes (de dénutrition, de sédentarité, etc.) pour mieux recommander.

Les outils de collecte retenus pour construire le profil de la personne sont les objets connectés et les questionnaires (Figure 3.4). Les données collectées peuvent être de différents types : brutes ou calculées. Chaque type de données nécessite une utilisation, une collecte et donc un traitement différent. Parmi ces traitements, des calculs peuvent être réalisés à partir de certaines données. Ainsi, à partir des données brutes collectées, un certain nombre de données peuvent être exploitées pour compléter le profil (Figure 3.3). Par exemple, le calcul de l'IMC à partir du poids et de la taille permet de personnaliser la collecte de données. Dans le cas d'un IMC inférieur à 18, le questionnaire MNA pourra être proposé à la personne âgée pour spécifier son profil nutritionnel.

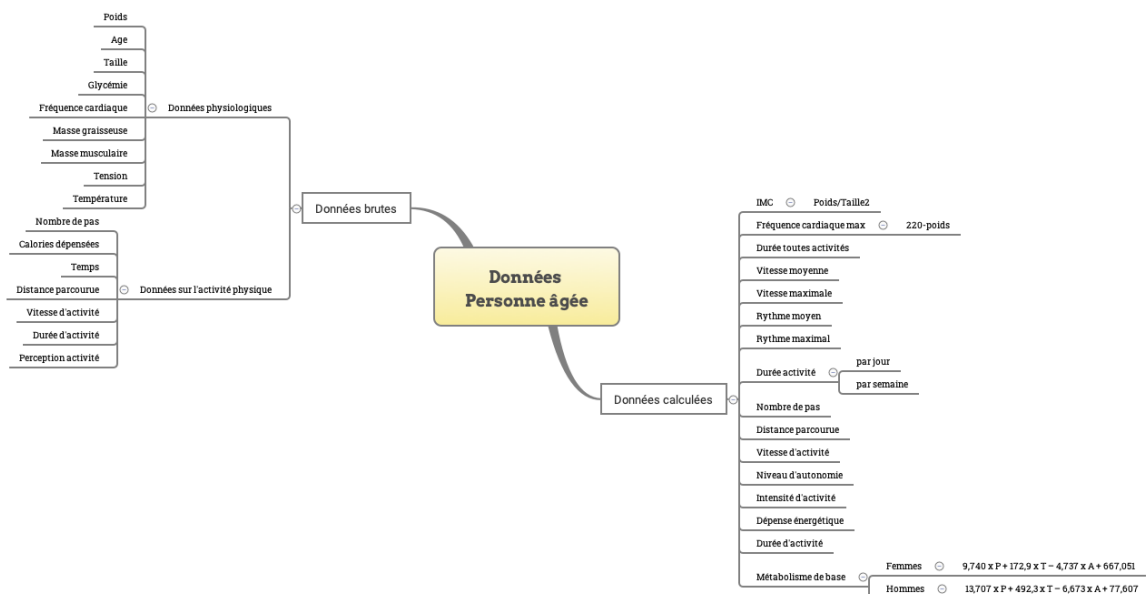


FIGURE 3.3 – Distinction entre données brutes et données calculées

Les données collectées auprès de la personne âgée sont de différents types : lorsqu'elles sont issues de capteurs, elles sont mesurées ; lorsqu'elles sont issues des questionnaires, elles sont déclarées ; et lorsqu'elles sont analysées, quelle que soit leurs sources, elles sont calculées. A partir des objets connectés, les données relatives au profil d'activité (nombre de pas, temps d'inactivité, etc.), au profil de santé (glycémie, etc.) sont collectées. A partir des questionnaires en ligne, toutes les données considérées utiles pour personnaliser les recommandations peuvent être prises en compte et sont relatives aux cas traités

dans l'étude. La Figure 3.1 présente un exemple d'utilisation de données selon les moyens d'acquisition. A partir de ces moyens de collecte, l'objectif est d'être en mesure d'exploiter ces données pour déduire de nouvelles connaissances et personnaliser la recommandation.

Entité	Moyen d'acquisition	Type d'acquisition	Exemple de données	Utilisation
Age	Lecteur carte vital	Déclaré	85	Calcul IMC, Recommandation activité plus personnalisée
Genre	Lecteur carte vital	Déclaré	Femme	Recommandation activité plus personnalisée
Taille	Lecteur carte vital	Déclaré	160	Calcul IMC
Poids	Pèse personne connecté/DMP	Mesuré	52	Calcul IMC, Calcul du métabolisme de base Calcul fréquence cardiaque max = 220-poids
IMC	Lecteur carte vital/DMP	Calculé	18.3	Maigreur
Glycémie	Glucomètre connecté	Mesuré	Glycémie faible	Suggestion possible de repas riches en glucose
Fréquence cardiaque	Montre connectée	Mesuré	98	Suggestion d'activités permettant de diminuer l'essoufflement, ...
Etat de santé	DMP	Signalé	Cardiopathie ischémique connue et profil bas risque	Suggestion d'activités physiques d'intensité légère à modérée
Distance parcourue	Montre connectée	Mesuré	0,2km/jour, 2km/sem.	Calcul du niveau d'activité
Calories dépensées par activité	Montre connectée	Mesuré	30 kcal	Calcul des dépenses énergétiques au total selon l'intensité et le rythme de l'activité exercée
Niveau d'autonomie	Serious Game	Déclaré	Normal	Suggestion d'aller chercher du pain à la boulangerie
Caractéristiques sociales	Serious Game	Déclaré	Vie seule, possède des petits enfants	Emmener ses petits enfants au jardin
Préférences	Serious Game	Déclaré	Aime les pommes	Suggestion de recette contenant des pommes et riche en glucose
Intolérance alimentaire	Serious Game / DMP	Déclaré	Lactose	La recette de tarte aux pommes ne contiendra pas de crème liquide (source de lactose)
Habitude alimentaire	Serious Game	Déclaré	1 verre d'eau/jour	Rappel de consommation d'eau

TABLEAU 3.1 – Exemple d'utilisation de données selon les moyens d'acquisition

Les habitudes alimentaires sont généralement difficiles à définir rapidement du fait que la majorité des questionnaires se réalisent sur une longue période. Nous pouvons néanmoins considérer les questionnaires traitant des portions pour orienter la recommandation. Par exemple, lorsque la personne consomme par jour une quantité d'aliments (viande, produits laitiers, etc.) inférieure aux recommandations, les recommandations incluront celles du PNNS.

Changement de l'usage applicatif et des recommandations fournies

Pour mieux vieillir, les facteurs de santé impliqués dans la modification de l'alimentation ou de l'activité doivent être considérés pour une meilleure prise en charge de l'état de santé des personnes âgées. Certaines pathologies (diabète, cancer, traitements, etc.) peuvent perturber le statut nutritionnel et ainsi causer une dénutrition, un amaigrissement ou au contraire une surcharge pondérale. Les pathologies demandant une modification du régime alimentaire ont pour cela été prises en compte dans notre conception pour personnaliser les recommandations. Les pathologies incluses sont principalement celles pour lesquelles une recommandation est disponible. Par exemple, les personnes diabétiques et hypertendues doivent limiter la consommation de produits riches en sel.

Pour retarder l'apparition de soucis de santé liés à l'âge (ostéoporose, hypertension, troubles digestifs, cholestérol, maladies cardio-vasculaires, etc.) et l'apparition de

troubles nutritionnels liés à des soucis de santé (dénutrition, troubles de la mastication, etc.), les besoins nutritionnels doivent être identifiés pour être comblés. Cependant, il n'est pas évident de faire adopter les recommandations nutritionnelles et changer le comportement et les habitudes alimentaires d'une personne de sorte à ce que son alimentation puisse couvrir ses besoins énergétiques (30 kcal/kg/jour) et nutritionnels (protéiques, glucidiques, vitamines et minéraux, etc.). Ainsi, la recommandation d'aliments riches en énergie ou en protéine permettra de répondre aux besoins nutritionnels de la personne âgée dénutrie.

La dénutrition est un processus insidieux et peu visible. Pour lutter contre la dénutrition, la détection précoce des signes de dénutrition est essentielle. Pour cela, la détection des causes (IMC faible, perte d'appétit, perte récente de poids, etc.) aide à orienter la recommandation. Les différentes stratégies de prise en charge consistent à : enrichir l'alimentation, adapter l'alimentation aux troubles digestifs, améliorer le contexte de la prise des repas. Nous nous sommes également intéressés à l'enrichissement de l'alimentation. Elle consiste à enrichir l'alimentation traditionnelle en augmentant les apports en protéines et calories des éléments constitutifs (entrées, plats, laitages, desserts, collations, boissons, etc.), par exemple après le repas sans en augmenter leur volume. Différents produits de base peuvent avoir cet effet (poudre de lait, lait concentré entier, fromage râpé, œufs, crème fraîche, beurre fondu, huile ou poudres de protéines industrielles, pâtes ou semoule enrichie en protéines, etc.). Par exemple, il est recommandé d'avoir en réserve une liste d'aliments pour enrichir les préparations ou collations (par exemple : poudre de lait, lait condensé, fromage, œufs, jambon, crème, huile, beurre, etc.), les repas (par exemple : yaourt à la crème, flan, crème dessert, oléagineux (noix, amandes, etc.), fruits secs, biscotte, pain grillé, charcuterie, etc.)¹.

De bonnes recommandations ne peuvent être obtenues que lorsqu'un modèle précis des utilisateurs et de leurs besoins est accessible. Pour répondre aux besoins nutritionnels des personnes âgées, ils doivent être dans un premier temps identifiés. Cela passe par une définition de la personne la plus complète possible. La caractérisation de la personne selon les différents aspects répondant aux besoins de l'application, consiste à définir l'ensemble des données utilisées pour construire le profil.

Les recommandations pour boire suffisamment peuvent également être étudiées dans les recommandations fournies, par exemple, au petit-déjeuner : 1 grand bol de café, de chicorée ou de thé et 1 verre de jus d'orange. Les repas sont au nombre de 4. Alors que les moments de la journée où il est préférable pour une personne déshydratée de boire sont plus nombreux : petit déjeuner, matinée, déjeuner, goûter, dîner, soirée. Tous ces éléments peuvent être définis dans le module Nutrition.

Pour personnaliser au mieux les recommandations et permettre une prise en charge personnalisée, il est indispensable de définir le profil de la personne âgée correspondant au mieux à son état de santé, ses envies, ses objectifs, ses motivations, ses freins et à son environnement. Pour cela, un certain nombre de données sont à collecter.

Les recommandations nutritionnelles que nous cherchons à fournir concernent des recommandations associées à un aliment, un groupe alimentaire du PNNS, un régime alimentaire. Ainsi, à partir des différentes sources exploitées, les recommandations retenues tiennent compte des problèmes de santé rencontrés chez les personnes âgées (cf. chapitre 2). La recommandation suggérée à l'utilisateur peut se présenter sous forme

1. <https://www.hug.ch/sites/interhug/files/documents/alimentation-hachee.pdf>

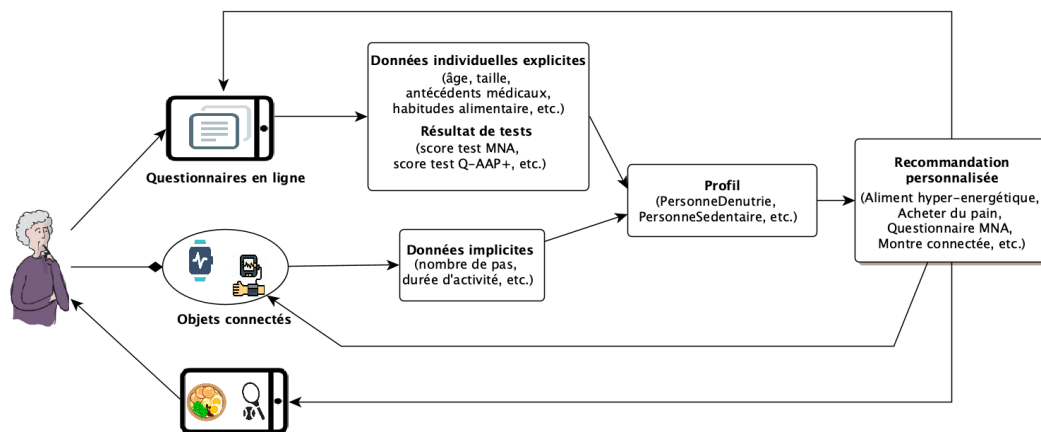


FIGURE 3.4 – Modèle d'exploitation des données

d'une liste :

- d'activité ou de programme d'activité : pour l'inciter à pratiquer des activités physiques en lien avec son état de santé, ses préférences, etc. ;
- d'aliments ou de régimes nutritionnels : pour enrichir son alimentation ;
- de questionnaires personnalisés provenant des tests d'auto-évaluation courts : pour cibler les déterminants de santé et construire le profil gérontologique de la personne âgée et tenant compte des caractéristiques individuelles.

Enfin, pour fournir une recommandation personnalisée adaptée aux besoins individuels de chaque individu, les données sur la personne âgée concernant les différents termes retrouvés dans les questionnaires d'évaluation : de l'état de santé, du statut nutritionnel et d'activité, sont exploitées pour compléter les données associées à la personne. Par exemple, dans le cas de la dénutrition, nous cherchons dans un premier temps à détecter les signes de la dénutrition pour pouvoir prévenir son apparition et personnaliser ainsi les recommandations. De ce fait, nous avons inclus dans nos usages applicatifs le comportement de l'application par rapport aux données collectées. Par exemple, lorsque la personne indique un poids et une taille permettant l'obtention d'un IMC inférieur à la normale, alors, le questionnaire MNA (composé de 6 questions) pourra être proposé à la personne âgée. Le questionnaire MNA est un outil validé et largement utilisé en clinique pour le dépistage et l'évaluation nutritionnelle. Il est utilisé en gériatrie pour détecter les patients en dénutrition ou à risque de dénutrition. Cet exemple montre comment la personnalisation de la collecte de données peut évoluer. L'inclusion du comportement du système à notre étude permet d'améliorer l'étape de collecte de données et la construction des questionnaires en ligne.

En résumé, les connaissances acquises pour l'identification des changements concernent :

- les besoins spécifiques des personnes âgées,
- les données et les éléments sur lesquels il est possible d'agir pour cibler la recommandation,
- les éléments permettant la construction du profil,

- les recommandations nutritionnelles exploitables et spécifiques pour une pathologie étudiée, un signe, un symptôme, etc.

3.3.2 Etape 2 : Suggestion des changements

La suggestion des changements permet d'identifier les changements à opérer sur les modules concernés par l'évolution. Pour réaliser cette étape, nous ne nous sommes pas limités au contenu disponible dans les documents analysés. Nous avons étendu notre réflexion pour tenir compte des connaissances préalablement modélisées dans MIAM. Nous exposons les suggestions de changements par la définition de questions de compétences et de scénarios d'usage.

3.3.2.1 Définition des questions de compétences et des scénarios d'usage

A partir de l'acquisition des besoins, des questions de compétences et des scénarios d'usage (GRÜNINGER et FOX, 1995) ont été définis. Les questions de compétences consistent en un ensemble de questions auxquelles une ontologie devrait pouvoir répondre selon un scénario d'usage (GRÜNINGER et FOX, 1995). Un large éventail de questions de compétences, leur utilité dans la création d'ontologies et leur intégration possible dans des outils de création ont été étudiés (REN et al., 2014, HOFER et al., 2015, DENNIS et al., 2017). Traditionnellement, les questions de compétences sont utilisées lors du développement d'ontologies pour rassembler les besoins fonctionnels des utilisateurs dans des cas d'utilisation spécifiques garantissant que toutes les informations pertinentes sont codées (JACOBSON et al., 1992). Les questions de compétences constituent ainsi un moyen pour déterminer les spécifications de l'ontologie à concevoir et d'évaluer un modèle de connaissances. Par exemple, une réponse à une question de compétences peut être obtenue simplement en parcourant la hiérarchie des concepts, ou en visualisant les relations entre les entités de l'ontologie, ou en déterminant l'héritage, ou par d'autres mécanismes (interrogation d'ontologie, etc.).

Les questions de compétences établies permettent d'explicitier les besoins précédemment identifiés pour faire évoluer l'ontologie et correspondent à des interrogations exprimées par la personne âgée ou les aidants la prenant en charge. Elles sont construites de sorte à répondre aux interrogations de la personne âgée concernant les aliments favorables à consommer ou au contraire à limiter selon les besoins de la personne âgée. Ces besoins peuvent être liés à des préférences sensorielles ou à un régime alimentaire à suivre. Ces différents aspects doivent donc être pris en considération pour aboutir à des recommandations pertinentes. Prenons comme exemple la question suivante : « Combien de fois par jour avez-vous consommé des produits laitiers (lait, yogourt sans sucre ou fromage faible en gras) ? ». Dans le cas d'une réponse inférieure à 2 fois par jour, nous devons pouvoir inclure dans les recommandations tous les produits de la catégorie ProduitsLaitiers pour que la personne puisse faire son choix selon ses propres préférences.

Les questions de compétences conduisant à une évolution de MIAM ont été catégorisées selon les thèmes abordés et ont permis de définir les modules concernés par les changements (Figure 3.5). Ce filtrage montre un aperçu des questions de compétences et des changements à opérer sur les différents modules de MIAM impliqués dans l'évolution. A partir des questions de compétences identifiées à partir de forums ou de problèmes

Questions concernant	Sources de connaissance	Changements
Portion journalière ou hebdomadaire	Recommandation PNNS4	Module Nutrition
		Module Aliment
Substitution de produits	Module Aliment	Pas de changement
Composition des aliments		Ajouts éventuels
	Table de composition des aliments	Module Nutrition
Caractère allergène des aliments		Module Aliment
	Module Nutrition	Module Aliment
Aliments recommandés pour le contrôle maladie		Module EtatDeSanté
	Aliments recommandés pour le contrôle maladie	Module Aliment

FIGURE 3.5 – Aperçu du filtrage des questions de compétences conduisant à une évolution de MIAM

fréquents, les scénarios d’usage ont été construits pour déterminer les spécifications des changements à appliquer à MIAM (Figure 3.6). Ces scénarios permettent de décrire ce que l’ontologie doit calculer pour répondre aux questions de compétences.

Lors de la phase d’exploration des différentes ressources disponibles pour répondre aux questions de compétences, nous pouvons rapidement nous apercevoir de la variété du vocabulaire employé pour désigner les mêmes groupes alimentaires. Pour lever cette ambiguïté menant à des incohérences, tous les termes ont été listés de sorte à pouvoir définir pour chaque terme préféré (en l’occurrence le terme employé par le PNNS), les alt-labels leur correspondant. Par exemple, le concept « Matières grasses » a été remplacé par le concept « Huile de colza, de noix, d’olive » nommé ainsi dans le PNNS.

Cette étape nous a permis d’identifier les relations sémantiques ainsi que les concepts sur lesquels nous devons appliquer les changements. Nous définirons dans le module Personne les connaissances nécessaires à la collecte de données (physiologiques, besoins nutritionnels, habitudes alimentaires, etc.) et la recommandation d’aliments (ou régimes alimentaires), d’activités (ou programmes d’activité) ou de questionnaires selon les besoins spécifiques des personnes âgées. Des modifications de modélisation de certains concepts de MIAM ont été essentielles afin de tenir compte notamment du changement du vocabulaire employé suite à l’évolution des recommandations nutritionnelles fournies par le PNNS en France.

Thème	Exemples de questions de compétences	Exemples de scénarios d’usage
Portion journalière ou hebdomadaire	Combien de fois dois-je manger du poisson par semaine ?	- Portion hebdomadaire par FamillePNNS
		- Lien avec le module Aliment <i>via</i> les catégories d’aliments
Substitution de produits	Je n’aime plus les produits laitiers, que puis-je faire ?	- Lien avec le module Aliment <i>via</i> aPourSubstituant
Composition des aliments		
	Quels aliments manger, si j’ai une carence en calcium ?	- Framework pour faire le lien avec Ciqual <i>via</i> module ConditionPhysiologique
Caractère allergène des aliments		
	J’ai une allergie aux arachides, dois-je les substituer avec d’autres aliments ?	- Lien avec le module ConditionPhysiologique <i>via</i> aPourSubstituant
Aliments recommandés pour le contrôle maladie		
	J’ai de l’hypertension, dois-je consommer des aliments particuliers ?	- <i>object property</i> entre Aliment et ConditionPhysiologique

FIGURE 3.6 – Exemples de questions de compétences et de scénarios d’usages selon le thème abordé

Ces deux premières étapes d'évolution ont permis de définir les modules de MIAM concernés par l'évolution.

3.3.3 Etape 3 : Suggestions des changements pour l'évolution de MIAM

3.3.3.1 Présentation de MIAM

MIAM est une ontologie modulaire composée d'un module noyau, le module Aliment, qui est utilisé par les 8 autres modules thématiques (Figure 3.7) : Nutrition (spécifique à la nutrition) ; Cuisine (relatif à la réalisation des recettes et des liens avec les types de plat) ; Préparation (relatif aux préparations de base associées à une recette) ; Unité (relatif aux métriques du domaine de la cuisine (cuillère à café, verre à moutarde) et métriques internationales (gramme, millilitre, etc.) ; Matériel (relatif aux matériels utilisés pour réaliser les recettes) et Sensoriel (relatif aux aspects sensoriels caractérisant les aliments et les recettes). Les concepts de base liant les concepts généraux de MIAM sont rassemblés dans "CORE-ONTO". De nombreux modules sont intéressants à réutiliser, comme le module Sensoriel, et feront l'objet d'une prochaine étude.

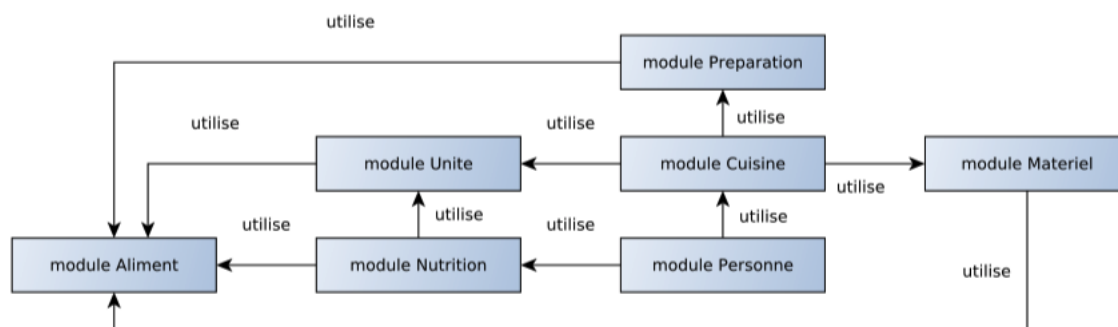


FIGURE 3.7 – Modules de MIAM (DESPRES, 2016)

3.3.3.2 Modules concernés par l'évolution

L'identification des changements a permis de définir les modules de MIAM concernés par l'évolution. Ainsi, pour répondre à des besoins en termes de recommandation d'activité physique et de collecte de données issues de questionnaires et d'objets connectés, l'évolution de MIAM (Figure 3.8) a nécessité :

- l'évolution de certains modules de MIAM : Nutrition, Aliment et Personne ;
- le développement de nouveaux modules : Activite, Pathologie, Signe, Symptome, ObjetConnecte et Mesure.

3.3.3.3 Importation des modules

L'importation des modules est réalisée dans Protégé. L'ensemble des modules importés (flèches bleues foncées) forment l'ontologie ONAFE (*Ontology of Nutrition and*

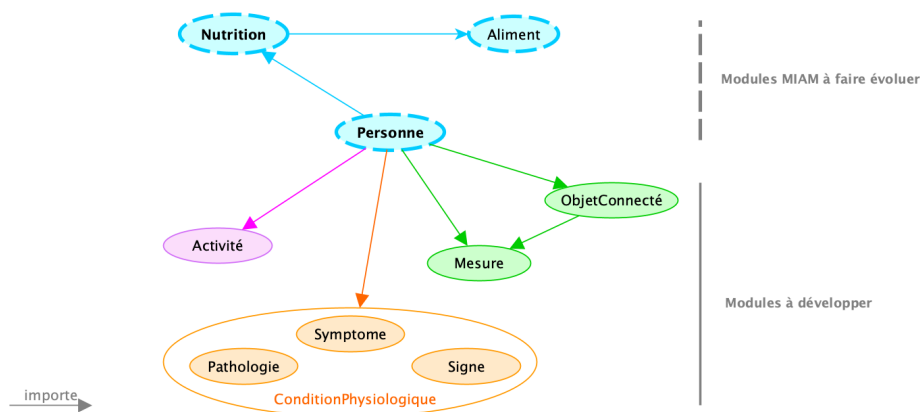


FIGURE 3.8 – Schéma montrant les modules de MIAM impliqués dans l'évolution de MIAM et les nouveaux modules à développer pour concevoir ONAFE

Activity For Elderly) (Figure 3.9). Le module *Personne* est un module auquel les autres modules font référence. Le sens des flèches indique quel module importe un autre module. La Figure 3.9 montre que la personne est au centre de notre conception. En effet, le module *Personne* importe tous les autres modules (sauf *Aliment*). Bien que le module *Aliment* soit concerné par l'évolution, il est importé uniquement par *Aliment*. Chaque module contient des concepts, des relations et des contraintes (inclus dans la T-Box) liés à chaque domaine et sont par la suite composés pour former une ontologie unique (T-box unique). L'orientation des flèches indique le sens de l'importation. Par exemple, le module *Nutrition* doit apporter les modules *Activité* et *Aliment*. Les relations faisant le lien avec les modules importés doivent être définies dans le module principal.

Les modules concernés par l'évolution ont dans un premier temps été extraits de MIAM. A partir des modules sélectionnés dans MIAM et nécessitant une évolution, nous avons défini les changements à effectuer sur les modules et les liens entre les modules. La nouvelle conceptualisation tient compte de l'évolution des changements.

3.3.3.4 Evolution du module Aliment et Nutrition

Le module *Aliment* n'a pas subi d'évolution majeure. Il a principalement servi à caractériser les aliments selon leurs apports nutritionnels (*AlimentHyperEnergetique*, *AlimentRicheEnProteine*, etc.) et définir quels sont les aliments "à favoriser", "à limiter" et "à proscrire" pour les inclure dans un régime alimentaire spécifique (*RegimeDenutri*, *RegimeDiabetique*, etc.). De la sorte, selon chaque besoin nutritionnel, un régime alimentaire sera suggéré. Nous présentons dans la Figure 3.10 un exemple montrant la classification d'aliment obtenu après les avoir qualifiés selon leur apports nutritionnels.

Les régimes alimentaires inclus dans le module *Nutrition* concernent actuellement les régimes relatifs aux problèmes de santé identifiés (dénutrition, diabète, hypertension, etc.) et pour lesquels une recommandation est personnalisable. L'évolution du module *Nutrition* ne peut se faire indépendamment du module *Personne*. Dans le cas de la dénutrition, les aliments riches en protéines et hyperénergétique pourront lui être recommandés. Dans le cas d'une personne diabétique, les aliments à indice

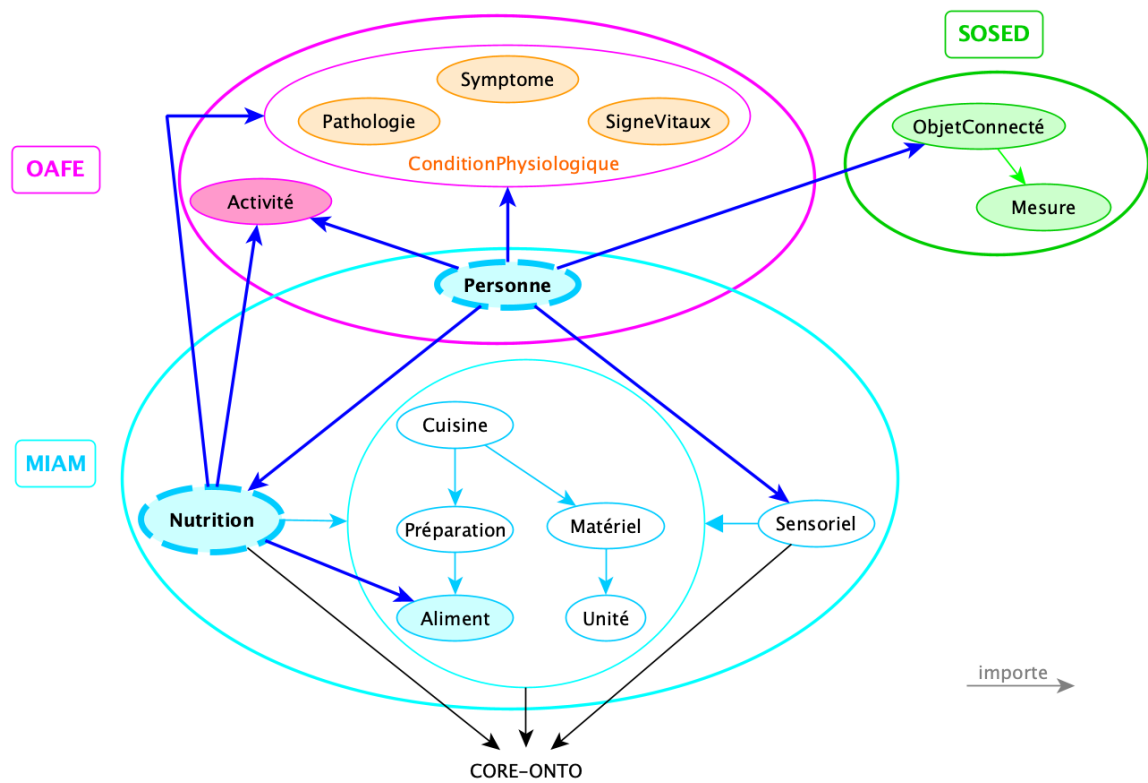


FIGURE 3.9 – Schéma montrant l'import des modules impliqués dans l'évolution de MIAM et les nouveaux modules à développer pour concevoir ONAFE. Les modules de MIAM ayant subi une évolution sont en pointillés, les modules utilisés sont ont un fond bleu.

glycémique bas pourront lui être recommandés. Le module Nutrition contient tous les concepts relatifs à la nutrition comme les allergènes alimentaires, la structure du repas (AllergeneAlimentaire, Repas, etc.). Ces concepts devront par la suite être intégrés au sein des régimes alimentaires.

Le module Nutrition contient également d'autres modules de MIAM non nécessaires à l'usage applicatif. Des décisions quant à la suppression d'un concept non inclus dans les nouveaux besoins sont prises au fur et à mesure, selon les cas. Par exemple, pour les GroupePNNS3 la décision dépend de la suite de l'évolution des besoins. Elle pourra servir par exemple pour déterminer un groupe alimentaire utilisé par d'autres recommandations s'appuyant sur les groupes alimentaires du PNNS 3.

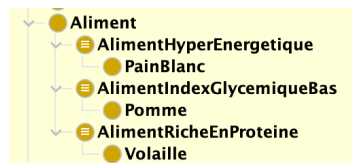


FIGURE 3.10 – Exemple de classification d'aliment selon leur apports nutritionnels

3.3.3.5 Evolution du module Personne

Le module Personne est un module auquel les autres modules font référence. Il est issu de l'ontologie MIAM et a évolué tout au long des étapes d'évolution pour inclure les connaissances relatives aux nouveaux besoins de collecte, de recommandation et de définition du profil de la personne.

Pour réaliser cette tâche d'évolution, la représentation de la personne âgée, à partir des données collectées la concernant, fournit un aperçu des termes liés à Personne et de la variété des modules impliqués dans son évolution (3.3.3.5).

Le module Personne a été complété pour caractériser la personne selon les données collectées et a évolué au fur et à mesure des réponses apportées aux questions de compétences pour répondre aux changements précédemment définis. Les concepts représentés dans Personne servent à construire le profil de la personne âgée et à personnaliser les recommandations. Les données collectées sont fixes ou variables. Les données fixes d'une personne (Nom, Taille, etc.) sont acquises une seule fois et ne varient pas en fonction du temps. Certaines propriétés peuvent soit varier au cours du temps (Poids, etc.) soit être contextuelles (FrequenceCardiaque, PressionArterielle, etc.). Ainsi, toutes les variables impliquées dans la définition de l'état de santé (Signe, Symptome, Pathologie), du profil de l'activité et de l'environnement social (CaracteristiqueSocial, etc.), des objectifs que la personne souhaite ou devrait atteindre (Objectif) ont été décrites. Cette modélisation a été enrichie en associant les freins liés à l'état de santé, au profil d'activité ou encore au profil social (ALFAIFI et al., 2017) de la personne âgée.

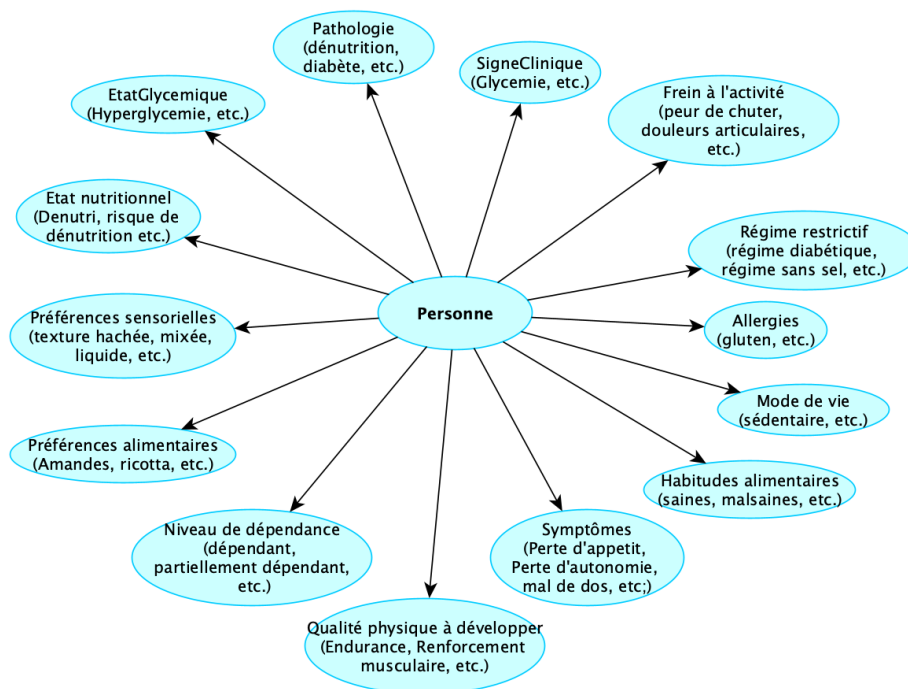


FIGURE 3.11 – Conceptualisation du module Personne pour répondre aux besoins d'évolution

3.3.4 Développement de nouveaux modules

3.3.4.1 Les modèles conceptuels existants d'activité

L'activité est un concept important dans de nombreux domaines (géographie, transports, psychologie, etc.), et un certain nombre d'ontologies liées à l'activité ont été développées. En conséquence, il existe un certain nombre de modèles conceptuels qui tentent d'acquérir la sémantique du concept Activité. Ces modèles sont adaptés à leurs cas d'utilisation et ne peuvent pas être facilement généralisés à d'autres applications. Ce travail vise à construire une ontologie de domaine de l'activité en s'appuyant sur les modèles génériques d'activité existante.

La représentation des connaissances d'un domaine particulier à partir de modèles conceptuels constitue une étape importante dans le processus de développement d'ontologies. La modélisation relative à l'identification d'activités a bénéficié ces dix dernières années de recherches approfondies s'appuyant sur deux types d'approches, une approche axée sur les données et une approche axée sur les connaissances.

La première approche permet d'analyser et de modéliser des données de capteurs sur des périodes à long terme et de traiter des données à grande échelle. Elle nécessite la génération d'un modèle pour chaque activité en utilisant des stratégies d'apprentissage probabilistes et statistiques (GAYATHRI et al., 2017).

La seconde approche, axée sur les connaissances, utilise des méthodes d'ingénierie des connaissances pour modéliser les activités en utilisant des connaissances de domaine et des techniques de raisonnement fondées sur l'intelligence artificielle pour être en mesure d'établir des inférences (YU et al., 2015). Cette approche est la plus utilisée et la

plus efficace pour représenter l'activité, gérer les données de capteurs considérés comme des ressources hétérogènes et raisonner sur les connaissances modélisées. En effet, les ontologies sont de plus en plus utilisées au sein de systèmes complexes pour modéliser et raisonner sur les activités, quelles soient simples ou composites. Certains auteurs (CHEN et NUGENT, 2009, RIBONI et BETTINI, 2011, WONGPATIKASEREE et al., 2013) utilisent les ontologies pour identifier les activités de la vie quotidienne (par exemple, faire le ménage, préparer un repas, etc.) dans le contexte de Smart Home qui dispose de capteurs (sensors) détectant des tâches contextuelles (remplir la casserole d'eau, allumer la plaque de cuisson, mettre la casserole au feu, etc.) *via* les définitions contenues dans l'ontologie. D'autres auteurs utilisent les ontologies pour la recommandation des activités (WIERINGA et al., 2011, PRAMONO et al., 2013, NASSABI et al., 2014, ZOLFAGHARI et al., 2016) ou pour la reconnaissance et la recommandation d'activités (FILOS et al., 2016, NI et al., 2016). Toutefois, la plupart échouent dans la personnalisation de la recommandation, car tous les paramètres influençant la recommandation ne sont pas pris en considération.

L'approche prédominante dans ces études consiste à dissocier l'ontologie noyau (core ontology) fournissant les concepts structurant du domaine et décrivant les relations entre ces concepts, de l'ontologie de domaine fournissant les concepts spécifiques du domaine d'étude ainsi que les relations entre ces concepts. Néanmoins, la réutilisation de ces ressources est complexe car aucune ontologie des auteurs précédemment cités n'est disponible pour être réutilisée.

Parmi les ontologies incluant le concept d'activité « Activity », certains auteurs mettent à disposition leur hiérarchie de classes sans donner les liens entre ces classes. Quelques-unes de ces ontologies légères et méta-thesaurus sont représentées dans la Figure 3.12. D'autres auteurs fournissent leur modélisation de l'activité : - ADL Ontology (CHEN et NUGENT, 2009) ; - Palspot Project Ontology (RIBONI et BETTINI, 2011) ; - Activity Log Ontology (WONGPATIKASEREE et al., 2013) ; - ODP Activity (ABDALLA et al., 2014) ; - ECOPPA (HODA et al., 2018). Dans la plupart des cas, ces représentations sont générales, incomplètes ou encore dépendantes d'un contexte. Le tableau 3.2 résume les principaux termes de premier niveau reliés au concept « Activity » qui ont été identifiés dans les différentes études modélisant l'activité. Les modélisations présentées diffèrent selon la nature de l'activité (statique ou dynamique) mais également selon le contexte d'étude, ce qui représente une limite et une difficulté majeure pour la réutilisation de ces modèles.

Nous nous sommes également intéressés aux relations explicitées entre les différents termes identifiés. Le tableau 3.3 synthétise les principaux termes en relation avec le concept « Person » trouvés dans les différentes études présentées précédemment. Nous pouvons observer que contrairement aux concepts, aucune relation n'est réutilisée et que (WONGPATIKASEREE et al., 2013) ainsi que la ressource DBpedia2 ne présentent dans leur modèle aucune relation. Nous observons dans le tableau 3.2 que le concept « Person » n'est pas toujours inclus dans les différentes représentations étudiées et que lorsque ce concept est considéré, sa représentation dépend du domaine étudié. Cependant, la recommandation d'activité ne peut être aboutie et personnalisée que si elle prend en considération le profil de la personne. Le tableau 3.4 résume les principaux termes de premier niveau reliés au concept « Activity ». Parmi les études citées, plusieurs termes sont utilisés pour conceptualiser la personne : Person, Profil, Agent, User. Nous

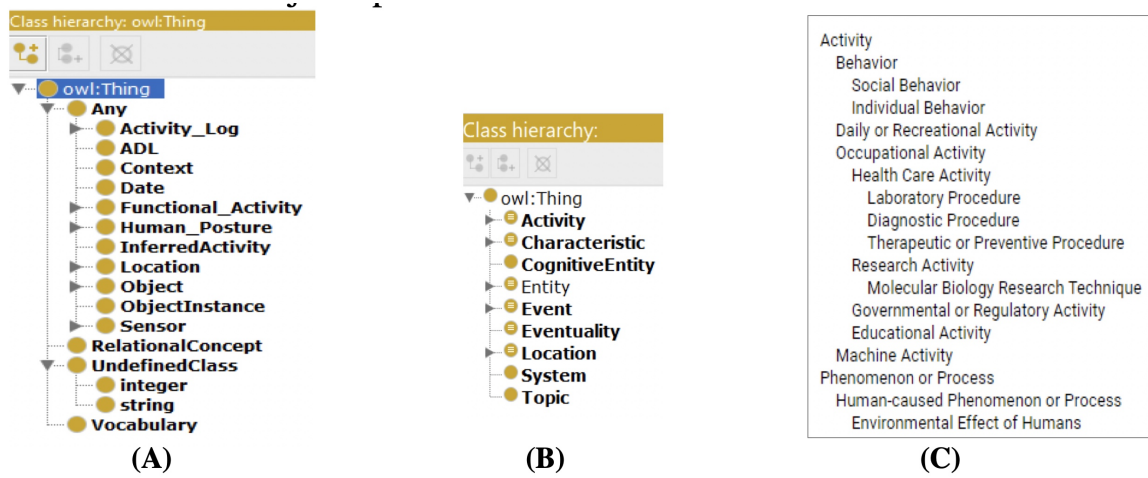


FIGURE 3.12 – Capture de l'ontologie « Home Activity » (A) (WONGPATIKASEREE et al., 2013), de l'ontologie DBpedia (B), et du méta-thésaurus UMLS (C) faisant notion au concept « Activity »

	Activity	Agent	Location	Effects	Environmental	Entity	Condition	Goal	DateTime	Duration	Action	Person	Context	Object	Sensor	Entity	Requirement	Outcome	Characteristic	CognitiveEntity	Event	Eventuality	System	Topic
1	x	x	x	x	x	x	x	x	x	x														
2	x								x		x	x												
3	x		x						x				x	x	x									
4	x	x														x								
5	x		x													x		x	x	x	x	x	x	x
6	x	x	x	x													x	x						
7	x	x	x		x				x	x														

TABLEAU 3.2 – Synthèse des principaux termes identifiés (niveau 1 de la ressource, énumérés par ordre de citation) reliés au concept « Activity ». 1 : CHEN et NUGENT, 2009 ; 2 : Riboni and Bettini, 2011 ; 3 : WONGPATIKASEREE et al., 2013 ; 4 : PROV-O, 2013 ; 5 : DBpedia, 2014 ; 6 : ABDALLA et al., 2014 ; 7 : HODA et al., 2018

observons également que les termes utilisés dépendent fortement du contexte étudié, d'où la nécessité de construire un profil permettant d'organiser les connaissances afin d'adapter la recommandation d'activités aux personnes âgées.

	superActivity	subActivity	typeOf	performed	atLocation	require	hasEntity	lastFor	achieve	resultedIn	performedDuring	impliesAction	performsActivity	hasArtifact	isInside	includeCommuni- cationRoute	actedOnBehalfOf	wasAssociated With	wasAttributedTo	used	wasGeneratedBy	startedAtTime	endedAtTime	wasDerivedFrom	wasInformedBy	takePlaceAt	hasEnd	hasDuration	produces	hasPart	hasParticipant	hasRequirement	hasTemporal Description	hasEnvironment	Entity	hasLocation	perform	
1	x	x	x	x	x	x	x	x	x	x																												
2											x	x	x	x	x																							
3																	x	x	x	x	x	x	x	x														
4																									x	x	x	x	x	x	x							
5																																			x	x	x	x

TABLEAU 3.3 – Synthèse montrant les relations utilisées reliés les différents termes définis par les auteurs cités dans le tableau 3.2 : CHEN et NUGENT, 2009 ; 2 : Riboni and Bettini, 2011 ; 4 : PROV-O, 2013 ; 6 : ABDALLA et al., 2014 ; 7 : HODA et al., 2018

	Person	Profile	Preferences	BasicPersonal Information	User	Role	Physical Characteristics	UsageStatistics	Medical Information	Culture	Religion	Risk	HealthCondition	LifeStyle	Goal	FitnessLevel	Plan	Interests	Identity	Physiological State	Location	Device
1		x	x	x		x													x			
2			x		x		x	x	x	x	x	x										
3													x									
4	x	x				x																
5			x	x			x							x	x	x	x			x	x	x

TABLEAU 3.4 – Synthèse des principaux termes identifiés dans les différentes études explorées associés au concept « Person ». 1 : PRIMO et al., 2012 ; 2 : AL-NAZER et al., 2014 ; 3 : YU et al., 2015 ; 4 : NI et al., 2016 ; 5 : HODA et al., 2018

3.3.4.2 Ontologies existantes

La seule ontologie de l'activité accessible depuis BioPortal est l'ontologie SMASH (Semantic Mining of Activity, Social, and Health data). Cette ontologie est destinée aux réseaux sociaux de santé qui est fondée sur l'étude YesiWell (PHAN et al., 2017). Elle est composée de trois autres ontologies et définit des concepts associés à une perte de poids durable, en particulier ceux liés à une intervention continue avec de fréquents contacts sociaux. Nous constatons que parmi toutes les ontologies citées, aucune ne rassemble toutes les activités pouvant être exercées.

Le groupe de travail du W3C sur les provenances (PROV) fournit quant à lui une ontologie PROV-O (PROVenance Ontology, 2013) minimale composée de trois

concepts centraux : « Activity », « Entity » et « Agent » et constitue une ressource intéressante à réutiliser (Figure 3.13). A ces trois concepts sont associés un ensemble de propriétés et de restrictions pouvant être utilisées pour représenter et échanger des informations de provenance générées dans différents systèmes et dans différents contextes. Il peut également être personnalisé pour créer de nouvelles classes et propriétés pour modéliser les informations de provenance de différentes applications et domaines. Un ODP de l'activité existe également et est intéressant à exploiter. Il fournit un modèle de conception d'ontologie générique pour modéliser le noyau commun du module Activité dans différents domaines. Ce modèle incorpore deux perspectives générales des activités : une perspective de « flux de travail », souvent observée dans les applications liées à la planification, et une perspective spatio-temporelle, souvent retrouvée dans l'analyse des activités géographiques.

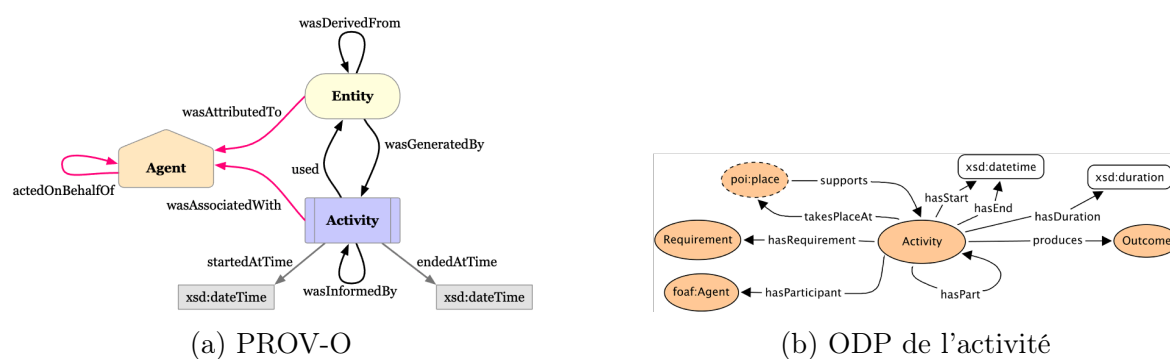


FIGURE 3.13 – Représentation des trois classes de PROV-O et des propriétés qui les relie (*PROV-O : The PROV Ontology 2013*) (a) et de l'ODP de l'activité pour le raisonnement sur l'activité (ABDALLA et al., 2014) (b)

Il existe un certain nombre d'ontologies (geriatrics, Human Disease Ontology, Ontology for General Medical Science (OGMS), Vital Sign Ontology (VSO)) en lien avec les pathologies, les signes et les symptômes. Human Disease Ontology (SCHRIML et al., 2018) a été développée en tant qu'ontologie standardisée pour les pathologies humaines dans le but de fournir à la communauté biomédicale des descriptions cohérentes, réutilisables et durables des pathologies humaines, des caractéristiques phénotypiques et des concepts liés aux pathologies du vocabulaire médical. Geriatrics définit le système de gériatrie avec des classes. Elle est intéressante à réutiliser car elle possède également des objetProperty et des dataProperty qui sont utilisées pour créer un système d'aide à la décision pour les personnes âgées. OGMS et VSO incluent les signes et les symptômes cliniques. VSO couvre les quatre signes vitaux humains consensuels : la pression artérielle, la température corporelle, la fréquence respiratoire, le pouls. VSO fournit également un vocabulaire structuré contrôlé pour décrire les données de mesure des signes vitaux, les divers processus de mesure des signes vitaux et les divers dispositifs et entité anatomique participant à ces mesures. Afin de mieux spécifier la signification de chaque classe, nous nous sommes référés à d'autres ontologies standardisées telles que la nomenclature systématique des termes cliniques de la médecine (SNOMED-CT) en rapport avec l'activité et des entités permettant de définir un état de santé pour la personne.

La décision quant à la réutilisation d'une ontologie ou un vocabulaire dépend de

plusieurs facteurs. GYRARD, ATEMEZING et al., 2021 définit les principales métriques utilisées pour choisir une ontologie de domaine à réutiliser. Ainsi, nous sommes passés par de nombreuses étapes de réflexions. Les ontologies citées n'ont pas été réutilisées soit en raison de leur contexte trop éloigné, soit en raison de la difficulté à réutiliser un vocabulaire nécessitant une traduction. Par exemple, l'ontologie "Human Disease Ontology" aurait pu être réutilisée pour définir les principales classes des pathologies, cependant, l'ontologie est volumineuse et nécessite un temps considérable. Il serait intéressant d'exploiter des outils permettant d'automatiser la traduction de toutes les entités de l'ontologie.

Au cours du scénario de réutilisation, seul l'ODP de l'activité a été choisi pour être réutilisé lors de la conception du module Activité (Figure 3.13b). A partir de la modélisation sélectionnée, le vocabulaire (concepts, relations et axiomes) n'a pas été réutilisé. La représentation de l'ODP a été réutilisée et les différents concepts et relations réutilisés (Activity) ont été traduits. Les autres ontologies citées n'ont pas été réutilisées soit en raison de leur contexte trop éloigné, soit en raison du manque de granularité du modèle, soit en raison de l'absence de la ressource ontologique.

3.3.4.3 Scénarios et questions de compétences

Nous présentons dans le tableau 3.5 une liste non exhaustive des scénarios issus principalement de cas d'utilisation identifiés à partir des discussions de forums. Nous dressons également une liste non exhaustive de questions de compétences issues de forums, de discussions sur des sites tels que l'EHPAD. L'ontologie doit être en mesure de répondre à des questions comme ceux énoncés dans le tableau 3.6. Chaque question de compétences est reliée au minimum à un scénario d'usage.

N°	Scénario d'usage	L'ontologie doit permettre de représenter
1	La personne âgée s'interroge sur les activités qu'elle pourrait pratiquer en fonction de son état de santé et de ses pathologies.	les activités en lien avec un état de santé et une pathologie donnée.
2	La personne âgée s'interroge sur les activités physiques qu'elle peut pratiquer en fonction de son niveau d'autonomie.	les activités physiques en décrivant le niveau d'autonomie nécessaire pour les pratiquer.
3	La personne âgée s'interroge sur les activités physiques qu'elle peut pratiquer en fonction de ses habitudes de vie.	les activités physiques en décrivant les liens avec le mode de vie de la PA.
4	La personne âgée s'interroge sur les activités qu'elle peut pratiquer en fonction des objectifs qu'elle se fixe (améliorer son endurance, etc.).	les activités physiques en caractérisant les bénéfices qu'elle procure (marcheNordique aPourGainPhysique Endurance).
5	La personne âgée s'interroge sur la localisation où se déroule une activité physique et le matériel qu'elle doit se procurer pour l'exercer.	les localisations et les matériels associés à chaque activité physique.

TABLEAU 3.5 – Exemples de scénarios d'usage

3.3.4.4 Module Activité

Le module « Activité » de l'ontologie de l'Activité pour les Personnes Agées (OAPA) ou OAFE (pour *Ontology of Activity For the Elderly*) (DANDAN et al., 2018) a été conçue pour modéliser le domaine de l'activité dans le but d'adapter la recommandation d'activité en fonction de l'état de santé, du niveau d'activité, du niveau d'autonomie, de l'environnement et des préférences de la personne âgée. L'ontologie décrit les caractéristiques propres à l'activité et le contexte dans lequel elle se déroule.

N°	QC	S
1	J'ai 85 ans et j'ai de l'hypertension artérielle depuis quelques années, j'aimerais savoir si la marche nordique est une activité qui peut avoir des bienfaits sur ma maladie ?	1,3
2	Quels sont les activités physiques qui pourraient participer à maintenir mon indépendance fonctionnelle sachant que je suis à la retraite depuis 3 ans et que mon activité est, jour après jour, de plus en plus réduite ?	2,4
3	J'ai 62 ans, il m'a été diagnostiqué de la polyarthrite rhumatoïde, quels exercices dois-je pratiquer pour maintenir un maximum de mobilité et préserver mes articulations ?	3
4	J'ai subi une intervention pour une hernie discale il y a un mois, y a-t-il des risques de faire du ski de fond sans avoir de récurrences ?	1,3
5	J'ai 65 ans et j'aimerais savoir s'il y a des bienfaits associés à la gymnastique douce si je l'exerce uniquement une fois par semaine et quels sont les matériels que je dois avoir à ma disposition ?	3,5

TABLEAU 3.6 – Exemples de questions de compétences (QC) associées aux scénarios d'usage (S) cf. 3.5

NeOn (SUÁREZ-FIGUEROA et al., 2012) constitue le cadre méthodologique dans lequel nous nous situons pour construire les différents modules. Les scénarios 1 (spécification, planification, conceptualisation, formalisation, implémentation) et 7 (réutilisation d'ODP) ont été appliqués. Le périmètre de l'ontologie est établi en adoptant la méthodologie de (USCHOLD et GRUNINGER, 1996).

L'étape d'acquisition des connaissances a été réalisée à partir de sources officielles (OMS, PNNS) et de sites spécialisés tels que les établissements d'hébergement pour personnes âgées dépendantes (EHPAD) fournissant des lignes directrices de recommandations d'activités physiques pour les personnes âgées ainsi que les bienfaits associés à ces activités.

Les déterminants utilisés le plus souvent pour caractériser la sédentarité sont la durée et la fréquence des moments passés en position assise. La mesure de la dépense énergétique est l'un des indicateurs les plus utilisés pour mesurer l'activité physique. Elle correspond au produit de l'intensité à la durée et par la fréquence d'activité. Elle peut être exploitée lorsque l'utilisateur indique quelle activité il a exercée, auquel cas, le MET ne pourra pas être défini. Depuis la fin des années 80, un compendium des activités physiques (Compendium of Physical Activities, CPA) pour adultes a été développé pour être utilisé dans des études épidémiologiques afin de normaliser l'attribution des intensités mesurées par l'échelle métabolique (Metabolic Equivalent of Task, MET) dans les questionnaires sur l'activité physique. Sa dernière mise à jour date de 2011 (AINSWORTH et al., 2011). Le MET (*Metabolic Equivalent of Task*) correspond à l'unité de mesure de la dépense énergétique et de l'intensité d'une activité. Ainsi, en fonction de la nature de l'activité, son intensité et sa durée, il est possible de connaître l'état d'activité d'une personne et de lui recommander une activité ayant un MET supérieur à celui de l'activité exercée.

L'acquisition de connaissances en plus de l'élaboration des scénarios et des questions de compétences a permis de construire un modèle conceptuel construit à partir d'une carte conceptuelle. Les activités représentées dans l'ontologie sont celles recommandées par le PNNS. Elles ont été qualifiées selon le vocabulaire CPA permettant d'associer l'intensité (mesurée à l'aide de l'échelle MET), le rythme (RythmeFaible, RythmeModere, etc.) et l'effort fourni (EffortLeger, EffortIntense, etc.) au moment de l'activité. Pour cela, l'intensité, le rythme et l'effort ont nécessité une grande attention. Par exemple, la mesure de l'effort est loin d'être une information évidente à définir et à modéliser.

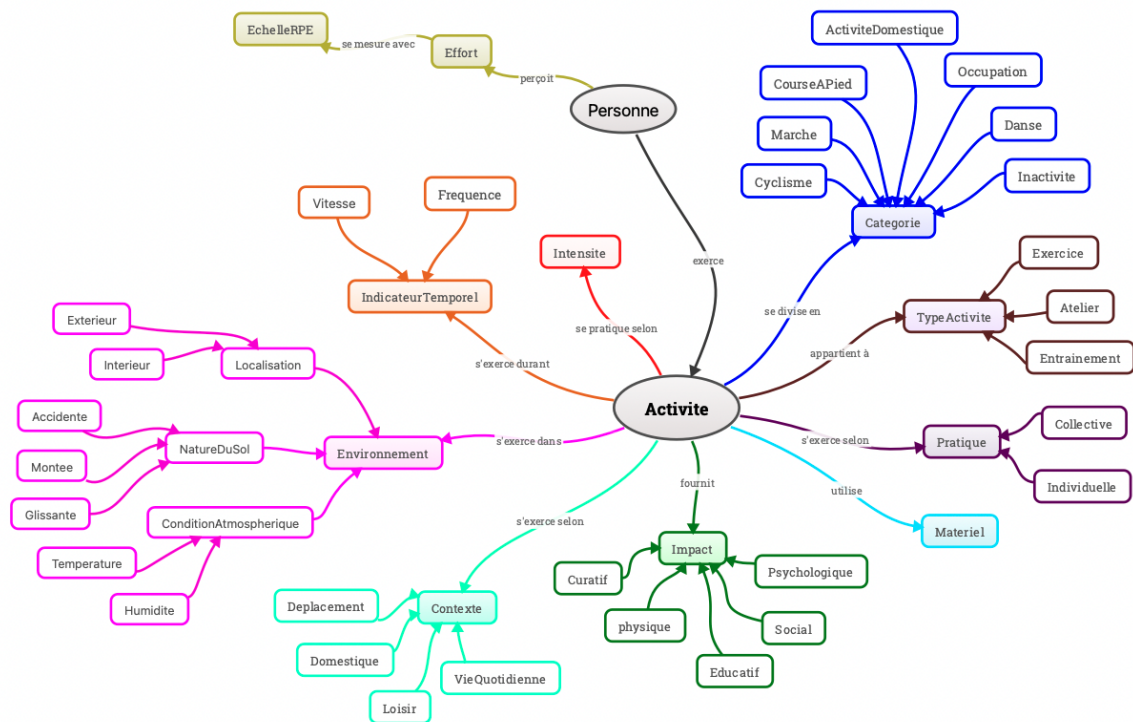


FIGURE 3.14 – Aperçu du modèle conceptuel du concept « Activité » montrant les relations de premier et de deuxième niveau

En effet, l’effort présente plusieurs dimensions, une dimension physiologique (effort physique), une dimension psychologique (effort mental) et une dimension plus large incluant l’effort perçu. Ceci implique la définition des déterminants de l’effort (débit d’oxygène (VO₂), fréquence cardiaque (Fc), etc.) mesurés à partir de l’échelle RPE (Rating of Perceived Exertion) de Borg pour la détermination de l’intensité de l’activité (BORG, 1998). La présence d’une linéarité entre l’intensité de l’exercice, les contraintes physiologiques (VO₂, Fc) et l’effort perçu nous permet de compléter la définition de l’intensité d’une activité.

Exemple de personnalisation de la recommandation

Les activités ont été qualifiées selon les qualités physiques qu’ils permettent de développer (QPD). Dans cet exemple, nous présentons la distinction entre : la relation aPourQualitePhysiqueDeveloppée qui relie l’activité (ActiviteEquilibre) à la QPD (Equilibre) ; et la relation aPourQualitePhysiqueADevelopper qui relie le profil (PersonnePeurDeChuter) à la QPD (Equilibre) (Figure 3.15). Nous pouvons ainsi obtenir par raisonnement l’inférence de l’activité recommandée dans le cas d’une personne sédentaire et qui a peur de chuter.

L’exemple présenté (Figure 3.16) schématise la représentation des concepts associés à la recommandation d’une ActiviteRenforcementMusculaire à une PersonneSedentaire et une personne ayant comme frein à l’activité (PourFreinActivite) la peur de chuter (PersonnePeurChuter). Lorsque la Personne aNombrePas une valeur inférieure à 4000,

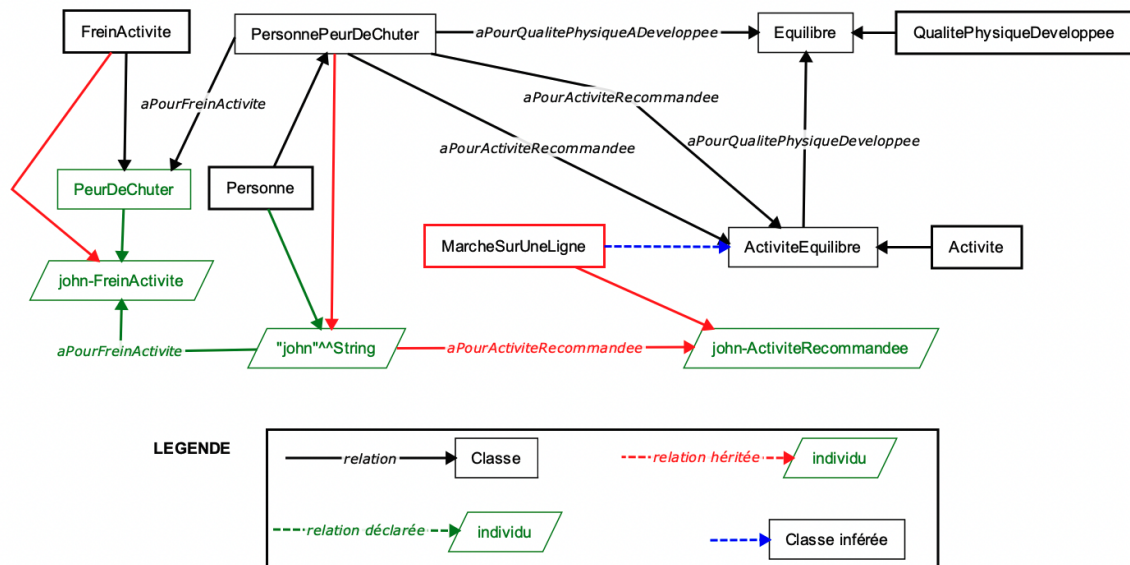


FIGURE 3.15 – Exemple de recommandation d’activité obtenue par raisonnement selon le profil de la personne âgée (qui a comme frein à l’activité une peur de chuter)

que son temps d’inactivité est supérieur à 4h, qu’elle aime (aPourLocalisationActivite-Preferee) les activités intérieures (ActiviteInterieure) alors la recommandation qui lui sera inférée consiste à s’entraîner à marcher sur une ligne (MarcheSurUneLigne) (Figure 3.16).

3.3.4.5 Module ConditionPhysiologique

Le module ConditionPhysiologique (CP) rassemble trois classes : Pathologie, Signe-Clinique, Symptome.

Les syndromes gériatriques sont des déterminants de santé utilisés par les médecins, notamment les gériatres et représentent les affectations fréquemment observées chez les personnes âgées. Selon (BELMIN et al., 2016), un syndrome gériatrique représente une situation de santé définie par quatre critères : sa fréquence augmente fortement avec l’âge (et/ou observée seulement chez des sujets âgés) ; elle résulte de facteurs multiples et divers, dont des facteurs favorisants (chroniques), incluant les effets du vieillissement et des facteurs précipitants (aigus ou intermittents) ; elle a pour conséquences fréquentes un risque de perte d’indépendance fonctionnelle et/ou d’entrée en institution ; sa prise en charge est multifactorielle (intervention multicomposante) et requiert une approche globale et holistique du patient. De part leurs critères, les syndromes gériatriques sont des éléments intéressants à prendre en compte dans la détection des signes et des symptômes cliniques et la caractérisation de l’état de santé du sujet âgé et personnaliser la recommandation. Un signe représente toute preuve objective de maladie observée ou détectable par le personnel de santé (patient, médecin, etc.). Par exemple, cela peut être une température ou une tension artérielle plus élevée ou plus basse que la normale. Un symptôme est un ressenti éprouvé par une personne qui sort de l’ordinaire en raison de la présence d’une maladie, d’une blessure ou d’un état comme une sensation de

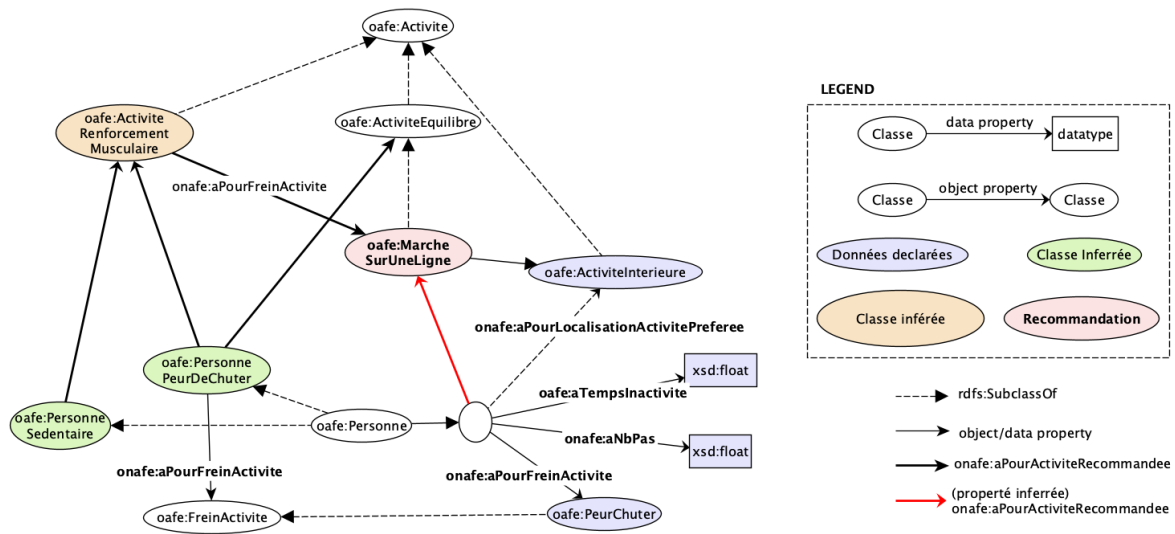


FIGURE 3.16 – Exemple de génération d’une recommandation pour un profil sédentaire

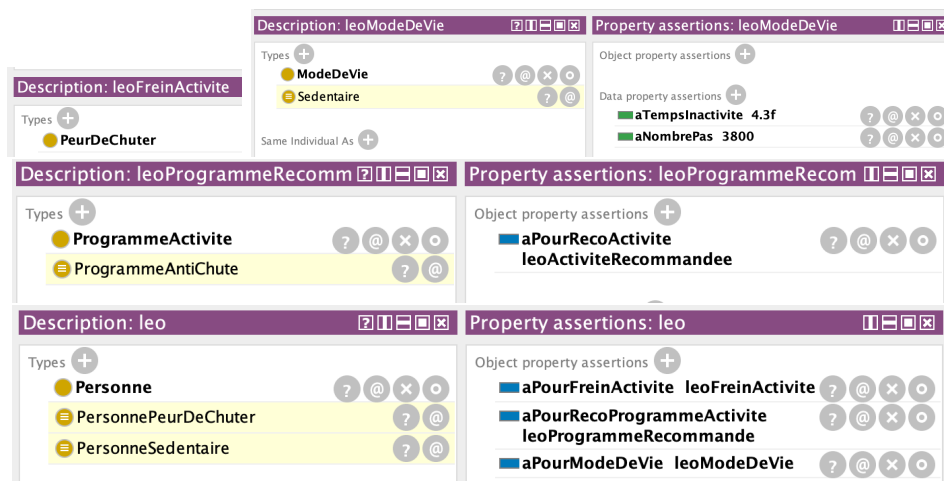


FIGURE 3.17 – Exemple montrant les inférences des recommandations obtenues par classification après lancement du raisonneur (classes en jaune).

fièvre, un mal de tête ou d'autres douleurs dans le corps. La MPE et les chutes répétées font partie des syndromes gériatriques difficiles à identifier par le médecin généraliste que par un gériatre. L'inclusion de ce concept dans l'ontologie n'est pas évidente en raison de sa définition large et des implications qu'elle cause. En effet, un syndrome gériatrique représente une manifestation définie qui dépend des facteurs étiologiques et des interactions multiples. Son inclusion pourrait être supportée par un gérontologue afin de spécifier le champ d'application. Ce concept est assez intéressant à inclure dans une prochaine modélisation, car il permet de cerner les situations médicales les plus spécifiques du soin gériatrique et donc de fournir une prise en charge.

Parmi les signes, nous distinguons les signes vitaux en raison de leurs caractères mesurables par les objets connectés disponibles actuellement sur le marché. Les signes vitaux sont les quatre signes qui peuvent donner une mesure immédiate du fonctionnement global et de l'état de santé du corps. Il s'agit de la température, de la fréquence cardiaque, de la fréquence respiratoire et de la pression artérielle (LOCKWOOD et al., 2004). Les plages de ces mesures varient en fonction de l'âge, du poids, du sexe et de l'état de santé général (*Vital Signs Table - Prohealthsys 2021*).

Le module CP contient les principales pathologies et symptômes rencontrés chez les personnes âgées. Les signes Vitaux sont définis selon des valeurs seuil permettant de les caractériser comme étant faible, normal, élevé (GlycémieFaible, TensionDiastoliqueElevee, etc.).

Pour être en mesure d'attribuer à une personne plusieurs pathologies et inclure les cas de personnes polyopathologiques, la classe Pathologie est une classe définie énumérée qui contient toutes les pathologies.

Pour définir un régime à une personne hyperglycémique, l'utilisation d'une restriction universelle (*only* ou *ObjectAllValuesFrom*) est nécessaire et permet d'obtenir par inférence les recommandations selon les assertions de propriétés définies (Figure 3.18e). Cette restriction décrit des classes d'individus qui, pour une relation donnée, n'ont de lien qu'avec les individus d'une classe spécifique. Par exemple, la classe des individus qui n'ont comme glycémie (aPourGlycémie) que les individus appartenant à la classe des GlycémieElevee. L'utilisation erronée de restrictions universelles plutôt qu'existentielles entraîne des inconsistances. Dans l'exemple de la Figure 3.18, nous présentons un exemple de recommandation d'Aliment et d'ObjetConnecte dans le cas d'un individu (nina) ayant une valeur Glycémie = 2 (Figure 3.18a). Pour cela, des assertions de propriétés ont été créées permettant de relier un individu à un autre (Figure 3.18d). L'exemple présenté ne permet de faire le lien entre l'état dénutri de l'individu et sa recommandation (RegimeDenutrition), car la description d'une telle recommandation dépend de données non considérées pour l'exemple de cet individu. Par la suite, d'autres cas pourront être traités comme ceux menant à des incohérences telle que la recommandation d'AlimentIndexGlycémiqueBas et d'AlimentIndexGlycémiqueElevee.

3.3.4.6 Modules ObjetConnecté et Mesure

Les modules ObjetConnecté et Mesure constituent l'ontologie SOSED (*ontology of Smart Objects for collected hEalth SEnsor Data*). Elle a été conçue pour décrire les objets connectés pouvant être utilisés par la personne âgée permettant le suivi des paramètres de santé et d'activité.

Le module ObjetConnecté inclut les objets connectés qui pourraient les accompagner

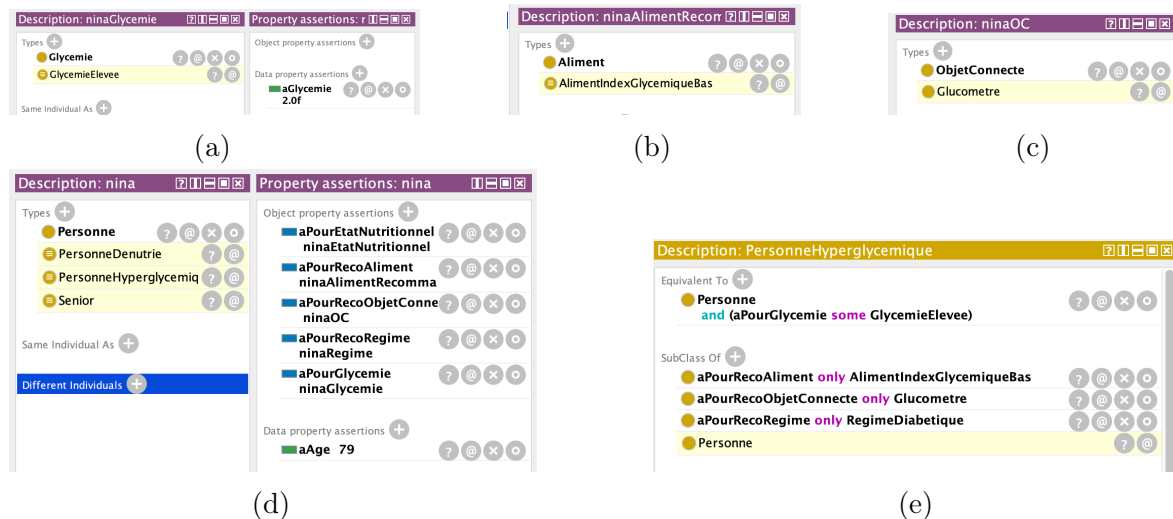


FIGURE 3.18 – Exemple montrant les inférences des recommandations obtenues par classification après lancement du raisonneur (classes en jaune). (a) Description de la Glycemie de nina. (b) Description de l’individu la recommandation d’AlimentIndiceGlycemiquebas dans le cas d’une PersonneHyperglycémique. (c) Description de l’individu ninaOC montrant l’ObjetConnecte dans le cas d’une PersonneHyperglycémique. (d) Description des assertions de propriétés permettant de caractériser l’EtatNutritionnel de nina, son âge et sa valeur de Glycemie. (e) Description de la classe PersonneHyperGlycémique.

dans leur quotidien et répondre au besoin de collecte de données et d’identification des signes de dénutrition et de sédentarité. Ainsi, des objets connectés pourront être proposés selon son profil établi. Par exemple, dans le cas d’une personne diabétique, un glucomètre connecté lui sera recommandé pour faciliter le suivi de sa glycémie et personnaliser les recommandations nutritionnelles en fonction de celle-ci. En général, les aliments à indice glycémique élevés sont “à limiter” et les aliments à indice glycémique bas sont “à favoriser”. Lorsque la glycémie de la personne est très élevée, les aliments à indice glycémique élevés devront dans ce cas être “à proscrire” et les aliments à indice glycémique moyen devront quant à eux être “à limiter”.

Le module Mesure inclut les mesures issues des objets connectés identifiés utiles à l’application. Par exemple, la montre connectée permet de mesurer le nombre de pas. Dans le cas d’une personne sédentaire, lorsqu’en moyenne le nombre de pas par jour est inférieur aux recommandations, des activités d’intensité faibles lui seront proposées.

3.3.5 Etape 4 : Évaluation de l’impact de l’évolution

Tout comme MIAM, les modules d’ONAFE ont été conçus de façon modulaire suivant une approche par composition. Chaque module (Figure 3.8) correspondant à un domaine a été conçu pour être auto-suffisant. Dans Protégé, la relation d’import a été utilisée pour construire ONAFE qui représente l’ensemble des modules impliqués dans les tâches d’évolution présentées.

La modélisation des différentes cartes conceptuelles a servi de base à l’évolution et à la construction des différents modules formalisée en OWL2. Les concepts du modèle

conceptuel sont représentés par des classes et les relations par des propriétés (« *object property* », « *data property* »). Chaque entité est annotée par un *label*, un *altlabel*, et une définition.

L'évaluation d'ONAFE ne s'est pas déroulée de manière itérative. L'application des scénarios d'usage préalablement établis ont permis de simplifier la gestion de l'ontologie en cours d'évolution, et de contrôler l'impact de son évolution. Au fur et à mesure de l'évolution et de la conception des nouveaux modules, des individus ont été introduits pour visualiser les classifications obtenues et les inférences attendues par le raisonneur. Les changements ont été gérés de manière à assurer la progression et la traçabilité entre les différentes versions de l'ontologie. En définitive, tous les changements opérés sur une ontologie ont été intégrés de sorte à maintenir sa structure et sa cohérence.

Pour évaluer ONAFE et s'assurer de sa consistance et de sa cohérence, une série de requêtes DL Query ont été réalisées pour répondre aux questions de compétences notamment celles citées dans le tableau 3.6. La Figure 3.19 présente un exemple de requêtes permettant de répondre à la question suivante : quelles sont les activités de faible intensité ayant un impact sur l'endurance et s'exerçant à l'extérieur ? Parmi les activités sélectionnées pour cet exemple, nous pouvons constater qu'à chaque restriction, le nombre de choix parmi les activités diminue. Ces résultats ont été obtenus grâce aux classes définies utilisées au moment de ces requêtes. Ainsi, les classes définies, comme *ActiviteEndurance*, permettent d'utiliser le raisonneur (Fact++) pour caractériser les activités selon les contraintes définies.

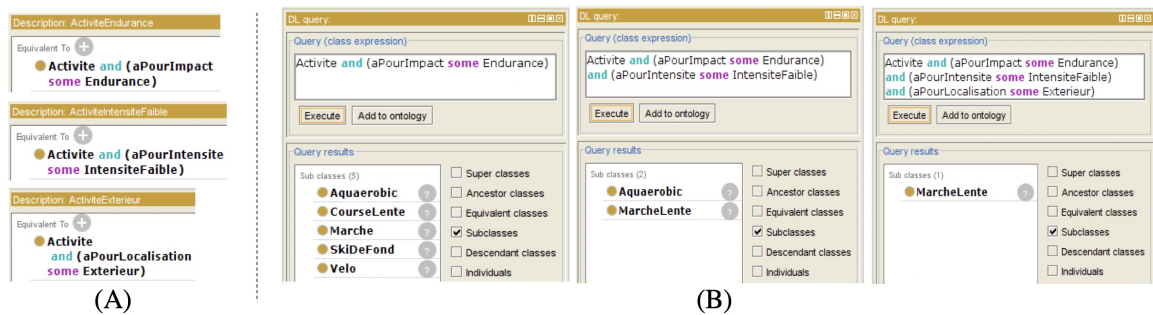


FIGURE 3.19 – Présentation de quelques exemples de classes (A) et de requêtes définies (DL Query) (B) pour obtenir les activités correspondantes pour la requête exécutée.

3.3.6 Etape 5 : Gestion des versions

La gestion des versions s'est déroulée de manière itérative. A chaque changement appliqué sur un module, une version opérationnelle est générée avec les annotations correspondantes aux changements opérés puis renommée selon la date de modification.

Le travail collaboratif nécessite de tenir compte des aspects importants liés au partage et à la mise à jour de documents et des modules relatifs à l'application de chaque étape du cycle d'évolution. L'utilisation d'espace de travail collaboratif en ligne (Git, Overleaf, etc.) facilite l'application de cette étape de gestion des versions.

3.4 Conclusion

Dans ce chapitre, nous avons présenté le processus d'évolution d'une ontologie modulaire permettant de fournir des recommandations personnalisées selon les besoins nutritionnels des personnes âgées. De plus en plus d'études soutiennent l'intérêt et le besoin de concevoir des ontologies modulaires, accessibles et compréhensibles par les experts du domaine. Le fait que chaque module soit conçu de sorte à être auto-suffisant nous a facilité l'évolution des modules et la conception de nouveaux modules, ceci en maximisant à la fois la réutilisabilité ainsi que la maintenabilité de l'ontologie globale. De ce fait, l'ajout de nouveaux modules et l'évolution d'un ou plusieurs modules a pu se faire de manière indépendante, en générant le minimum d'effets de bord et en n'affectant pas la hiérarchie des concepts des autres modules.

Bien que nous nous sommes appuyés sur une approche globale d'évolution d'une ontologie, l'évolution de MIAM présente un certain nombre d'aspect.

1. à l'architecture modulaire de l'ontologie à faire évoluer. Chaque module correspond à un espace de domaine défini. L'avantage offert par la modularité des ontologies est qu'elle favorise la réutilisation et le maintien de l'ontologie tout au long de son évolution ;
2. à l'intégration d'autres ontologies au moment de l'évolution de MIAM. Chaque ontologie modulaire représente des concepts de domaine pouvant s'adapter à des besoins particuliers ;
3. à la conception et évolution centrée sur la personne et sur les recommandations à fournir. Les exigences spécifiques des utilisateurs sont prises en compte dans la définition des questions de compétence et des scénarios d'usage. Cette méthodologie soutient l'évolution des modules et l'évaluation de l'ontologie finale ;
4. à l'application des étapes d'évolution de manière itérative, selon les besoins identifiés tout au long du processus, les problèmes de conception rencontrés et la validité de la version générée à la fin de chaque tâche.

La maintenance d'ontologies lors de leurs phases d'enrichissement et de peuplement peut être facilitée par l'utilisation d'outils. Par exemple, FOE (Flat OWL Editor)(DESPRES et GUEZENNEC, 2017) peut être utilisé pour soutenir les étapes nécessitant un traitement automatisé comme pour enrichir l'ontologie avec les connaissances de le tableau de composition de Ciqual disponibles sous format de tableaux.

Dans ce chapitre, nous avons également présenté la modélisation et le développement de nouveaux modules pour la recommandation d'activités personnalisées, ciblées et incitatives pour les personnes âgées sédentaires, dépendants ou en perte d'autonomie. La rigueur dans la représentation des activités et de la personne est fondamentale pour une interopérabilité, un échange et une réutilisation significative des données.

L'ontologie SOSED répond pour le moment au périmètre de conception. Elle pourra néanmoins évoluer par la suite pour inclure d'autres modules selon l'évolution des besoins applicatifs. Les ontologies en lien avec les objets connectés sont des ontologies IoT tenant compte de l'architecture du système IoT et n'entrent pas dans notre domaine d'application actuel mais pourrait évoluer par la suite pour inclure une conception et une problématique plus vaste. Elle pourra également tenir compte de la sémantique des

données de capteurs pour faciliter leur intégration au sein d'un écosystème vu au sens large.

Le travail présenté nous a permis d'exploiter les relations entre les concepts modélisés pour raisonner sur les différents concepts représentés dans ONAFE. En rendant ONAFE accessible pour tous et en la traduisant dans d'autres langues, notre travail participera à l'évolution de cette thématique en pleine expansion et prometteuse pour les années à venir. L'exploitation des connaissances de ONAFE et son intégration au sein d'une plateforme dédiée aux personnels de soins constituera un outil d'aide à la prescription de repas et d'activités physiques adaptés (suivant une structure et une composition de repas, ou un type, une durée, une fréquence, une intensité d'activité) selon les caractéristiques physiologiques et personnelles de chaque individu. La prévention de la sédentarité par les professionnels de santé pourra passer une prescription d'activité physique qui suit une posologie, un type, une durée, une fréquence, une intensité. L'ontologie évoluée s'intègre au sein d'un écosystème ayant une vision plus globale sur la santé et le bien-être des utilisateurs.

Les principaux défis soulevés lors de notre étude sur l'évolution d'ontologie modulaire sont liés aux aspects suivants :

- la complexité : un modèle d'ontologie est riche et, par conséquent, une ontologie a une structure entrelacée. Chaque changement entraîne une solution de contournement spécifique au changement. Même lorsque les effets d'un changement sont mineurs, l'effet cumulé de tous les changements réalisant la demande d'un utilisateur peut être énorme.
- la dépendance : les ontologies réutilisent et étendent souvent d'autres ontologies. Les changements dans une ontologie peuvent affecter les ontologies qui s'appuient sur elles. Par conséquent, les changements entre les ontologies dépendantes sont interdépendants et la synchronisation immédiate entre les ontologies dépendantes est requise. De toute évidence, la complexité de l'évolution des ontologies augmente avec le nombre d'ontologies dépendantes en cours d'évolution.
- la gestion des versions : le développement d'ontologies est un processus décentralisé et collaboratif. Par conséquent, la distribution physique des ontologies dépendantes doit être prise en compte. L'évolution de l'ontologie nécessite de suivre les changements appliqués à une ontologie et de diffuser l'ensemble des changements lorsqu'une requête explicite se présente.

Support pour la collecte de données et la présentation de la recommandation

Sommaire

4.1	Introduction	83
4.2	Conception des questionnaires en ligne.....	83
4.3	Apport des <i>Serious Games</i>	84
4.4	Approches de conception du <i>Serious Games</i>	86
4.4.1	Définition et catégorisation des <i>Serious Games</i>	86
4.4.2	Méthodologies de conception de <i>Serious Games</i>	87
4.4.3	Etapas de développement d'un <i>Serious Games</i>	88
4.4.4	Principes de scénarisation	88
4.5	Conception du <i>Serious Games</i>	89
4.5.1	Caractéristiques des utilisateurs cibles	89
4.5.2	Caractéristiques du <i>Serious Games</i>	89
4.5.3	Objectifs de conception.....	90
4.5.4	Scénarisation du <i>Serious Games</i>	90
4.5.5	Activités du <i>Serious Games</i>	90
4.5.6	Représentation iconique des aliments.....	92
4.5.7	Choix du périphérique	94
4.5.8	Développement du <i>Serious Games</i>	94
4.6	Construction d'un jeu de données fictif	95
4.7	Conclusion.....	96

L'objet de ce chapitre est de présenter notre approche de collecte de données, dans un premier temps *via* des questionnaires en ligne. A partir de cette approche, nous décrirons l'intérêt d'exploiter un autre type de support technologique pour répondre à des besoins plus large que la collecte de données. Nous décrivons ainsi l'intérêt d'exploiter les *Serious Game* comme support de collecte de données, mais également comme moyen de tenir compte des aspects de *coaching* éducatif, de la présentation et de la personnalisation de la recommandation.

4.1 Introduction

Nous avons inclus dans notre étude la prise en compte du support technologique afin de répondre dans un premier temps à un besoin de collecte de données. Cette étape clé représente la première phase de personnalisation (cf. chapitre 1). Pour commencer, la stratégie adoptée consiste à collecter les données des personnes âgées au moyen de questionnaires en ligne de sorte à l’interroger au fur et à mesure sur tous les éléments complétant son profil. Nous présentons dans la section 4.2 la conception des questionnaires en ligne.

Après avoir exploré les questionnaires en ligne comme moyen de collecte des données, nous nous sommes aperçus qu’ils ne répondaient pas à tous nos besoins en termes de collecte et de présentation de la recommandation. En effet, les utilisateurs ont généralement besoin d’être satisfaits des éléments qui leur sont proposés tout en fournissant le moins d’effort possible. La prise en compte de l’effort cognitif de l’utilisateur pour interagir et fournir ses informations personnelles au système est importante. Ce compromis effort-précision est déterminant dans la tâche de collecte, mais également dans la prise en considération de la recommandation fournie.

Pour approcher au mieux cette personnalisation, la conception centrée sur l’utilisateur est une approche de développement de systèmes qui implique les utilisateurs finaux potentiels et permet de mieux comprendre leurs comportements, leurs tâches et leurs besoins. Les *Serious Games* font partie de ces supports applicatifs qui présentent de nombreux avantages pour le développement d’un système centré « utilisateur » et répondant aux besoins à la fois en termes :

- de collecte des données explicites et implicites en lien avec ses connaissances et ses habitudes nutritionnelles et d’activité ;
- d’expression des besoins de la personne âgée ;
- de visualisation et d’adoption de la recommandation ;
- de stimulation ou de *coaching* éducatif nutritionnel et d’activité physique pour l’encourager à exercer une activité physique ou favoriser ou éviter de consommer un aliment en particulier ;
- et de personnalisation de tous ces éléments.

4.2 Conception des questionnaires en ligne

Pour concevoir la structure des questionnaires en ligne, nous nous sommes appuyés sur les questionnaires d’évaluation sélectionnés et présentés dans la section 2.2.3. Un questionnaire est constitué d’un ensemble de questions et chaque question est composée d’un ensemble de propositions. Selon le type de la question (text field, checkbox, etc.), l’utilisateur fournit ou choisit la ou les réponses qui lui correspondent.

L’objectif consiste à poser des questions à la personne de manière successive en fonction des données initialement collectées. L’affichage des questionnaires en ligne suit les étapes suivantes :

- dans un premier temps, une interface introductive est présentée pour collecter des données essentiels (sexe, âge, caractéristiques physiologiques, etc.) pour la construction d’un premier modèle de profil général (Figure 4.1).

- dans un second temps, des questionnaires plus spécifiques sont suggérés à l'utilisateur pour collecter des données selon 3 principales catégories : santé, nutrition et activité.
- à partir des données collectées concernant ces catégories, des questionnaires personnalisés et spécifiques au profil préalablement établi pourront être proposés. Par exemple, dans le cas où la personne présente un IMC inférieur à 18,5 (seuil de maigreur), le questionnaire MNA lui sera proposé.

Questions générales

Quel âge avez-vous ?*

88

Avez-vous une pathologie ?

Hypertension

Diabète

Quel est votre mode de vie ? *

Très actif

Actif

Sédentaire

Qu'est ce qui vous empêche de faire de l'activité physique ?

Douleurs articulaires

Peur de chuter

Répondre

FIGURE 4.1 – Aperçu d’une interface de questionnaires en ligne présentant des questions d’ordre général

Pour faciliter la génération des questionnaires en ligne, une interface dédiée à l’expert a été conçue pour permettre la génération de questionnaires selon les besoins de collecte (Figure 4.2). A partir de cette interface, l’expert du domaine a la possibilité de saisir les questions selon leurs catégories, tout en ayant la possibilité de définir les critères de choix, le score pour chaque proposition, le type de question (choix unique, choix multiples, etc.). Le caractère obligatoire de la réponse attendue a été évité pour ne pas contraindre la personne âgée à répondre sans qu’elle ne l’ait souhaitée.

A la suite de ce travail, nous avons exploré l’apport des *Serious Game* pour une conception répondant à des besoins plus importants que les questionnaires en ligne.

4.3 Apport des *Serious Games*

Les *Serious Games* suscitent depuis de nombreuses années un intérêt croissant pour le changement de comportements et d’attitudes plus largement dans les domaines de la santé, des politiques publiques et de la publicité ainsi que de l’éducation et de la formation. La diversité des secteurs d’application et l’étendue des champs d’applications possibles expliquent les fortes perspectives de croissance pour ce marché au niveau

Paramètres - Catégorie : "Activité"

Questionnaire - Nom et description

Titre
Activité

Description
Profil général

Groupe de question N°1 [Delete from the Database]

+ Questions

Question n°3 [Supprimer]

Boutons rai

Quel est votre mode de vie ?

Champ obligatoire: Oui Non

+ Proposition

[X Supprimer la proposition de la base de données]
Très actif
0

[X Supprimer la proposition de la base de données]
Actif
0

[X Supprimer la proposition de la base de données]
Sédentaire
1

FIGURE 4.2 – Aperçu d’une interface administrateur pour la génération de questionnaires en ligne

mondial. Un nombre considérable de *Serious Game* ont été développés au cours des dix dernières années (ALMEIDA et SIMOES, 2019) et leur exploitation dans les domaines de la santé est actuellement en pleine expansion.

L’intérêt des *Serious Game* a également été étudié auprès de la population âgée. Communément, lorsque le thème du jeu vidéo est abordé, il n’est pas intuitivement associé aux personnes âgées. Toutefois, 62 % des 50-64 ans et 50 % des personnes de plus de 65 ans déclarent être des joueurs (Ifop, 2018), et les nombreux *Serious Game* pour personnes âgées déjà existants consolident ce fait. Les *Serious Game* font actuellement partie des nouvelles mesures thérapeutiques utilisées dans le milieu gériatrique participant au maintien d’une bonne santé et à l’autonomie des personnes âgées. En plus d’être thérapeutiques, ils ont également une visée pédagogique, diagnostique. Ils permettent de mesurer, de contrôler et de stimuler les capacités physiques et cognitives des personnes âgées tout en restant dans des conditions ludiques. Utilisés dans les EHPAD, ils constituent également un moyen pour entretenir le lien social. Par exemple, ils sont proposés aux patients atteints de maladies neurologiques (Alzheimer, etc.) pour fournir des solutions alternatives de traitement, de stimulation et de réadaptation (VALLEJO et al., 2017, GHASSEMI et al., 2019, etc.). Récemment, leurs intérêts ont été étudiés pour accroître les compétences chirurgicales des étudiants en médecine et améliorer l’apprentissage et l’éducation (GORBANEV et al., 2018, RAVYSE et al., 2017). Une étude sur l’utilisation des jeux dans la rééducation des chutes à domicile a été réalisée pour démontrer que les jeux pouvaient être utilisés avec plus de cohérence par rapport à la rééducation conventionnelle (UZOR et BAILLIE, 2014). Des études psychologiques ont montré qu’en réalité, les utilisateurs ne peuvent connaître réellement

leurs préférences que s'ils sont confrontés à un ensemble de produits ou à des alternatives d'éléments (HONICKE et BROADBENT, 2016).

Avec l'émergence du paradigme de l'IoT, une nouvelle direction pour les *Serious Game* devrait apparaître, où la collecte de données à partir de l'environnement physique peut être utilisée pour de nouvelles applications novatrices. Dans la littérature, des travaux ont été réalisés pour exploiter les *Serious Game* et améliorer la personnalisation notamment de parcours pédagogiques (STREICHER et SMEDDINCK, 2016). Bien que des avancées majeures aient été réalisées en IA ces dernières années, comme en témoignent les avancées importantes dans l'apprentissage en profondeur avec les réseaux de neurones (MIKKULAINEN et al., 2019), l'ajout de fonctionnalités de personnalisation aux *Serious Game* de manière entièrement automatisée n'est pas encore facilement réalisable. Dans ce sens, la puissance offerte par les *Serious Game* contribuent ainsi à mettre la technologie au service de la santé personnalisée. A travers leur interface interactive et immersive, ils offrent au-delà du divertissement, un moyen pour atteindre les utilisateurs sur le plan émotionnel, de favoriser un engagement actif de leur part et d'adapter le soin en fonction des intérêts, des capacités, des besoins et des contextes. En effet, les théories des jeux sur lesquels s'appuie la conception des *Serious Game* font naître chez l'utilisateur le sentiment de réussite et de contrôle qui contribuent à la motivation et au changement de comportement (CONNOLLY et al., 2012).

4.4 Approches de conception du *Serious Games*

4.4.1 Définition et catégorisation des *Serious Games*

Plusieurs définitions existent au sujet des *Serious Game*. Selon le ministère de l'Éducation nationale, le *Serious Game* est un « outil de formation, communication et de simulation qui est une déclinaison utile du jeu vidéo au service des professionnels ». Selon MARFISI-SCHOTTMAN et al., 2010, un *Serious Game* est un jeu qui dépasse le divertissement. C'est une application qui utilise les principes théoriques des jeux vidéo comme une interface ludique, un scénario et des personnages qui interagissent avec l'utilisateur, le tout regroupé au sein d'une application dans le but d'atteindre un objectif professionnel, de transmettre ou d'évaluer une ou plusieurs compétences. Ainsi, en plus de leur caractère ludique, les *Serious Game* peuvent avoir trois types de visées : de formation ou d'évaluation, de marketing ou de promotion et d'information ou de sensibilisation.

JANTKE et GAUDL, 2011 catégorisent les *Serious Game* selon quatre dimensions : le contenu éducatif principal fourni par le jeu (la santé, l'interaction sociale, l'éducation scolaire, etc.) ; les principes d'apprentissage primaires (la pratique des compétences, la résolution de problèmes, etc.) ; la tranche d'âge visée par le jeu ; et la plate-forme sur laquelle le jeu est joué. ALVAREZ et al., 2012 proposent une catégorisation des *Serious Game* selon leur finalité : les *advergames*, les *edugames* (*Learning Game*), les *exergames*, les *newsgames*, les *socialgames*. La combinaison de différentes catégories au sein d'un seul et même *Serious Game* est possible selon la finalité attendue.

La classification des *Serious Game* selon leurs dimensions « ludique » et « sérieuse » de manière simultanée peut se faire selon 3 critères (DJAOUTI, 2011) :

- le « *gameplay* » du *Serious Game* : ce critère renseigne sur la dimension « ludique » et apporte des informations sur le type de structure utilisé pour créer le jeu ;
- la finalité du *Serious Game* : ce critère renseigne sur la ou les « vocations dépassant le simple divertissement » souhaité par le concepteur ;
- le secteur, le domaine d'application du *Serious Game* et le public visé.

L'intérêt d'un *Serious Game* est d'apporter un certain changement au joueur comme pour faire subir un changement dans les connaissances, l'attitude, la capacité physique, la capacité cognitive, la santé ou le bien-être mental. Cet aspect fait la distinction entre les *Serious Game* et les jeux qui ont été conçus pour le divertissement.

La conception d'un *Serious Game* peut prendre différents cheminements, parmi eux la "gamification" ou la "ludification" du *Serious Game* (GORIA, 2014). La "gamification" est un processus de transformation de ce qui n'est pas un jeu en un jeu, ceci en y ajoutant ou en retirant des éléments structurels. Ces éléments se fondent sur des points essentiels afin de motiver la personne à accomplir une tâche, comme la valorisation des actes de la personne qui est implémentée dans certains jeux à travers leurs histoires et leurs scénarios ou la gratification et la récompense sur les actions effectuées à travers l'attribution de badges, d'insignes, de plaques ou d'emblèmes.

4.4.2 Méthodologies de conception de *Serious Games*

Le domaine du *Serious Game* étant assez novateur et les concepteurs experts assez rares, de nombreuses méthodes ont été élaborées pour soutenir la conception des *Serious Game*. Une des méthodes de conception du *Serious Game* consiste à décrire le *Serious Game* selon 6 facettes (MARNE et al., 2011) qui finalement permettent de se poser un certain nombre de questions :

1. Objectif pédagogique : que cherche t-on à transmettre au joueur ?
2. Simulation du domaine : comment symboliser les éléments du jeu à présenter du joueur ?
3. Interactions avec la simulation : comment rendre les interactions stimulantes ?
4. Problèmes et progressions : quels problèmes sont à faire résoudre au joueur et comment mesurer sa progression ?
5. Décorum : par quels éléments scénaristiques et média (tablette, smartphones, etc.) les objectifs pédagogiques peuvent-ils être atteints ?
6. Condition d'utilisation : quelles sont les conditions qui permettent d'exploiter le *Serious Game* tout en conservant ses qualités pédagogiques et ludiques ?

La conception d'un *Serious Game* implique la participation et la collaboration de différents acteurs et experts, à la fois des *Serious Game*, de développeurs, de designers, de "story boarders" mais également ceux du domaine d'application (dans notre cas, les personnes âgées, la nutrition, l'activité physique, etc.). Dans la section qui suit, l'objectif est de fournir des éléments de conception du *Serious Game* pour répondre aux besoins.

Lors de la conception du *Serious Game* et au vu du cadre de conception n'impliquant pas d'acteurs externes, nous nous sommes principalement appuyés sur les facettes de conception.

4.4.3 Etapes de développement d'un *Serious Games*

Les étapes de développement d'un *Serious Game* suivent les mêmes étapes de conception d'un jeu vidéo, à savoir : le « *Game Design* », le « *Level Design* » et le « *Game Development* » (ADAMS, 2014). Le « résultat » de ce processus est le *Serious Game*. Le *Game Design* est un processus visant à détailler de manière très précise le fonctionnement du jeu. Dans un premier temps, le but est d'avoir une vision globale du jeu, en d'autres termes réaliser la carte d'identité du jeu. Cela revient à trouver un titre au jeu, déterminer le type de jeu souhaiter ainsi que le public visé et enfin, la ou les plateformes sur lesquelles tournera le jeu.

Dans un deuxième temps, un "*game designer*" devra inventer les règles et l'univers du jeu, cela implique donc de définir le concept ou l'idée générale du jeu. A partir de là, les conditions et règles qui régissent les actions du joueur sont définies en s'appuyant sur quatre éléments de base : le challenge, les choix de l'utilisateur, le changement et enfin la chance. Tous ces éléments devront être bien pensés afin d'avoir un jeu suffisamment attrayant pour que le joueur ne se lasse pas et décide de ne plus y jouer.

La finalité de cette étape est la rédaction du document *Game Design* qui résume tous les points cités précédemment. Le design est une part importante du projet, car en plus de devoir accroître l'immersion, il doit réussir à transmettre par de simples modèles graphiques les idées adéquates. Chaque design doit être validé par l'expert pédagogique. Le *Level Design* est complémentaire au *Game Design*. Il est consacré à la création des différents niveaux du jeu ainsi que le scénario de chaque niveau (DJAOUTI, 2011). Chaque *level* doit être pensé pour maintenir le scénario. Plus explicitement, le *level design* doit permettre de varier la tension du jeu de manière à garder constant l'intérêt du joueur en agissant au niveau des thèmes ludiques, de la difficulté et autres variables qui demande au joueur une adaptation constante pour la mise en pratique de ses connaissances et compétences.

4.4.4 Principes de scénarisation

La scénarisation ne consiste pas seulement à construire un simple récit ou une mise en contexte. Le scénario conçu doit orienter vers des situations pertinentes, pouvoir expliquer pourquoi elles surviennent, et fournir des éléments d'évaluation et d'auto-évaluation.

Plusieurs critères peuvent déterminer la construction d'un scénario. Le scénario comporte 3 niveaux scénaristiques : microscénarisation, mésoscénarisation, macroscé-narisation (CARPENTIER, 2015). La microscénarisation correspond à « la situation en cours », et permet de contrôler directement la difficulté ou d'apporter une modification quelconque sur ce qui se déroule. La mésoscénarisation permet le « contrôle de la session d'apprentissage ». Il correspond à la gestion d'une utilisation continue du dispositif d'apprentissage, soit une session de jeu. La macroscé-narisation définit le « parcours d'apprentissage », donc le déroulement du jeu en lui-même, de l'entrée du joueur dans l'environnement jusqu'au moment où le joueur a acquis l'ensemble des connaissances véhiculées par le *Serious Game*. La scénarisation peut être scriptée, générative en fonction des réponses de l'utilisateur, ou hybride (CARPENTIER, 2015). La méthode hybride a été retenue pour ce projet. Elle permet de scripter et donner un fil conducteur solide au scénario, et en même temps de rendre l'expérience d'apprentissage personnalisable avec

le côté génératif. Le scénario doit s’articuler d’objectifs. Chaque objectif peut être de nature explicite ou implicite, et de type pédagogique ou narratif. Ce sont ces chemins vers les objectifs que doivent explorer les différents scénarios.

4.5 Conception du *Serious Games*

4.5.1 Caractéristiques des utilisateurs cibles

La condition physique d’une personne âgée et son autonomie peuvent varier du tout au tout et intégrer des paramètres que le *Serious Game* ne peut pas prendre en compte. Par exemple, une personne non autonome a besoin d’une personne proche ou d’un assistant pour l’aider dans ses tâches quotidiennes. Cette personne ne jouant pas à ce jeu, peut influencer sur les menus de l’utilisateur. De même, jouer ne garantit pas forcément que les conseils sont appliqués et que les habitudes sont prises. C’est pourquoi, le *Serious Game* actuel cible une personne âgée autonome.

4.5.2 Caractéristiques du *Serious Games*

En tant que joueurs, les personnes âgées sont principalement intéressées par les puzzles et les jeux éducatifs (CHESHAM et al., 2017). Les aspects préférés d’un jeu dans cette tranche d’âge sont la compétition de scores, la possibilité de jouer seul, et que le jeu soit intellectuellement stimulant en regard de la qualité des graphiques ou du côté fantaisiste (BLOCKER et al., 2014). Connaissant les préférences des personnes âgées vis-à-vis des jeux vidéo, le *Serious Game* conçu a pris la direction d’un puzzle.

Ce jeu doit être capable de formuler des suggestions nutritionnelles personnalisées et doit réussir à récupérer *via* son scénario les préférences alimentaires de l’utilisateur.

En appliquant la méthode des 6 facettes du jeu sérieux, on obtient ce descriptif :

- Objectif pédagogique : donner au joueur de bonnes habitudes de nutrition.
- Simulation du domaine : utiliser la symbolique actuelle, conseiller pour détromper des idées reçues.
- Interactions avec la simulation : jeux de logiques sur les groupes alimentaires, composition de menus, correction de menus préétablis, quizz.
- Problèmes et progressions : faire comprendre les enjeux d’équilibre alimentaire, de planification, avec des niveaux au scoring important.
- Decorum : il inclue le coach, le personnage non joueur, et la possibilité d’avoir un avis d’expert.
- Condition d’utilisation : parties en solo, ne requiert pas d’être assisté, se joue n’importe où, environ 20 à 30 minutes par session.

A partir de ces facettes, quatre paliers d’apprentissages se distinguent, et marquent les quatre objectifs pédagogiques explicites :

1. Classer : selon les groupes alimentaires du PNNS (fruits et légumes, matières grasses ajoutées, etc) ;
2. Doser : selon les repères nutritionnels adaptés aux personnes âgées ;

3. Composer : pour l’obtention d’un menu quotidien (notion d’équilibre alimentaire) ;
4. Planifier : pour l’obtention de plusieurs menus sur une durée prolongée (notion d’équilibre alimentaire approfondie).

Chaque palier comporte son lot d’objectifs pédagogiques implicites, comme le fait de reconnaître les produits selon leurs attributs (groupe PNNS, riche en sel, etc.).

4.5.3 Objectifs de conception

Le champs d’application pour la conception d’un *Serious Game* répondant aux besoins applicatifs est assez large. Nous avons pris le choix de nous intéresser à certains aspects de collecte de données. Le *Serious Game* a été conçu de sorte à répondre, dans un premier temps, à un besoin de collecte de données comme présenté pour le questionnaire en ligne. Nous avons cependant pu améliorer cette approche en appliquant les principes de scénarisation du *Serious Game*. En plus de répondre au besoin de la collecte de données de manière explicite, nous cherchons par la suite à connaître ses préférences alimentaires. Pour cela nous avons mis en place un certain nombre d’activités répondant aux objectifs pédagogiques.

4.5.4 Scénarisation du *Serious Games*

Pour ce type de *Serious Game*, où l’interaction en phase d’apprentissage se fait par des activités de logique essentiellement de type puzzle et réflexion, l’importance est de varier la fréquence d’apparition des activités de sorte à faire travailler des compétences différentes et éviter la répétition et l’ennui.

Le principe de scénarisation présenté dans la section 4.4.4 a été appliqué pour structurer les questionnaires en ligne et les présenter de manière successive selon les données préalablement collectées.

Les niveaux mésoscénaristique et microscénaristique ont été inclus dans la scénarisation. La mésoscénaristique correspond à un enchaînement de missions à accomplir sous la supervision du *coach*. Ces missions peuvent être de tous types et sans rapport direct avec la nutrition, mais cela dépend du choix de la macrosécénaristique qui n’est pas encore définie, et qui permet d’ étoffer le décorum. Pour réaliser ces missions, l’utilisateur devra venir à bout de plusieurs mini-jeux. Chaque activité ou ensemble d’activités est amené et justifié par les différentes microscénaristiques de l’arc couvert par la mésoscénaristique du moment. La figure 4.3 montre une possibilité de mésoscénaristiques hybrides. Elles sont hybrides car la mésoscénaristique a un objectif précis à remplir, ce qui permet de donner une directive scriptée pour y parvenir. Le choix des activités et de l’objectif à atteindre sont dépendants de la progression du joueur et de l’accomplissement des objectifs pédagogiques implicites.

4.5.5 Activités du *Serious Games*

Le choix des activités est crucial car c’est à travers elles que le joueur devra travailler les compétences et interagir avec la simulation. Le design a été pensé pour utiliser la sélection et le *drag and drop*.

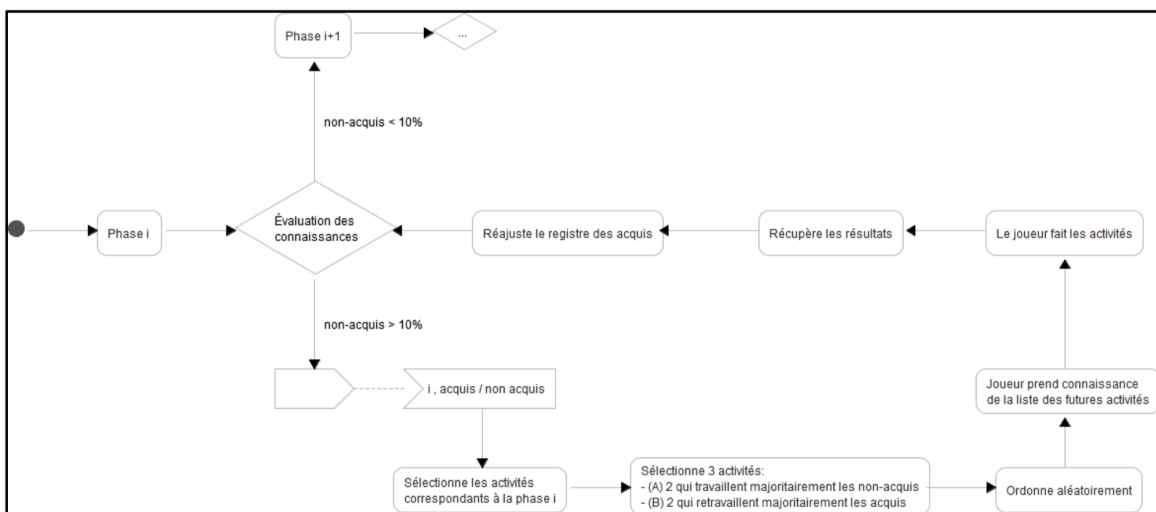


FIGURE 4.3 – Organisation des phases d'apprentissages

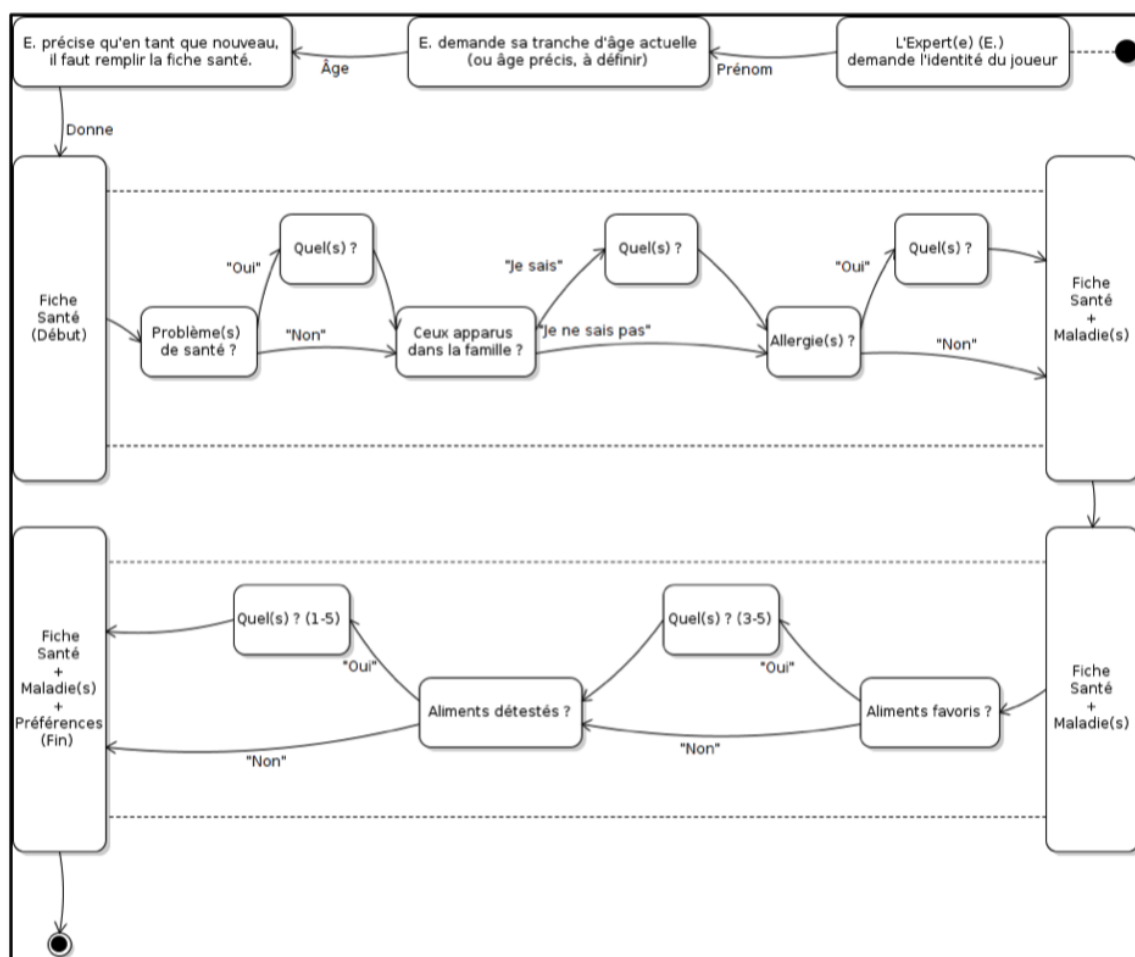


FIGURE 4.4 – Exemple de scénario pour présentation de la « fiche santé » au joueur

L'ensemble des activités permettent de couvrir l'ensemble des connaissances que nous cherchons à transmettre à l'apprenant. Les activités créées permettent de faire atteindre l'objectifs pédagogiques attendues. Chaque activité permet de travailler un sous-objectif implicite et d'intégrer des éléments d'évaluation et de satisfaction pour le joueur.

Nous présentons quatre de ces activités (Figure 4.5) :

1. **ClassifiX** : cette activité est pensée pour aider à la classification des aliments par thèmes, et sert au premier objectif pédagogique. Les thèmes peuvent être simples en ne reprenant que les groupes alimentaires du PNNS, ou plus complexes en devenant transversaux. Une fois le thème affiché, le joueur peut choisir de conserver ou non les aliments qui lui sont présentés. Si le choix est bon, le joueur voit un « +1 » s'afficher et s'ajouter à son score, et sinon « -1 ». ClassifiX fait réfléchir le joueur entre l'action de *drag et de drop* et permet d'installer le doute qui aura tendance à s'atténuer avec la montée du score et la rapidité d'exécution.
2. **Find It** : cette activité aide le joueur à atteindre le second objectif pédagogique. Find It se joue un peu à la manière de ClassifiX, à la différence que la sélection positive remplace le *drag and drop* et que les thèmes sont centrés sur les repères principaux des groupes alimentaires. Ainsi un thème possible de Find It serait « À limiter » ou « 1 fois par jour ». Dans cette activité, le score est affiché tout le long de la partie, ce qui rajoute une couche de tension en regard de ClassifiX qui ne donne le score qu'en fin de partie. Le *drag and drop* cette fois absent, il n'y a pas de place au doute.
3. **Cuisine** : dans cette activité, le joueur est amené à composer un menu de plusieurs plats pour atteindre le troisième et le quatrième objectif pédagogique. Dans l'idéal, le plat composé recevrait une note de Nutri-Score pour avoir une visualisation simple et actuelle de la qualité du plat proposé.
4. **Be Coach** : dans cette activité, le joueur se met à la place de l'expert et doit juger et corriger, si besoin, des menus préétablis. Cela permet d'appliquer directement les connaissances acquises sur les groupes alimentaires, leur dosage, et la composition sur une durée d'une journée. Pour corriger les éléments affichés, le joueur est amené à barrer ce qu'il pense être en excès. Une version plus difficile demande au joueur de substituer les aliments si nécessaire. Il est possible de retourner sur ses pas en appuyant de nouveau sur "*strike*" qui devient "*erase*" devant les lignes barrées précédemment.

A partir des activités ClassifiX et Cuisine, une interface de « Composition de menu » a été conçue (cf. section 4.5.8). Ainsi, la question de la représentation iconique des aliments s'est posée.

4.5.6 Représentation iconique des aliments

Pour représenter les aliments à l'utilisateur, plusieurs critères ont été pris en compte pour choisir la forme la plus adaptée, tels que la représentativité, la disponibilité, l'uniformité et l'apparence des aliments. Vu l'importance de l'apparence et du design pour une telle application, nous avons choisi de mettre les aliments sous forme d'image. Les images sont plus attractives qu'un texte, elles influent énormément sur l'expérience

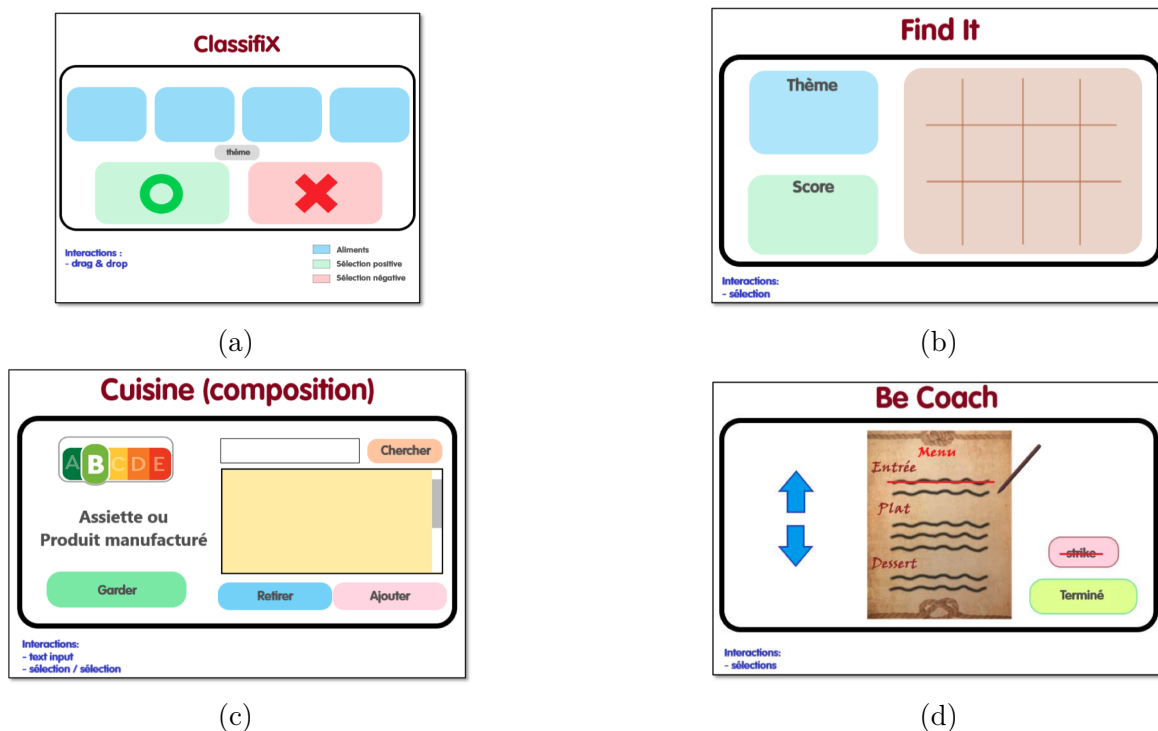


FIGURE 4.5 – Les activités du *Serious Game*

utilisateur. L'image est aussi en quelque sorte un langage universel qui permet d'élargir la tranche d'âge des utilisateurs. Parmi les formats d'images existants, le format PNG est le plus intéressant pour les icônes, car il permet d'apporter de la transparence à l'image (image sans arrière-plan) par rapport au format JPG. De plus, lorsqu'il s'agit des icônes ou des images sans arrière plan, le format PNG prend moins d'espace de stockage en mémoire que le format JPG, et l'image se charge plus rapidement dans l'interface.

La bibliothèque choisie pour définir les images des aliments est Flaticon¹, une base de données d'icônes gratuites et personnalisables. Dans cette bibliothèque, les images d'aliments en format PNG sont de type clipart. Ce type d'image est facilement modifiable pour les rendre uniformes. Un certain nombre de photos d'aliments (ou image réelle) sont disponibles sur Google Image, mais contrairement à la bibliothèque utilisée, certaines images ne sont pas libres de droit. De plus, les tailles d'images sont hétérogènes et il faut ajuster chaque image téléchargée pour pouvoir les utiliser.

Plus de 2500 images mises sur Flaticon sont libres de droits. De plus, cette plateforme nous donne la possibilité de modifier les images grâce à un éditeur intégré. Cependant, certains aliments, tels que la crème fraîche ou la pâte d'arachide par exemple, ne sont pas forcément représentés dans la bibliothèque d'images. Ce qui peut être envisagé pour résoudre cette complication est d'utiliser l'image la plus neutre et la plus représentative à la fois pour y ajouter les informations manquantes. Par exemple, dans le cas de la crème fraîche, un contenant neutre a été choisi et un texte a été ajouté pour y préciser le contenu.

1. <https://www.flaticon.com/fr/>

4.5.7 Choix du périphérique

Le choix entre les périphériques de jeu est large (PC standard, smartphone, tablette, console connectée à un téléviseur, etc.). L'avantage offert par un périphérique mobile (smartphone ou tablette) est qu'il contient des capteurs et des fonctionnalités intéressantes à exploiter dans le *Serious Game* (GPS, écran tactile, capteurs d'accélération). Le jeu vidéo est souvent blâmé en tant que contributeur à la sédentarité, en particulier chez les enfants et les adolescents. Cependant, *Serious Game* de type Exergames (danse ou sport sur Wii Fit, etc.) peuvent contribuer à cette lacune. Contrairement aux PC de bureau, les *Serious Game* proposés sur smartphones peuvent contribuer à favoriser les activités physiques en plein air comme la marche et la course. De ce fait, le *Serious Game* prend la forme d'une application Web accessible sur tout support mobile (smartphone, tablette). L'application doit être "*responsive*", et les interactions adaptées à la personne âgée.

4.5.8 Développement du *Serious Games*

En réalité, la question du choix de langage n'est pas aussi cruciale que les précédentes étapes de conceptions. Quel que soit le choix, il doit permettre de développer les conceptualisations du *Serious Game*. Les technologies web sont au service des *Serious Game*, que ce soit pour le cœur du *Serious Game* ou pour aider au développement. Ils laissent une liberté quant au produit final (application mobile, Web, ou multi-plateformes). S'ajoute à ces langages de nombreux framework (PIXI.js, etc.) qui permettent de dessiner et d'intégrer au bon vouloir n'importe quel élément graphique dans un canvas HTML5.

Interface interactive pour la composition de menu

L'interface de « composition de menu » a été conçue pour récupérer les préférences de l'utilisateur selon les catégories d'aliments retrouvés dans le PNNS (CIV, CENTRE D'INFORMATION DES VIANDES, 2011). L'objectif explicite est de pousser l'utilisateur à choisir des aliments pour construire son repas, selon ces préférences individuelles. Le principe de cette interface consiste à faire glisser des aliments d'une zone (zone de *drag*) vers une autre (zone de *drop*) qui contient un plateau contenant les différents éléments d'un menu (entrée, plat, dessert, boisson) retrouvés dans les recommandations du PNNS. Chaque élément du plateau est considéré comme étant une zone de dépôt. L'assiette du plat principal est plus grande que ceux de l'entrée et du dessert, le compartiment boisson est représenté par un verre. Les données collectées à partir de cette interface concernent dans un premier temps ses préférences alimentaires. Il sera par la suite également possible d'intégrer d'autres critères de préférences (sensorielles, mode de cuisson, etc.). Des objectifs implicites peuvent être également imaginés pour exploiter l'idée de la catégorisation des aliments au sein de groupes alimentaires. En effet, cette catégorisation permet d'attirer l'attention de l'utilisateur sur l'appartenance d'un aliment à une catégorie et pourrait pousser l'utilisateur à varier les aliments choisis en fonction de ces groupes.

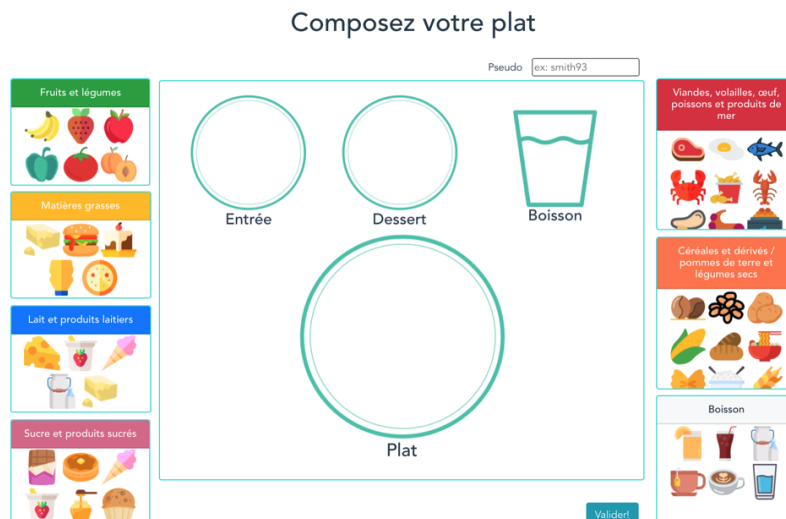


FIGURE 4.6 – Interface de composition de plat

Interfaces de coaching éducatif et présentation

La collecte de données peut être facilitée par l'intermédiaire de questions/jeux interactifs et l'adoption de la recommandation par la personne âgée peut être facilitée *via* une interface qui lui offre une suggestion de choix. Un exemple d'interface est proposé (Figure 4.7) pour permettre à l'utilisateur de répondre à des questions concernant un thème basé sur l'activité physique afin de récupérer des « étoiles ». L'attribution d'étoiles est différente en fonction des réponses données, de plus, après chaque question, un petit message s'affiche, apportant des informations supplémentaires permettant de sensibiliser la personne âgée sur des thématiques en lien avec ses compétences et ses besoins.



FIGURE 4.7 – Interface de *coaching* éducatif sur des notions d'activité physique et de présentation de la recommandation

4.6 Construction d'un jeu de données fictif

Pour être en mesure de tester cette approche, il est nécessaire d'avoir un jeu de données. Actuellement, la question de la construction d'un jeu de données en santé représente une réelle problématique (cf. section 2.3.1) et constitue une tâche à part

entière. Bien que de nombreuses bases de données médico-administratives (Sniiram, PMSI, etc.) sont exploitées pour réaliser des études en santé. Dans notre cas d'étude, la difficulté d'obtenir des données longitudinales (à la fois sur les habitudes alimentaires, sur l'activité issue d'objets connectés, etc.) est bien plus considérable. Actuellement, la majorité des jeux de données issues de capteurs d'activité, disponibles ne sont pas exploitables. Lorsqu'un jeu de données est disponible, les difficultés rencontrées pour l'exploiter sont liées à : la représentativité des données selon les problématiques étudiées ; la spécificité des données selon le domaine de recherche ; le format de données à traiter pour être exploité.

Au vu du constat exposé sur la difficulté associée à la collecte de données en santé et à la constitution d'un jeu de données, nous avons fait le choix de construire un jeu de données fictif. Cette solution nous a permis d'avoir des données plus spécifiques aux problèmes traités et de nous abstraire des difficultés associées à la confidentialité des données.

Le jeu de données a été construit sur la base des différentes sources de données (DMP, objets connectés, *Serious Game*) et des éléments permettant la construction d'un profil de santé. Nous avons cherché à ajouter des données représentatives des problèmes de santé abordés. La figure 4.8 présente un exemple des éléments constituant le jeu de données fictif et qui sera utilisé par la suite pour appliquer l'approche d'intégration de données présentée dans le chapitre suivant. Les données des utilisateurs représentées dans le jeu ont ensuite été intégrées à la base de données par simulation d'utilisateur. Pour cela, les interfaces programmées ont enfin été utilisées pour alimenter, à partir du jeu de données construit, la base de données. Ces données seront ensuite traitées dans la suite du processus de création du profil de santé (cf. chapitre 5).

DMP								SG				OC							
id	aAge	aPoids	aTaille	aPathologie	aTypePathologie	Cause Hospitalisation	Allergie	id	Caractéristique sociale	Frein Psychologique à l'activité	Préférences Sensorielles	Objectif/Cible	id	Nb de pas/jour	Activité Détectée	Distance Parcours (km)	Durée Activité/Moderée (min/jour)	Vitesse Moy	Fc Moy
1	85	42	1,62	Diabète	Chronique	Deshydratation	Arachides	1	Seule	PeurDeTomber	AimeSucré	Renforcement musculaire	1	3500	Marche	0,2	1	1	64
2	97	67	1,76	Hypertension	Chronique	Vertiges	Blé	2	En couple	ManqueMotivation	AimeSalé	Endurance	2	8000	Course	3	40	8	120
3	63	82	1,9	Arthrose	Chronique	Hernie discale	Fruits de mer	3	Réseau Ami	Fatigue	AimeAmer	Equilibre	3	2000	Velo	5	30	20	90
4	95	59	1,77	Ulcère	Chronique	Pneumonie	Graines de sésame	4	Celibataire	ManqueMateriel	AimeSucré	Souplesse	4	4000	Marche	7	50	3	73
5	85	63	1,98	Alzheimer	Chronique	AVC	Lait	5	Co-location	MeteoDefavorable	AimeSalé	CardioRespiratoire	5	3000	Velo	9	33	15	99
6	76	89	1,8	Parkinson	Chronique	Chute	Noix	6	Couple	AntecedentHabitueSedentaire	AimeAmer	Endurance	6	8000	Marche	4	3	4	127
7	90	97	1,61	Insuffisance rénale	Chronique	Catarate	oeufs	7	FamilleNombreuse	EtatDeStress	AimeSucré	Equilibre	7	14000	Course	4,9	45	12	60
8	69	137	1,88	Asthme	Chronique	Crise cardiaque	Soja	8	MonoParental	ManqueInteret	AimeSalé	MobiliteArticulaire	8	17000	Velo	2	17	13	72
9	82	48	1,57	Syndrôme dépressif	Chronique	Accident véhicule	AINS	9	PresenceDenfants	Peurblessure	AimeAmer	RenforcementMusculaire	9	9000	Marche	22	88	5	88
10	77	76	1,82	Trouble de la mastication	Chronique	Dénutrition Déshydratation	Pollen	10	Recomposee	PeurDeTomber	AimeAmer	RenforcementOssieux	10	10000	Course	44	29	10,1	77

FIGURE 4.8 – Aperçu du jeu de données fictif construit

4.7 Conclusion

A travers ce chapitre, nous avons présenté des éléments de conception du *Serious Game*. En effet, penser le support de collecte et de visualisation lors de la conception d'un système personnalisable permet d'appréhender les problématiques qui y sont liés et de centrer le développement selon les besoins de l'utilisateur. Le *Serious Game* participe également à la sensibilisation de la personne âgée en lui exposant les risques face à un comportement particulier détecté. L'interface de composition de menu conçue

permet de répondre à un besoin de collecte de données qu'il n'est pas possible d'obtenir par l'intermédiaire d'un questionnaire en ligne. Par exemple, la collecte des habitudes alimentaires représente une réelle difficulté. Elle nécessite de poser un grand nombre de questions sur une longue durée. L'utilisation d'une telle interface permet à l'utilisateur de rendre ludique une telle tâche et de collecter de manière implicite d'autres informations à son sujet (connaissances sur la nutrition, etc.).

La conception des *Serious Game* répondant aux besoins et aux objectifs définis nécessitent l'implication d'un bon nombre d'experts dans la conception de *Serious Game* et dans le domaine d'application (les personnes âgées, la nutrition, l'activité physique). Le choix des icônes présentant les informations affichées ou les actions à réaliser ont présenté un réel défi que ce soit pour le choix d'une source d'icônes ou pour le choix d'une icône par rapport à une perception attendue et présumée instantanée par le joueur. En effet, l'icône choisie doit être représentative d'une réalité pour augmenter l'immersion de l'utilisateur.

Pour avoir à notre disposition un jeu de données représentatif de notre cas d'étude, nous avons initialement prévu de prendre contact avec des EHPAD pour proposer aux résidents d'utiliser les interfaces conçues pour collecter de données en lien avec leurs habitudes alimentaires, leur statut de dénutrition, leur état de santé général, leurs préférences en termes d'alimentation et d'activités physiques. Nous prévoyons également d'explorer les possibilités de ces travaux avec un gérontologue ayant accès à des structures d'hébergement favorisant l'autonomie des personnes âgées institutionnalisées. Cet objectif a été revu en raison de l'impossibilité d'accéder aux EHPAD suite à la situation sanitaire de 2020.

Pour conclure, à partir de ce travail, nous avons pu nous apercevoir des impératifs technologiques nécessaires pour réduire le processus de collecte de données et de développement combinant l'exploitation des *Serious Game* et des objets connectés. En effet, jusqu'à ce jour, il subsiste un gap dans les études réalisées à ce sujet. Nous abordons en perspectives de ce travail (cf. section 5.5) l'orientation des solutions possibles pour répondre à ce problème.

Intégration de données guidée par des ressources sémantiques

Sommaire

5.1	Introduction	99
5.1.1	Approches d'intégration de données issues d'une base de données	100
5.1.2	Outils d'intégration de données issues d'une base de données	100
5.1.3	Objectif	101
5.2	Méthodologie proposée : l'approche DIKG2	101
5.2.1	Collecte de données	102
5.2.2	Extraction et structuration des données	103
5.2.3	Enrichissement des données et liaison des entités	104
5.2.4	Enrichissement du PHKG pour le profilage et l'inférence de recommandations	106
5.3	Interrogation des ressources sémantiques	106
5.4	Implémentation de l'approche DIKG2.....	109
5.5	Conclusion.....	109

L'objet de ce dernier chapitre est de répondre au problème d'intégration sémantique des données. Pour palier le manque d'outils accessibles et facilement exploitables pour une structure simplifiée de base de données, une approche a été proposée et s'appuie sur la mise en correspondance des données collectées avec les éléments de l'ontologie et l'évolution du graphe RDF suivant l'évolution du profil de l'utilisateur tout en gérant les conflits liés à la gestion et à la duplication des données. Nous présenterons pour cela notre approche d'intégration sémantique des données (DIKG2) à partir de laquelle nous exploitons les connaissances représentées dans notre ontologie de domaine pour concevoir un PHKG représentant le profil de santé de la personne âgée. A travers cette approche, nous présentons un exemple d'intégration dynamique de données stockées dans une BDR et issues de formulaires Web.

Texte de référence

- Dandan, R., Despres, S. : "DIKG2 : A Semantic Data Integration Approach for Knowledge Graphs Generation from Web Forms", 34th International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems, IEA/AIE 2021, Lecture Notes in Computer Science 12798, Springer, p. 255-260, Kuala Lumpur, Malaysia, 2021.

- Ghor, T. A. ; Agrawal, E. ; Alam, M. ; Alqawasmeh, O. ; D’amato, C. ; Annane, A. ; Azzam, A. ; Berezovskyi, A. ; Biswas, R. ; Bonduel, M. ; Brabant, Q. ; Bucur, C.-i. ; Camossi, E. ; Carriero, V. A. ; Chari, S. ; Fraga, D. C. ; Ciroku, F. ; Cochez, M. ; Curien, H. ; Cutrona, V. ; **Dandan, R.** ; Dess, D. ; Carlo, V. D. ; Djebri, A. E. A. ; Erp, M. V. ; Falakh, F. M. ; Izquierdo, A. F. ; Futia, G. ; Gangemi, A. ; Gasperoni, S. ; Grall, A. ; Heling, L. ; Henri, P. ; Herradi, N. ; Issa, S. ; Jozashoori, S. ; Juniarta, N. ; Kaffee, L.-a. ; Keles, I. ; Khare, P. ; Kovtun, V. ; Leone, V. ; Li, S. ; Lieber, S. ; Lisena, P. ; Makhalova, T. ; Marinucci, L. ; Minier, T. ; Moreau, B. ; Loustaunau, A. M. ; Nandini, D. ; Ozdowska, S. ; Moura, A. P. D. ; Padhee, S. ; Palma, G. ; Jimnez, P. D. P. ; Presutti, V. ; Reda, R. ; Rizza, E. ; Rosales-mndez, H. ; Rudolph, S. ; Sack, H. ; Sciuillo, L. ; Simanjuntak, H. ; Stomeo, C. ; Thanapalasingam, T. ; Tietz, T. ; Varanka, D. ; Vidal, M.-e. ; Wolowyk, M. ; Zocholl, M. : **Linked Open Data Validity – A Technical Report from ISWS 2018**, In : CoRR abs/1903.1255, **2019**.

5.1 Introduction

Au cours des deux dernières décennies, l’intégration des données s’appuyant sur la sémantique est devenue l’un des principaux challenges liés aux données. L’utilisation des standards du Web sémantique et des données liées, notamment les graphes de connaissances, fournit un support pour répondre aux problèmes liés à l’accès et à l’intégration des données.

Les ontologies se trouvent au coeur de l’intégration des données sémantiques. Les systèmes d’intégration de données fondées sur l’ontologie (Ontology Based-Data Integration, OBDI) rassemblent trois éléments : une ontologie, un ensemble de sources de données, la correspondance entre les deux. Cette correspondance a pour rôle de réconcilier la structure entre les deux niveaux et d’interroger les données en utilisant les éléments de l’ontologie comme prédicats. La correspondance n’est pas présente dans ce paramètre simplifié souvent appelé : réponse aux requêtes fondée sur l’ontologie (Ontology-Based Query Answering, OBQA). Les sources de données sont des référentiels où sont stockées les données concernant le domaine. Ces sources peuvent être nombreuses, hétérogènes et chacune gérée et entretenues indépendamment les unes des autres. Cette distinction permet la spécification précise entre les données contenues dans les sources de données et les éléments de l’ontologie. Nous abordons les différentes sources de données dans la section 2.3. Jusqu’à présent, un certain nombre de travaux ont montré l’intérêt de s’appuyer sur les ontologies pour la définition et la modélisation du domaine étudié et les graphes de connaissances pour l’intégration de données lors de la conception d’un système de recommandations (ESTEVA, 2010, KADADI et al., 2015, COLLARANA et al., 2017, CALVANESE et al., 2018).

L’une des principales difficultés rencontrées dans l’intégration de données est l’accès transparent aux formes complexes de données stockées dans diverses sources. La représentation sémantique des données représentant les connaissances du domaine et appuyée par une ontologie facilite l’exploitation et l’interrogation des données.

Dans le cadre de notre travail, nous cherchons à intégrer les données collectées issues du *Serious Game* ou de formulaires Web. Pour cela, il est indispensable d’avoir un outil ou une approche pour réaliser cette étape.

5.1.1 Approches d'intégration de données issues d'une base de données

Actuellement, une vaste quantité de données sont stockées dans des BDR. Une multitude d'approches et de techniques ont été proposées ces dernières années pour l'alignement (*mapping*) des données stockées dans des BDR en RDF. Pour permettre l'accessibilité des données hébergées dans des BDR au Web sémantique, des approches ont été proposées ces dernières décennies. Le but étant de convertir des données relationnelles en RDF ou d'exposer des données relationnelles pour les rendre interrogeables, ce processus est nommé (*RDB-to-RDF*). Le W3C publie en 2021 la première recommandation R2RML, un langage standard permettant de décrire les alignements entre une base de données relationnelle et une représentation RDF équivalente. Cette recommandation encourage les développeurs d'outils *RDB-to-RDF* à se conformer à un langage de *mapping* standard. Cependant, les choix effectués dans la spécification R2RML impliquent certaines limitations sur les types d'alignements qui peuvent être exprimés. De plus, un langage d'alignement indépendant de l'implémentation tel que R2RML ne répond pas à certaines des questions courantes qui se posent lors de la traduction de données relationnelles existantes en RDF, telles que le choix de réutiliser les vocabulaires existants ou la manière dont les données sont accessibles ou interrogeables. De nombreuses approches et outils "*RDB-to-RDF*" ont été proposés ces dernières années et présentent de nombreuses similarités (MICHEL et al., 2014). Cependant, la majorité d'entre eux sont principalement appliqués à un domaine de recherche et dépendent fortement de la structure de la base de données. De ce fait, l'implémentation de ces approches "*RDB-to-RDF*" n'est pas facilement applicable ou adaptable à l'objectif final recherché.

5.1.2 Outils d'intégration de données issues d'une base de données

Le groupe de travail BDR2RDF du W3C a proposé, il y a quelques années, deux recommandations standard : "Direct Mapping" et "R2RML mapping language". Depuis, Direct Mapping est devenu l'approche dominante dans la majorité des études. Cette approche utilise des règles simples pour convertir des données issues de BDR en format RDF. Cependant, la génération d'instances à partir de cette approche ne permet pas de prendre en compte la sémantique des données. Parce que le schéma conceptuel des données n'est pas considéré, l'utilisation de cette approche pour la création d'un graphe de connaissances complique la tâche du processus d'alignement des différents schémas de la BDR et de l'ontologie. Certaines études sont apparues ces derniers temps pour adapter les approches s'appuyant sur la méthode BDR2RDF. Par exemple, OLDAKOWSKI, 2011 proposent un système d'accès aux bases de données relationnelles sous forme virtuelle et la génération de graphe RDF (D2RQ Platform) et ZHAO et al., 2019 proposent une approche (R2ML) facilitant l'implémentation de SPARQL Endpoint dans BDR. Cependant, la diversité des outils et les langages d'alignements associés complique la tâche des fournisseurs de données souhaitant publier leurs données dans un format lisible par les machines. Ceci rend également leur exploitation encore plus complexe.

5.1.3 Objectif

Dans ce chapitre, nous proposons l’approche DIKG2 (Data Integration for Knowledge Graphs Generation) permettant la structuration de données semi-structurées provenant de formulaires Web et stockés dans une RDB. L’intégration dynamique de données dans des PHKG permet pour lier les utilisateurs à leurs données. L’application de cette approche s’appuie sur l’ontologie ONAFE présentée dans la section 3 et représentant les connaissances nécessaires à l’élaboration d’une recommandation personnalisée d’activité selon le profil de la personne âgée. En exploitant un vocabulaire commun entre les formulaires Web et les entités de notre ontologie de domaine, ONAFE, nous décrivons le processus d’alignement réalisé entre le schéma de la base de données et le schéma conceptuel de l’ontologie.

L’objectif de ce chapitre est de proposer une approche permettant la structuration de données issues de formulaires Web et facilitant l’intégration dynamique de données dans des PHKG en s’appuyant sur les connaissances décrites et modélisées dans ONAFE (cf. section 3). A cette étape, notre seul but est d’effectuer des ajouts, des modifications, des consultations et des suppressions. Les BDR assurent pour cela l’optimisation du traitement et de la sauvegarde des données. En exploitant un vocabulaire commun entre les formulaires Web et les entités de ONAFE, nous décrivons dans cette approche l’alignement réalisé entre le schéma de la base de données et le schéma conceptuel de ONAFE. Le PHKG construit et enrichi avec les connaissances de ONAFE permet la construction du profil de l’utilisateur à partir duquel la recommandation est personnalisée.

5.2 Méthodologie proposée : l’approche DIKG2

L’intégration sémantique de données nécessite la structuration des données collectées et stockées dans la BDR. Cette étape est essentielle pour faciliter l’intégration de données au sein de PHKG.

Les différentes étapes menant à la construction du PHKG sont représentées dans la Figure 5.1 sont les suivantes :

- collecte des données (étape 1) ;
- stockage des données (étape 2) ;
- extraction des données (étape 3) ;
- mise en correspondance des vocabulaires (étape 4) ;
- liaison des entités (étape 5) ;
- alignement entre schéma (étape 6) ;
- inférence des entités (étape 7).

Chacune de ces étapes sont détaillée dans les sections qui suivent.

Les différentes étapes de l’approche DIKG2 (étape 3 à 7) consistent à réaliser un alignement entre le schéma de données BDR et le schéma conceptuel de ONAFE. La structuration des données présentée dans cette approche dépend principalement de la structure des questionnaires.

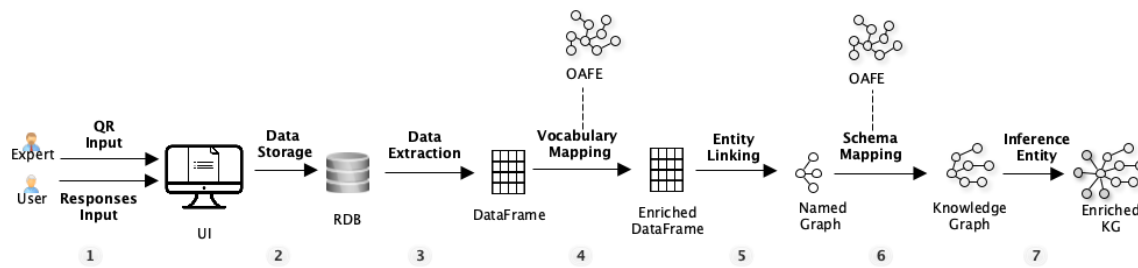


FIGURE 5.1 – Présentation du processus d’intégration sémantique de données menant à la génération du PHKG

Pour faciliter l’alignement entre le vocabulaire des questionnaires avec celui de ONAFE, les données ont été structurées sous forme de triplet RDF (Figure 5.2), dans lequel :

- le sujet représente l’identifiant de l’utilisateur ;
- le prédicat représente la question posée à l’utilisateur ;
- l’objet représente réponse à la question posée.

La construction d’un triplet RDF, à partir des questions et des réponses, consistera donc à associer une réponse à son utilisateur.

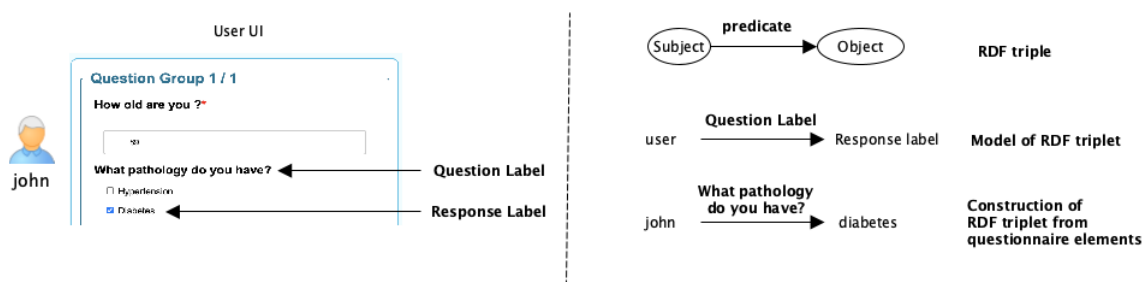


FIGURE 5.2 – Génération de triplets RDF à partir de la structure du formulaire Web

A partir des questionnaires sélectionnés par l’expert du domaine et présentés à l’utilisateur par l’intermédiaire de formulaires Web, les réponses de l’utilisateur sont stockées dans une BDR.

La Figure 5.3 montre le processus suivi conduisant à la génération d’un PHKG à partir des données textuelles du questionnaire. Le modèle du graphe est construit à partir des éléments du questionnaire et l’exploitation de ONAFE est nécessaire à l’enrichissement du graphe construit. Les sections suivantes détaillent chaque étape de la Figure 5.1.

5.2.1 Collecte de données

L’étape de collecte de données, présentée dans le chapitre 4), est un composant fondamental de tout pipeline de construction de graphe de connaissances (Figure 5.1

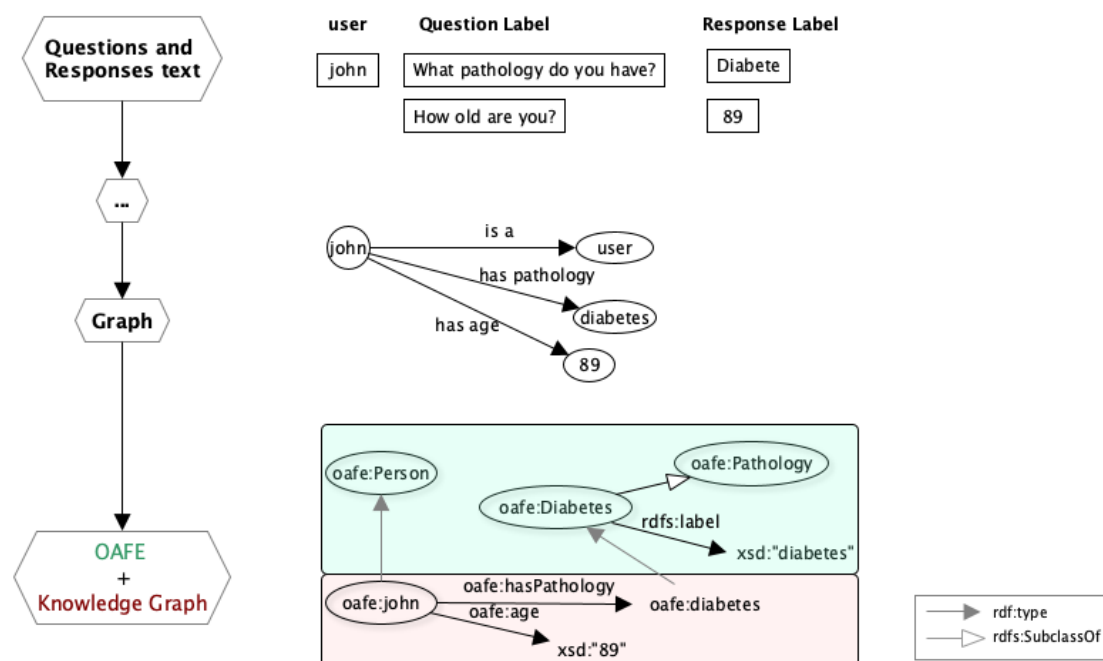


FIGURE 5.3 – Résumé des étapes menant à la construction du PHKG à partir des questions et des réponses issues du questionnaire

(étape 1)). Elle permet d’acquérir un maximum d’informations sur l’utilisateur qui sont essentielles à la construction de son profil et à l’élaboration d’une recommandation personnalisée. Elle consiste à intégrer des questionnaires (officiels et officieux) à partir d’une interface Web conçue à cet effet et à collecter, *via* une interface Web, les réponses des utilisateurs.

5.2.2 Extraction et structuration des données

Les données collectées à l’issue de l’étape 1 sont stockées dans BDR (Figure 5.4). L’utilisateur aura ainsi la possibilité de modifier, de mettre à jour et de supprimer à tout moment ses données. Notre approche sépare le stockage des données du traitement nécessaire à leur intégration. De ce fait, la modification des données collectées n’affectera pas le processus de traitement. Pour s’abstraire des contraintes liées à la structure de la BDR, les données sont extraites puis structurées dans des DataFrames (tables). Les données collectées sont ensuite nettoyées puis alignées dans un unique DataFrame (DataFrame 1).

Un DataFrame est une structure de données bidimensionnelle, les données sont alignées de façon tabulaire en lignes et en colonnes. L’intérêt d’un DataFrame est qu’il offre une structure à une (sous forme de séries) ou deux dimensions (sous forme d’objets) et facilite la manipulation des données notamment grâce aux index. Lorsque les éléments du DataFrame sont des séries d’objets (entités), il devient une structure à 3 dimensions. Nous employons ainsi ce terme pour maintenir cette notion tridimensionnelle des tables manipulées.

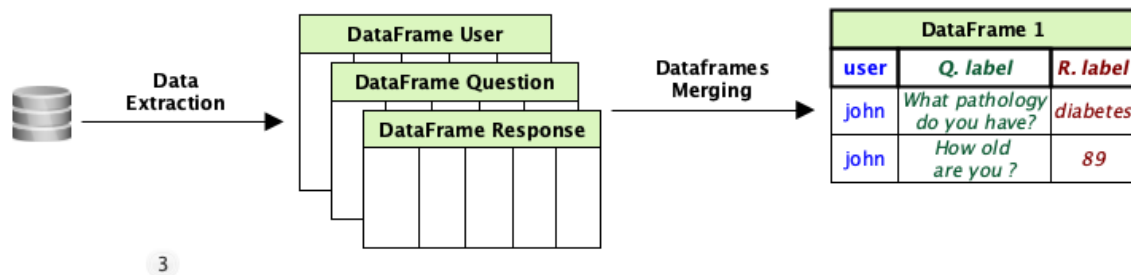


FIGURE 5.4 – Etapes d’intégration de données issues de formulaires Web au sein d’un BDR puis traitement (nettoyage, fusion) et structuration de données du formulaire Web dans le DataFrame 1

5.2.3 Enrichissement des données et liaison des entités

La génération du PHKG nécessite de s’appuyer sur le vocabulaire de ONAFE pour enrichir les données collectées. Pour cela, les éléments d’intérêt de ONAFE (classes, relations, labels) ont été extraits et structurés dans des DataFrames (Figure 5.5). Cette étape permet de récupérer les URI et les labels associés à chaque entité de l’ontologie. Les DataFrame "Class" et "OP/DP" contiennent respectivement toutes les classes et les "Object Properties" et les "Data Properties" (OP/DP) de ONAFE ainsi que leurs labels correspondants.

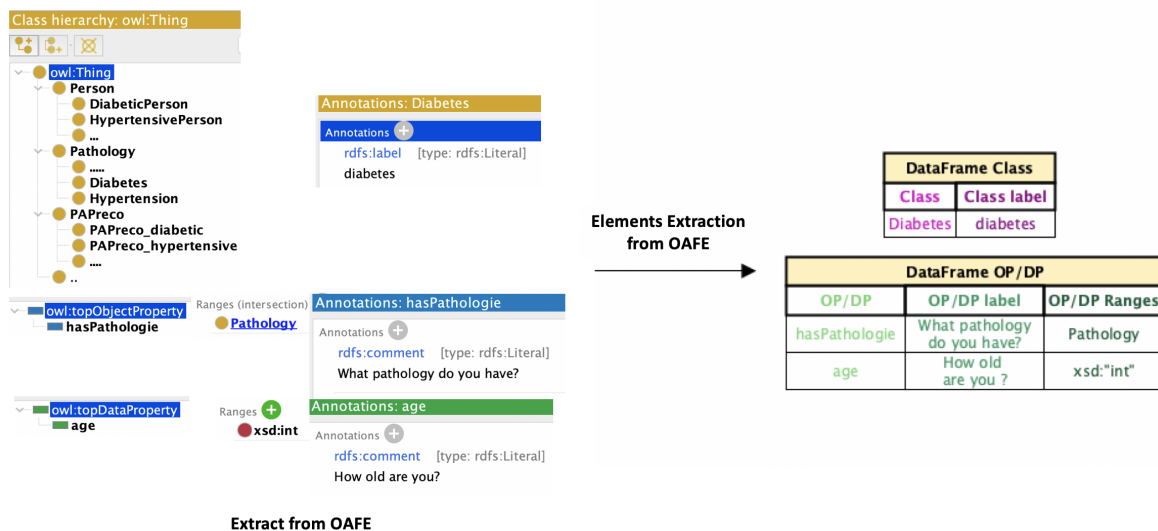


FIGURE 5.5 – Extraction des éléments d’intérêt de ONAFE dans des DataFrames

Pour permettre l’intégration et la modélisation des données dans un graphe, une série de traitements tabulaires et d’alignements entre le vocabulaire des questionnaires et celui de ONAFE ont été réalisés sur le DataFrame 1 (Figure 5.6). Les étapes 4, 5 et 6 de la Figure 5.1 menant à la génération du PHKG sont détaillées dans la (Figure 5.6).

L’étape 4 permet de construire un DataFrame enrichi par le vocabulaire de ONAFE.

Dans un premier temps (étape 4), deux alignements sont réalisés :

- entre le label des questions et le label des OP et DP (étape 4a) ;
- et entre le label des réponses et le label des Classes (étape 4b).

Par exemple, à la suite de cet alignement, la question "Quelle pathologie avez-vous ?" a été associée au label de l'OP "hasPathologie" et la question « Quel âge avez-vous ? » a été associée au label de la DP "Age".

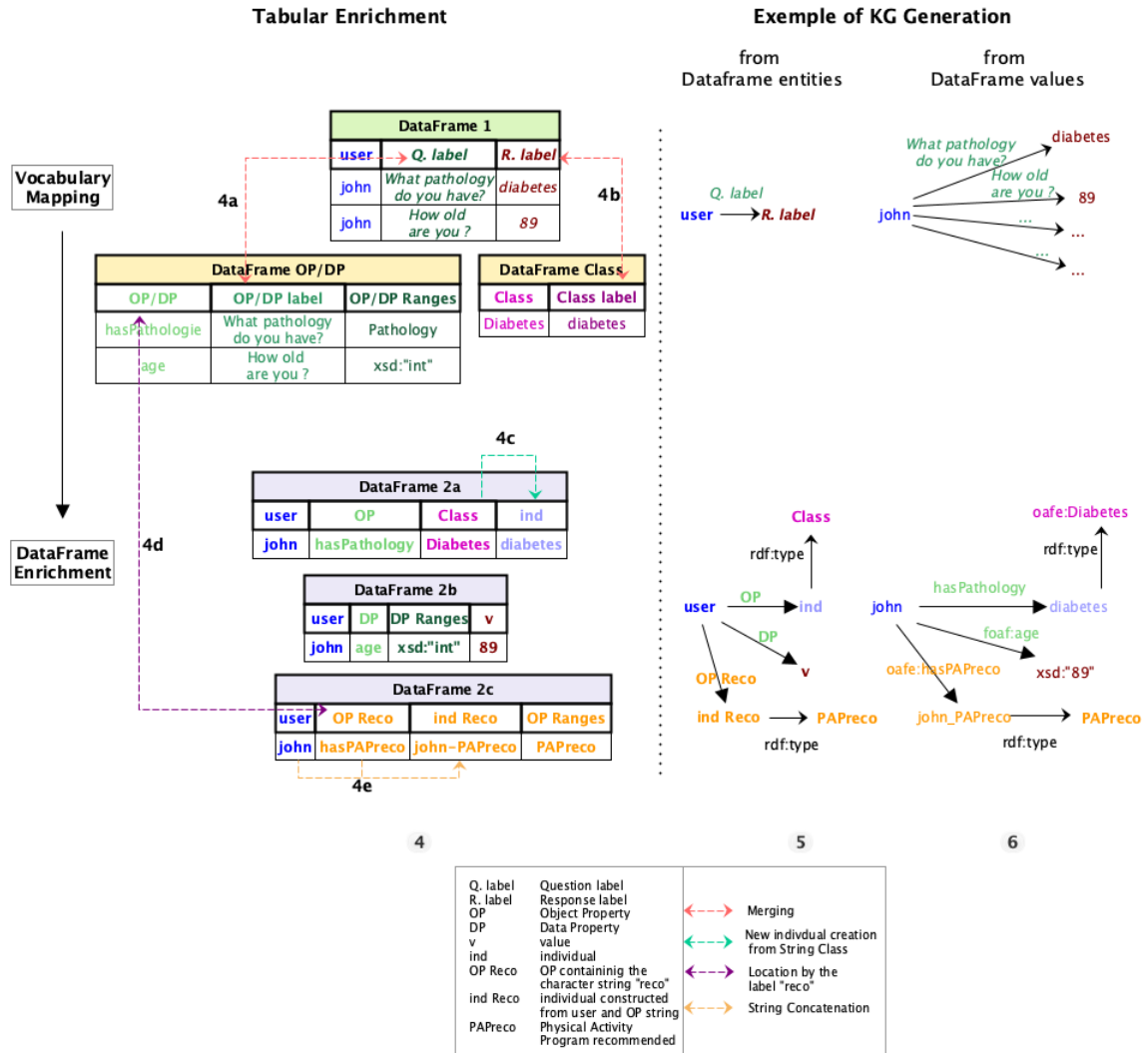


FIGURE 5.6 – Processus d’enrichissement des données avec le vocabulaire de ONAFE (étape 4), et de liaison des entités du DataFrame 2 entre elles (étape 5) menant à la génération du PHKG (étape 6)

Dans la Figure 5.6, les DataFrames 2a, 2b et 2c représentent les DataFrames enrichis par le vocabulaire ONAFE et rassemblent toutes les entités permettant la création des triplets entre l'utilisateur et ses données.

Par la suite, nous avons cherché à relier un utilisateur à une classe par une OP/DP exprimée en OWL2. Pour pouvoir le faire, il est nécessaire de passer par la création

d'un individu appartenant à la classe faisant le lien avec l'individu. Ainsi, un nouvel individu (ind) a été créé à partir de chaque classe du DataFrame 2a (étape 4c). Pour reprendre l'exemple cité, lorsque le label de la question correspond à celui du label de l'OP (par exemple « Quelle pathologie avez-vous ? »), la réponse (ex : diabetes) est utilisée pour créer un nouvel individu nommé "diabete" (DataFrame 2a). Cette étape permet de créer le triplet : "john hasPathology diabetes".

Pour relier un utilisateur à sa recommandation d'activité, les OP (OP Reco) contenant la chaîne de caractère "reco" (étape 4d) ont été rassemblées dans un même DataFrame (DataFrame 2c) et un nouvel individu a été créé à partir des chaînes de caractères (user et OPreco). Cette étape permet de créer le triplet : "john hasPAPreco john-papreco".

A la suite de ce processus d'enrichissement des données, les différentes entités des DataFrames 2a, 2b et 2c sont utilisées à l'étape 5 pour lier les données à leurs utilisateurs (étape 6). Les données structurées dans ces DataFrames sont ensuite utilisées à l'étape 6 pour la génération d'un premier PHKG. La section suivante présente l'évolution de ce PHKG représentant une structure de graphe facilitant son alignement avec le schéma de ONAFE.

5.2.4 Enrichissement du PHKG pour le profilage et l'inférence de recommandations

L'intégration des données et l'évolution du PHKG nécessitent l'exploitation des axiomes de ONAFE. L'objectif consiste à extraire un PHKG à partir de l'ontologie pour y intégrer des données. A partir de l'approche présentée, un schéma de données commun reliant l'utilisateur à ses données et à ses recommandations a pu être défini. La Figure 5.7 montre les relations créées entre le schéma de données et le schéma de l'ontologie ONAFE.

Les noeuds blancs représentés permettent de schématiser le modèle commun de données. L'instanciation des individus et la création des assertions de propriétés au niveau de la A-Box (flèches bleues) permettent de peupler ONAFE. De ce fait, le raisonnement sur cette nouvelle structure permet d'inférer de nouvelles relations (flèches rouges) entre les instances créées et les classes de ONAFE. Ces inférences permettent d'associer à la personne, la recommandation de programme d'activité physique selon son profil déduit.

5.3 Interrogation des ressources sémantiques

Les PHKG construits sont interrogés à l'aide de langages de requêtage (SPARQL, etc.). Pour illustrer l'attribution d'une recommandation à un profil santé d'un individu, nous présentons quelques exemples de construction de requêtes SPARQL pour l'interrogation des ressources sémantiques construites, soit : les ontologies et le PHKG. La première étape consiste à construire des patrons de graphes représentant un profil particulier. En fonction de l'ontologie, si elle contient ou pas la recommandation, la construction de ces recommandations peut se faire de différentes manières, soit : - en construisant un patron de graphe permettant d'insérer la recommandation dans

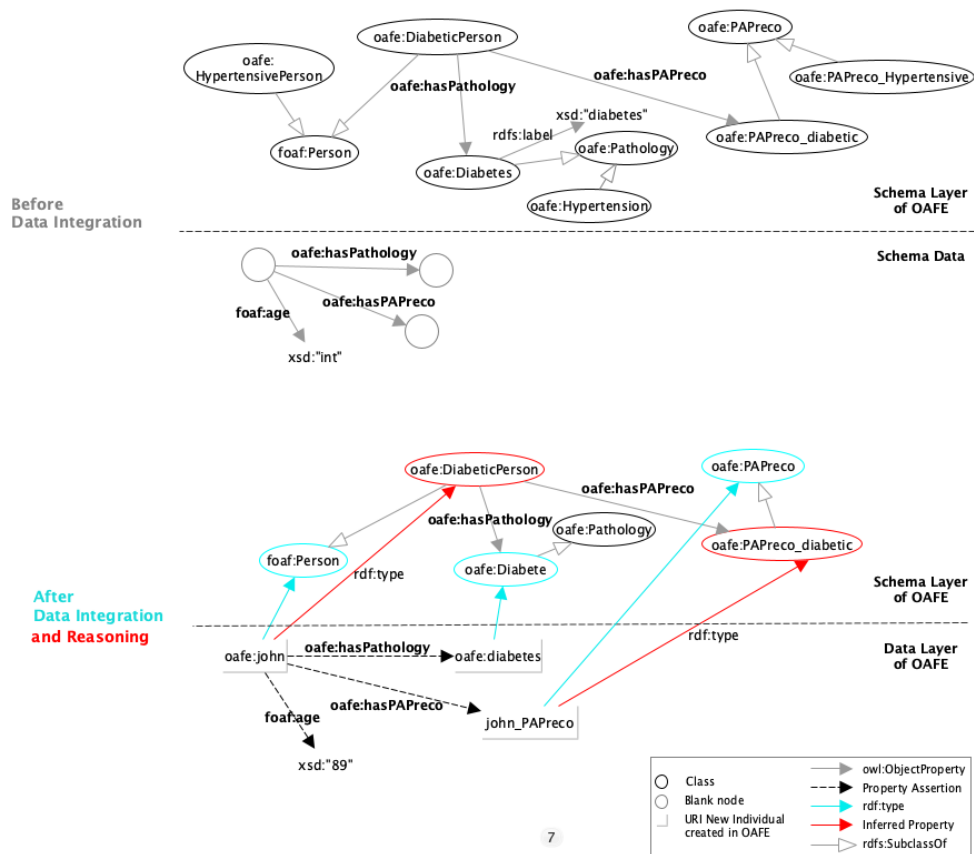


FIGURE 5.7 – Exemple de raisonnement sur les données pour générer une recommandation

l'ontologie (Figure 5.8) - en réalisant un appariement de graphe pour récupérer (Figure 5.9);

Lorsque la recommandation n'a pas été représentée dans l'ontologie, des règles d'inférence peuvent venir compléter les recommandations incluses dans l'ontologie. Les profils utilisateurs sont ensuite ciblés pour leur attribuer une recommandation spécifique selon les recommandations officielles (Figures 5.9 et Figure 5.10).

Chaque ontologie inférée incluant le PHKG de l'utilisateur est ensuite interrogée pour récupérer et lister tous les utilisateurs et les labels des recommandations inférées (IAR) selon leur profil (Figure 5.10). Par exemple, la recommandation pour les personnes sédentaires est d'augmenter le niveau d'activité. Ainsi, la recommandation consiste à suggérer des activités à intensité faible avec une durée recommandée de 30 minutes.

```

1 PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
2 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
3 PREFIX owl: <http://www.w3.org/2002/07/owl#>
4 PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
5 PREFIX ONAFE: <http://www.semanticweb.org/rahmadandan/ONAFE#>
6
7
8 # Insertion de la recommandation pour personne sédentaire dans le PHKG
9 INSERT {
10 ?personne ONAFE:aRecoDureeActivite "30"^^xsd:int .
11 ?y rdf:type ONAFE:ActiviteEndurance .
12 ?z rdf:type ONAFE:IntensiteFaible .
13 }
14 WHERE {
15 ?personne rdf:type ONAFE:PersonneSedentaire ;
16 ONAFE:aPourRecoActivite ?y ;
17 ONAFE:aPourIntensiteRecommandee ?z .
18 ?y rdf:type owl:NamedIndividual , ONAFE:Activite .
19 ?z rdf:type owl:NamedIndividual , ONAFE:Intensite .
20 }

```

FIGURE 5.8 – Exemple de requêtes SPARQL - Ajout de recommandation

```

1 PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
2 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
3 PREFIX owl: <http://www.w3.org/2002/07/owl#>
4 PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
5 PREFIX ONAFE: <http://www.semanticweb.org/rahmadandan/ONAFE#>
6
7
8 # Appariement de graphes RDF
9 SELECT DISTINCT ?x ?IAR
10 WHERE {
11 #description des patrons de graphes à trouver
12 ?x rdf:type ONAFE:Personne .
13 ?x ONAFE:aPourProgrammeRecommande ?pr .
14 ?pr rdf:type ONAFE:ProgrammeActivite .
15 ?pr ONAFE:aPourActiviteRecommande ?ar .
16 ?ar rdf:type ONAFE:Activite .
17 #description des individus
18 ?pr rdf:type ?y .
19 ?ar rdf:type ?z .
20 #récupération des labels
21 ?y rdfs:label ?IAR .
22 ?z rdfs:label ?IAR
23 }

```

FIGURE 5.9 – Exemple de requêtes SPARQL - Appariement de graphes pour la récupération des labels des activités recommandées

```

1 PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
2 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
3 PREFIX owl: <http://www.w3.org/2002/07/owl#>
4 PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
5 PREFIX ONAFE: <http://www.semanticweb.org/rahmadandan/ONAFE#>
6
7 #2/ construction à partir de l'ontologie ONAFE inférée
8 #faisant la correspondance entre l'utilisateur et son profil
9
10 CONSTRUCT {
11   ?personne ONAFE:aPourRecoActivite ONAFE:ActiviteEndurance .
12   ?personne ONAFE:aPourIntensiteActivite ONAFE:ActiviteIntensiteFaible .
13   ?personne ONAFE:aRecoDureeActivite "30"^^xsd:int
14 }
15 WHERE {
16   ?personne rdf:type ONAFE:PersonneSedentaire
17 }

```

FIGURE 5.10 – Exemple de requêtes SPARQL - Construction de la recommandation à partir des recommandations insérées dans ONAFE

5.4 Implémentation de l'approche DIKG2

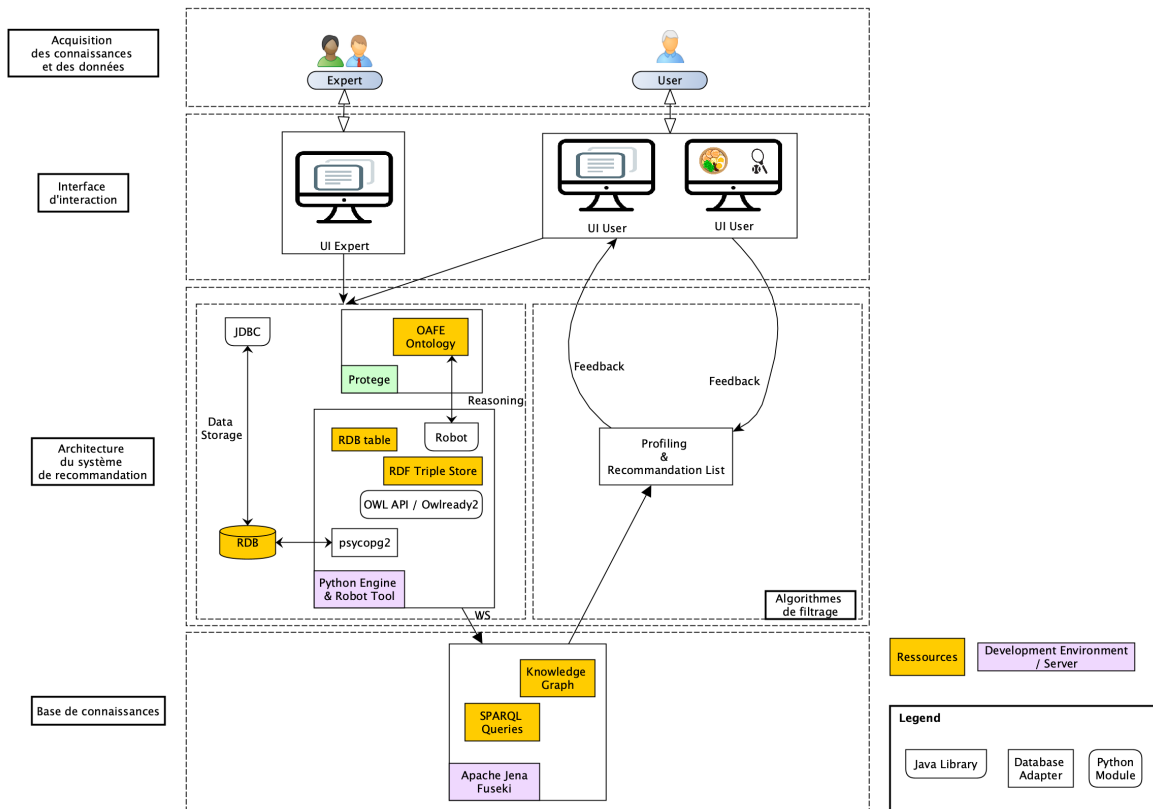
L'approche présentée a été implémentée en utilisant un certain nombre d'outils, notamment l'outil OWL API (HORRIDGE et BECHHOFER, 2011) pour la manipulation des ontologies, l'instanciation des individus, la création des assertions et l'inférence de nouvelles connaissances et l'outil Robot d'OBO (JACKSON et al., 2019) et Owlready2 (LAMY, 2017) qui s'appuie sur les principes d'OWL API HORRIDGE et BECHHOFER, 2011 et pour l'instanciation des individus, la création des assertions, et l'inférence de nouvelles connaissances. Le traitement des données a été réalisé à l'aide de bibliothèques Python. La bibliothèque Pandas a permis la manipulation et la structuration et l'analyse des données. Les opérations de manipulation de tables numériques et de séries temporelles nous ont permis d'aligner les données avec les entités de l'ontologie. Ainsi, en transformant les DataFrames bidimensionnels en tridimensionnels incluant des séries de données, les données ont pu être rattachées aux URI des entités représentées dans l'ontologie.

Le serveur Fuseki d'Apache Jena, un *framework* Java gratuit et *open source* pour la création d'applications de Web sémantique et de données liées, nous a permis l'importation des différentes ressources sémantiques construites (ontologies, PHKG) sous forme de jeu de données (dataset) et de graphe interrogeable en SPARQL.

Cette implémentation nous a permis de tester notre approche et de nous assurer de la consistance logique de ONAFE après l'intégration des données. À partir du jeu de données construit, l'exploitation du schéma de ONAFE a permis d'inférer à l'utilisateur son profil et par conséquent de lui attribuer les recommandations personnalisées qui lui correspondent. Dans la figure 5.11 nous présentons l'architecture du *framework* permettant l'articulation des différentes ressources et outils utilisés pour permettre l'intégration sémantique des données et construire le profil de l'utilisateur.

5.5 Conclusion

Dans ce chapitre, nous avons abordé le problème d'intégration de données par la création de PHKG s'appuyant sur les connaissances représentées dans notre ontologie de domaine pour la personnalisation des recommandations d'activité physique. La création


 FIGURE 5.11 – Architecture du *framework* pour implémenter l'approche DIKG2

d'un PHKG à partir d'informations sur le Web, même structurées nécessite l'exploitation de nombreux outils. Pour faciliter cette tâche, nous proposons une approche permettant de répondre au problème d'intégration sémantique de données au sein de PHKG prenant en compte la vision de OBDA. A partir de l'approche présentée, nous montrons l'intérêt de définir un schéma commun de données pour faciliter l'alignement entre le schéma de données et celui d'une ontologie de domaine. Par conséquent, l'intégration et l'enrichissement sémantique des données deviennent plus faciles à aborder.

L'approche présentée s'appuie sur l'application de règles simples adaptées pour des données issues de formulaires Web et stockées dans une BDR. Notre approche présente des similarités avec les principes du standard BDR2RDF. Contrairement à la majorité des études réalisées pour construire un profil utilisateur, notre BDR modélise la structure d'un questionnaire et non pas celui d'un profil utilisateur. Par ailleurs, le processus d'intégration sémantique des données est indépendant de la structure de la BDR, ce qui rend l'approche généralisable pour tout système utilisant des formulaires Web pour la collecte de données. Ainsi, l'ajout de nouveaux questionnaires et l'évolution du profil de l'utilisateur n'affecteront pas le traitement et l'intégration des données dans le PHKG. Cependant, les données collectées peuvent être obsolètes, imprécises ou tout simplement fausses. Il donc est nécessaire de tenir compte de la validité des données lors de la manipulation et l'intégration des données (GHOR et al., 2019).

Généralement, les sources de données qui alimentent les systèmes peuvent être multiples. Dans ce chapitre, nous avons uniquement traité des données provenant des

formulaires Web et stockées dans une BDR. Cependant, quelle que soit la source de données, la structure de la base de données est telle que chaque donnée pourra être associée à son utilisateur suivant le même schéma de la base de données. De plus, le fait que le traitement nécessaire à l'intégration de données soit séparé de celui de la gestion des données facilite son adoption, quelle que soit la source de données. Notre approche peut facilement évoluer pour tenir en compte des différentes structures de données. De ce fait, le même raisonnement présenté pourra s'appliquer de la même manière en s'appuyant sur les données issues d'objets connectés et stockées par exemple dans des fichiers CSV.

L'intégration des données peut également être soutenue par l'exploitation d'un vocabulaire commun entre celui utilisé pour collecter les données et celui représenté dans l'ontologie. Actuellement, l'annotation des questionnaires dans l'ontologie est manuelle. Néanmoins, le processus d'annotation des questionnaires et d'évolution de l'ontologie selon ces annotations peut être automatisable, notamment en utilisant le langage RDFa. Ainsi, l'expert du domaine aura la possibilité d'annoter les questionnaires au fur et à mesure de leurs ajouts dans les questionnaires en ligne.

Notre approche de création d'un profil santé par l'intermédiaire des PHKG présente des avantages quant à la gestion du coût et de la complexité lié au stockage des profils généralement observés dans la majorité des approches de personnalisation. En créant un PHKG à chaque utilisateur, les problèmes de complexité rencontrés lors du peuplement d'ontologies sont évités.

L'approche présentée sépare le stockage des données du traitement nécessaire à l'intégration de données et à la construction du profil. Ainsi, la modification des données collectées n'affectera pas le processus de traitement nécessaire à la génération du PHKG. Pour faciliter l'intégration et l'exploitation des données, il est donc primordial de garder à l'esprit les problèmes d'interopérabilité pour l'intégration et l'exploitation des données. En effet, lorsque les sources d'informations n'ont pas été conçues pour prendre en compte les problèmes liés à l'interopérabilité des données, le problème d'intégration de données devient un défi de taille.

Conclusion et perspectives

Conclusion

L'objectif global de la thèse est de proposer des solutions et des outils pouvant s'intégrer au sein d'un système de recommandation sémantique appliqué à la santé dans le but d'améliorer la personnalisation des recommandations et par ailleurs la prise en charge des personnes âgées. À travers cette thèse, nous abordons la problématique d'intégration sémantique de données. Les solutions proposées consistent à exploiter les ontologies et les PHKG pour établir un raisonnement entre les connaissances représentées dans l'ontologie et les données collectées.

Les données exploitées dans ce travail s'appuient sur le vocabulaire de l'ontologie conçue et appliquent les principes FAIR "I". Ces principes déclarent que les données et les métadonnées doivent être décrites dans un langage formel, accessible, partagé et largement applicable pour la représentation des connaissances. L'utilisation d'ontologies renforce l'application des principes FAIR "I" pouvant être considérés comme les plus difficiles à accomplir et restent les caractéristiques clés les plus importantes pour FAIR. Dans les systèmes appliqués à la santé, la construction d'un profil utilisateur est centrale pour la personnalisation des recommandations. L'approche d'intégration sémantique de données présentée favorise également l'application des principes FAIR depuis l'étape de collecte des données, pour faciliter l'accès aux données, jusqu'à l'étape de création du PHKG rendant les données en santé interopérables pour une meilleure prise en charge des problèmes de santé.

Parce que les besoins de tout un chacun varient et évoluent, les solutions proposées pour prévenir la dénutrition et la sédentarité doivent s'adapter à l'évolution des nouvelles technologies pour répondre aux problèmes spécifiques de chaque individu. Les plateformes de recommandation futur doivent apporter des solutions pour consommer des produits qui plaisent tout en les variant.

L'intégration des solutions proposées dans la présente thèse permet de soutenir le développement d'applications innovantes telles que des applications de coach personnalisés numérisés qui peuvent tenir les patients informés et aider à gérer leur maladie chronique, et permettre aux médecins de prendre des décisions efficaces sur les problèmes de santé ou de recevoir des alertes en temps opportun selon les besoins grâce à une surveillance continue.

Un système de recommandation bien conçu se doit de suggérer du contenu dans toutes les circonstances. La complexité des systèmes de recommandation ne réside pas

dans la quantité d'algorithmes qu'ils contiennent mais dans leur capacité à exploiter les données des utilisateurs et à être intégré au sein d'un écosystème cohérent. L'utilisation de ce système de recommandation pourra être intégré dans la vision future de conception de maison intelligente (*Smart Home*) ou dans les écosystèmes intégrant le programme d'assistance à l'autonomie à domicile (AAL). En effet, le développement des systèmes AAL, s'inscrivant dans les systèmes à intelligence ambiante, est devenu le centre de nombreux projets de recherche ces dernières années et qui ne cesse d'évoluer. Un système d'intelligence ambiante est composé d'un ensemble d'appareils communicants disséminés dans l'environnement. L'une des initiatives intéressantes de ces systèmes consiste à repousser les limites technologiques pour aider les personnes âgées à vivre de façon autonome plus longtemps chez elles. L'évolution de ces technologies ont bouleversé les approches classiques de recommandation. L'exploitation des connaissances du domaine pour une meilleure personnalisation devient alors une nécessité pour répondre aux besoins des utilisateurs et trouver d'autres approches pour répondre aux problèmes d'interopérabilité des données.

Aujourd'hui, les recommandations disponibles sont générales et traitent uniquement des cas les plus fréquents de sorte à correspondre à un large public. Dans le futur, l'amélioration des recommandations pour une meilleure description des cas particuliers à prendre en charge et la participation des experts du domaine dans la construction des recommandations intégrées dans les ontologies conçues permettront d'aboutir à des recommandations personnalisées. Une telle personnalisation trouvera tout son sens dans le cas de recommandations pour des personnes polysymptomatiques ou polypathologiques. La prévention de la sédentarité par les personnels pourra passer par la prescription d'activité physique sur mesure qui suit une posologie, un type, une durée, une fréquence, une intensité. Cette plateforme pourra constituer le « *check-up* du Bien vieillir » .

Bien que les objets connectés soient un moyen efficace pour collecter des données en temps réel, en continue et sans que la personne âgée ne soit obligée de fournir une charge cognitive importante, le manque d'interopérabilité de leur système présente un réel frein quant à leur exploitation dans notre projet. Depuis 2010, l'IoT fait partie de la nouvelle révolution du numérique en permettant l'interaction du monde physique avec celui du numérique. Il permet de répondre aux besoins de recueil de données. Pour alimenter la prolifération des IoT, le principal défi concerne leur interopérabilité. Le manque d'interopérabilité vient de l'hétérogénéité des OS, des périphériques, des protocoles réseau, des applications et des données fournies, transmises et exploitées par ces objets connectés. Passer par un développement de services permet de répondre en partie à cette interopérabilité et de répondre aux besoins du système futur. Les ontologies représentent également une solution efficace pour gérer l'intégration des technologies et des ressources utilisées. Depuis quelques années, la communauté sémantique a développé un certain nombre d'ontologies (IoT-Lite du W3C, etc.) pour décrire les concepts et les relations entre eux dans de nombreux domaines d'application notamment celui de l'IoT. Cependant, la normalisation, l'uniformité et la facilitation de la conception des systèmes sont devenues une nécessité pour satisfaire les besoins croissants et évolutifs des utilisateurs.

Perspectives

Le principal challenge dans le déploiement d'un tel environnement complexe et qu'ils soient capables de gérer des services, de traiter, de raisonner et de décider de manière autonome sur ces sources de données et de connaissances distribuées mais aussi de coopérer entre eux de manière efficace et flexible. Il est donc devenu nécessaire d'apporter des solutions pour répondre aux problèmes d'interopérabilité des données et des systèmes.

Dans la continuité des approches de conception d'interfaces (de collecte ou de présentation de la recommandation) personnalisables, la solution proposée consiste à créer des services Web capables de collecter des données à caractère hétérogène et distribué provenant de sources variées (objets connectés, smartphone, DMP, etc.). Cette approche est avantageuse pour le développement de système sémantique de recommandation et d'interfaces du *Serious Game* personnalisable car elle fournit un cadre permettant de séparer logiquement le moteur de requête (*back-end*) de la conception d'interfaces (*front-end*). Les services Web offrent plusieurs avantages. Les services Web constituent la force motrice derrière les architectures orientées services (SOA). En effet, ils fournissent la couche logicielle supplémentaire nécessaire pour coupler de manière flexible un service à un agent. En plus d'être un moyen de communication entre différents environnements, ils constituent un moyen rapide de distribution de l'information entre clients, fournisseurs, partenaires commerciaux et leurs différentes plates-formes.

Les avantages fournis par les services Web renforcent le besoin d'héberger le moteur de requêtes côté serveur, contrairement aux implémentations actuelles qui contraignent le client à effectuer la majorité des calculs pour fournir la réponse aux différentes requêtes. L'intégration est sans doute le facteur essentiel qui favorise l'utilisation des services Web notamment pour faire le lien entre les applications qui n'ont pas forcément la même technologie pour le transfert et la réception de données. Ainsi, subdiviser un tel projet en plusieurs services Web permet d'isoler le fonctionnement de processus complexes qui peuvent à leur tour être interrogés par des sources différentes de façon indépendante. Pour cela, il suffit de respecter les structures d'échanges. Les services Web offrent également des avantages en terme de sécurité. En effet, les différents services Web peuvent être placés sur plusieurs serveurs ainsi la charge et le dimensionnement des serveurs peuvent être fait proportionnellement à leur réel utilisation. Enfin, la mise à jour d'une brique du projet peut être indépendante du reste du projet. Le seul point à respecter est de garder la même structure de communication entrante et sortante.

Dans l'optique de mettre en place un écosystème modulaire à partir de services Web pour répondre aux problèmes liés à l'interopérabilité des systèmes, une telle solution permet d'avoir une méthode générique d'accès aux fonctionnalités fournies par les agents. L'analyse contextuelle et la satisfaction des demandes, ainsi que l'interaction dynamique avec d'autres agents montrent l'intérêt d'utiliser des applications distribuées au sein d'environnements complexes. Une telle vision permettra à n'importe quel fournisseur possédant des données à intégrer - quelles que soient leurs sources (OS), leurs formats ou le langage de développement utilisé pour leur cheminement ou leur structuration - aura la possibilité de se greffer à cet écosystème. Cela permet aux développeurs d'avoir une flexibilité quant au choix des langages et des outils pour accomplir un

projet. De ce fait, la communication, l'échange et l'intégration de données en plus de la génération de requêtes personnalisées selon l'évolution des recommandations et des besoins nutritionnels des personnes âgées sera ainsi facilité. La conception d'une telle plateforme de services numériques pour la prise en charge personnalisée quel que soit l'âge de la personne, représentera une application innovante centrée sur l'utilisateur et sur laquelle n'importe quel fournisseur numérique pourra s'y greffer de sorte à gérer efficacement l'intégration des capteurs et des données collectées.

Bibliographie

- ABDALLA, A. ; HU, Y. ; CARRAL, D. ; LI, N. ; JANOWICZ, K. (jan. 2014) :** « An ontology design pattern for activity reasoning ». In : *CEUR Workshop Proceedings* 1302 (cité p. 68-71).
- ADAMS, E. (2014) :** *Fundamentals of Game Design*. New Riders (cité p. 88).
- ADOMAVICIUS, G. ; TUZHILIN, A. (2005) :** « Personalization technologies : a process-oriented perspective ». In : *Communications of the ACM* 48, p. 83-90 (cité p. 28).
- AE, P. ; LA, S. ; DJ, T. ; AS, A. ; RE, G. (jan. 2011) :** « Starting the conversation performance of a brief dietary assessment and intervention tool for health professionals ». en. In : *American journal of preventive medicine* 40.1. Publisher : Am J Prev Med (cité p. 46).
- AINSWORTH, B. E. ; HASKELL, W. L. ; HERRMANN, S. D. ; MECKES, N. ; BASSETT, D. R. ; TUDOR-LOCKE, C. ; GREER, J. L. ; VEZINA, J. ; WHITT-GLOVER, M. C. ; LEON, A. S. (août 2011) :** « 2011 Compendium of Physical Activities : A Second Update of Codes and MET Values ». en. In : *Medicine & Science in Sports & Exercise* 43.8, p. 1575-1581 (cité p. 73).
- ALFAIFI, Y. ; GRASSO, F. ; TAMMA, V. (2017) :** « Towards an Ontology to Identify Barriers to Physical Activity for Type 2 Diabetes ». In : *Proceedings of the 2017 International Conference on Digital Health*. DH '17. New York, NY, USA : ACM, p. 16-20 (cité p. 66).
- ALMEIDA, F. ; SIMOES, J. (avr. 2019) :** « The Role of Serious Games, Gamification and Industry 4.0 Tools in the Education 4.0 Paradigm ». In : *Contemporary Educational Technology* 10.2, p. 120-136 (cité p. 85).
- ALOTAIBI, Y. K. ; FEDERICO, F. (déc. 2017) :** « The impact of health information technology on patient safety ». In : *Saudi Medical Journal* 38.12, p. 1173-1180 (cité p. 42).
- ALVAREZ, J. ; DJAOUTI, D. ; RAMPNOUX, O. (nov. 2012) :** *Introduction au Serious Game - Serious Game an Introduction (French and English)* (cité p. 86).
- AMDOUNI, E. ; JONQUET, C. (avr. 2021) :** « FAIR or FAIRer? An integrated quantitative FAIRness assessment grid for semantic resources and ontologies » (cité p. 21).
- AMEEN, A. ; HEAD (mars 2019) :** « Knowledge based Recommendation System in Semantic Web-A Survey ». In : *International Journal of Computer Applications* 182 – No. 43, p. 975-8887 (cité p. 27).

- ARNDT, N. ; NAUMANN, P. ; RADTKE, N. ; MARTIN, M. ; MARX, E. (jan. 2019) : « Decentralized Collaborative Knowledge Management Using Git ». In : *Journal of Web Semantics. Managing the Evolution and Preservation of the Data Web* 54, p. 29-47 (cité p. 18).
- BALOG, K. ; KENTER, T. (sept. 2019) : *Personal Knowledge Graphs : A Research Agenda*, p. 220 (cité p. 22).
- BEDI, P. ; KAUR, H. ; MARWAHA, S. (2007) : « Trust Based Recommender System for Semantic Web ». In : *IJCAI* (cité p. 27).
- BELMIN, J. ; CHASSAGNE, P. ; FRIOCOURT, P. ; GONTHIER, R. ; JEANDEL, C. ; NOURHASHEMI, F. ; PFITZENMEYER, P. (sept. 2016) : *Gériatrie : pour le Praticien*. Google-Books-ID : iuJ1DwAAQBAJ. Elsevier Health Sciences (cité p. 35, 75).
- BERGAMASCHI, S. ; GUERRA, F. ; ORSINI, M. ; SARTORI, C. ; VINCINI, M. (août 2011) : « A semantic approach to ETL technologies ». In : *Data & Knowledge Engineering* 70.8, p. 717-731 (cité p. 21).
- BERNERS-LEE, T. ; HENDLER, J. ; LASSILA, O. (mai 2001) : « The Semantic Web : A New Form of Web Content That is Meaningful to Computers Will Unleash a Revolution of New Possibilities ». In : *ScientificAmerican.com* (cité p. 10).
- BIZER, C. ; HEATH, T. ; BERNERS-LEE, T. (juill. 2009) : « Linked Data : The Story so Far ». In : *International Journal on Semantic Web and Information Systems* 5, p. 1-22 (cité p. 10).
- BLOCKER, K. A. ; WRIGHT, T. J. ; BOOT, W. R. (2014) : « Gaming Preferences of Aging Generations ». In : *Gerontechnology : International Journal on the Fundamental Aspects of Technology to Serve the Ageing Society* 12.3, p. 174-184 (cité p. 89).
- BORG, G. (1998) : *Borg's Perceived Exertion and Pain Scales*. Champaign, IL : Human Kinetics (cité p. 74).
- BREWSTER, C. ; NOUWT, B. ; RAAIJMAKERS, S. ; VERHOOSSEL, J. (nov. 2019) : « Ontology-based Access Control for FAIR Data ». In : *Data Intelligence* 2, p. 66-77 (cité p. 20).
- BURKE, R. (mai 2000) : « Knowledge-Based Recommender Systems ». In : *Encyclopedia of library and information systems* 69 (cité p. 27).
- BURKE, R. ; FELFERNIG, A. ; GÖKER, M. H. (juin 2011) : « Recommender Systems : An Overview ». en. In : *AI Magazine* 32.3, p. 13-18 (cité p. 26).
- CALBIMONTE, J.-P. ; DUBOSSON, F. ; HILFIKER, R. ; COTTING, A. ; SCHUMACHER, M. (oct. 2017) : *The MedRed Ontology for Representing Clinical Data Acquisition Metadata*, p. 47 (cité p. 18).
- CALVANESE, D. ; DE GIACOMO, G. ; LEMBO, D. ; LENZERINI, M. ; ROSATI, R. (2018) : « Ontology-Based Data Access and Integration ». In : *Encyclopedia of Database Systems, Second Edition*. Sous la dir. de L. LIU ; M. T. ÖZSU. Springer (cité p. 29, 99).
- CANTADOR, I. ; BELLOGÍN, A. ; CASTELLS, P. (2008) : « A multilayer ontology-based hybrid recommendation model ». In : *AI Communications* 21.2-3, p. 203-210 (cité p. 30).

- CARPENTIER, K. (jan. 2015)** : « Scénarisation personnalisée dynamique dans les environnements virtuels pour la formation ». These de doctorat. Compiègne (*cité p. 88*).
- CHEN, L. ; NUGENT, C. (nov. 2009)** : « Ontology-based Activity Recognition in Intelligent Pervasive Environments ». en. In : *International Journal of Web Information Systems* 5.4, p. 410-430 (*cité p. 68-70*).
- CHESHAM, A. ; WYSS, P. ; MÜRI, R. M. ; MOSIMANN, U. P. ; NEF, T. (avr. 2017)** : « What Older People Like to Play : Genre Preferences and Acceptance of Casual Games ». In : *JMIR serious games* 5.2, e8 (*cité p. 89*).
- CIV, CENTRE D'INFORMATION DES VIANDES (2011)** : *Livret CIV - Bien dans mon assiette, bien dans mes baskets*. Rapp. tech. (*cité p. 94*).
- CLELAND, C. ; FERGUSON, S. ; ELLIS, G. ; HUNTER, R. F. (déc. 2018)** : « Validity of the International Physical Activity Questionnaire (IPAQ) for assessing moderate-to-vigorous physical activity and sedentary behaviour of older adults in the United Kingdom ». In : *BMC Medical Research Methodology* 18.1, p. 176 (*cité p. 46*).
- COLLARANA, D. ; GALKIN, M. ; TRAVERSO-RIBON, I. ; LANGE, C. ; VIDAL, M.-E. ; AUER, S. (2017)** : « Semantic Data Integration for Knowledge Graph Construction at Query Time ». In : *2017 IEEE 11th International Conference on Semantic Computing (ICSC)*. IEEE (*cité p. 99*).
- CONNOLLY, T. ; BOYLE, E. ; MACARTHUR, E. ; HAINEY, T. ; BOYLE, J. (sept. 2012)** : « A systemic literature review of empirical evidence on computer games and serious games ». In : *Computers & Education* 59, p. 661-686 (*cité p. 86*).
- D'AQUIN, M. ; SCHLICHT, A. ; STUCKENSCHMIDT, H. ; SABOU, M. (2009)** : « Criteria and Evaluation for Ontology Modularization Techniques ». In : *Modular Ontologies : Concepts, Theories and Techniques for Knowledge Modularization*. Lecture Notes in Computer Science. Springer, p. 67-89 (*cité p. 51*).
- DANDAN, R. ; DESPRES, S. ; NOBÉCOURT, J. (2018)** : « OAFE : an Ontology for the Description of Elderly Activities ». In : *14th International Conference on Signal-Image Technology & Internet-Based Systems, SITIS 2018, Las Palmas de Gran Canaria, Spain, November 26-29, 2018*. Sous la dir. de G. S. di BAJA ; L. GALLO ; K. YÉTONGNON ; A. DIPANDA ; M. C. SANTANA ; R. CHBEIR. IEEE, p. 396-403 (*cité p. 72*).
- DE LEENHEER, P. ; MENS, T. (jan. 2008)** : « Ontology Evolution ». In : p. 131-176 (*cité p. 54*).
- DE MOOR, A. ; DE LEENHEER, P. ; MEERSMAN, R. (2006)** : « DOGMA-MESS : A Meaning Evolution Support System for Interorganizational Ontology Engineering ». In : *Conceptual Structures : Inspiration and Application*. Sous la dir. de H. SCHÄRFE ; P. HITZLER ; P. ØHRSTRØM. Lecture Notes in Computer Science. Berlin, Heidelberg : Springer, p. 189-202 (*cité p. 17*).
- DEBRUYNE, C. ; TRAN, T.-K. ; MEERSMAN, R. (juin 2013)** : « Grounding Ontologies with Social Processes and Natural Language ». In : *Journal on Data Semantics* 2.2, p. 89-118 (*cité p. 17*).
- DENNIS, M. ; DEEMTER, K. van ; DELL'AGLIO, D. ; PAN, J. Z. (2017)** : « Computing Authoring Tests from Competency Questions : Experimental Validation ».

- English. In : *The Semantic Web – ISWC 2017 - 16th International Semantic Web Conference, Proceedings*. Springer (cité p. 61).
- DESPRES, S. (2016)** : « Construction d'une ontologie modulaire. Application au domaine de la cuisine numérique ». Français. In : *Revue d'Intelligence Artificielle* 30.5, p. 509-532 (cité p. 51, 63).
- DESPRES, S. ; GUEZENNEC, G. (2017)** : « Flat OWL Editor : un outil utilisant des feuilles de calcul pour découpler les tâches des acteurs impliqués dans la gestion d'une ontologie ». Français. In : *TOTH 2017* (cité p. 80).
- DJAOUTI, D. (nov. 2011)** : « Serious Game Design : considérations théoriques et techniques sur la création de jeux vidéo à vocation utilitaire ». Thèse de doct. Université de Toulouse, Université Toulouse III - Paul Sabatier (cité p. 86, 88).
- DREVET, S. ; GAVAZZI, G. (oct. 2019)** : « Dénutrition du sujet âgé ». In : *La Revue de Médecine Interne* 40.10, p. 664-669 (cité p. 2, 35).
- ELALLIOUI, Y. ; EL BEQQALI, O. (mars 2012)** : « User Profile Ontology for the Personalization Approach ». In : *International Journal of Computer Applications* 41, p. 31-40 (cité p. 30).
- ESTEVA, J. (2010)** : *Ontology For Data Integration* (cité p. 99).
- ETTORRE, A. ; ROCHA RODRÍGUEZ, O. ; FARON, C. ; MICHEL, F. ; GANDON, F. (2020)** : « A Knowledge Graph Enhanced Learner Model to Predict Outcomes to Questions in the Medical Field ». In : *Knowledge Engineering and Knowledge Management*. Sous la dir. de C. M. KEET ; M. DUMONTIER. Lecture Notes in Computer Science. Cham : Springer International Publishing, p. 237-251 (cité p. 22).
- FARID, M. ; ELGOHARY, R. ; MOAWAD, I. ; ROUSHDY, M. (oct. 2018)** : *User Profiling Approaches, Modeling, and Personalization*. SSRN Scholarly Paper ID 3389811. Rochester, NY : Social Science Research Network (cité p. 29).
- FERNÁNDEZ-LÓPEZ, M. ; GOMEZ-PEREZ, A. ; JURISTO, N. (mars 1997)** : « METHONTOLOGY : from ontological art towards ontological engineering ». In : *Engineering Workshop on Ontological Engineering (AAAI97)* (cité p. 18).
- FERRY, M. (2010)** : « Nutrition, vieillissement et santé ». In : *Gérontologie et société* 33 / n° 134.3, p. 123 (cité p. 36).
- FILOS, D. ; TRIANTAFYLIDIS, A. ; CHOUVARDA, I. ; BUYS, R. ; CORNELISSEN, V. ; BUDTS, W. ; WALSH, D. ; WOODS, C. ; MORAN, K. ; MAGLAVERAS, N. (2016)** : « PATHway : Decision Support in Exercise Programmes for Cardiac Rehabilitation ». eng. In : *Studies in Health Technology and Informatics* 224, p. 40-45 (cité p. 68).
- FLOURIS, G. ; MANAKANATAS, D. ; KONDYLAKIS, H. ; PLEXOUSAKIS, D. ; ANTONIOU, G. (juin 2008)** : « Ontology Change : Classification and Survey ». In : *The Knowledge Engineering Review* 23, p. 117-152 (cité p. 54).
- FLOURIS, G. ; PLEXOUSAKIS, D. ; ANTONIOU, G. (2006)** : « Evolving Ontology Evolution ». In : *SOFSEM 2006 : Theory and Practice of Computer Science*. Sous la dir. de D. HUTCHISON ; T. KANADE ; J. KITTLER ; J. M. KLEINBERG ; F. MATTERN ; J. C. MITCHELL ; M. NAOR ; O. NIERSTRASZ ; C. PANDU RANGAN ; B. STEFFEN ; M. SUDAN ; D. TERZOPOULOS ; D. TYGAR ; M. Y. VARDI ; G. WEIKUM ; J. WIEDERMANN ; G. TEL ; J. POKORNÝ ; M. BIELIKOVÁ ; J. ŠTULLER. T. 3831. Berlin, Heidelberg : Springer Berlin Heidelberg, p. 14-29 (cité p. 52).
-

- GAUCH, S. ; SPERETTA, M. ; CHANDRAMOULI, A. ; MICARELLI, A. (jan. 2007) : *User Profiles for Personalized Information Access*. T. 4321, p. 89 (cité p. 29).
- GAYATHRI, K. S. ; EASWARAKUMAR, K. S. ; ELIAS, S. (avr. 2017) : « Probabilistic Ontology Based Activity Recognition in Smart Homes Using Markov Logic Network ». In : *Knowledge-Based Systems* 121, p. 173-184 (cité p. 67).
- GENTILE, A. ; GRUHL, D. ; RISTOSKI, P. ; WELCH, S. (oct. 2019) : « Personalized Knowledge Graphs for the Pharmaceutical Domain ». In : p. 400-417 (cité p. 22).
- GEZER, C. ; TASKIN, E. (mai 2016) : *An overview of oneM2M standard*, p. 1708 (cité p. 45).
- GHASSEMI, M. ; TRIANDAFILOU, K. ; BARRY, A. ; STOYKOV, M. E. ; ROTH, E. ; MUSSA-IVALDI, F. A. ; KAMPER, D. G. ; RANGANATHAN, R. (fév. 2019) : « Development of an EMG-Controlled Serious Game for Rehabilitation ». In : *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 27.2, p. 283-292 (cité p. 85).
- GHOR, T. A. ; AGRAWAL, E. ; ALAM, M. ; ALQAWASMEH, O. ; D'AMATO, C. ; ANNANE, A. ; AZZAM, A. ; BEREZOVSKIY, A. ; BISWAS, R. ; BONDUDEL, M. ; BRABANT, Q. ; BUCUR, C.-i. ; CAMOSSO, E. ; CARRIERO, V. A. ; CHARI, S. ; FRAGA, D. C. ; CIROKU, F. ; COCHEZ, M. ; CURIEN, H. ; CUTRONA, V. ; DANDAN, R. ; DESS, D. ; CARLO, V. D. ; DJEBRI, A. E. A. ; ERP, M. V. ; FALAKH, F. M. ; IZQUIERDO, A. F. ; FUTIA, G. ; GANGEMI, A. ; GASPERONI, S. ; GRALL, A. ; HELING, L. ; HENRI, P. ; HERRADI, N. ; ISSA, S. ; JOZASHOORI, S. ; JUNIARTA, N. ; KAFFEE, L.-a. ; KELES, I. ; KHARE, P. ; KOVTUN, V. ; LEONE, V. ; LI, S. ; LIEBER, S. ; LISENA, P. ; MAKHALOVA, T. ; MARINUCCI, L. ; MINIER, T. ; MOREAU, B. ; LOUSTAUNAU, A. M. ; NANDINI, D. ; OZDOWSKA, S. ; MOURA, A. P. D. ; PADHEE, S. ; PALMA, G. ; JIMNEZ, P. D. P. ; PRESUTTI, V. ; REDA, R. ; RIZZA, E. ; ROSALES-MNDEZ, H. ; RUDOLPH, S. ; SACK, H. ; SCIULLO, L. ; SIMANJUNTAK, H. ; STOMEIO, C. ; THANAPALASINGAM, T. ; TIETZ, T. ; VARANKA, D. ; VIDAL, M.-e. ; WOLOWYK, M. ; ZOCHOLL, M. (2019) : *Linked Open Data Validity – A Technical Report from ISWS 2018* (cité p. 110).
- GIUNCHIGLIA, F. ; DUTTA, B. ; MALTESE, V. ; FARAZI, F. (mai 2012) : « A Facet-Based Methodology for the Construction of a Large-Scale Geospatial Ontology ». In : *Journal on Data Semantics* 1 (cité p. 18).
- GOLDMAN, N. ; GLEI, D. A. ; CHANG, M.-C. (jan. 2004) : « The role of clinical risk factors in understanding self-rated health ». In : 14, p. 49-57 (cité p. 41).
- GOMEZ PEREZ, A. (1999) : *Développements récents en matière de conception, de maintenance et d'utilisation des ontologies*. Rapp. tech., p. 9-20 (cité p. 17).
- GOMEZ-PEREZ, A. ; SUÁREZ-FIGUEROA, M. C. (jan. 2009) : « NeOn Methodology for Building Ontology Networks : a Scenario-based Methodology ». In : (cité p. 17).
- GORBANEV, I. ; AGUDELO-LONDOÑO, S. ; GONZÁLEZ, R. A. ; CORTES, A. ; POMARES, A. ; DELGADILLO, V. ; YEPES, F. J. ; MUÑOZ, Ó. (jan. 2018) : « A systematic review of serious games in medical education : quality of evidence and pedagogical strategy ». In : *Medical Education Online* 23.1, p. 1438718 (cité p. 85).

- GORIA, S. (sept. 2014)** : « Stratégie de développement d'un serious game : entre processus de gamification et de disengagement ». In : *SG 2014, AIM Serious Games et innovation*. T. 2. Telecom Ecole de Management, p. 3 (cité p. 87).
- GOUVERNEMENT DES TERRITOIRES DU NORD-OUEST (2017)** : *Évaluation Des Habitudes Alimentaires* (cité p. 47).
- GRUBER, T. R. (juin 1993)** : « A translation approach to portable ontology specifications ». In : *Knowledge Acquisition* 5.2, p. 199-220 (cité p. 10).
- GRÜNINGER, M. ; FOX, M. S. (1995)** : « Methodology for the Design and Evaluation of Ontologies ». In : *Department of Industrial Engineering* (cité p. 61).
- GUYATT, G. H. ; JOAN EAGLE, D. ; SACKETT, B. ; WILLAN, A. ; GRIFFITH, L. ; McILROY, W. ; PATTERSON, C. J. ; TURPIE, I. (déc. 1993)** : « Measuring quality of life in the frail elderly ». In : *Journal of Clinical Epidemiology* 46.12, p. 1433-1444 (cité p. 46).
- GYRARD, A. ; ATEMEZING, G. A. ; SERRANO, M. (2021)** : « PerfectO : An Online Toolkit for Improving Quality, Accessibility, and Classification of Domain-based Ontologies ». In : *Semantic IoT : Theory and Applications Interoperability, Provenance and Beyond* (cité p. 22, 72).
- GYRARD, A. ; GAUR, M. ; SHEKARPOUR, S. ; THIRUNARAYAN, K. ; SHETH, A. (2018)** : « Personalized Health Knowledge Graph ». In : p. 7 (cité p. 22).
- HAASE, P. ; STOJANOVIC, L. (mai 2005)** : *Consistent Evolution of OWL Ontologies*. T. 3532, p. 197 (cité p. 52, 54).
- HAUSSMANN, S. ; SENEVIRATNE, O. ; CHEN, Y. ; NE'EMAN, Y. ; CODELLA, J. ; CHEN, C.-H. ; MCGUINNESS, D. ; ZAKI, M. (oct. 2019)** : « FoodKG : A Semantics-Driven Knowledge Graph for Food Recommendation ». In : p. 146-162 (cité p. 22).
- HIRSCH, B. ; ABERNETHY, A. (déc. 2012)** : « Structured Decision-Making : Using Personalized Medicine to Improve the Value of Cancer Care ». In : *Journal of Personalized Medicine* 3, p. 1-13 (cité p. 28).
- HODA, M. ; MONTAGHAMI, V. ; AL OSMAN, H. ; EL SADDIK, A. (2018)** : « ECOPPA : Extensible Context Ontology for Persuasive Physical-Activity Applications ». In : *Proceedings of the International Conference on Information Technology & Systems (ICITS 2018)*. Sous la dir. de Á. ROCHA ; T. GUARDA. T. 721. Cham : Springer International Publishing, p. 309-318 (cité p. 68-70).
- HOFER, P. ; NEURURER, S. ; HAUFFE, H. ; INSAM, T. ; ZEILNER, A. ; GOBEL, G. (2015)** : « Semi-Automated Evaluation of Biomedical Ontologies for the Biobanking Domain Based on Competency Questions ». In : *Studies in Health Technology and Informatics*, p. 65-72 (cité p. 61).
- HOGAN, A. ; BLOMQVIST, E. ; COCHEZ, M. ; D'AMATO, C. ; DE MELO, G. ; GUTIERREZ, C. ; GAYO, J. E. L. ; KIRrane, S. ; NEUMAIER, S. ; POLLERES, A. ; NAVIGLI, R. ; NGOMO, A.-C. N. ; RASHID, S. M. ; RULA, A. ; SCHMELZEISEN, L. ; SEQUEDA, J. ; STAAB, S. ; ZIMMERMANN, A. (juill. 2021)** : « Knowledge Graphs ». In : *ACM Computing Surveys* 54.4, p. 1-37 (cité p. 21).
- HONICKE, T. ; BROADBENT, J. (fév. 2016)** : « The influence of academic self-efficacy on academic performance : A systematic review ». In : *Educational Research Review* 17, p. 63-84 (cité p. 86).
-

- HORRIDGE, M. ; BECHHOFFER, S. (2011)** : « The OWL API : A Java API for OWL ontologies ». In : *Semantic Web 2.1*, p. 11-21 (*cit e p. 109*).
- INSEE (2020)** : *Tableaux de l' conomie fran aise -  dition 2020* (*cit e p. 1*).
- JACKSON, R. C. ; BALHOFF, J. P. ; DOUGLASS, E. ; HARRIS, N. L. ; MUNGALL, C. J. ; OVERTON, J. A. (juill. 2019)** : « ROBOT : A Tool for Automating Ontology Workflows ». In : *BMC Bioinformatics 20.1* (*cit e p. 109*).
- JACOBSON, I. ; CHRISTERSON, M. ; STAFF, A. P. ; JONSSON, P. ;  VERGAARD, G. (1992)** : *Object-oriented Software Engineering : A Use Case Driven Approach*. ACM Press (*cit e p. 61*).
- JANTKE, K. ; GAUDL, S. (jan. 2011)** : *Taxonomic contributions to digital games science*, p. 8 (*cit e p. 86*).
- JARRAR, M. (mai 2005)** : « Towards methodological principles for ontology engineering ». Th se de doct. Vrije Universiteit Brussel (*cit e p. 51*).
- KADADI, A. ; AGRAWAL, R. ; NYAMFUL, C. ; ATIQ, R. (jan. 2015)** : « Challenges of Data Integration and Interoperability in Big Data ». In : *Proceedings - 2014 IEEE International Conference on Big Data, IEEE Big Data 2014*, p. 38-40 (*cit e p. 99*).
- KAISER, M. J. ; BAUER, J. M. ; RAMSCH, C. ; UTER, W. ; GUIGOZ, Y. ; CEDERHOLM, T. ; THOMAS, D. R. ; ANTHONY, P. ; CHARLTON, K. E. ; MAGGIO, M. ; TSAI, A. C. ; GRATHWOHL, D. ; VELLAS, B. ; SIEBER, C. C. ; MNA-INTERNATIONAL GROUP (nov. 2009)** : « Validation of the Mini Nutritional Assessment short-form (MNA-SF) : a practical tool for identification of nutritional status ». eng. In : *The Journal of Nutrition, Health & Aging 13.9*, p. 782-788 (*cit e p. 46*).
- KARINKANTA, S. ; PIIRTOLA, M. ; SIEV NEN, H. ; UUSI-RASI, K. ; KANNUS, P. (juill. 2010)** : « Physical therapy approaches to reduce fall and fracture risk among older adults ». In : *Nature reviews. Endocrinology 6*, p. 396-407 (*cit e p. 39*).
- KHAN, Z. ; KEET, C. (jan. 2021)** : « Structuring Abstraction to Achieve Ontology Modularisation ». In : *Advanced Concepts, Methods, et Applications in Semantic Computing*, p. 72-92 (*cit e p. 51*).
- KHATTAK, A. ; BATOOL, R. ; PERVEZ, Z. ; KHAN, A. ; LEE, S. (sept. 2013)** : « Ontology Evolution and Challenges ». In : *Journal of Information Science and Engineering 29*, p. 851-871 (*cit e p. 54*).
- KHATTAK, A. M. ; LATIF, K. ; LEE, S. ; LEE, Y.-K. (2009)** : « Ontology Evolution : A Survey and Future Challenges ». en. In : *U- and E-Service, Science and Technology*. Sous la dir. de D.  LEZAK ; T.-h. KIM ; J. MA ; W.-C. FANG ; F. E. SANDNES ; B.-H. KANG ; B. GU. Communications in Computer and Information Science. Berlin, Heidelberg : Springer, p. 68-75 (*cit e p. 54*).
- KLEIN, M. ; A, D. ; NOY, N. (juin 2003)** : « A Component-Based Framework for Ontology Evolution ». In : (*cit e p. 53*).
- KOTIS, K. ; VOUIROS, G. A. (juill. 2006)** : « Human-centered ontology engineering : The HCOME methodology ». In : *Knowledge and Information Systems 10.1*, p. 109-131 (*cit e p. 17*).
- KOTIS, K. I. ; VOUIROS, G. A. ; SPILIOPOULOS, D. (2020)** : « Ontology engineering methodologies for the evolution of living and reused ontologies : status, trends, findings and recommendations ». In : *The Knowledge Engineering Review 35* (*cit e p. 17, 54*).

- LAËTITIA, L. G. (2020) : « Réponses rapides dans le cadre du COVID-19 - Pathologies chroniques et risques nutritionnels en ambulatoire ». In : HAS santé, p. 11 (*cité p. 2*).
- LAMY, J.-B. (juill. 2017) : « Owlready : Ontology-oriented programming in Python with automatic classification and high level constructs for biomedical ontologies ». In : *Artificial Intelligence in Medicine* 80, p. 11-28 (*cité p. 109*).
- LOCKWOOD, C. ; CONROY-HILLER, T. ; PAGE, T. (2004) : « Vital Signs ». en. In : *JBI Reports* 2.6, p. 207-230 (*cité p. 77*).
- MARFISI-SCHOTTMAN, I. ; GEORGE, S. ; TARPIN-BERNARD, F. (jan. 2010) : « Tools and Methods for Efficiently Designing Serious Games ». In : *4th European Conference on Games Based Learning 2010, ECGBL 2010* (*cité p. 86*).
- MARNE, B. ; HUYNH-KIM-BANG, B. ; LABAT, J.-m. (mai 2011) : *Articuler motivation et apprentissage grâce aux facettes du jeu sérieux* (*cité p. 87*).
- MEIJ, B. S. van der ; WIJNHOFEN, H. A. H. ; LEE, J. S. ; HOUSTON, D. K. ; HUE, T. ; HARRIS, T. B. ; KRITCHEVSKY, S. B. ; NEWMAN, A. B. ; VISSER, M. (2017) : « Poor Appetite and Dietary Intake in Community-Dwelling Older Adults ». In : *Journal of the American Geriatrics Society* 65.10, p. 2190-2197 (*cité p. 36, 37, 41*).
- MELVILLE, P. ; SINDHWANI, V. (2010) : « Recommender Systems ». en. In : *Encyclopedia of Machine Learning*. Sous la dir. de C. SAMMUT ; G. I. WEBB. Boston, MA : Springer US, p. 829-838 (*cité p. 26*).
- MICHEL, F. ; MONTAGNAT, J. ; FARON-ZUCKER, C. (mai 2014) : « A survey of RDB to RDF translation approaches and tools ». In : (*cité p. 100*).
- MIDDLETON, S. E. ; SHADBOLT, N. ; ROURE, D. D. (2004) : « Ontological user profiling in recommender systems ». In : *ACM Trans. Inf. Syst.* 22.1, p. 54-88 (*cité p. 30*).
- MIKKULAINEN, R. ; LIANG, J. ; MEYERSON, E. ; RAWAL, A. ; FINK, D. ; FRANCON, O. ; RAJU, B. ; SHAHRZAD, H. ; NAVRUZGAN, A. ; DUFFY, N. ; HODJAT, B. (jan. 2019) : « Chapter 15 - Evolving Deep Neural Networks ». In : *Artificial Intelligence in the Age of Neural Networks and Brain Computing*. Sous la dir. de R. KOZMA ; C. ALIPPI ; Y. CHOE ; F. C. MORABITO. Academic Press, p. 293-312 (*cité p. 86*).
- MINGIONI, M. (nov. 2016) : *L'alimentation des personnes âgées : étude sensorielle des fruits et légumes enrichis*. Angers (*cité p. 2*).
- MOBASHER, B. ; DAI, H. ; LUO, T. ; NAKAGAWA, M. (jan. 2002) : « Discovery and Evaluation of Aggregate Usage Profiles for Web Personalization ». In : *Data Mining and Knowledge Discovery* 6.1, p. 61-82 (*cité p. 28*).
- NARULA, G. S. ; YADAV, U. ; DUHAN, N. ; JAIN, V. (2018) : « Evolution of FOAF and SIOC in Semantic Web : A Survey ». In : *Big Data Analytics*. Sous la dir. de V. B. AGGARWAL ; V. BHATNAGAR ; D. K. MISHRA. Advances in Intelligent Systems and Computing. Singapore : Springer, p. 253-263 (*cité p. 17, 18*).
- NASSABI, M. H. ; AKKER, H. op den ; VOLLENBROEK-HUTTEN, M. (jan. 2014) : « An Ontology-Based Recommender System to Promote Physical Activity for Pre-Frail Elderly ». In : *Mensch & Computer 2014 - Workshopband* (*cité p. 68*).

- NATH, K. ; DHAR, S. ; BASISHTHA, S. (2014)** : « Web 1.0 to Web 3.0 - Evolution of the Web and its various challenges ». In : *2014 International Conference on Reliability Optimization and Information Technology (ICROIT)* (cité p. 27).
- AL-NAZER, A. ; HELMY, T. ; AL-MULHEM, M. (2014)** : « User's Profile Ontology-Based Semantic Framework for Personalized Food and Nutrition Recommendation ». en. In : *Procedia Computer Science* 32, p. 101-108 (cité p. 70).
- NI, Q. ; PAU DE LA CRUZ, I. ; GARCÍA HERNANDO, A. B. (jan. 2016)** : « A Foundational Ontology-Based Model for Human Activity Representation in Smart Homes ». In : *Journal of Ambient Intelligence and Smart Environments* 8.1. Sous la dir. de L. SHEN ; A. MUÑOZ ; T. ZHANG, p. 47-61 (cité p. 68, 70).
- NINOT, G. ; SOYEZ, F. ; FIOCCO, S. ; NASSIH, K. ; MORIN, A. ; PRÉFAUT, C. (mai 2010)** : « The VQ11, a short health-related quality of life questionnaire for routine practice in COPD patients ». In : *Revue des maladies respiratoires* 27, p. 472-81 (cité p. 46).
- NOWACK, B. (2009)** : *The Semantic Web - Not a piece of cake... - benjamin nowack's blog* (cité p. 5).
- NOY, N. F. ; KLEIN, M. (juill. 2004)** : « Ontology Evolution : Not the Same as Schema Evolution ». In : *Knowledge and Information Systems* 6.4, p. 428-440 (cité p. 53).
- OBEID, C. ; LAHOUD, I. ; KHOURY, H. E. ; CHAMPIN, P. (2018)** : « Ontology-based Recommender System in Higher Education ». In : *Companion of the The Web Conference 2018 on The Web Conference 2018, WWW 2018, Lyon , France, April 23-27, 2018*. Sous la dir. de P. CHAMPIN ; F. L. GANDON ; M. LALMAS ; P. G. IPEIROTIS. ACM, p. 1031-1034 (cité p. 30).
- OLDAKOWSKI, R. (fév. 2011)** : « D2RQ Platform – Treating Non-RDF Databases as Virtual RDF Graphs ». In : *Nature Precedings* (cité p. 100).
- ORGANIZATION, W. H. (2009)** : *Global health risks : mortality and burden of disease attributable to selected major risks* (cité p. 2).
- PAOLUCCI, M. ; KAWAMURA, T. ; PAYNE, T. (juin 2002)** : *Semantic Matching of Web Services Capabilities*. T. 2002, p. 347 (cité p. 10).
- PHAN, N. ; DOU, D. ; WANG, H. ; KIL, D. ; PINIEWSKI, B. (avr. 2017)** : « Ontology-Based Deep Learning for Human Behavior Prediction with Explanations in Health Social Networks ». en. In : *Information Sciences* 384, p. 298-313 (cité p. 70).
- POUYET, V. ; GIBOREAU, A. ; CUVELIER, G. ; BENATTAR, L. (nov. 2015)** : « Les préférences culinaires des personnes âgées vivant en institution : facteurs d'appréciation sensoriels et cognitifs ». In : *Cahiers de Nutrition et de Diététique* 50.5, p. 271-279 (cité p. 2).
- POVEDA-VILLALÓN, M. ; ESPINOZA-ARIAS, P. ; GARIJO, D. ; CORCHO, O. (jan. 2020)** : *Coming to Terms with FAIR Ontologies* (cité p. 20).
- PRAMONO, D. ; SETIAWAN, N. Y. ; SARNO, R. ; SIDIQ, M. (2013)** : « Physical Activity Recommendation for Diabetic Patients Based on Ontology ». In : *In 7th International Conference on Information & Communication Technology and Systems*, p. 27-32 (cité p. 68).
-

- PRATIWI, P. S.; XU, Y.; LI, Y.; TROST, S. G.; CLANCHY, K. M.; TJONDRONEGORO, D. (2018) : « User Profile Ontology to Support Personalization for E-Coaching Systems ». en. In : p. 7 (*cit e p. 30*).
- PRIMO, T.; SILVA, J. L. T.; RIBEIRO, A. M.; VICARI, R. M.; BOFF, E. (2012) : « Towards Ontological Profiles in Communities of Practice ». In : *IEEE Multidisciplinary Engineering Education Magazine* 7, p. 13-22 (*cit e p. 70*).
- PROV-O : *The PROV Ontology* (2013) (*cit e p. 71*).
- RAVYSE, W. S.; SEUGNET BIGNAUT, A.; LEENDERTZ, V.; WOOLNER, A. (mars 2017) : « Success factors for serious games to enhance learning : a systematic review ». In : *Virtual Reality* 21.1, p. 31-58 (*cit e p. 85*).
- REBELE, T.; SUCHANEK, F.; HOFFART, J.; BIEGA, J.; KUZHEY, E.; WEIKUM, G. (oct. 2016) : « YAGO : a Multilingual Knowledge Base from Wikipedia, Wordnet, and Geonames ». In : *The 15th International Semantic Web Conference*. Kobe, Japan (*cit e p. 18*).
- REGMI, P. R.; WAITHAKA, E.; PAUDYAL, A.; SIMKHADA, P.; VAN TEIJLINGEN, E. (d ec. 2016) : « Guide to the design and application of online questionnaire surveys ». In : *Nepal Journal of Epidemiology* 6.4, p. 640-644 (*cit e p. 46*).
- REN, Y.; PARVIZI, A.; MELLISH, C.; PAN, J. Z.; DEEMTER, K. van; STEVENS, R. (mai 2014) : « Towards Competency Question-Driven Ontology Authoring ». en. In : *The Semantic Web : Trends and Challenges*. Lecture Notes in Computer Science. Springer, Cham, p. 752-767 (*cit e p. 61*).
- RIBONI, D.; BETTINI, C. (juin 2011) : « OWL 2 Modeling and Reasoning with Complex Human Activities ». en. In : *Pervasive and Mobile Computing* 7.3, p. 379-395 (*cit e p. 68*).
- RICCI, F.; ROKACH, L.; SHAPIRA, B. (2015) : « Recommender Systems : Introduction and Challenges ». In : *Recommender Systems Handbook*. Sous la dir. de F. RICCI; L. ROKACH; B. SHAPIRA. Boston, MA : Springer US, p. 1-34 (*cit e p. 28*).
- RIECKEN, D. (ao ut 2000) : « Personalized Views of Personalization ». In : *Commun ACM* 43 (*cit e p. 25*).
- RIMITHA, S.; ABBURU, V.; KIRANMAI, A.; C, M.; CHANDRASEKARAN, K. (d ec. 2019) : « Improving Job Recommendation Using Ontological Modeling and User Profiles ». In : *2019 Fifteenth International Conference on Information Processing (ICINPRO)*. IEEE (*cit e p. 30*).
- SAFYAN, M.; QAYYUM, Z. U.; SARWAR, S.; IQBAL, M.; CASTRO, R. G.; AL-DULAIMI, A. (ao ut 2019) : « Ontology evolution for personalised and adaptive activity recognition ». In : *IET Wireless Sensor Systems* 9.4, p. 193-200 (*cit e p. 54*).
- SALATINO, A.; THANAPALASINGAM, T.; MANNOCCI, A.; OSBORNE, F.; MOTTA, E. (jan. 2018) : « The Computer Science Ontology : A Large-Scale Taxonomy of Research Areas ». In : p. 187-205 (*cit e p. 18*).
- SANT E (HAS), H. A. de (nov. 2007) : *Strat egie de prise en charge en cas de d enutrition prot eino- nerg etique chez la personne  g ee*, p. 92-96 (*cit e p. 37, 40*).
- (2020) : *Diagnostic de la d enutrition de la personne  g ee*. RECOMMANDER LES BONNES PRATIQUES, p. 8 (*cit e p. 2, 35*).
-

- SANTÉ (OMS), O. mondiale de la (2020)** : *Lignes directrices de l'OMS sur l'activité physique et la sédentarité : en un coup d'oeil*. Organisation mondiale de la Santé (cité p. 39).
- SANTÉ PUBLIQUE FRANCE, (2020)** : *Activité physique et sédentarité dans la population française. Situation en 2014-2016 et évolution depuis 2006-2007*. 15 (cité p. 2, 39, 40).
- SANTOS-EGGIMANN, B. ; CUÉNOUD, P. ; SPAGNOLI, J. ; JUNOD, J. (avr. 2009)** : « Prevalence of Frailty in Middle-Aged and Older Community-Dwelling Europeans Living in 10 Countries ». In : *The journals of gerontology. Series A, Biological sciences and medical sciences* 64, p. 675-81 (cité p. 35).
- SCHRIML, L. ; MITRAKA, E. ; MUNRO, J. ; TAUBER, B. ; SCHOR, M. ; NICKLE, L. ; FELIX, V. ; JENG, L. ; BEARER, C. ; LICHENSTEIN, R. ; BISORDI, K. ; CAMPION, N. ; HYMAN, B. ; KURLAND, D. ; OATES, C. ; KIBBEY, S. ; SREEKUMAR, P. ; LE, C. ; GIGLIO, M. ; GREENE, C. (nov. 2018)** : « Human Disease Ontology 2018 update : classification, content and workflow expansion ». In : *Nucleic acids research* 47 (cité p. 71).
- SILVEIRA, T. ; ZHANG, M. ; LIN, X. ; LIU, Y. ; MA, S. (2019)** : « How good your recommender system is? A survey on evaluations in recommendation ». In : *Int. J. Mach. Learn. Cybern.* tex.ids= silveira2019a (cité p. 30).
- SOLIDARITÉS ET DE LA SANTÉ, M. des (avr. 2019)** : *2019, priorité à l'alimentation, la nutrition, l'activité physique* (cité p. 2, 35).
- STOJANOVIC, L. ; MOTIK, B. (2002)** : « Ontology Evolution within Ontology Editors ». en. In : 13th International Conference on Knowledge Engineering and Knowledge Management EKAW 2002, p. 10 (cité p. 52).
- STOJANOVIC, L. (jan. 2004)** : « Methods and Tools for Ontology Evolution / ». In : (cité p. 52, 54).
- STREICHER, A. ; SMEDDINCK, J. D. (2016)** : « Personalized and Adaptive Serious Games ». In : *Entertainment Computing and Serious Games : International GI-Dagstuhl Seminar 15283, Dagstuhl Castle, Germany, July 5-10, 2015, Revised Selected Papers*. Sous la dir. de R. DÖRNER ; S. GÖBEL ; M. KICKMEIER-RUST ; M. MASUCH ; K. ZWEIG. Lecture Notes in Computer Science. Cham : Springer International Publishing, p. 332-377 (cité p. 86).
- SUÁREZ-FIGUEROA, M. C. ; GÓMEZ-PÉREZ, A. ; FERNÁNDEZ-LÓPEZ, M. (2012)** : « The NeOn Methodology for Ontology Engineering ». en. In : *Ontology Engineering in a Networked World*. Sous la dir. de M. C. SUÁREZ-FIGUEROA ; A. GÓMEZ-PÉREZ ; E. MOTTA ; A. GANGEMI. Springer Berlin Heidelberg, p. 9-34 (cité p. 18, 73).
- SULMONT-ROSSÉ, C. ; VAN WYMELBEKE, V. (juin 2019)** : « Les déterminants d'un apport protidique faible chez les personnes âgées dépendantes ». In : *Cahiers de Nutrition et de Diététique* 54.3, p. 180-189 (cité p. 37).
- SURE, Y. ; STAAB, S. ; STUDER, R. (2004)** : « On-To-Knowledge Methodology (OTKM) ». In : *Handbook on Ontologies*. Sous la dir. de S. STAAB ; R. STUDER. International Handbooks on Information Systems. Berlin, Heidelberg : Springer, p. 117-132 (cité p. 18).

- TALEBI, S. ; .K, M. ; KUMAR, G. (jan. 2019) : « Building Knowledge Graph Based on User Tweets : Proceedings of DAL 2018 ». In : *Lecture Notes in Networks and Systems*, p. 433-443 (cité p. 30).
- TAMMA, V. ; BENCH-CAPON, T. (mai 2001) : « A Conceptual Model to Facilitate Knowledge Sharing in Multi-Agent Systems ». In : (cité p. 56).
- THOMPSON, M. ; BURGER, K. ; KALIYAPERUMAL, R. ; ROOS, M. ; SILVA SANTOS, L. O. B. da (jan. 2020) : « Making FAIR Easy with FAIR Tools : From Creolization to Convergence ». In : *Data Intelligence* 2.1-2, p. 87-95 (cité p. 21).
- TOMMASINI, R. ; SEDIRA, Y. ; DELL'AGLIO, D. ; BALDUINI, M. ; ALI, M. I. ; PHUOC, D. ; DELLA VALLE, E. ; CALBIMONTE, J.-P. (oct. 2018) : *VoCaLS : Vocabulary Ebersand ; Catalog of Linked Streams* (cité p. 18).
- TURNER, J. E. ; LIRA, V. A. ; BRUM, P. C. (2017) : « New Insights into the Benefits of Physical Activity and Exercise for Aging and Chronic Disease ». In : *Oxidative Medicine and Cellular Longevity* 2017, p. 2503767 (cité p. 39).
- USCHOLD, M. ; GRUNINGER, M. (juin 1996) : « Ontologies : Principles, Methods and Applications ». en. In : *The Knowledge Engineering Review* 11.02, p. 93 (cité p. 73).
- USCHOLD, M. ; KING, M. (1995) : « Towards a Methodology for Building Ontologies ». In : *In Workshop on Basic Ontological Issues in Knowledge Sharing, held in conjunction with IJCAI-95* (cité p. 18).
- UZOR, S. ; BAILLIE, L. (avr. 2014) : « Investigating the long-term use of exergames in the home with elderly fallers ». In : *Conference on Human Factors in Computing Systems - Proceedings* (cité p. 85).
- VALLEJO, V. ; WYSS, P. ; RAMPA, L. ; MITACHE, A. V. ; MÜRI, R. M. ; MOSIMANN, U. P. ; NEF, T. (mai 2017) : « Evaluation of a novel Serious Game based assessment tool for patients with Alzheimer's disease ». In : *PLOS ONE* 12.5, e0175999 (cité p. 85).
- Vital Signs Table - Prohealthsys* (2021) (cité p. 77).
- VRANDEČIĆ, D. ; PINTO, H. S. ; DR. TEMPICH, C. ; SURE-VETTER, Y. (oct. 2005) : « The DILIGENT knowledge processes ». In : *J. Knowledge Management* 9, p. 85-96 (cité p. 17).
- WAAGMEESTER, A. ; STUPP, G. ; BURGSTALLER-MUEHLBACHER, S. ; GOOD, B. M. ; GRIFFITH, M. ; GRIFFITH, O. ; HANSPERS, K. ; HERMJAKOB, H. ; HUDSON, T. S. ; HYBISKE, K. ; KEATING, S. M. ; MANSKE, M. ; MAYERS, M. ; MIETCHEN, D. ; MITRAKA, E. ; PICO, A. R. ; PUTMAN, T. ; RIUTTA, A. ; QUERALT-ROSINACH, N. ; SCHRIML, L. M. ; SHAFEE, T. ; SLENTER, D. ; STEPHAN, R. ; THORNTON, K. ; TSUENG, G. ; TU, R. ; UL-HASAN, S. ; WILLIGHAGEN, E. ; WU, C. ; SU, A. I. (fév. 2020) : *Wikidata as a FAIR knowledge graph for the life sciences* (cité p. 21).
- WARDHANA, H. ; ASHARI, A. ; SARI, A. (mars 2018) : « Review of Ontology Evolution Process ». In : *International Journal of Computer Applications* 179, p. 26-33 (cité p. 52, 54).
- WASHBURN, R. A. ; MONTOYE, H. J. (avr. 1986) : « The assessment of physical activity by questionnaire ». In : *American Journal of Epidemiology* 123.4, p. 563-576 (cité p. 41).

- WERNER, D. ; CRUZ, C. ; NICOLLE, C. (avr. 2012) :** « Ontology-based Recommender System of Economic Articles ». In : *8th International Conference on Web Information Systems and Technologies*. SciTePress - Science and Technology Publications 725-728. Porto, Portugal, ISBN 978-989-8565-08-2 (*cité p. 30*).
- WIERINGA, W. ; OP DEN AKKER, H. ; JONES, V. M. ; OP DEN AKKER, R. ; HERMENS, H. J. (2011) :** « Ontology-Based Generation of Dynamic Feedback on Physical Activity ». In : *Artificial Intelligence in Medicine*. Sous la dir. de M. PELEG ; N. LAVRAČ ; C. COMBI. T. 6747. Berlin, Heidelberg : Springer Berlin Heidelberg, p. 55-59 (*cité p. 68*).
- WILKINSON, M. ; KUZNIAR, A. ; KALIYAPERUMAL, R. ; THOMPSON, M. ; DUMONTIER, M. ; BURGER, K. ; SANTOS, L. (sept. 2016) :** « FAIR Data Points Supporting Big Data Interoperability ». In : p. 10 (*cité p. 19*).
- WONGPATIKASEREE, K. ; IKEDA, M. ; BURANARACH, M. ; SUPNITHI, T. ; LIM, A. O. ; TAN, Y. (2013) :** « Location-Based Concept in Activity Log Ontology for Activity Recognition in Smart Home Domain ». In : *Joint International Semantic Technology Conference (JIST) : Gy*. Sous la dir. de D. HUTCHISON ; T. KANADE ; J. KITTLER ; J. M. KLEINBERG ; F. MATTERN ; J. C. MITCHELL ; M. NAOR ; O. NIERSTRASZ ; C. PANDU RANGAN ; B. STEFFEN ; M. SUDAN ; D. TERZOPOULOS ; D. TYGAR ; M. Y. VARDI ; G. WEIKUM ; H. TAKEDA ; Y. QU ; R. MIZOGUCHI ; Y. KITAMURA. T. 7774. Berlin, Heidelberg : Springer Berlin Heidelberg, p. 326-331 (*cité p. 68, 69*).
- YEN, P.-Y. ; MCALEARNEY, A. S. ; SIECK, C. J. ; HEFNER, J. L. ; HUERTA, T. R. (sept. 2017) :** « Health Information Technology (HIT) Adaptation : Refocusing on the Journey to Successful HIT Implementation ». In : *JMIR Medical Informatics* 5.3, e28 (*cité p. 1*).
- YU, H. Q. ; ZHAO, X. ; DENG, Z. ; DONG, F. (2015) :** « Ontology Driven Personal Health Knowledge Discovery ». In : *Knowledge Management in Organizations*. Sous la dir. de L. UDEN ; M. HERIČKO ; I.-H. TING. T. 224. Cham : Springer International Publishing, p. 649-663 (*cité p. 67, 70*).
- ZABLITH, F. ; ANTONIOU, G. ; D'AQUIN, M. ; FLOURIS, G. ; KONDYLAKIS, H. ; MOTTA, E. ; PLEXOUSAKIS, D. ; SABOU, M. (jan. 2015) :** « Ontology Evolution : A Process-Centric Survey ». In : *The Knowledge Engineering Review* 30, p. 45-75 (*cité p. 52-55*).
- ZABLITH, F. ; SABOU, M. ; D'AQUIN, M. ; MOTTA, E. (mai 2009) :** *Ontology Evolution with Evolva* (*cité p. 53*).
- ZHAO, Z. ; HAN, S. ; KIM, J. (fév. 2019) :** « R2LD : Schema-based Graph Mapping of relational databases to Linked Open Data for multimedia resources data ». In : *Multimedia Tools and Applications* 78.20, p. 28835-28851 (*cité p. 100*).
- ZOLFAGHARI, S. ; ZALL, R. ; KEYVANPOUR, M. R. (avr. 2016) :** « SOnAr : Smart Ontology Activity Recognition Framework to Fulfill Semantic Web in Smart Homes ». In : *Second International Conference on Web Research (ICWR)*, p. 139-144 (*cité p. 68*).

Liste des abréviations, des sigles et des symboles

AAL *Active Assisted Living*

ALD Affection Longue Durée

ANC Apport Nutritionnel Conseillé

APA Activité Physique Adaptée

API *Application Programming Interface*

BDD Base De Données

BDR Base de Données Relationnelle

CEN Comité Européen de Normalisation

CIV Centre d'Information des Viandes

CNO Compléments Nutritionnels Oraux

DIKG2 *Data Integration Approach for Knowledge Graphs Generation*

DMP Dossier Médical Patient

DOI *Digital Object Identifier*

DP *Data Properties*

EHPAD Etablissements d'Hébergement pour Personnes âgées Dépendantes

EOSC *European Open Science Cloud*

ETL *Extract-Transform-Load*

FAIR Findable, Accessible, Interoperable, Reusable

FOAF *Friend Of a Friend*

FOE *Flat OWL Editor*

GKG *Google Knowledge Graph*

HAS Haute Autorité de Santé

HTML HyperText Markup Language
HTTP Hypertext Transfer Protocol
HTTPS HyperText Transfer Protocol Secure
IA Intelligence artificielle
IC Ingénierie des Connaissances
IEEE *Institute of Electrical and Electronics Engineers*
IG Indice Glycémique
IMC Indice de Masse Corporelle
IO Ingénierie d'Ontologie
IOT *Internet Of Things*
IPAQ *International Physical Activity Questionnaire*
ISO *International Organization for Standardization*
JPG *Joint Photographic Experts Group*
LOD *Linked Open Data*
MET *Metabolic Equivalent of Task*
MIAM Ontologie
MIO Méthodologies d'ingénierie d'ontologies
MNA *Mini Nutritional Assessment*
MPE Malnutrition Protéinoénergétique
OAFE *Ontology of Activity For the Elderly*
OAPA Ontologie de l'Activité pour les Personnes Agées
OBDA *Ontology-Based Data Access*
OBQA *Ontology-Based Query Answering*
ODBA *Ontology-Based Data Access*
ODP *Ontology Design Pattern*
OMS Organisation Mondiale de la Santé
ONAFE *Ontology of Nutrition and Activity For Elderly*
OP *Object Properties*
OS *Operating System*
OWL *Web Ontology Language*
PHKG *Personnal Health Knowledge Graph*
PID *Persistent Identifier*

PKG Personal Knowledge Graph

PMSI Programme de Médicalisation des Systèmes d'Information

PNG Portable Network Graphics

PNNS Programme National Nutrition Santé

PROV – O PROV Ontology (W3C)

R2RML RDB to RDF Mapping Language

RD2RDF RDB to RDF Mapping Language

RDB Relational Database

RDF Resource Description Framework

RTO Ressources Termino-Ontologique

SEGA Short Emergency Geriatric Assessment

SGBDR Systèmes de Gestion de Base de Données Relationnels

SKOS Simple Knowledge Organization System

SMASH Semantic Mining of Activity, Social, and Health data

SNOMED – CT Systematized Nomenclature Of Medicine Clinical Terms

SOA Service-Oriented architecture

SOSED Ontology of Smart Objects for collected hEalth SEnsor Data

SPARQL SPARQL Protocol and RDF Query Language

TIS Technologies de l'Information sur la Santé

UMLS Unified Medical Language System

URL Uniform Resource Locator

URN Uniform Resource Name

W3C World Wide Web Consortium

XML Extensible Markup Language

Table des matières

Préambule	ix
Synopsis de la thèse	ix
Guide de lecture du manuscrit de thèse	1
Introduction	1
Objectifs	5
Contributions de la thèse	5
Liste des publications scientifiques internationales	6
Structuration de la thèse	7
Chapitre 0 Préliminaires	9
0.1 Ontologies et Web sémantique	10
0.2 RDF et SPARQL	11
0.3 OWL2	13
0.4 Méthodologies d'ingénierie d'ontologie	16
0.5 Ontologies et principes FAIR	19
0.6 Graphes de connaissances	21
I Etat de l'art et positionnement	23
Chapitre 1 Personnalisation de la recommandation	24
1.1 Qu'est-ce que la personnalisation ?	25
1.2 Aperçu des approches de personnalisation	26
1.2.1 Les approches de recommandation	26
1.2.2 Impact du Web dans l'évolution des approches de recommandation	26
1.2.3 Les sources de connaissances	27
1.2.4 En résumé	28
1.3 Les phases de la personnalisation	28
1.4 Positionnement	31
Chapitre 2 Application : prise en charge personnalisée pour les personnes âgées	33
2.1 Introduction	34

2.2	Bien vieillir : un objectif de santé publique	34
2.2.1	Problématiques liées à la dénutrition	35
2.2.2	Problématiques liées à la sédentarité	38
2.2.3	Outils d'évaluation actuels	41
2.3	Outils de collecte de données en santé	42
2.3.1	Le DMP	42
2.3.2	Les objets connectés	42
2.3.3	Les questionnaires en lignes	46
II Contributions		48
Chapitre 3 Évolution des ressources sémantiques modulaires		49
3.1	Introduction	51
3.2	Synthèse des approches d'évolution existantes	52
3.3	Évolution de l'ontologie MIAM	55
3.3.1	Etape 1 : Détection des besoins pour l'évolution	55
3.3.2	Etape 2 : Suggestion des changements	61
3.3.3	Etape 3 : Suggestions des changements pour l'évolution de MIAM	63
3.3.4	Développement de nouveaux modules	67
3.3.5	Etape 4 : Évaluation de l'impact de l'évolution	78
3.3.6	Etape 5 : Gestion des versions	79
3.4	Conclusion	80
Chapitre 4 Support pour la collecte de données et la présentation de la recommandation		82
4.1	Introduction	83
4.2	Conception des questionnaires en ligne	83
4.3	Apport des <i>Serious Games</i>	84
4.4	Approches de conception du <i>Serious Games</i>	86
4.4.1	Définition et catégorisation des <i>Serious Games</i>	86
4.4.2	Méthodologies de conception de <i>Serious Games</i>	87
4.4.3	Etapas de développement d'un <i>Serious Games</i>	88
4.4.4	Principes de scénarisation	88
4.5	Conception du <i>Serious Games</i>	89
4.5.1	Caractéristiques des utilisateurs cibles	89
4.5.2	Caractéristiques du <i>Serious Games</i>	89
4.5.3	Objectifs de conception	90
4.5.4	Scénarisation du <i>Serious Games</i>	90
4.5.5	Activités du <i>Serious Games</i>	90
4.5.6	Représentation iconique des aliments	92
4.5.7	Choix du périphérique	94
4.5.8	Développement du <i>Serious Games</i>	94
4.6	Construction d'un jeu de données fictif	95
4.7	Conclusion	96
Chapitre 5 Intégration de données guidée par des ressources sémantiques		98

TABLE DES MATIÈRES

5.1	Introduction	99
5.1.1	Approches d'intégration de données issues d'une base de données	100
5.1.2	Outils d'intégration de données issues d'une base de données	100
5.1.3	Objectif	101
5.2	Méthodologie proposée : l'approche DIKG2	101
5.2.1	Collecte de données	102
5.2.2	Extraction et structuration des données	103
5.2.3	Enrichissement des données et liaison des entités	104
5.2.4	Enrichissement du PHKG pour le profilage et l'inférence de recommandations	106
5.3	Interrogation des ressources sémantiques	106
5.4	Implémentation de l'approche DIKG2	109
5.5	Conclusion	109
	Conclusion et perspectives	112
	Bibliographie	129
	Glossaire	130
	Liste des abréviations, des sigles et des symboles	133
	Table des matières	136
	Index des figures	139
	Index des tableaux	140
	Résumé	141
	Abstract	142

Table des figures

1	Guide de lecture du manuscrit de thèse	x
2	Les axes sur lesquels s'appuient les objectifs de la thèse (d'après la pile des technologies et des langages du Web sémantique de NOWACK, 2009) .	5
3	Nuage de mots reflétant les thématiques et les problématiques abordées dans la thèse	8
1	Exemple d'un triplet RDF	11
2	Exemple d'un PHKG	22
1.1	Phases de la personnalisation	29
1.2	Positionnement vis-à-vis des approches de personnalisation	31
2.1	Spirale de la dénutrition (FERRY, 2010)	36
2.2	Exemple de déduction de nouvelles connaissances et de personnalisation des recommandations à partir des données collectées	47
3.1	D'après le cycle d'évolution proposé par ZABLITH, ANTONIOU et al., 2015	53
3.2	Comparaison des étapes d'évolution en référence avec les étapes du cycle proposé par ZABLITH, ANTONIOU et al., 2015	55
3.3	Distinction entre données brutes et données calculées	57
3.4	Modèle d'exploitation des données	60
3.5	Aperçu du filtrage des questions de compétences conduisant à une évolution de MIAM.....	62
3.6	Exemples de questions de compétences et de scénarios d'usages selon le thème abordé.....	62
3.7	Modules de MIAM (DESPRES, 2016).....	63
3.8	Schéma montrant les modules de MIAM impliqués dans l'évolution de MIAM et les nouveaux modules à développer pour concevoir ONAFE....	64
3.9	Schéma montrant l'import des modules impliqués dans l'évolution de MIAM et les nouveaux modules à développer pour concevoir ONAFE. Les modules de MIAM ayant subi une évolution sont en pointillés, les modules utilisés sont ont un fond bleu.	65
3.10	Exemple de classification d'aliment selon leur apports nutritionnels.....	66
3.11	Conceptualisation du module Personne pour répondre aux besoins d'évolution	67

3.12	Capture de l'ontologie « Home Activity » (A) (WONGPATIKASEREE et al., 2013), de l'ontologie DBpedia (B), et du méta-thésaurus UMLS (C) faisant notion au concept « Activity »	69
3.13	Représentation des trois classes de PROV-O et des propriétés qui les relie (PROV-O : The PROV Ontology 2013) (a) et de l'ODP de l'activité pour le raisonnement sur l'activité (ABDALLA et al., 2014) (b) .	71
3.14	Aperçu du modèle conceptuel du concept « Activité » montrant les relations de premier et de deuxième niveau.....	74
3.15	Exemple de recommandation d'activité obtenue par raisonnement selon le profil de la personne âgée (qui a comme frein à l'activité une peur de chuter)	75
3.16	Exemple de génération d'une recommandation pour un profil sédentaire .	76
3.17	Exemple montrant les inférences des recommandations obtenues par classification après lancement du raisonneur (classes en jaune).....	76
3.18	Exemple montrant les inférences des recommandations obtenues par classification après lancement du raisonneur (classes en jaune). (a) Description de la Glycemie de nina. (b) Description de l'individu la recommandation d'AlimentIndiceGlycémiquebas dans le cas d'une PersonneHyperglycémique. (c) Description de l'individu ninaOC montrant l'ObjetConnecte dans le cas d'une PersonneHyperglycémique. (d) Description des assertions de propriétés permettant de caractériser l'EtatNutritionnel de nina, son âge et sa valeur de Glycemie. (e) Description de la classe PersonneHyperGlycémique.	78
3.19	Présentation de quelques exemples de classes (A) et de requêtes définies (DL Query) (B) pour obtenir les activités correspondantes pour la requête exécutée.....	79
4.1	Aperçu d'une interface de questionnaires en ligne présentant des questions d'ordre général	84
4.2	Aperçu d'une interface administrateur pour la génération de questionnaires en ligne.....	85
4.3	Organisation des phases d'apprentissages.....	91
4.4	Exemple de scénario pour présentation de la « fiche santé » au joueur....	91
4.5	Les activités du <i>Serious Game</i>	93
4.6	Interface de composition de plat.....	95
4.7	Interface de <i>coaching</i> éducatif sur des notions d'activité physique et de présentation de la recommandation	95
4.8	Aperçu du jeu de données fictif construit	96
5.1	Présentation du processus d'intégration sémantique de données menant à la génération du PHKG	102
5.2	Génération de triplets RDF à partir de la structure du formulaire Web...102	
5.3	Résumé des étapes menant à la construction du PHKG à partir des questions et des réponses issues du questionnaire	103
5.4	Etapes d'intégration de données issues de formulaires Web au sein d'un BDR puis traitement (nettoyage, fusion) et structuration de données du formulaire Web dans le DataFrame 1	104

TABLE DES FIGURES

5.5	Extraction des éléments d'intérêt de ONAFE dans des DataFrames.....	104
5.6	Processus d'enrichissement des données avec le vocabulaire de ONAFE (étape 4), et de liaison des entités du DataFrame 2 entre elles (étape 5) menant à la génération du PHKG (étape 6)	105
5.7	Exemple de raisonnement sur les données pour générer une recommandation	107
5.8	Exemple de requêtes SPARQL - Ajout de recommandation	108
5.9	Exemple de requêtes SPARQL - Appariement de graphes pour la récupération des labels des activités recommandées	108
5.10	Exemple de requêtes SPARQL - Construction de la recommandation à partir des recommandations insérées dans ONAFE	109
5.11	Architecture du <i>framework</i> pour implémenter l'approche DIKG2	110

Liste des tableaux

1	Tableau synoptique du manuscrit de thèse	ix
3.1	Exemple d'utilisation de données selon les moyens d'acquisition.....	58
3.2	Synthèse des principaux termes identifiés (niveau 1 de la ressource, énumérés par ordre de citation) reliés au concept « Activity ». 1 : CHEN et NUGENT, 2009 ; 2 : Riboni and Bettini, 2011 ; 3 : WONGPATIKASEREE et al., 2013 ; 4 : PROV-O, 2013 ; 5 : DBpedia, 2014 ; 6 : ABDALLA et al., 2014 ; 7 : HODA et al., 2018	69
3.3	Synthèse montrant les relations utilisées reliés les différents termes définis par les auteurs cités dans le tableau 3.2 : CHEN et NUGENT, 2009 ; 2 : Riboni and Bettini, 2011 ; 4 : PROV-O, 2013 ; 6 : ABDALLA et al., 2014 ; 7 : HODA et al., 2018	70
3.4	Synthèse des principaux termes identifiés dans les différentes études explorées associés au concept « Person ». 1 : PRIMO et al., 2012 ; 2 : AL-NAZER et al., 2014 ; 3 : YU et al., 2015 ; 4 : NI et al., 2016 ; 5 : HODA et al., 2018	70
3.5	Exemples de scénarios d'usage	72
3.6	Exemples de questions de compétences (QC) associées aux scénarios d'usage (S) cf. 3.5.....	73

Résumé

La santé des personnes âgées est un enjeu majeur de santé publique des prochaines décennies. Un intérêt a été porté aux moyens adaptés pour améliorer le "mieux vieillir" des générations à venir. La personnalisation des recommandations nutritionnelles et d'activités physiques représente une solution efficace pour répondre aux problèmes de santé des personnes âgées. La personnalisation dépend des moyens de collecter puis d'intégrer les données, quelles que soient leurs sources. Cette étape est essentielle pour pouvoir exploiter les données et construire un profil utilisateur. Pour répondre au besoin de collecte de données, faire participer la personne âgée permet de l'impliquer dans son parcours de santé et dans l'intégration de ses besoins. Ainsi, la conception d'un Serious Game pour une visée plus large que la collecte de données a été incluse dans le processus de personnalisation. Le Serious Game construit présente des avantages pour la stimulation des capacités cognitives et physiques et améliore les aspects de personnalisation. Pour répondre au problème d'intégration des données, les solutions proposées dans la thèse consistent à utiliser les ontologies pour éliminer les hétérogénéités sémantiques. Elles fournissent un support de représentation, de définition d'un vocabulaire unifié et d'inférences logiques permettant d'établir un raisonnement entre les connaissances représentées dans l'ontologie et les données collectées. Pour que l'ontologie puisse tenir compte des nouvelles exigences d'intégration de données et de recommandations nutritionnelles et d'activités, une attention particulière a été apportée au processus d'évolution. Les étapes d'un cycle d'évolution d'ontologie ont été appliquées pour faire évoluer une ontologie modulaire du domaine de la nutrition impliquant la construction de nouveaux modules thématiques. Pour pouvoir exploiter les données, une approche d'intégration sémantique de données issues d'une base de données relationnelle a été conçue puis implémentée. L'approche s'appuie sur la construction d'un graphe de connaissances personnel de santé (PHKG), comme moyen pour représenter les données de la personne et d'établir un alignement entre le schéma conceptuel de données et celui de l'ontologie. L'enrichissement du PHKG avec le vocabulaire de l'ontologie permet de construire un « profil santé » de la personne âgée et d'établir la correspondance entre le profil et la recommandation. Les graphes de connaissances appliqués au domaine de la santé ont récemment été introduits pour exprimer l'idée d'une vision personnalisée et demeurent jusqu'à ce jour un sujet quelque peu sous-exploré. Les solutions proposées dans la thèse consistent à utiliser l'expressivité de l'ontologie pour personnaliser la prise en charge non thérapeutique des personnes âgées et répondre ainsi au problème initial de santé publique.

Mots clés : Ingénierie des ontologies, Intégration de données, Graphes de connaissances, Profilage, Acquisition des connaissances, Personnalisation, E-santé.

Abstract

The health of the elderly is a major public health issue for the coming decades. A particular interest was in ways to improve the "better aging" of future generations to come. Personalizing nutritional and physical activity recommendations is an effective way to address the health concerns of older adults. Effective personalization depends on the ways of data collection and integration, regardless of its source. This step is essential in order to exploit the data and build a user profile. To address the need for data collection, involving the elderly allows them to be involved in their health care journey and the integration of their needs. Thus, designing a Serious Game for a broader purpose than data collection was included in the personalization process. The constructed Serious Game has advantages for the stimulation of cognitive and physical abilities and improves the personalization aspects. To address the data integration problem, the proposed solutions in the thesis consist in using ontologies to eliminate semantic heterogeneities. They provide support for representation, definition of a unified vocabulary and logical inferences allowing to establish reasoning between the knowledge represented in the ontology and the collected data. For the ontology to take into consideration the new requirements of data integration and nutritional and activity recommendations, special attention was paid to the evolution process. The steps of an ontology evolution cycle have been applied to evolve a modular ontology of the nutrition domain involving the construction of new thematic modules. To exploit the data, a semantic data integration approach from a relational database was designed and implemented. The approach relies on the construction of a personal health knowledge graph (PHKG), as a means to represent the person's data and to establish an alignment between the conceptual data schema and the ontology. The enrichment of PHKG with the ontology vocabulary allows one to build a "health profile" of the elderly and to establish the correspondence between the profile and the recommendation. Knowledge graphs applied to the health domain have recently been introduced to express the idea of a personalized vision and are still a somewhat under-explored topic. The solutions proposed in this thesis consist in using the expressiveness of ontology to personalize the non-therapeutic care of elderly people and thus work towards improving the initial public health problem of the health of the aging population.

Keywords : Ontology Evolution, Ontology Engineering, Data Integration, User Profiling, Knowledge Acquisition, Personalization, E-Health.