



Vol.31, No.1, 2022-Table of Contents

[Open Access](#)

INTRODUCTION

Plumejeaud-Perreau, C, Gravier, J, Masson, E, Nahassia, L, Rodier, X, Saux, E, Fargette, M, Libourel, T, Mathian, H, Nuninger, L, Sanders, L

Revue Internationale de Géomatique, Vol.31, No.1, pp. 7-19, 2022, DOI:10.3166/rig31.7-19

Numérique versus symbolique

Hélène Mathian¹, Lena Sanders²

Revue Internationale de Géomatique, Vol.31, No.1, pp. 21-45, 2022, DOI:10.3166/RIG31.21-45

Appréhender le changement des catégories pour l'étude d'une dynamique spatiale sur le temps long

Christine Plumejeaud-Perreau¹, Lucie Nahassia², Julie Gravier²

Revue Internationale de Géomatique, Vol.31, No.1, pp. 47-80, 2022, DOI:10.3166/RIG31.47-80

Explorer les processus de mobilité passée Raisonnement ontologique fondé sur la connaissance des pratiques socioculturelles et des vestiges archéologiques

Laure Nuninger¹, Thérèse Libourel², Xavier Rodier³, Rachel Opitz⁴, Philip Verhagen⁵, Catherine Fruchart¹, Clément Laplaige³, Samuel Leturcq³, Nathanael Levoguer³

Revue Internationale de Géomatique, Vol.31, No.1, pp. 81-110, 2022, DOI:10.3166/RIG31.81-110

Analyse des empreintes guidée par un modèle de connaissances pour la compréhension des dynamiques de navigation maritime côtière

Wissame Laddada¹, Éric Saux²

Revue Internationale de Géomatique, Vol.31, No.1, pp. 111-133, 2022, DOI:10.3166/RIG31.111-133

La genèse systémique d'empreinte pour une maîtrise de l'observation de la Terre

Mireille Fargette¹, Maud Loireau², Najet Raouani³, Thérèse Libourel⁴

Revue Internationale de Géomatique, Vol.31, No.1, pp. 135-197, 2022, DOI:10.3166/RIG31.135-197

Une ontologie orientée objet pour la modélisation des socio-écosystèmes

Éric Masson

Revue Internationale de Géomatique, Vol.31, No.1, pp. 199-230, 2022, DOI:10.3166/RIG31.199-230

INTRODUCTION

Est-il possible de dégager une démarche méthodologique pour étudier des dynamiques spatiales, une démarche qui s'appuierait sur les apports respectifs des approches symboliques et des approches numériques ? Partant d'expériences méthodologiques et scientifiques différentes (en géographie, archéologie, histoire, agronomie, écologie, etc.), ce numéro thématique tente d'identifier les rôles respectifs de ces deux approches, leurs points de jonctions, pour dégager cette fameuse démarche reproductible et transposable dans plusieurs contextes disciplinaires, voire transdisciplinaires.

1. Des ontologies comme support de cadre symbolique

Le lecteur pourra se rendre compte que l'approche symbolique qui a été privilégiée dans les exemples déclinés tout le long des articles repose sur l'usage d'ontologies. Des ontologies, oui, mais desquelles parle-t-on ? Ce numéro spécial est issu du travail d'une action de prospective du GdR CNRS MAGIS regroupant des scientifiques de différentes disciplines, géographes, informaticiens, archéologues, historiens, biologistes et agronomes. Le premier travail du groupe fut de s'entendre sur ce que recouvrait la notion d'ontologie pour chacun. Un premier travail « ontologique » (osons le dire) qui nous conduit à constater que, parmi la multiplicité des définitions du terme « ontologie » à disposition, nous l'avons abordée sous les différentes acceptions/définitions suivantes :

L'ontologie est « la science de ce qui est, des types et des structures des objets, propriétés, événements, processus et relations en tout domaine de la réalité.../... de ce qui pourrait exister » ni celle qu'en donne une des définitions du CNRTL¹ (ce dictionnaire indique aussi la polysémie du terme) : « Partie de la philosophie qui a pour objet l'étude des propriétés les plus générales de l'être, telles que l'existence, la possibilité, la durée, le devenir. » La définition qu'en propose Barry Smith (Smith, 2003) recoupe tout à fait la définition du CNRTL, qui sera souvent mobilisée dans ce numéro pour formaliser des dynamiques spatiales, suivant en cela également l'approche de Pierre Livet (2010) pour qui « l'ontologie consiste à analyser un domaine, en identifiant les entités pertinentes (objets, propriétés, relations, événements et processus), et les opérations qui peuvent être opérées sur ces entités. »

1. <https://www.cnrtl.fr/definition/ontologie>

Du côté des informaticiens, une ontologie n'est plus concernée par la question philosophique de l'existence mais répond à un besoin de normalisation des communications et des échanges. Tom Gruber définit une ontologie comme « une spécification d'une conceptualisation » (Gruber, 1993). Cette définition est une tentative de réponse au problème récurrent de l'incompréhension qui émerge lorsque plusieurs groupes collaborant entre eux utilisent des termes parfois identiques mais ayant une définition et un sens qui leur sont propres. Ainsi, ce consensus commun autour du sens des termes favorise leur réutilisation et le partage de la connaissance (Ushold et Gruninger, 1996). De manière plus pragmatique, Nicola Guarino précise la définition précédente en relevant les différences qui existent entre une « conceptualisation » et une « ontologie » (Guarino, 1998). Une ontologie se distingue alors selon ses niveaux de conceptualisation :

- le domaine d'étude est appréhendé selon une approche logique rigoureuse (niveau logique) ;
- les concepts pertinents qui définissent le vocabulaire du domaine d'étude sont identifiés, définis et structurés en un réseau sémantique (niveau épistémologique) ;
- ces concepts sont reliés entre eux par des relations sémantiques (niveau ontologique).

Une ontologie au sens informatique se différencie selon les trois niveaux définis par Nicola Guarino. D'un côté si l'on ne considère que les deux niveaux épistémologique et ontologique, les ontologies sont un outil informatique composé d'un ensemble structuré des termes et concepts représentant le sens d'un champ d'informations, que ce soit par les métadonnées d'un espace de noms, ou les éléments d'un domaine de connaissance. L'ontologie constitue en soi un modèle de données représentatif d'un ensemble de concepts dans un domaine, ainsi que des relations entre ces concepts. Elles sont utilisées dans un but de partage de l'information dédiée à un domaine ou une application particulière, et tentent par conséquent de répondre au problème de l'interopérabilité entre différentes sources. Si l'on considère en plus le niveau logique, les ontologies sont un outil informatique composé d'un ensemble structuré de concepts et de relations définis à l'aide d'un formalisme logique. Ce formalisme permet une modélisation théorique d'une sémantique formelle. Dans les articles 2 et 4, nous les appelons « ontologies axiomatisées ». Ce formalisme logique conduit au mécanisme de raisonnement qui permet, dans un premier temps, d'évaluer la consistance de la théorie, mais également d'inférer de nouvelles connaissances à partir des connaissances initiales par l'usage de raisonneurs.

Certains disent que l'« ontologie est aux données ce que la grammaire est au langage ». Les articles de ce numéro thématique parlent tous au moins des ontologies comme d'un cadre conceptuel formalisé de niveau symbolique (au sens des niveaux épistémologiques et ontologiques proposés par Guarino).

2. Le domaine numérique

Concernant les approches numériques qui sont à l'œuvre, nous devons d'abord déterminer ce qui, selon les auteurs, relève du domaine numérique. Il y a les « données » et les méthodes (statistiques, classification, apprentissage, *deep learning*, etc.) pour les exploiter. Nos exemples sont illustrés par des méthodes statistiques – ou analyse de données, suivant la conception qu'en donne Benzécri (1973) – dans les deux premiers articles, par des méthodes d'inférence et de raisonnement par logique formelle du premier ordre dans le quatrième article ou des méthodes de classification et apprentissage dans les articles 5 et 6, orientée objet ou non.

Concernant les « données », une courte digression dans cette introduction s'impose. Face au déluge de données (*big data*, réseaux sociaux, Web de données, capteurs, etc.), les scientifiques sont mobilisés pour tirer des conclusions *via* diverses méthodes numériques, mais souvent sous l'emprise d'un effet de mode qui peut les conduire à utiliser celles qu'ils comprennent le moins bien. Dans tous les domaines d'application, on observe une explosion des sujets de thèse et de master incluant l'usage du *deep learning* (apprentissage par réseaux de neurones artificiels), une méthode issue de l'intelligence artificielle qui prend sa revanche sur les approches symboliques qui ont pu prévaloir dans cette matière (Cardon *et al.*, 2018). Ce courant touche tous les domaines d'application disciplinaires. Outre les méthodes, ce sont aussi les modalités de compilation et d'organisation des données qui sont bouleversées, en témoigne le développement des « data lake »², ou lacs de données (Madera, 2018) qui, notamment grâce à des technologies *noSQL* dans le *cloud*, offriraient la possibilité de recombinaison à volonté l'information pour en tirer des enseignements, sans les structurer au préalable suivant une question de recherche. En témoigne également l'émergence des nouveaux métiers autour de la « data science » ou « science des données ».

Cependant la donnée n'est pas une information, et n'est pas donnée en tant que telle (Ollion et Boelaert, 2015). Elle est issue d'un processus de construction imaginé par les producteurs pour répondre à leurs questions, qui ne sont pas celles que le chercheur a forcément lorsqu'il ré-utilise les données dans un autre but. C'est là un point crucial que discutent plusieurs de nos articles, en particulier le deuxième (Mathian et Sanders). Les données sont obtenues, comme dit (Latour, 1993) : « Décidément, on ne devrait jamais parler de “données”, mais toujours d’“obtenues”. » Il est important d'utiliser les bons mots, car, par exemple en substituant « données » et « information » de façon interchangeable, alors qu'au niveau information, un processus cognitif proprement humain intervient par l'usage implicite d'un cadre symbolique de référence, on arrive à diffuser le message que les ordinateurs sont intelligents (Holmes, 2001). Or ils ne manipulent que des données. Il s'ensuit une cascade de conséquences ennuyeuses, qui doivent être appréhendées

2. <https://www.lebigdata.fr/data-lake-definition>

en amont, en particulier dans le cadre d'études de dynamiques spatiales avec des enjeux interdisciplinaires. Les articles 2 et 5 de ce numéro spécial discutent abondamment de cette question.

Par ailleurs, concernant les approches numériques pour traiter les données, il y a les habitudes (dont les mauvaises) : en effet, la pratique d'approches statistiques et de l'analyse de données nécessite une grande rigueur et de l'expérience (donc du temps de pratique) car, comme (DeVeaux *et al.*, 2008) le disent si bien, « Math is Music; Statistics is Literature ». Par ailleurs, nombreux sont ceux qui confondent corrélation et causalité, choses bien distinctes comme l'explique le petit livre de H. Krivine (2016) de manière particulièrement amusante et pédagogique. Enfin, l'habitude de plus en plus répandue parmi la communauté académique de mobiliser les méthodes statistiques comme un livre de recettes qui garantirait la saveur du plat (la publication scientifique en l'occurrence) a conduit la société américaine de statistique à bannir la (mauvaise) pratique de l'interprétation à la hussarde de la *p-value* pour les tests statistiques (Wasserstein et Lazar, 2016). Plus fondamentalement, le non-contrôle des hypothèses des modèles statistiques quand nécessaire est général (Rodgers, 2010). Or les données ne signifient rien à l'état brut. C'est un modèle qui permet de les interpréter : la démarche scientifique fait des hypothèses sur le rapport des données avec la réalité et donne du sens aux données pour confirmer/infirmier ce modèle, et c'est précisément ce chemin cognitif que tente de modéliser au mieux l'article 5. Même en se plaçant dans le cadre de la statistique exploratoire comme dans l'article 1 qui vise à s'abstraire d'un modèle statistique, le choix de la méthode relève lui-même d'un cadre théorique qui conditionne l'interprétation des analyses. Aujourd'hui, plus que jamais, les données doivent être interprétées avec précaution. Pour ce faire, il faut disposer d'une connaissance précise du domaine étudié.

3. Intérêt et écueils des approches symboliques fondées sur des ontologies

L'attrait des ontologies tient ici précisément au fait qu'elles permettent d'intégrer plus de connaissance(s) sur la donnée dans le modèle de calcul par une spécification formelle du domaine abordé. Ce type d'approche qui exige de formaliser des implicites sur les modèles et les construits des données permet également de rapprocher des acteurs qui ne parlent pas le même langage pour s'entendre sur les concepts importants à modéliser dans le domaine d'étude. Les avantages semblent indéniables, et ainsi des auteurs (Martinez-Cruz *et al.*, 2012) promeuvent une définition de cette formalisation ontologique, pour le moins prometteuse :

“Ontologies provide a restriction-free framework to represent a machine readable reality, even in the Web. This framework assumes an open world in which information can be explicitly defined, shared, reused or distributed. Moreover, information can also be interchanged and used to make deductions or queries.”

Leur argument va dans le sens de la définition informatique d'ontologie axiomatisée, et ils expliquent que l'avantage apporté par les possibilités de

déduction dans le graphe de relations construites via le modèle rend l'approche infiniment supérieure à celle des bases de données relationnelles. L'article 4 de ce numéro fait une démonstration complète de ce potentiel afin de proposer des routes de navigation à suivre, à partir d'une formalisation des connaissances de navigation maritime côtière. Cependant, certains de nos articles prennent une distance avec ce discours parfois un peu technologique, car l'approche ontologique décrite est implémentée dans des bases de données et des thésaurus, comme dans l'article 2. Si la capacité d'inférence automatisée est certes perdue, l'avantage de l'interopérabilité disciplinaire est démontré de façon indiscutable, mais au prix d'un effort de discussion plutôt que d'un effort de développement informatique.

Il est certain que les ontologies ont été désignées comme ayant un rôle essentiel dans la mise en place du Web sémantique (*i.e.* d'un partage interopérable de données rendues interprétables par des machines) car elles offrent la capacité de prendre en charge la description de la sémantique comme le raisonnement sur les données. Par ce fait, les ontologies présentent des données structurées avec une formalisation sémantique qui est indispensable à la génération de ressources accessibles, partageables, liées (*linked data*), manipulables et interprétables par des machines comme le conçoit Tim Berners-Lee (Berners-Lee, 1998). Inversement, le mouvement pour le Web sémantique a mis en place des standards, et indirectement des technologies et des outils (par exemple les différents raisonneurs Pellet, Hermit, etc.) fondés sur ces standards, connus sous le nom de « piles du web sémantique » gérés par le World Wide Web Consortium (W3C). Ce dernier a standardisé l'ensemble des piles comprenant les langages et les protocoles dont le langage OWL qui permet une modélisation ontologique des descriptions formelles écrites en logiques de description. Notons l'étroite collaboration entre l'Open Geospatial Consortium (OGC) et le W3C pour la définition d'un modèle ontologique pour la représentation de la dimension spatiale à savoir, l'ontologie fondée sur le standard de GeoSPARQL. GeoSPARQL est une ontologie standard pour la représentation et l'interrogation de données géospatiales.

Cependant, ontologie ne rime pas forcément avec interopérabilité (Isaac, 2011). Les modèles ontologiques sont d'abord une projection des représentations mentales de leur concepteur et de leur besoin. Ainsi dans le domaine de l'environnement, plusieurs ontologies existent pour modéliser le même besoin, celui de capturer des observations, comme *Observation and Measurement* (Cox, 2016), OBOE (Madin *et al.*, 2007), mais ces ontologies, comme d'autres, ne sont pas complètement équivalentes et donnent lieu à des travaux d'alignement³. C'est également un point largement développé dans l'article 2.

Concernant le thème des dynamiques spatiales, savoir raisonner sur des dimensions temporelles et spatiales pour inférer des faits non explicités à partir des données est bien entendu intéressant. Mais là aussi, le dialogue interdisciplinaire reste central pour créer des modèles spatiotemporels réellement interopérables, au

3. https://www.w3.org/2015/spatial/wiki/Alignment_to_OBOE

risque sinon d'une certaine maltraitance de la dimension sémantique des données. Dans ce numéro thématique nous avons donc voulu insister sur les limites d'un modèle ontologique axiomatisé, en montrant la nécessité d'allers-retours constants entre l'objet thématique questionné et la mise en œuvre informatique. Nous prenons l'exemple de (Harbelot, 2018) qui, avec le modèle Continuum, propose la modélisation des objets qui se transforment (les endurants au sens de l'ontologie DOLCE – cf. Masolo, 2003) dans un système spatiotemporel grâce à l'usage de tranche de « vie », aussi appelée l'approche des *fluents*. Ce modèle décrit de façon ingénieuse la dimension spatiale et temporelle attachée à chaque état stable d'une entité observée, mais n'est pas prescripteur de la dimension descriptive de l'entité (le quoi observé). Or avec son usage dans de récents travaux en agro-écologie (Tran, 2015), on a constaté que les tranches de vie stables d'objets modélisés (des parcelles agricoles ou des oiseaux) étaient particulièrement mal renseignées et déconnectées des protocoles ayant conduit à la collecte des données de suivi, ce qui a pu donner lieu à de mauvaises interprétations des résultats. Évidemment la critique touche l'usage du modèle Continuum et non le modèle lui-même, et il n'aurait tenu qu'aux auteurs de proposer une description interopérable et plus juste des parcelles agricoles et du suivi de biodiversité attaché. En effet, nous soupçonnons que parfois, l'attrait pour une mise en œuvre de la démarche très centrée sur des questions technologiques afin de produire rapidement des résultats publiables a pu empiéter sur la nécessité de prendre le temps requis pour l'élaboration d'un cadre conceptuel solide, partagé, interopérable et propice à l'exploration de nouvelles pistes de recherches.

Il est vrai que la question technologique autour des ontologies est en fait un problème central qui a des impacts sur la calculabilité et la gestion des volumes de données pour la résolution des inférences. Le foisonnement technologique dans le secteur de la recherche et de la R&D industrielle n'a pas encore permis d'atteindre un niveau de maturation tel que ces problèmes soient abordés frontalement. En effet, l'utilisateur lambda qui veut déployer un système d'information basé sur des ontologies doit d'abord hésiter entre quelques choix de technologies pour leur implémentation (quel *triplestore*, quel raisonneur, comment exprimer les contraintes ?) mais surtout, ses développements d'un jour ont peu de chance d'être capitalisés pour le lendemain du fait de l'évolution (très) rapide des technologies associées, alors que le coût de formation à ces technologies est important. Ce problème a ainsi conduit à l'expression d'un agenda de recherche concernant les ontologies spatiotemporelles (Claramunt, 2020) qui soulève ces défis, qui doivent être pris à bras le corps, au risque sinon d'un abandon de cette recherche au profit d'approches purement numériques.

Enfin, le tout sémantique d'un jour n'est pas forcément à jour le lendemain. Quelles garanties et pratiques pour l'exhaustivité, la qualité, la pérennité, le volume, l'indépendance, les possibilités de diffusion, la fraîcheur, la confiance et non-redondance des ressources sémantiques mobilisées ou créées dans le cadre de l'étude de dynamiques spatiales ? Notre numéro spécial n'en peut rien dire encore. Il

se concentre plutôt sur la démonstration de la complémentarité des approches numériques et symboliques.

4. Complémentarité des approches numériques et symboliques

Les approches symboliques, par la mise en œuvre de modèles, la mobilisation de concepts et de connaissances expertes (proprement humaines et non disponibles sur support numérique), permettent de guider l'analyse d'une dynamique spatiale, et définissent ce qui serait reconnaissable comme empreinte dans l'observable. C'est le cas par exemple de l'article 5 exemplifié par l'analyse d'une oasis par un agronome utilisant les techniques de télédétection. C'est également le cas de l'article 3 qui restitue deux études de cas portant sur la compréhension des mobilités humaines et leur articulation avec les réseaux de transport par l'analyse de vestiges et relevés LiDAR. Cependant, le défaut d'une approche symbolique seule est de restreindre la découverte à ce que l'on cherche ou ce que l'on s'attend à trouver. Par ailleurs, lorsque le monde et les sociétés modélisées par les ontologies évoluent, modifiant les catégories bien établies des types observés, les ontologies qui servent à étudier cette évolution doivent aussi évoluer, s'adapter ou devenir plus généralistes. Ce problème de l'impermanence des catégories et de biais contextuel est discuté dans les articles 2 et 3 qui proposent des cas d'étude de dynamiques spatiales sur le temps long (plusieurs siècles). Enfin, la logique formelle des ontologies s'accommode encore mal de l'incertitude autour des objets de l'analyse de dynamiques spatiales, même si ce champ fait l'objet de travaux de recherche qui ne sont pas discutés dans ce numéro.

Les approches numériques par l'analyse de données numériques, si elles utilisent des modèles parfois implicites, comme expliqué dans l'article 1, et si leur choix est souvent guidé par l'*habitus* de la discipline (statistique, écologie, géographie, etc.), offrent la possibilité de compléter et enrichir de façon itérative les modèles conceptuels. En effet, la discussion de leur résultat permet de questionner la validité des modèles (le choix de tel agent dans le modèle de simulation est-il vraiment dicté par telle règle par exemple ?), permettent de découvrir des inattendus, qui restent à expliquer. L'interprétation doit cependant prendre en compte le fait que la donnée a été produite pour répondre à une question ou un usage spécifique avec un protocole particulier, et pas forcément pour la question scientifique posée ultérieurement lors d'une ré-exploitation de ces données diffusées de façon libre. L'interprétation doit aussi tenir compte des sources de bruit contenu dans la donnée (article 5), non seulement celles relatives aux conditions de collecte, mais surtout aux sources pointées comme non pertinentes dans le cadre critique de l'ontologie systémique construite pour la question posée. Un va-et-vient perpétuel entre le monde symbolique et numérique est donc une nécessité, une richesse pour alimenter le corpus de connaissances scientifiques, et ce numéro spécial à travers tous les articles essaie de le démontrer.

5. Organisation des articles

Chacun des articles souhaite donc illustrer la complémentarité d'approches numériques et symboliques pour l'analyse des dynamiques spatiales, et ce, dans une optique transdisciplinaire. Par ailleurs, ces articles tirent **une synthèse réflexive** d'expériences variées de modélisation de dynamiques spatiales en naviguant entre différentes approches, théorique et opérationnelle. Il faut donc noter ici que le niveau de discussion, en s'appuyant sur de précédentes études et jeux de données associées qui ont déjà été publiés, est d'un niveau conceptuel assez élevé, et parfois ardu. Surtout, il s'engage sur l'expression argumentée d'opinions relatives à l'usage d'ontologies dans les domaines de spécialité des auteurs. Le contenu de ces chapitres a tout le moins été intensément débattu entre les membres de l'Action Prospective du GdR CNRS Magis, de 2015 à 2020, lors de leurs rencontres. Mais il a aussi fait l'objet de relectures par des relecteurs scientifiques extérieurs au groupe, spécialistes des domaines abordés. Nous les remercions de leurs inestimables contributions et leur patience pour la relecture de chapitres requérant une bonne concentration. Nous avons toutefois défendu et maintenu certaines de nos opinions qui nous semblaient importantes, et nous espérons que les lecteurs apprendront ainsi de nos expériences et seront inspirés par nos propositions.

L'article 1 par exemple s'appuie sur l'ontologie proposée par Denis Phan pour montrer son rôle de médiation entre une pratique numérique (l'analyse de données - AD) et une pratique symbolique (simulation multi-agents - SMA) qui se complètent et se nourrissent de manière itérative. Mais cet article s'attache également à démontrer la compatibilité des modèles conceptuels thématiques (que ce soit celui de la « rationalité intentionnelle » des acteurs selon Boudon, ou d'acteurs sous l'influence d'un collectif selon Bourdieu) avec les modèles méthodologiques mobilisés (AD et SMA), pour l'analyse des dynamiques de ségrégations spatiales scolaires. Cela autorise enfin le déploiement d'une pratique interprétative des faits observés en toute rigueur.

Les articles 2 et 3 mobilisent des exemples d'analyses sur le temps long pour montrer le rôle de médiation qu'ont joué des cadres ontologiques, d'une part, en guidant les analyses numériques, et, d'autre part, en matérialisant les discours conceptuels, et ce faisant en autorisant des débats épistémologiques autour des comparaisons effectuées entre disciplines et acteurs de la recherche très divers dans leur pratique et leur culture.

A travers l'étude de peuplements du territoire à différentes échelles, l'article 2 illustre comment un cadre conceptuel partagé concernant la description des différents artefacts que le temps a laissé de ces occupations humaines dans les sources, que ce soit les archives du sol ou les archives cartographiques, autorise leur comparaison, mais dans certaines limites qui sont exposées et argumentées. Le souci de rigueur et d'objectivité est bien présent, à travers notamment un retour sur l'historicité de ce cadre, et l'influence des pratiques des acteurs scientifiques.

L'article démontre clairement comment ce cadre conceptuel de niveau symbolique conditionne les résultats d'analyses numériques quantitatives qui en découlent.

L'article 3 est une proposition originale autour d'un graphe ontologique, le « track graph », qui pourrait être le support générique d'analyses spécialisées des réseaux de mobilités formels et informels qui s'entremêlent, mais dont le temps long ne nous laisse que des traces imparfaites. Après avoir montré à travers deux cas d'études très différents publiés dans deux précédentes publications comment ce *track graph* peut s'instancier pour modéliser un réseau de mobilité, les auteurs font le lien avec deux approches pré-existantes et complémentaires. La première concerne la modélisation de réseaux formels, *i.e.* consciemment conçus par la collectivité pour favoriser les déplacements humains d'un lieu à l'autre, tandis que la seconde concerne les réseaux informels de déplacement, *i.e.* émergeant d'*habitus* culturels variés menant à des tracés spontanés de chemins de passage. Cet article propose donc un pont entre ces vocabulaires et ces deux approches pour *in fine* faire une proposition de recherche purement ontologique qui serait potentiellement un bon support transdisciplinaire pour l'analyse des mobilités passées, actuelles et à venir.

Dans l'article 4, l'étude des dynamiques spatiales est abordée sous le prisme de la représentation des connaissances dans un contexte de navigation maritime côtière. Ce contexte, très proche de la problématique de « wayfinding » en environnement maritime, place la notion de repère au tout premier plan de l'étude. En plus de proposer un vocabulaire propre à ce domaine, cet article propose une description formelle des concepts (« ontologie axiomatisée ») ainsi qu'un certain nombre de règles liées aux conditions de visibilité des repères. La démarche adoptée est « classique » au sens de la représentation des connaissances pour la définition d'un modèle de connaissances : conceptualisation, formalisation des concepts et des rôles en logiques de description, modélisation ontologique puis raisonnement. Après avoir retranscrit la connaissance des experts en matière de définition de trajectoire maritime côtière au sein d'une base de connaissances, le système informatique infère des routes de navigation à partir des faits qui sont présents dans la base de connaissances. Deux stratégies d'exploitation des routes inférées sont envisagées. L'une concerne leur exploitation au sein d'un système de routage pour proposer au navigateur des routes de navigation à suivre, l'autre concerne leur exploitation à des fins d'analyse des empreintes de positionnement de bateaux afin de capturer les portions ayant une sémantique pour le modèle et celles dépourvues de sémantique et sur lesquelles une analyse doit être approfondie. Dans le cadre de ce numéro spécial dédié à l'étude de dynamiques spatiales, les auteurs ont opté pour cette seconde stratégie d'exploitation du modèle de connaissances.

Dans ce numéro spécial nous avons tendus de façon progressive à proposer une démarche de modélisation ontologique qui monte en niveau, de plus en plus « méta », dans cette optique de proposer un cadre général théorique interdisciplinaire pour modéliser les dynamiques spatiales. Ce cadre, qui n'est pas

encore fixé, mais dont des esquisses assez avancées sont proposées dans les deux derniers chapitres.

En effet, dans un contexte de demande sociétale forte pour une modélisation des enjeux des crises en cours ou à venir (sanitaires, économiques et sociales) dans un monde en transition, écologique et climatique, la mise au point de dispositifs informatisés de combinaison guidée et multi-échelle intelligente des données de suivi serait utile pour tous les socio-éco-systèmes. Ceux-ci sont complexes par définition, du fait des interactions et rétroactions entre parties du système (Collins *et al.*, 2011), et requièrent le support de modèles symboliques de connaissances, qui montent en niveaux de généralité. Si la convention d'Arrhus de 1998 a souligné à juste titre la nécessité d'un accès libre à l'information environnementale (Prieur, 1999), ces données environnementales, économiques, sociales qui doivent faire partie des communs de la connaissance (Hess et Ostrom, 2007) ne sont pas si facilement mises en musique pour faire sens. En effet, les infrastructures de recherche de type *cloud* de données avec services sur données liées (*Linked Open Data*) existantes (comme le European Open Science Cloud⁴) permettent la collecte et la centralisation généralisée des données, quelles que soient les disciplines. Cependant ces infrastructures ne suffisent pas à répondre à leur raison d'être, qui est notamment de permettre à chacun d'appréhender l'impact de son mode de vie, des politiques publiques, sur son environnement, proche ou lointain, et *in fine* sur ses conditions de vie et d'agir suivant son libre arbitre. Déjà certains travaux montrent que l'écosystème de ces infrastructures de données peut lui-même être modélisé comme un système complexe et dynamique qui mérite d'être monitoré, afin d'en assurer une meilleure gouvernance (Noucher *et al.*, 2019). Quant aux observatoires, en se référant au modèle théorique de l'observatoire scientifique des relations Société-Milieu en appui à la gestion des territoires (observatoire de type OSAGE), ils montrent des préoccupations similaires dans les enjeux de société qui leurs sont confiés et la restitution qu'ils font de leurs résultats (Loireau et Fargette, 2022).

Il s'agit donc ici de redonner une motivation et une raison d'être à ces données, qui, même si sémantiquement sont bien décrites, se doivent d'être appréhendées dans le cadre de la compréhension du fonctionnement de socio-éco-systèmes complexes. La problématique du croisement de données multi-échelles et multisources est ici abordée dans un cadre théorique symbolique, qui s'ouvre largement vers l'interdisciplinarité. Le cadre proposé nous invite, comme Philippe Descola dans la *Composition des Mondes*, à faire un pas de côté dans notre exercice de modélisation, de « composition du monde » dirait-il. Il s'agit donc de bien établir que nous percevons le monde et l'empreinte des systèmes en jeu à travers le prisme de nos connaissances et de nos facultés naturelles ou technologiques de perception. L'article 5, tout en proposant des illustrations essentiellement tournées vers la biologie et l'agronomie montre que cet écart entre monde perçu et monde réel, monde pensé et « vérité », peut être modélisé ontologiquement pour intégrer une

4. <https://eoscobservatory.eosc-portal.eu/>

démarche d'observation et d'analyse des socio-éco-systèmes. Enfin, l'article 6 s'ancre dans une démarche ontologique orientée objet pour proposer une position de recherche autour de six concepts ayant une portée transdisciplinaire et ayant pour ambition de modéliser le métabolisme des socio-écosystèmes. Ici, loin de l'ingénierie sémantique, nous montrons la puissance qu'une approche ontologique peut avoir pour faire avancer le front des connaissances transdisciplinaires.

En effet et finalement, ce numéro est le fruit de **l'interdisciplinarité**, celle qui a présidé à nos discussions scientifiques. Chacun des articles est une mise en évidence d'un processus à double sens allant du concret (l'empreinte, la trace, la source d'information) vers le concept immatériel et abstrait, c'est-à-dire une connaissance acquise et recherchée (l'interprétation, l'explication d'une dynamique spatiale) et vice-versa, où des connaissances d'experts vont guider l'analyse des dynamiques du phénomène d'intérêt. C'est même souvent un processus itératif avec ces allers-retours entre connaissances et traces numérisées, exprimé selon un vocabulaire diversifié. De plus, certains articles privilégient un sens, ou l'autre, dans leur focus et leurs exemples. Ceci illustre parfaitement la diversité d'*habitus* disciplinaires selon les articles, qui eux-mêmes sont le plus souvent écrits en interdisciplinarité à plusieurs auteurs.

Bibliographie

- Berners-Lee T. (1998). *The World Wide Web: A Very Short Personal History*.
- Cardon D., Cointet J.-P., Mazières A. (2018). La revanche des neurones: L'invention des machines inductives et la controverse de l'intelligence artificielle. *Réseaux*, vol. n° 211, n° 5, p. 173-220. doi: 10.3917/res.211.0173
- Claramunt C. (2020). Ontologies for geospatial information: Progress and challenges ahead. *JOSIS*, n° 20, p. 35-41, doi: 10.5311/JOSIS.2020.20.666
- Collins S. L. *et al.* (2011). An integrated conceptual framework for long-term social-ecological research. *Frontiers in Ecology and the Environment*, vol. 9, n° 6, p. 351-357, doi: 10.1890/100068.
- Cox S. J. D. (2016). Ontology for observations and sampling features, with alignments to existing models. *Semantic Web*, vol. 8, n° 3, p. 453-470, doi: 10.3233/SW-160214.
- Descola P. (2017) *La composition des mondes*, Champs essais, Paris.
- De Veaux R. D., College W. et Velleman P. F. (2008). Math is music; statistics is literature – or Why are there no six year old Novelists? *Amstat News*, Cornell University, 2008.
- Gruber T. R. (1995). Toward principles for the design of ontologies used for knowledge sharing? *International Journal of Human-Computer Studies*, vol. 43, n° 5-6, p. 907-928, doi: 10.1006/ijhc.1995.1081.
- Guarino N (1998). Formal ontology in information systems. *Proceedings of FOIS'98*, Guarino N. (Ed.), vol. 46, Trento, Italy. IOS Press, p. 3-15.
- Harbelot B., Arenas H., Cruz C. (2014). A semantic model to query spatialtemporal data. *Information Fusion and Geographic Information Systems* (IF and GIS 2013),

- V. Popovich, C. Claramunt, M. Schrenk et K. Korolenko (Ed.), Springer Berlin Heidelberg, Berlin, Heidelberg, p. 75-89. doi: 10.1007/978-3-642-31833-7_5
- Hess C., Ostrom E. (2006). *Understanding Knowledge as a Commons*, MIT Press.
- Holmes H. (2001). The great term robbery [computer jargon]. *Computer*, vol. 34, n° 5, p. 94-96, doi: 10.1109/2.920619.
- Isaac A. H. J. C. A. (2011). Entre thésaurus et ontologies : une affaire d'interopérabilité et d'alignement. *Documentaliste - Sciences de l'information*, 48.
- Krivine H. (2016). *Petit traité de Hasardologie*, Cassini, Paris.
- Latour B. (1993). Le "pédofil" de Boa-Vista : montage photo-philosophique. *La clef de Berlin. Petites leçons de sociologie des sciences*, Paris, La Découverte, p. 171-225.
- Livet P., Muller J.-P., Phan D. et Sanders L. (2010). Ontology, a mediator for agent-based modeling in social science. *Journal of Artificial Societies and Social Simulation*, vol. 13, n° 1, p. 3. DOI: 10.18564/jasss.1538
- Loireau M. et Fargette M. (2022). Science paysagère au service de l'observatoire scientifique Sociétés-Milieus en appui à la gestion territoriale. *Journal of Interdisciplinary Methodologies and Issues in Sciences*, vol. Scientific observatories..., n° Geographical Information..., p. 8762. doi: 10.46298/jimis.8762.
- Madera C. (2018). *L'évolution des systèmes et architectures d'information sous l'influence des données massives : les lacs de données*. Thèse en informatique, Université Montpellier. <https://tel.archives-ouvertes.fr/tel-02138983>
- Madin J., Bower S., Schildhauer M., Krivov S., Pennington D. et Villa F. (2007). An ontology for describing and synthesizing ecological observation data. *Ecological Informatics*, vol. 2, n° 3, p. 279-296, doi: 10.1016/j.ecoinf.2007.05.004.
- Martinez-Cruz C., Blanco I. J. et Vila M. A. (2012). Ontologies versus relational databases: Are they so different? A comparison. *Artificial Intelligence Review*, vol. 38, n° 4, p. 271-290.
- Masolo C., S. Borgo A. Gangemi N. Guarino, A. Oltramari et L. Schneider (2003). The WonderWeb Library of Foundational Ontologies and the DOLCE ontology. WonderWeb Deliverable D18, Final Report (vr. 1.0, 31-12-2003).
- Noucher M. F. Gourmelon et C. Claramunt (2019). Pour un observatoire des données géographiques du Web : Expérimentation à partir des infrastructures de données géographiques françaises. *Revue internationale de géomatique*, vol. 29, n° 1, p. 9-30, janvier, doi: 10.3166/rig.2019.00074.
- Ollion É. et Boelaert J. (2015). Au delà des big data. Les sciences sociales et la multiplication des données numériques. *Sociologie*, vol. 6, n° 3, p. 295-310.
- Prieur M. (1999). La Convention d'Aarhus, instrument universel de la démocratie environnementale. *Revue juridique de l'Environnement*, vol. 24, n° 1, p. 9-29, doi:10.3406/rjenv.1999.3592.
- Rodgers J. L. (2010). The epistemology of mathematical and statistical modeling: A quiet methodological revolution. *American Psychologist*, vol. 65, n° 1, p. 1-12, doi:10.1037/a0018326.

- Smith B. (2003). Ontology. *Blackwell Guide to the Philosophy of Computing and Information*, Floridi L. (Ed.), Blackwell, Oxford, p. 155-166.
- Tran B.-H., Plumejeaud-Perreau C., Bouju A., et Bretagnolle V. V. (2015). A semantic mediator for handling heterogeneity of spatio-temporal environment data. *Metadata and Semantics Research*, Manchester, United Kingdom, doi: 10.1007/978-3-319-24129-6.
- Ushold M. et Gruninger M. (1996). Ontology: Principles, methods and applications. *Knowledge Engineering Review*, vol. 11 n° 2, p. 93-126. doi:10.1017/S0269888900007797
- Wasserstein R. L. et Lazar N. A., (2016). The ASA Statement on p -Values: Context, process, and purpose. *The American Statistician*, vol. 70, n° 2, p. 129-133. <https://doi.org/10.1080/00031305.2016.1154108>.

Christine PLUMEJEAUD-PERREAU
Migrinter, UMR 7301 CNRS &
univ. de Poitiers

Julie GRAVIER
UMR 8504 Géographie-cités

Éric MASSON
Lab. TVES ULR 4477, univ. de Lille

Lucie NAHASSIA
UMR 8504 Géographie-cités

Xavier RODIER
CNRS/univ. de Tours -
CITERES/MSH Val de Loire

Éric SAUX
Institut de recherche de l'École
navale, BCRM Brest

Mireille FARGETTE
IRD, UMR 228 Espace-Dev, Montpellier

Thérèse LIBOUREL
Univ. de Montpellier, UMR 228
Espace-Dev

Hélène MATHIAN
UMR 5600 Environnement Ville Société

Laure NUNINGER
CNRS/univ. of Bourgogne Franche-
Comté ; Chrono-Environnement/MSHE

Lena SANDERS
UMR 8504 Géographie-cités

Numérique *versus* symbolique

Dialogue ontologique entre deux approches

Hélène Mathian¹, Lena Sanders²

1. UMR 5600 Environnement Ville Société ; 2. UMR 8504 Géographie-cités

RÉSUMÉ. L'objectif de cet article est de comparer une approche statistique, l'analyse des données (AD) et une approche de simulation, les systèmes multi-agents (SMA). Ces deux familles de méthodes sont a priori considérées comme représentatives d'une approche numérique, respectivement symbolique, de la modélisation spatiale. Le cas d'application qui est mobilisé tout au long de l'article est celui de la ségrégation de l'espace scolaire en Île-de-France. En premier lieu sont explicitées et discutées les différentes étapes menant d'une question thématique à l'opérationnalisation d'une méthodologie d'analyse statistique ou de simulation destinée à analyser cette question. Pour effectuer cette comparaison, on développe un cadre conceptuel à l'interface entre les deux, qui permet de vérifier la compatibilité entre les arrières plans théoriques associés aux domaines thématiques et de modélisation en jeu. Ce cadre conceptuel prend appui sur une démarche ontologique qui est ensuite présentée. Celle-ci permet d'identifier les complémentarités entre AD et SMA et de montrer comment ces deux méthodes peuvent dialoguer dans le cadre d'une même recherche. Nous montrons combien les aspects numériques et symboliques sont finalement étroitement imbriqués au sein même de chacune de ces méthodes. Cette imbrication permet de construire une « spirale d'interactions » entre les deux familles de méthodes dont l'intérêt est illustré par les va et vient entre les phases d'analyse de structure et de simulation dynamique dans le cas de la ségrégation scolaire.

ABSTRACT. The aim of this article is to compare a statistical approach, "geometric data analysis" (GDA), and a simulation approach, the multi-agent systems (MAS), considered as representative, respectively, of a numerical and a symbolic approach of modelling. The case study concerns segregation of scholar space in the Parisian area. First the different steps leading from a thematic question to the development of an operational model to analyze this question are presented. The central and essential role of a conceptual framework at the interface of both is shown. Indeed, before operationalisation, it is necessary to verify the compatibility between the theoretical backgrounds associated to the thematic hypotheses and the model considered. An ontological approach is then presented and used to compare GDA and MAS in order to identify their complementarities and show how these approaches can dialogue in the same research. The close interweaving between numerical and symbolic aspects in each of these approaches is shown. This leads to the construction of a "spiral of interactions" between GDA and MAS which interest is illustrated by the back and forth between modeling structure and dynamics in the case of scholar segregation.

MOTS-CLÉS : ontologie, numérique, symbolique, analyse des données, système multi-agents, ségrégation scolaire.

KEYWORDS: Ontology, Numerical, Symbolic, Data analysis, Multi-agent system, School segregation.

DOI:10.3166/RIG.31.21-45 © 2022 Lavoisier

1. Introduction

L'objectif commun des articles de ce numéro spécial est de mettre en évidence les formes de complémentarité entre les approches symboliques et numériques. Dans cet article, cet objectif prend la forme d'une comparaison entre les hypothèses, prérequis et production de connaissances apportés par une approche statistique (l'analyse des données) et une approche modélisatrice à base d'agents (les systèmes multi-agents), en prenant appui sur une approche ontologique. Dans les deux cas il s'agit de *représenter/modéliser* un phénomène d'intérêt pour décrire comment il est structuré et comprendre les processus qui sous-tendent son organisation dans l'espace. La comparaison de l'analyse des données (AD) et des systèmes multi-agents (SMA) est particulièrement stimulante dans la mesure où chacune de ces familles de méthodes a créé, lors de son introduction et diffusion au sein de la recherche en sciences sociales, un fort engouement et quasiment amené un changement de paradigme dans la modélisation dans ces domaines. Dans les années 1970, l'AD a été adoptée avec enthousiasme car elle permettait, grâce à l'ordinateur, de traiter des données abondantes issues d'enquêtes et d'observations de terrain, de mettre en évidence des régularités et des structures, sans que des hypothèses *a priori* sur les distributions statistiques de ces données ne soient imposées. Dans les années 1990, ce sont les SMA qui ont rapidement diffusé, leur succès étant dû à la possibilité d'explorer par la simulation les processus sous-jacents aux phénomènes étudiés. Ils sont fondés sur un formalisme à base de règles et permettent d'approcher les moteurs d'un changement socio-spatial. Ces deux familles de méthodes ont donc des objectifs bien différenciés, et de ce fait, leur complémentarité peut paraître d'emblée évidente. En revanche, pour évaluer dans quelle mesure ces deux approches peuvent dialoguer et se compléter sur un même cas d'étude, une comparaison de leurs fondements épistémologiques est nécessaire. Suivant la compatibilité de ces fondements, il sera en effet plus ou moins aisé de cheminer d'une approche à l'autre. La démarche la plus simple consisterait à simplement juxtaposer les résultats de chacune d'entre elles lors d'une étape finale. Une démarche plus ambitieuse repose sur un va et vient entre ces deux approches tout au long des étapes du travail de recherche. Une compatibilité épistémologique forte est alors nécessaire pour que ce va et vient prenne sens à chaque niveau où il est opéré, qu'il s'agisse, en amont, de celui des entités élémentaires en entrée de la modélisation, ou de celui, en aval, auquel les résultats de la modélisation peuvent être interprétés. A chacun de ces niveaux, ainsi qu'aux niveaux intermédiaires, nous montrons que les deux qualités « symboliques » et « numériques » interfèrent.

Pour comparer ces deux approches nous commençons par discuter des caractères numérique et symbolique de chacune d'entre elles en prenant appui sur un cas d'application, celui de la ségrégation de l'espace scolaire en Île-de-France, cas qui est mobilisé tout au long de cet article (section 1). Dans un deuxième temps, il s'agit de montrer comment sont articulées les différentes étapes constituant la chaîne de traitements à effectuer face à un questionnement thématique. A une extrémité de cette chaîne se trouvent les données et les connaissances empiriques, à l'autre,

l'opérationnalisation d'une méthodologie d'analyse statistique ou de simulation. Nous montrons notamment l'intérêt d'explicitier le cadre conceptuel de modélisation qui permet de joindre ces deux extrémités, et évoquons pour cela la place d'une démarche ontologique. L'ontologie utilisée, proposée par Pierre Livet dans la continuité des travaux de Barry Smith (2003), est ensuite présentée. Elle a été développée pour expliciter quels sont les éléments nécessaires pour étudier un phénomène spatiotemporel et modéliser son évolution, en s'assurant de la bonne compatibilité entre les concepts, les données empiriques et le cadre méthodologique choisi. Cette ontologie est utilisée ici pour comparer les deux familles de méthodes mobilisées, l'analyse de données en statistiques et les systèmes multi-agents dans le domaine de la simulation informatique (section 2). La dernière étape consiste à examiner les formes de complémentarité et de dialogue possibles entre AD et SMA pour construire de nouvelles connaissances. Le cas de la ségrégation de l'espace scolaire donnera lieu au développement de différents cadrages ontologiques et servira à illustrer comment une « spirale d'interactions » peut être construite entre les différents niveaux associés aux deux approches mobilisées (section 3).

1. Symbolique *versus* numérique, de quoi parle-t-on ?

Une première étape consiste à justifier en quoi nous considérons *a priori* ces deux familles de méthodes comme des représentantes respectives des approches numérique et symbolique, afin de pouvoir évaluer, ensuite, les implications de cette différence. D'après les définitions du CNRTL¹, le numérique « concerne des nombres, se présente sous la forme de nombres ou de chiffres, ou concerne des opérations sur des nombres » alors que le symbolique « utilise des symboles, procède par symboles ». Dans son article « Introduction à l'approche symbolique en analyse des données », Edwin Diday (1989) précise ces notions à propos des « objets » :

« La distinction entre objets "symboliques" et "numériques" est claire dès lors que l'on considère qu'un objet est "numérique" s'il peut être représenté et utilisé comme un point de l'espace R_p considéré comme un espace vectoriel muni des opérations habituelles et qu'il est "symbolique" si ce n'est pas le cas (autrement dit, s'il est nécessaire de définir une sémantique propre au domaine d'application car la sémantique des nombres ne convient pas). »

Son propos concerne l'analyse des données, méthode classiquement utilisée pour traiter des données numériques et qu'il propose d'étendre au traitement de données plus complexes. Il évoque ainsi des objets décrits par des données composées non de simples modalités associées à des variables quantitatives ou qualitatives, mais de ce qu'il nomme une « conjonction de propriétés ». Tel est le cas, par exemple, lorsqu'une variable se présente sous la forme d'un intervalle de valeurs ou d'un

1. CNRTL (Centre national de ressources textuelles et lexicales) : www.cnrtl.fr

ensemble de modalités jointes par la conjonction « ou ». Dans l'application qui est développée par la suite, à savoir la ségrégation scolaire, la variable « établissement fréquenté par l'élève » est de type simple, qualitative, et peut donner lieu à des comptages classiques. En revanche, la variable « établissement préféré par l'élève » qui est primordiale dans la dynamique de choix des établissements, peut s'exprimer sous la forme {collège Picasso ou collège Ravel}, introduisant une information de forme plus complexe.

L'étude de la ségrégation scolaire est propice pour mobiliser à la fois des analyses de nature numérique et symbolique et les données dont nous disposons relèvent également de ces deux types :

- des données issues de la DEPP (Direction de l'évaluation, de la prospective et de la performance du ministère de l'Éducation nationale) portant sur les élèves des collèges de l'Île-de-France, notamment l'établissement fréquenté par l'enfant et la catégorie sociale de ses parents. Ces données sont fondamentalement numériques, même si les catégories initiales concernées sont qualitatives. Leur croisement permet par exemple de construire un tableau de contingence², cas le plus classique d'application de l'analyse des correspondances (AFC).

- des informations et des connaissances sur le comportement des élèves et des chefs d'établissements, en matière de choix d'établissement pour les premiers, de stratégies de recrutement des élèves pour les seconds. Ces informations, construites à partir d'entretiens (Poupeau et François, 2008), permettent par exemple de comprendre la forme des préférences des uns et des autres, informations qui peuvent être formalisées sous forme symbolique par des règles.

La nature des données étant précisée, il s'agit de s'interroger sur la forme numérique ou symbolique des traitements qui leur sont appliqués ainsi que sur celle dans laquelle sont exprimés les résultats obtenus. Diday propose ainsi de distinguer quatre cas, reposant sur le croisement des propriétés des données et de celles des traitements. Le cas le plus classique consiste à appliquer une analyse numérique sur des données numériques, cas typique des analyses factorielles (Diday, 1989 ; Balbo *et al.*, 2009). En AD, les opérations classiquement effectuées sur un tableau de contingence reposent ainsi sur des calculs de fréquences (en lignes ou en colonnes) ou de fonctions de fréquences (pour calculer, par exemple, la similarité entre les profils de lignes ou de colonnes d'un tableau de contingence). Les sorties sont sous la forme de valeurs numériques, comme par exemple les coordonnées sur les axes factoriels. Dans ce cas, les entrées, les opérations et les sorties sont toutes sous forme numérique.

Dans notre contribution, l'approche symbolique considérée est la simulation avec systèmes multi-agents (SMA), approche pour laquelle tant les données que les analyses qui en sont faites, peuvent prendre une forme relativement complexe, non résumable sous une forme numérique. Dans cette approche, en effet, les objets

2. A l'intersection de chaque ligne i et colonne j du tableau figure le nombre d'élèves de la catégorie sociale j fréquentant l'établissement i .

considérés sont localisés dans l'espace et des actions conduisant à des changements sont simulées à chaque pas de temps au niveau des objets, en fonction des caractéristiques de son environnement et des échanges qui s'opèrent avec des objets situés en d'autres lieux. L'expression de tels mécanismes nécessite un formalisme symbolique. Les principales opérations se présentent sous forme de règles qui déterminent les actions d'agents, en fonction de leur situation initiale, de leurs propriétés et des interactions qu'ont ces agents entre eux et avec leur environnement. Ces règles s'expriment souvent sous la forme symbolique « Si... Alors... Sinon... » qui permet par exemple de traduire qu'un élève de telle catégorie aura une préférence pour un établissement ayant un certain profil social. Quant aux sorties, elles peuvent être exprimées sous forme symbolique (tel élève résidant dans tel contexte et préférant tel établissement est affecté à tel autre) ou numériques (nombre d'élèves de telle catégorie sociale par établissement).

Le caractère *a priori* numérique *versus* symbolique de ces deux approches étant établi, l'objectif est d'en comparer les fondements épistémologiques en prenant appui sur une approche ontologique.

2. Un cadre ontologique pour comparer AD et SMA

Avant de mettre en œuvre la comparaison entre ces deux approches, nous explicitons l'ensemble des étapes qui rythment tout travail de modélisation, qu'il soit statistique ou informatique, en mobilisant la démarche proposée par Denis Phan (2014) qui consiste à expliciter le cadre conceptuel permettant de mettre en lien le domaine thématique et celui de la modélisation. Ce cadre conceptuel qui varie suivant les questionnements et les méthodes choisis est en effet souvent implicite. L'enjeu consiste à l'explicitier pour l'AD et les SMA. L'ontologie occupe une place centrale dans ce cadre conceptuel où elle joue un rôle de « médiation ». Nous en proposons donc une définition que nous appliquons ensuite pour comparer les approches par AD et par SMA.

2.1. Un cadre conceptuel pour lier un phénomène thématique à sa modélisation

Si on se situe du point de vue du thématicien, le point de départ d'une recherche est une question thématique. Il s'agit dans notre cas de comprendre comment émerge et s'organise une ségrégation socio-spatiale dans le domaine scolaire. Le choix d'une méthode est ensuite opéré en fonction des savoir-faire et habitudes des chercheurs, de la nature des données disponibles et des particularités de la question posée. La mise en lien de la question, des données et de la méthode peut être réalisée de différentes façons et l'absence d'explicitation des choix qui sont faits peut être source de malentendus dans un contexte où des chercheurs de disciplines différentes collaborent. La figure 1, adaptée de Phan (2014), représente les positionnements respectifs des différents domaines impliqués dans une démarche modélisatrice : à un

bout est représenté le « domaine de référence empirique » au sein duquel est formulée la question de recherche, à l'autre celui relatif aux méthodes de traitements des données associées au domaine d'intérêt, qu'elles relèvent des statistiques ou de l'informatique. Le rectangle central est le lieu du dialogue interdisciplinaire et de la mise en cohérence entre le « domaine de référence empirique » ancré dans une thématique précise et le « cadre opérationnel de développement » qui fait référence aux aspects d'opérationnalisation et d'implémentation informatique des traitements.

Le « domaine de référence empirique » (à l'extrémité gauche de la figure 1) est celui du questionnement du chercheur sur un aspect du monde empirique. Il peut s'agir d'un phénomène spatiotemporel, social ou socio-écologique. Le cas qui servira de référence dans cet article concerne l'organisation de l'espace scolaire francilien : cet espace est-il différencié ? Suivant quelle géographie ? Comment émergent ces différenciations ? Quels en sont les mécanismes producteurs ? Tels sont des exemples de questions que se posent les géographes. La nature des questions posées par les chercheurs est influencée par leur formation disciplinaire et celles posées par des sociologues, des économistes ou des psychologues seront différentes, bien que portant sur un même objet, l'espace scolaire. Ce « domaine de référence empirique » est ainsi fortement ancré dans un point de vue particulier (Müller et Diallo, 2012) qui influe sur la manière de construire le « cadre conceptuel thématique » figurant en partie centrale du schéma. Il ne s'agit pas de saisir le phénomène étudié dans toute sa « réalité » mais de mettre en œuvre une approche permettant de cibler les questions posées suivant le point de vue adopté et la nature des données disponibles.

Le « cadre opérationnel de développement » (à l'extrémité droite de la figure 1) correspond à la mise en œuvre de la démarche de modélisation mise au point dans le cadre central avec les outils appropriés, qu'ils soient informatiques, mathématiques ou statistiques. Il s'agit ici d'une part, de l'AD, dans l'esprit où elle a été développée par J.-P. Benzécri, et d'autre part, d'une approche de simulation par SMA. Ces deux approches émanent de cultures différentes, et si elles sont combinées dans certains travaux, c'est en en général la première au service de la seconde (que ce soit pour traiter des données en amont d'une modélisation SMA ou les données issues d'un tel modèle), et elles sont rarement mises en perspective face à un même objectif. Il est de ce fait intéressant de tester le rôle de médiation que peut jouer une approche ontologique pour les faire dialoguer. Chacune de ces approches méthodologiques peut en effet être utilisée pour répondre aux questions posées dans le « domaine de référence empirique » et prend elle-même appui sur un champ théorique bien défini qui réfère au « cadre conceptuel de modélisation » (à droite dans le rectangle central). Il s'agit alors de s'assurer de la compatibilité entre les grandes lignes (contraintes, présupposés) associées à ce champ théorique de modélisation et celles associées au champ disciplinaire correspondant au domaine de référence empirique. Tel est le rôle du « cadre conceptuel » développé en partie centrale de la figure 1.

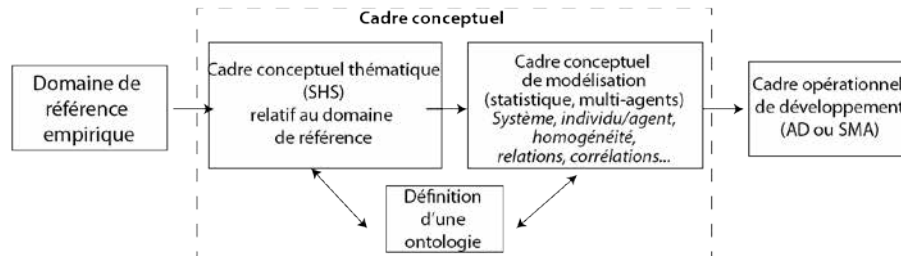


Figure 1. L'ontologie comme médiateur d'un dialogue conceptuel
(d'après Phan, 2014)

Ce « cadre conceptuel » central met ainsi en relation les cadres conceptuels attachés respectivement au champ thématique associé au « domaine de référence empirique » étudié et au champ de modélisation qui sera opérationnalisé lors de l'application (respectivement l'AD et les SMA). Le cadre conceptuel thématique concerne la manière dont l'espace scolaire est appréhendé. Les questions formulées plus haut renvoient à des différenciations socio-spatiales qu'il s'agit de mettre en évidence et d'en comprendre l'émergence. Cette question est abordée différemment suivant le positionnement épistémologique des chercheurs. François *et al.* (2014) ont ainsi comparé, face à cette même question, les approches de différents chercheurs, les uns prenant appui sur l'individualisme méthodologique et prônant une « rationalité intentionnelle » des élèves comme des chefs d'établissements dans leurs choix respectifs, les autres mettant en avant l'effet des structures sociales sur les choix des élèves et chefs d'établissement. En sciences sociales, ces deux visions correspondent respectivement à celles issues des travaux de Boudon (1973) et de Bourdieu (1979). Les facteurs explicatifs, qui seront ensuite conceptualisés dans le « cadre conceptuel de modélisation », vont ainsi différer suivant le cadre conceptuel thématique adopté. Dans le premier cas, le comportement des agents, élèves comme chefs d'établissement, renvoie à un modèle décisionnel de type micro-économique alors que dans le second ces comportements dépendent des structures spatiales et sociales (François *et al.*, 2014). Avant d'analyser comment ces deux positions épistémologiques différentes peuvent s'insérer dans une modélisation par AD ou SMA, il est nécessaire d'examiner à leur tour les implications théoriques de chacune de ces familles de modèles.

Le « cadre conceptuel de modélisation » associé aux SMA implique une formalisation où les concepts de système complexe, d'émergence, d'agents et d'interactions sont centraux. Les entités élémentaires modélisées sont ainsi des agents qui interagissent entre eux et avec leur environnement, ces interactions entraînant l'émergence d'une structure observable à un niveau d'organisation supérieur, que ce soit celui des établissements ou de l'espace géographique. Les

éléments constituant la ségrégation de l'espace scolaire peuvent s'exprimer et faire sens dans ce cadre conceptuel. Les agents sont alors de deux sortes, représentant respectivement les élèves et les chefs d'établissement. Ces agents interagissent, et les interactions peuvent se faire entre d'une part, les agents-élèves à travers des mécanismes d'imitation et d'autre part, les agents-élèves et les agents-chefs d'établissement à travers des mécanismes de préférences (choix d'établissement pour l'agent-élève, choix d'élèves pour l'agent chef d'établissement). Les structures qui émergent de ces interactions sont observables au niveau d'une part, des établissements, à partir de leur composition sociale, et d'autre part, de l'espace étudié (l'Île-de-France dans notre cas d'application), à partir de la répartition des établissements eux-mêmes dans l'espace en fonction de leur profil social. Le cadre de modélisation des SMA est ainsi parfaitement compatible avec le cadre conceptuel thématique à ce niveau très général. Cette compatibilité vaut aussi bien, que celui-ci s'inscrive dans l'individualisme méthodologique, ou bien dans une approche structurelle qui met en avant l'effet des structures sociales et spatiales sur les choix des individus, tel que cela a été décrit plus haut à propos du cadre conceptuel thématique. Si en revanche le formalisme de modélisation choisi avait été les équations différentielles, ceci aurait impliqué de formaliser des règles d'évolution directement au niveau d'entités agrégées³. Il y aurait eu une rupture entre les ancrages épistémologiques en jeu, alors que cette même approche aurait pu être parfaitement adaptée face à un autre questionnement thématique.

Le cadre conceptuel de l'AD est quant à lui attaché au champ de la statistique exploratoire, domaine que J.-P. Benzécri a développé dans les années 1970, en prenant ses distances relativement à la statistique classique reposant sur les méthodes inférentielles. Dès la première page de son livre sur l'*Analyse des correspondances* (Benzécri, 1973), il énonce son premier principe sur l'analyse des données, « Statistique n'est pas probabilité », critiquant par là la statistique en tant que discipline « riche en hypothèses qui ne sont jamais satisfaites dans la pratique. » Accordant une forte importance aux faits empiriques, il souligne : « La conception probabiliste est-elle conforme aux phénomènes étudiés ? La connaissance que nous avons de ces phénomènes permet-elle de faire des calculs ? ». Sa critique concernait tout l'arrière-plan théorique associé à la statistique inférentielle, présupposant que les variables d'intérêt suivent des lois bien définies, en général gaussiennes, alors que la réalité peut être toute autre. Il apparaît ainsi que le point de départ sur lequel il base sa conception de l'analyse des données est celui d'une incompatibilité entre le cadre conceptuel des statistiques mathématiques (avec leurs lois et leurs tests d'hypothèses) et celui des faits auxquels on s'intéresse qui sont décrits par de larges ensembles de données. En reprenant les termes de la figure 1, cette critique correspond ainsi à une incompatibilité entre le « cadre conceptuel thématique »

3. C'est par exemple le cas d'un modèle de croissance d'une population (P) en fonction du temps lorsque la capacité de l'environnement (C) est limitée. La croissance de la population suit alors une courbe logistique qui correspond à l'équation : $dP/dt = a P (1 - P/C)$ où « a » est un paramètre.

(relatif aux « phénomènes étudiés » de Benzécri) et le « cadre conceptuel du modèle » (relatif aux lois de probabilité et « hypothèses qui ne sont jamais satisfaites dans la pratique »). Le second principe qu'il énonce, « Le modèle doit suivre les données, non l'inverse », rend compte de son choix pour y remédier. Il consiste à réduire au minimum les contraintes que pourraient représenter le « cadre conceptuel de la modélisation ». Il s'agit de faire émerger (de rendre visible) la structure qui existe au sein des données, invisible en première analyse de par leur masse importante. Les seules contraintes qu'il préconise d'imposer et qui s'inscriraient dans le carré « cadre conceptuel de modélisation » sont celles de l'exhaustivité (il s'agit de prendre en compte toutes les dimensions caractérisant le phénomène d'intérêt) et de l'homogénéité de la nature de ces données. Le tableau de contingence constitué des comptages d'élèves appartenant aux différentes catégories sociales et fréquentant les différents établissements vérifie la propriété d'exhaustivité à partir du moment où l'ensemble des établissements franciliens sont considérés et où la nomenclature des catégories sociales est complète et cohérente. Les entités élémentaires comptées, les élèves, sont par ailleurs parfaitement homogènes de par leur statut même qui implique une certaine classe d'âge et l'exercice d'une pratique scolaire ainsi que par le caractère centralisé et harmonisé des données issues de la DEPP.

Cette analyse rapide des cadres conceptuels de modélisation associés à l'AD et aux SMA illustre la diversité des présupposés associés à une méthode de modélisation donnée, l'une faisant référence aux systèmes complexes et l'autre à une méthode exploratoire multidimensionnelle. Dans les deux cas cependant, une interrogation sur la nature et le sens des entités en jeu s'impose. Une approche ontologique est de ce fait appropriée pour aller plus loin dans la comparaison et l'articulation des deux méthodes de modélisation.

2.2. Un cadre ontologique pour comparer AD et SMA

Nous avons choisi de mobiliser la définition de Pierre Livet (2010) pour qui :

« l'ontologie consiste à analyser un domaine, en identifiant les entités pertinentes (objets, propriétés, relations, événements et processus), et les opérations qui peuvent être opérées sur ces entités. »

Cette définition se place au croisement de celle du philosophe Barry Smith⁴ et de l'informaticien Grüber⁵. Elle est centrée sur l'explicitation des entités qui caractérisent le domaine d'intérêt en y ajoutant un élément particulièrement

4. L'ontologie est « la science de ce qui est, des types et des structures des objets, propriétés, événements, processus et relations en tout domaine de la réalité.../... de ce qui pourrait exister » (Smith, 2003).

5. L'ontologie est « une spécification de la conceptualisation d'un domaine donné » (Grüber 1993).

intéressant quand il s'agit de lier le cadre conceptuel thématique à celui de la modélisation (cf. figure 1). Il s'agit des « opérations » qui peuvent être menées sur les entités en jeu. Cette approche de l'ontologie permet de s'assurer de la pertinence du passage du cadre conceptuel thématique au sein duquel est formulé le questionnement du chercheur en sciences sociales (ici sur la ségrégation de l'espace scolaire) au cadre conceptuel de la modélisation. La notion de « test ontologique » (Livet et Sanders, 2014), fondée sur l'opérationnalisation de la définition présentée plus haut, permet de fluidifier la circulation entre ces cadres conceptuels et de vérifier leur compatibilité lors d'une application.

Dans la suite de cette section nous prenons appui sur cette définition de l'ontologie pour comparer les deux cadres conceptuels de modélisation, l'AD d'un côté, les SMA de l'autre. Dans un premier temps nous explicitons quelles sont les *entités* en jeu, puis nous examinons quelles sont les *opérations* appliquées à ces entités. Les deux premiers termes de la définition font référence aux « objets » et aux « propriétés ». D'un point de vue thématique, les objets correspondent aux entités élémentaires qu'il s'agit d'observer et d'étudier et les propriétés à leurs caractéristiques. Examinons comment « objets » et « propriétés » sont appréhendés en AD et dans les SMA.

De manière très générale, les objets traités dans le cadre de l'AD sont de nature très variée (arbres, bâtiments, humains, établissements scolaires, villes) mais c'est toujours en nombre important que ces objets sont intéressants à analyser. C'est en effet à partir de leur diversité que l'AD peut extraire une structure qui fait sens, la faire émerger dans « l'océan des faits ». Pour Gibrat leur objectif est effectivement de « traiter simultanément de grands ensembles de faits et les confronter en vue de découvrir l'ordre global », le terme de « fait » référant au couple (individu, variable) dans le vocabulaire statistique, et correspondant ici au couple (objet, propriété). Lors de l'application d'une analyse des correspondances (AFC) au cas scolaire, les objets réfèrent d'emblée à deux échelles :

- celle élémentaire des « individus statistiques » (les élèves), qui sont les entités de base sur lesquelles se font les observations et le recueil des données ;
- celle des agrégats de ces individus statistiques au niveau des établissements scolaires.

Dans les SMA, les objets font référence aux entités de base que sont les « agents », qui ont une représentation de l'environnement dans lequel ils sont situés ainsi que des autres agents qui s'y trouvent, et qui ont la capacité d'agir. Alors que dans le cas de l'AD il s'agit toujours d'un objet qui peut être observé, qu'il s'agisse du résultat d'une expérimentation ou d'un fait existant dans le domaine empirique, dans celui des SMA, l'objet est plutôt un concept, un construit référant à un objet existant dans le domaine empirique mais en en représentant une abstraction. L'agent-élève représente ainsi un élève qui fait des choix suivant certaines règles et dont les actions dépendent à la fois de ses caractéristiques (« propriétés ») et de celles de son environnement. Il en est de même pour l'agent-établissement.

Dans le domaine statistique les propriétés font référence aux « variables » qui décrivent les « individus statistiques », c'est-à-dire les objets étudiés. En AD ces variables sont la plupart du temps nombreuses afin de cerner de la manière la plus exhaustive possible le domaine auquel on s'intéresse. Dans une analyse en composantes principales (ACP) les statuts respectifs des variables et des individus statistiques sont clairement distingués alors que dans le cas d'une analyse des correspondances (AFC) appliquée sur un tableau de contingence, les lignes et colonnes du tableau de données ont des statuts plus neutres et sont interchangeables. C'est l'interprétation qui introduit des différences sémantiques en termes d'individus statistiques (objets) et de variables (propriétés). Dans le cas, par exemple, du tableau de contingence résultant du croisement de l'établissement scolaire fréquenté par les élèves et de la catégorie sociale de leurs parents, le géographe tendra à considérer les établissements scolaires comme des individus statistiques et les catégories sociales comme des variables alors que l'inverse aurait tout autant du sens d'un point de vue statistique et ne changerait pas les résultats des calculs effectués. Le rôle des propriétés est différent dans le cadre des SMA où elles sont généralement en plus petit nombre. Il s'agit le plus souvent de quelques caractéristiques qui sont susceptibles de différencier les comportements des agents et il est intéressant pour le modélisateur de repérer d'éventuels effets seuils pouvant conduire le système à évoluer différemment suivant ces comportements.

Le terme suivant de la définition de l'ontologie mobilisée concerne les « relations ». Dans l'AD de Benzécri il s'agit de « découvrir sans parti pris, sans a priori, quels courants de lois traversent l'océan des faits » (Benzécri, 1976, cité par Cibois, 1999), ou dit autrement, de faire émerger une structure à partir des relations qui existent entre ces faits. Dans une ACP ces relations sont simplement mesurées par les corrélations entre les variables (quantitatives) caractérisant les individus statistiques. Dans une AFC, elles font référence aux similarités respectives entre les profils des lignes et des colonnes du tableau de données. Il s'agit ainsi de mesurer le degré de covariation des différentes colonnes ou des différentes lignes. Dans les deux cas (ACP et AFC) c'est une vision statistique (numérique) des relations entre des distributions qu'il faut identifier et rendre visible à partir de l'analyse des données. Dans les SMA en revanche, les relations sont d'une toute autre nature et renvoient aux interactions entre les agents et leur environnement ou entre les agents eux-mêmes. Ces interactions ne sont pas le *résultat* d'une analyse. Elles jouent au contraire un rôle moteur dans le fonctionnement de la simulation et s'expriment en général sous une forme symbolique. L'imitation, amenant un agent à agir de la même façon qu'un autre avec lequel il interagit, est un exemple simple de ce type de relation.

Les dernières entités relatives à la définition de l'ontologie concernent les événements et les processus.

Le fonctionnement des SMA repose sur la mise en œuvre de processus *bottom-up* : les mécanismes régissant les interactions entre les agents ou entre les agents et

leur environnement sont formalisés par des règles opérant au niveau même des agents. Au cours de la simulation les actions des uns vont influencer sur les actions des autres et de l'ensemble de ces interactions va émerger une ou des structures observables à un niveau d'organisation supérieur. Dans le cas scolaire (François *et al.*, 2014), il s'agit par exemple de l'émergence d'une forme d'organisation spatiale plus ou moins ségréguée des établissements scolaires en fonction de leur profil social. Ces règles qui constituent le moteur de la dynamique du système modélisé sont au cœur même d'une formalisation par SMA. La dimension temporelle est ainsi intrinsèque à l'approche, chaque agent ayant la capacité d'agir à chaque pas de temps. En revanche, pour rendre compte d'un changement dans le cadre de l'AD, celui-ci doit être inscrit dans les données mêmes. C'est à travers elles seulement qu'un processus pourra être appréhendé. On peut ainsi avoir des individus statistiques associés à des dates (par exemple, lorsque ces individus sont des établissements scolaires, le collège Ravel en 2000, 2002 et 2004 peuvent constituer trois individus statistiques sur lesquels on récolte des données) et l'analyse permet d'identifier la trajectoire de cette entité dans les plans factoriels. Une autre manière de procéder consiste à construire des variables porteuses du changement (taux d'évolution par exemple) et l'analyse permet alors d'identifier les structures du changement. Alors que dans un SMA, il s'agit de formaliser le moteur même du processus grâce à des règles qui conduisent à une transformation des entités à chaque pas de temps de la simulation, dans le cas de l'AD il s'agit de rendre visible comment se structurent des changements les uns par rapport aux autres.

Une fois les différentes « entités » décrites, il reste à détailler à quoi correspondent les « opérations ». Celles-ci sont en effet indispensables pour faire « parler » des données ou mettre en œuvre les règles dans une simulation. Il s'agit de l'ensemble des manipulations effectuées sur les entités, d'une part, sur les objets et leurs propriétés dans le but de les analyser (AD) ou de les faire évoluer (SMA), d'autre part, pour mesurer ou exprimer les relations qui existent au sein du système étudié ou enfin pour caractériser la dynamique de ce système. Il peut s'agir de calculs (tableau de contingence, matrice de distance entre lignes et entre colonnes, diagonalisation de matrice, corrélations, calculs de contributions, etc.) ou de la formulation de règles (« si telle propriété et tel contexte, alors telle action de la part de l'agent »). Dans les deux cas, les opérations effectuées font émerger une structure :

- la structure de l'espace des objets et de leurs propriétés dans le cas de l'AD, rendue visible grâce à l'analyse factorielle ;
- la structure issue des interactions entre agents, observable à un niveau d'organisation supérieur dans le cas des SMA.

Dans les deux cas il s'agit de l'émergence d'une organisation, cachée dans les données pour l'AD, aboutissement du déroulé d'une suite d'actions et d'interactions pour les SMA.

3. Une spirale d'interactions entre numérique et symbolique

Après avoir comparé les cadres conceptuels de modélisation associés à l'AD et aux SMA, nous nous proposons de montrer que conceptions numériques et symboliques peuvent être articulées à plusieurs niveaux. D'une part, elles cohabitent au sein même de chacun de ces cadres, le symbolique s'invitant dans l'AD et la modélisation par SMA étant composée de plusieurs étapes numériques (sections 3.1 et 3.2). D'autre part, AD et SMA peuvent être étroitement associés et combinés au sein d'une même recherche (section 3.3).

La question de la ségrégation sociale permet d'illustrer les différentes formes d'articulations qui lient ces différentes approches. Sur un plan thématique il s'agit de tester si la ségrégation sociale au sein du système scolaire reproduit simplement la ségrégation sociale résidentielle ou si elle constitue un lieu de réduction ou d'aggravation de cette dernière. Ce questionnement est par essence dynamique et un enjeu consiste à tester les conditions d'émergence et d'évolution de différentes structures de ségrégations sociales au sein du système scolaire. Il est notamment intéressant d'explorer (avec un SMA) comment une certaine situation initiale (abstraite à partir d'une structure mise en évidence par une AD des données observées) peut être transformée suivant les stratégies mises en place par les différents acteurs du système scolaire (chefs d'établissement et familles).

D'un point de vue conceptuel, le système étudié est composé d'élèves et d'établissements scolaires (environ 1 120 collèges) qui sont considérés comme des « objets » en relation. Nous déclinons successivement les formalisations associées à chacune, en reprenant à dessein les termes utilisés pour la description des ontologies, de manière à les mettre en résonance d'un point de vue conceptuel.

3.1. Articulation des dimensions numérique et symbolique dans l'AD de la ségrégation scolaire

L'analyse des données nécessite de construire un « tableau » décrivant « l'ensemble des objets du système analysé ». Ici le tableau croise les élèves (en ligne) avec leur catégorie sociale (en colonne) et l'établissement fréquenté. L'objet de l'AD (ici une ACM⁶) est de décrire la structure d'association/relation entre les propriétés (caractéristiques) des objets (individus). Dans ce formalisme les n individus sont décrits chacun par une série de valeurs, identifiant par des 0 (absence) et des 1 (présence) sa catégorie d'appartenance pour chacune des deux variables, à savoir la catégorie sociale de sa famille (p_1 modalités) et le collège qu'il fréquente (p_2 modalités). Chaque individu est un point d'un espace de dimension $R^{p_1+p_2}$. Chaque modalité de variable est un point d'un espace de dimension R^n . Les opérations associées sont basées sur le calcul de la distance (chi2) entre les points et

6. Analyse des correspondances multiples.

celui d'un système d'axes (diagonalisation de la matrice d'inertie) permettant de synthétiser au mieux la forme du nuage de points. Une projection du nuage de points (individus et modalités) sur les axes principaux permet de rendre compte de ces proximités et d'identifier la structure globale de différenciation des élèves et par symétrie la structure d'association entre « catégories sociales » et « établissements ». Les catégories sociales les plus fréquemment co-localisées dans les mêmes établissements seront proches dans les plans d'axes et réciproquement les établissements fréquentés par les mêmes « profils sociaux » seront proches.

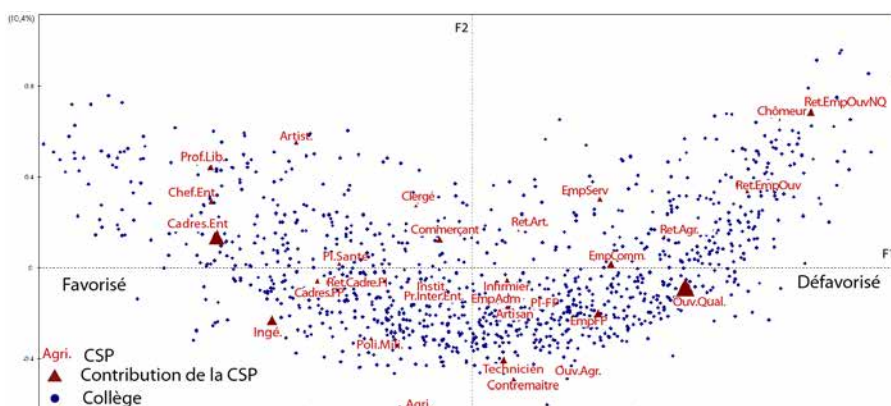


Figure 2. Résultats de l'analyse factorielle des correspondances sur les origines sociales détaillées des élèves des collèges franciliens. Lecture : les points bleus représentent les collèges (1125), les triangles rouges représentent les modalités de la variable « catégories sociales » en fonction de leur poids (effectif). (source : rapport H. Ciesielsky 2012 - Éducation nationale, 2012)

Dans cette analyse, les établissements sont formalisés comme une modalité de variable (un élève appartient ou n'appartient pas à tel établissement). Une propriété majeure de l'ACM et de la distance utilisée pour rendre compte de la structure d'association entre les modalités des catégories est la propriété d'équivalence distributionnelle. Cette propriété induit, entre autres, que le résultat obtenu à partir du tableau décrit ci-dessus serait équivalent à celui qui serait obtenu si l'on travaillait directement sur le tableau de contingence croisant les établissements (en lignes) et les catégories sociales (en colonnes) et qui décrirait donc les « établissements » par le nombre d'élèves dans chacune des catégories sociales, c'est-à-dire par son profil social. La figure 2 illustre le résultat de cette analyse. La différenciation sur le premier plan factoriel montre un grand continuum de situations sociales des établissements (points) s'organisant le long d'un arc dessiné par les différentes catégories sociales (triangles). Le long de cet arc, les catégories s'organisent selon une hiérarchie sociale, illustrant une ségrégation forte. D'un côté de ce continuum on trouve des collèges ayant une surreprésentation d'élèves de catégories telles que « chef d'entreprise », « cadre d'entreprise », « profession

libérale » alors que de l'autre côté on trouve des collèges fréquentés essentiellement par des élèves issus des catégories sociales « ouvriers ».

D'un point de vue numérique, le fait de passer d'« individus-élèves » à des « individus-établissements » relève d'une simple opération d'agrégation. Dans ce contexte, l'établissement est vu comme un contenant du point de vue du « cadre conceptuel de modélisation » (*cf.* figure 1). Le principe d'équivalence distributionnelle souligne bien la neutralité de ce statut. Dans le « cadre conceptuel thématique », en revanche, l'établissement représente bien plus qu'un simple contenant. C'est un objet complexe, lieu d'interactions entre élèves, et entretenant des relations avec les autres établissements par un jeu de concurrence ou complémentarité. Ces éléments figurent sous forme symbolique dans la représentation qu'a le thématicien du système scolaire lors de son interprétation des résultats de l'analyse statistique.

Si l'on s'intéresse maintenant à la ségrégation scolaire d'un point de vue dynamique, l'AD permet de l'explorer non pas en termes de processus, mais d'évolution des structures entre deux dates. Pour ce faire, on peut construire un tableau de données superposant les deux tableaux décrivant les profils sociaux d'un même ensemble d'établissements à chacune des deux dates. Un établissement scolaire apparaîtra ainsi deux fois, comme deux individus distincts dans le plan factoriel, et on pourra interpréter l'espacement entre leurs positions comme reflétant l'évolution entre deux états différents d'un même établissement. Cette évolution s'interprétera par rapport à la structure d'association des catégories sociales, qui définit dans ce cas une structure « moyenne » sur les deux dates. C'est ce qu'illustre la figure 3 où sont identifiés un certain nombre de collèges. Lorsque les points-dates d'un même collègue sont proches, il y a peu de changement entre les deux dates. S'il y a une certaine distance entre eux, alors on interprète le déplacement comme un changement dans la spécificité sociale de cet établissement relativement à tous les autres, qu'il se dirige vers une plus grande diversité de profils (déplacement vers le centre du graphique) ou au contraire une spécialisation plus forte vers les classes défavorisées ou favorisées selon la direction du déplacement. C'est une évaluation numérique (espacement des points) qui permet d'interpréter de manière symbolique le changement entre deux dates.

Ainsi, dans la mise en œuvre d'une AD, alors que les données sont numériques et que la méthode repose sur des opérations numériques, les résultats (sorties) de l'analyse sont symboliques, et cela de deux points de vue. D'une part, c'est par l'interprétation des associations des modalités d'une variable (propriétés) mises en évidence par l'AD que le thématicien peut qualifier la structure de la division sociale (ici le gradient s'ordonne selon une certaine hiérarchie sociale). D'autre part, il va identifier des types d'établissements, au regard de leurs compositions sociales, et pouvoir les confronter aux caractéristiques des espaces résidentiels où ils sont localisés. Le résultat d'une telle analyse s'interprète donc de manière symbolique et « l'analyse des données apparaît ici comme une pratique interprétative, laissant une

place à un raisonnement analogique encadré par une interprétation statistique rigoureuse » (Lebaron, 2010). Partant d'une description numérique, l'analyse permet de mettre à jour des structures d'associations sur lesquelles vont se fonder des catégorisations abstraites, qui ne peuvent se faire qu'à l'interface des domaines conceptuels thématique et de modélisation (figure 1). Le raisonnement en termes de dynamique s'effectue aussi à cette interface comme l'illustre le propos de Bourdieu parlant des résultats de l'ACM : « elle permet de porter au jour la structure des oppositions, ou, ce qui revient au même, la structure de la distribution des pouvoirs et des intérêts spécifiques qui détermine, et explique, les stratégies des agents ». Ainsi dans notre exemple, les établissements sont appréhendés comme des individus (objets), dont la composition sociale est le résultat de processus impliquant différents acteurs et jeux d'intérêts (au niveau des familles comme au niveau des chefs d'établissements). Dans cette approche, la formalisation de la dynamique n'est pas explicite, elle est « portée au jour » par l'interprétation qui est faite du résultat des analyses en mobilisant une connaissance symbolique externe.

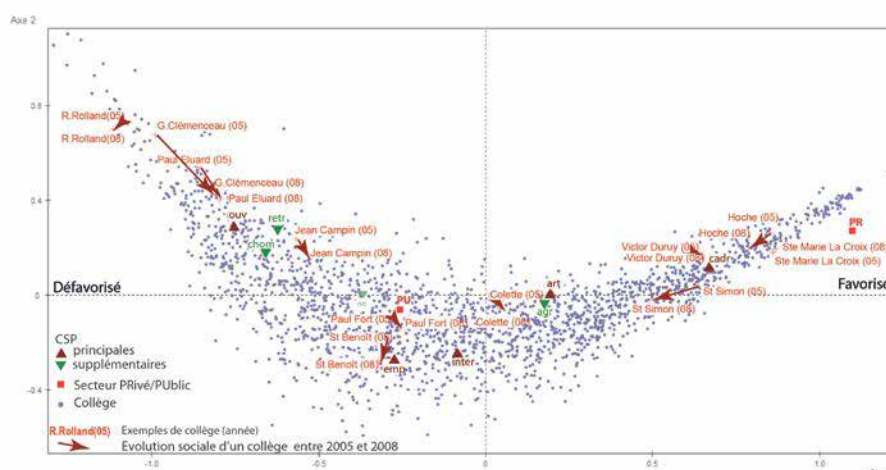


Figure 3. Évolution des profils : résultats de l'analyse factorielle des correspondances (AFC) sur les origines sociales (5 catégories) des élèves des collèges franciliens, à 2 dates, 2005 et 2008 (source : DEPP)

3.2. Articulation des dimensions numérique et symbolique dans un modèle SMA sur la ségrégation scolaire

Un modèle à base d'agents a été conçu afin de mieux comprendre comment émerge un espace scolaire ségrégué (Muller et Diallo, 2012 ; François *et al.*, 2014). Comme exposé en section 2.2, l'approche à base d'agents repose essentiellement sur des connaissances symboliques. Ainsi « le modèle conceptuel du modèle est une

représentation formalisée directe du discours théorique du thématicien » (Muller et Diallo, 2012).

La base du modèle repose sur la définition des agents, objets ayant la capacité d'agir. Ici ce sont les « élèves » et « les chefs d'établissement », agents actifs, et « les établissements » qui sont des agents passifs (figure 4). Les propriétés des élèves sont leur âge, leur capital culturel, leur capital économique, leur niveau scolaire, leur nationalité et leur lieu de résidence. Les chefs d'établissement dirigent un établissement et sont dotés d'une propriété caractérisant leur comportement (« républicain », « élitiste », « pragmatique ») en termes de stratégie de recrutement des élèves. Les établissements, quant à eux, ont des caractéristiques propres : leur capacité d'accueil, leur statut « public » ou « privé », et des caractéristiques qui reflètent la stratégie du chef d'établissement, à savoir la liste des différentes filières proposées. Ils ont également des caractéristiques agrégées à partir du profil des élèves qui le fréquentent (économique, culturel, niveau scolaire et origine des élèves). La figure 4 présente le modèle conceptuel associé. On y retrouve les éléments décrits ci-dessus. Les modalités des propriétés sont des catégories symboliques. Le schéma montre bien les différents jeux de relations entre l'ensemble des entités, qu'il s'agisse d'objets, d'agents ou de propriétés. A un instant donné, un.e élève a un certain profil (économique, culturel, origine et niveau scolaire), et réside en un lieu, connaît l'existence d'un certain nombre d'établissements, en préfère certains et finalement fréquente un établissement donné. Les caractéristiques agrégées des établissements vont varier au cours du temps en fonction des choix et de l'affectation des élèves.

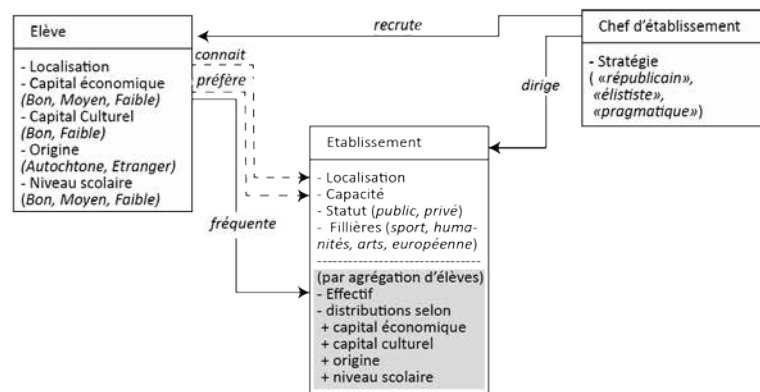


Figure 4. Objets, propriétés et relations de l'application à la ségrégation scolaire

Les relations entre les élèves et les établissements sont définies au niveau du cadre conceptuel thématique du modèle et déterminent la forme que va prendre leurs évolutions respectives. Ainsi la relation « élève fréquente un établissement »

représentée sur le schéma est le résultat d'un processus de décision qui met en relation élèves et chefs d'établissement, et qui s'appuie sur un calcul de préférences du côté des élèves, et un système de sélection du côté des chefs d'établissements.

Le modèle thématique postule que chaque élève a une connaissance partielle des établissements, et les préférences sont évaluées entre chaque élève et l'ensemble des établissements « connus » de lui. Deux règles permettent de formaliser le processus de choix d'établissement de l'élève :

- La connaissance est basée sur la proximité géographique, le rayon de connaissance variant d'un élève à l'autre en fonction de son « capital économique » et son « capital culturel ».
- La préférence se fonde sur un degré d'appétence différencié d'un type d'élève pour un type d'établissement, ce dernier étant défini par les filières proposées (sport, arts, langues, humanités) et son statut (public ou privé), l'ensemble reflétant la stratégie du chef d'établissement.

L'implémentation de la première « règle » implique essentiellement un calcul de distance entre lieux de résidence et établissements (valeur numérique). Pour la seconde règle, le choix qui a été fait est d'associer à chaque type d'élève une suite de scores reflétant ses différentes appétences pour chaque valeur possible des caractéristiques d'établissements. Le type de l'élève est déterminé par les combinaisons possibles des caractéristiques de l'élève, par exemple un capital économique moyen, un capital culturel fort, un niveau scolaire fort, et de nationalité étrangère. Le score correspond à une évaluation de l'appétence d'un type d'élève pour une filière. C'est le résultat d'un tirage aléatoire de lois normales centrées sur des valeurs d'appétences moyennes de 1, 0, ou -1. Un établissement étant défini par une liste de présences ou d'absences de filières, et un statut, la préférence d'un type d'élève pour un type d'établissement sera la somme des scores obtenus (valeur numérique) pour chacune des caractéristiques.

Concernant la formalisation du temps, il est intrinsèque au modèle, à la différence de l'AD où l'évolution est une interprétation *a posteriori* d'un jeu de données spécifique (« superposition » de données décrivant de la même manière des états à des dates différentes). Ici, à chaque itération du modèle, les profils des établissements sont recalculés, comme les préférences et les affectations à chaque nouvelle génération d'entrants au collège.

De manière générale, les règles se construisent tout d'abord à partir d'une formalisation de connaissances (discours du thématicien) sous forme de schéma conceptuel. Il est ensuite nécessaire pour l'implémenter de mobiliser un formalisme mathématique ou statistique ayant recours à des données numériques. L'exemple développé ci-dessus montre comment à partir d'une connaissance symbolique, on construit une règle à base de mesures numériques de « préférences ». Cette construction n'est pas une évidence, et la description qui vient d'en être faite peut sembler exagérément simplificatrice. Les catégories et scores utilisés sont là à titre

exploratoire et c'est à travers des tests successifs que l'on pourra avoir une idée de leur éventuel pouvoir explicatif.

Enfin, le modèle produit tout un ensemble de données numériques. Il est possible d'extraire des suivis au niveau de chaque élève, de chaque établissement et des propriétés observées au niveau macroscopique du système scolaire. La figure 5 illustre ainsi un exemple d'analyse de l'évolution de la mixité au cours du temps (axe des X) calculée à partir des données sur les établissements, extraites des sorties du modèle. Il s'agit de l'indice de Hoover, mesurant la concentration moyenne de chaque modalité (sociales, scolaire, origine) à l'intérieur des établissements. Un tel modèle produit des données numériques qu'il n'aurait pas été possible d'observer de manière empirique. Ces différentes visualisations numériques d'expérimentations à base symbolique permettent d'explorer le résultat des mécanismes mis en œuvre et leurs interactions.

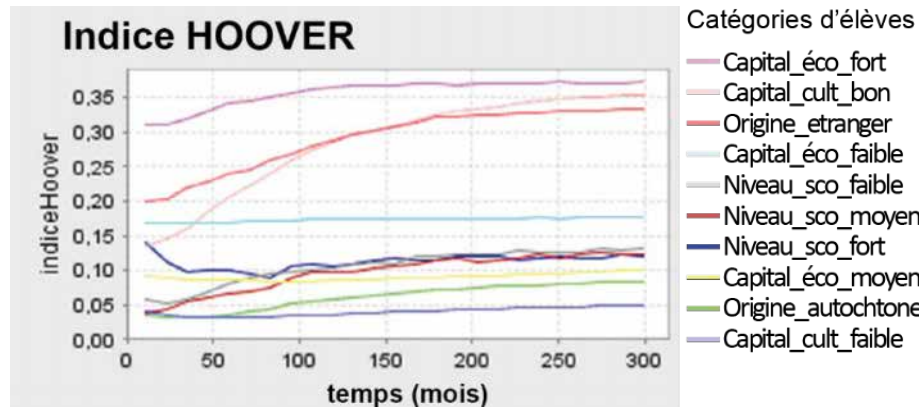


Figure 5. Exemple d'analyse de la mixité à l'intérieur des établissements à partir des sorties du modèle (l'indice de Hoover exprime pour chaque catégorie la part d'élèves de cette catégorie qu'il faudrait déplacer d'un établissement à l'autre pour aboutir à l'équirépartition)

Au terme de cette comparaison entre l'AD et les SMA, nous avons illustré comment des approches, étiquetées respectivement numérique et symbolique, font en fait cohabiter des considérations numériques et symboliques, et ce, que l'on considère les **entrées** du modèle, les **traitements** qui sont faits ou les **sorties**. Le tableau 1 synthétise cette comparaison.

Tableau 1. Forme numérique vs symbolique des entrées, traitements et sorties correspondant à une AD et à un SMA, appliqués au cas de la ségrégation scolaire

	Entrées	Traitements	Sorties
AD	Observables numériques	Numériques : – calcul des similarités entre établissements, – calcul des axes factoriels	Numériques : coordonnées sur les axes factoriels, contributions, plans factoriels ; Symboliques : interprétation des structures de différenciation
SMA	Connaissances symboliques (discours formalisé des experts)	Symboliques : règles régissant les comportements des agents ; Numériques : paramétrages pour implémenter les règles (calcul de scores...)	Numériques : trajectoires individuelles et agrégées Symboliques : interprétation des différentiels entre scénarios

3.3. Une spirale d'interactions entre analyse des données et système multi-agent

L'intérêt de mettre en relation ces deux approches est de montrer quelles formes peut prendre le dialogue entre les deux approches tant du point de vue opérationnel que thématique. Ces formes sont schématisées graphiquement dans la figure 6, en intégrant les trois étapes fondamentales évoquées plus haut, les entrées (E), les traitements (T) et les sorties (S). Leur enchaînement est matérialisé de manière linéaire par des flèches. A gauche on trouve l'approche AD et à droite l'approche SMA. Dans chaque approche sont identifiés systématiquement les statuts numériques/symboliques des « matériaux » en jeu, le symbole de « tableau de données » désignant le numérique, et celui du « personnage » faisant référence à un niveau d'abstraction et d'interprétation qui renvoie à des connaissances symboliques. Le temps est signifié par une flèche positionnée à l'étape où il est formalisé et où il intervient : dans les entrées-sorties pour l'AD, intrinsèque à l'étape de traitement dans les SMA.

Le fait que chaque approche articule du numérique et du symbolique permet d'envisager un dialogue entre les deux approches. C'est l'expérience qui a été faite sur la ségrégation scolaire illustrée figure 6. Les flèches en tirets désignent le lien entre le résultat que produit une approche et son utilisation à une étape de l'autre approche. L'interprétation de l'évolution de la ségrégation sociale (en bas à gauche), par exemple, peut servir à formaliser des règles sur les stratégies des chefs d'établissement (en haut à droite). Réciproquement, les données issues du modèle

SMA (en bas à droite) peuvent être analysées par une AD (en haut à gauche) pour interpréter le scénario en cours.

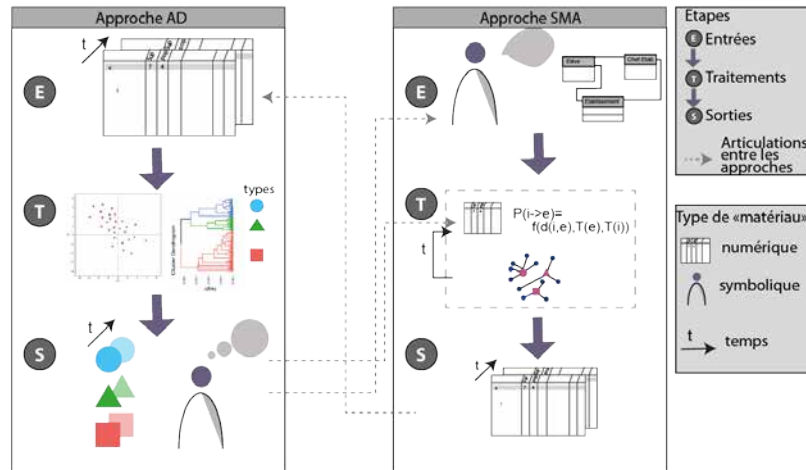


Figure 6. Cohabitation du numérique (AD) et du symbolique (SMA) à l'intérieur de chaque approche et articulation entre les deux approches

La figure 7 présente un exemple concret d'articulation entre les deux approches. À gauche on a le résultat d'une typologie sociale de l'espace résidentiel (AFC + CAH des catégories de revenus des résidents par commune en Île-de-France). Les classes de la typologie sont constituées d'agrégats de communes se ressemblant. La signification de chaque classe est définie sur une base numérique, à savoir le profil moyen des communes constituant la classe (figure 7a). L'interprétation symbolique de ces classes a permis de repérer les grandes lignes des divisions socio-spatiales d'un espace métropolitain. Celles-ci ont servi pour construire la situation initiale stylisée sur laquelle appliquer les règles formalisées au sein du SMA. Les agents-élèves ont ensuite été distribués dans l'espace résidentiel, en fonction de leurs propriétés (figure 7b). Deux propriétés de l'agent-élève, « capital culturel » et « capital économique », non directement observables, ont pu être identifiées lors de l'interprétation des résultats de l'AFC effectuée sur les catégories sociales (figure 3). Elles sont davantage porteuses de sens dans les différences de comportement de choix des élèves que la variable « catégorie sociale ». Pierre Bourdieu exprime ainsi la dimension heuristique de l'AFC : « j'utilise beaucoup l'analyse des correspondances, parce que je pense que c'est une procédure relationnelle dont la philosophie exprime pleinement ce qui selon moi constitue la réalité sociale. » (Bourdieu⁷ in Lebaron 2015). Dans un SMA, on va

7. dans la préface à l'édition allemande du *Métier de sociologue*.

ainsi pouvoir mobiliser des propriétés (variables) jugées explicatives dans la « réalité sociale » telle qu'elle est interprétée par l'expert thématique, même si elles ne sont pas observables dans le domaine empirique. Tel est ainsi le cas du « capital culturel » et du « capital économique » des élèves dont la combinaison est, suivant l'interprétation que l'expert thématique a fait à partir d'AFC sur les catégories sociales, motrice dans la dynamique de la ségrégation scolaire.

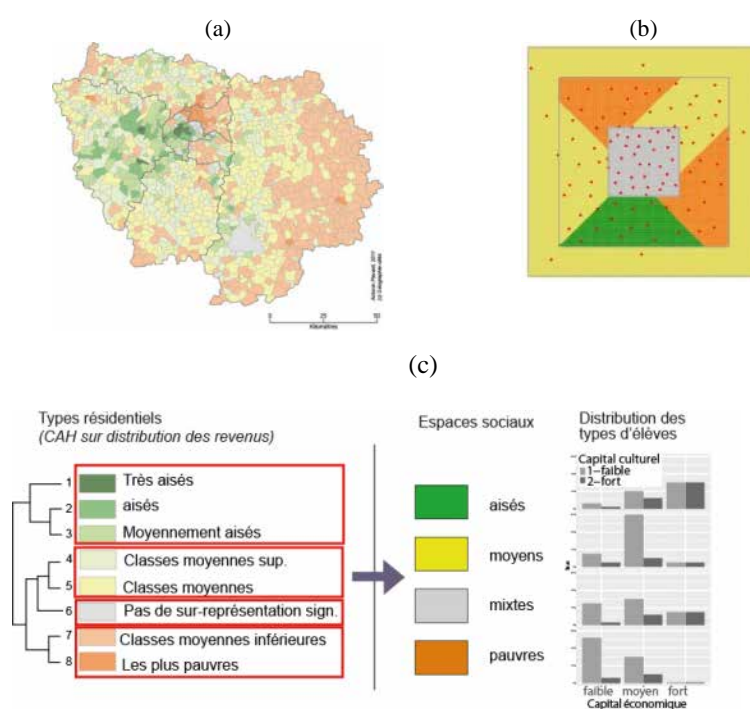


Figure 7. Articulation entre les résultats d'une typologie effectuée par AD et la construction d'une situation spatiale initiale stylisée en entrée du modèle SMA

Conclusion

Si l'on repositionne l'ensemble de la démarche relativement à la figure 1, nous avons montré comment s'articule le cadre conceptuel thématique avec le cadre conceptuel de modélisation de chacune des deux méthodes AD et SMA. Nous avons ainsi précisé au §2.1 le cadre épistémologique sur lequel sont fondées les hypothèses du modèle conceptuel du thématique, prenant appui sur les travaux de Bourdieu, et qui mettent en avant le rôle des structures sociales sur les choix des élèves et des chefs d'établissement. Ce positionnement épistémologique contribue à la bonne articulation entre les deux approches. Si d'autres choix épistémologiques avaient été faits, l'interprétation symbolique des résultats d'une typologie n'aurait pas

forcément pu servir de référentiel à la construction abstraite des entités du modèle SMA. C'est ainsi la compatibilité conceptuelle entre les différents éléments représentés sur la figure 1 qui permet de valider l'articulation entre des méthodes différentes et non la convergence des résultats obtenus. Glaser et Strauss (1967) le soulignent de manière très claire :

« La production de théorie ne s'appuie pas sur le fait, mais sur la catégorie conceptuelle (ou une propriété conceptuelle de la catégorie) qui en a été extraite. Un concept peut être produit à partir d'un fait qui devient alors simplement un élément dans un univers de nombreux indicateurs possibles pour le concept et de données pouvant lui être associées » (Glaser et Strauss, 1967).

Finalement, l'articulation entre les approches AD et SMA telle qu'elle a été formalisée et illustrée, permet de dessiner une spirale méthodologique, qui fait alterner approche numérique et approche symbolique. L'expérience de cette spirale est particulièrement riche dans la mesure où elle permet de tester deux formes d'exploration qui se répondent : d'une part, des structures d'associations pour en comprendre les mécanismes et d'autre part, des mécanismes qui conduisent à ces structures. Ainsi l'AD permet « d'aider à trouver et départager les candidats à l'explication sociologique » (Lebaron, 2010). Ces structures d'associations sont réinterprétées dans le cadre conceptuel thématique pour être ensuite reliées à des processus. De même, en SMA, les résultats obtenus par expérimentation de différentes règles ou paramètres permettent de cibler des « candidats à l'explication ». Le but d'un modèle SMA est en effet d'identifier un « candidat explicatif » possible pour rendre compte de la genèse, de la structure et de la dynamique du phénomène étudié. Cependant, plutôt que de prétendre qu'il s'agit de la seule explication possible, la démarche consiste à tester quelles sont les « conditions de possibilités » du phénomène étudié (Livet *et al.*, 2014). Les deux méthodes, pourtant ancrées dans les arrières plans théoriques très différents de la statistique et de la simulation informatique, ont ainsi en commun cette ambition de proposer un « candidat explicatif » au phénomène étudié tout en maintenant une position expérimentale ouverte de révision et de possibilité d'évolution des modèles explicatifs proposés. Ce cheminement de l'une à l'autre méthode, suivant une approche dynamique en spirale permet de relier les deux points de vue et d'approcher progressivement les processus sous-tendant la dynamique étudiée. Chacune des approches permet en quelque sorte de valoriser la connaissance apportée par l'autre.

Remerciements

Les auteures souhaitent vivement remercier les deux évaluateurs/évaluatrices pour leurs suggestions et questions, celles-ci leur ayant permis d'explicitier un certain nombre de points.

Bibliographie

- Balbo F., Saunier J., Diday E., Pinson S. (2009). De l'utilisation de l'analyse de données symboliques dans les systèmes multi-agents. *Actes de la conférence « Extraction et gestion des connaissances » (EGC'09)*, Strasbourg, 27-30 janvier.
- Glaser B., Strauss A. (1967). *The discovery of grounded theory*, trad. de J.-L. Fabiani, 1995, *Revue Enquête*.
- Benzécri J.-P. (1973). *L'analyse des données, tome 2 : L'analyse des correspondances*, Bordas.
- Boudon R. (1973). *L'inégalité des chances: la mobilité sociale dans les sociétés industrielles*, Paris, Éditions Armand Colin.
- Bourdieu P. (1979). *La Distinction : Critique sociale du jugement*, Paris, Éditions de Minuit.
- Cibois Ph. (1999). Modèle linéaire contre modèle logistique en régression sur données qualitatives. *Bulletin de Méthodologie Sociologique*, n° 64, p. 5-24.
- Diday E. (1989). Introduction à l'approche symbolique en analyse des données. *RAIRO, Recherche opérationnelle*, tome 23, n° 2, p. 193-236.
- François J.-C., Mathian H., Sanders L., Bulle N., Waldeck R., Phan D. (2014). Modélisation par SMA de la structuration sociale de l'espace scolaire : une ontologie intégrée mais non réductrice de plusieurs points de vue méthodologiques. *Ontologies et modélisation par SMA en SHS*, Phan D. (dir.), Londres/Paris, Hermès-Lavoisier, p. 461-475 .
- Gruber T.R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, vol. 5, n° 2, p. 199-220
- Lebaron F. (2010). L'analyse géométrique des données dans un programme de recherche sociologique: les cas de la sociologie de Bourdieu. *Revue MODULAD*, n° 42.
- Lebaron F. (2015). L'espace social. Statistique et analyse géométrique des données dans l'œuvre de Pierre Bourdieu. Chapitre 3. *La méthodologie de Pierre Bourdieu en action. Espace culturel, espace social et analyse des données*, F. Lebaron éd., Paris, Dunod, coll. « Psycho Sup », p. 43-58.
- Livet P., Müller J.-P., Phan D., Sanders L. (2010). Ontology, a mediator for agent-based modeling in social science. *Journal of Artificial Societies and Social Simulation*, vol. 13, n° 1, 3, <http://jasss.soc.surrey.ac.uk/13/1/3.html>
- Livet P., Sanders L. (2014). Le « test ontologique » : un outil de médiation pour la modélisation agent. *Ontologies et modélisation par SMA en SHS*, Phan D. (dir.), Hermès-Lavoisier, Londres-Paris, p. 95-110.
- Livet P., Phan D., Sanders L. (2014). Diversité et complémentarité des modèles multi-agents en sciences sociales. *Revue française de sociologie*, 2014/4, vol. 55, p. 689-729. <https://www.cairn-int.info/revue-francaise-de-sociologie-2014-4-page-689.htm>
- Mathian H., Sanders L. (2014). *Objets géographiques et processus de changement. Approches spatio-temporelles*, ISTE Éditions. ISBN 978-1-78405-031-3.
- Müller J.-P., Diallo A. (2012). Vers une méthode multi-point de vue de modélisation multi-agent. *Systèmes multi-agents : ouverture, autonomie et co-évolution. Actes des JFSMA'12*

(*Journées francophones sur les systèmes multi-agents*), 17-19 octobre, Honfleur (France), Toulouse : Cépaduès, p. 33-42. ISBN 978-2-36493-037-7.

Phan D., Müller J-P., Sibertin Blanc C., Ferber J., Livet P. (2014). Introduction à la modélisation par SMA en SHS : comment fait-on une ontologie ? *Ontologies et modélisation par SMA en SHS*, sous la dir. de Phan D., Traité RTA, série « Informatique et systèmes d'information », Paris, Hermès Science Publications, p. 21-51. ISBN 978-2-7462-3207-5.

Poupeau F., François J. C. (2008). *Le sens du placement. Ségrégation résidentielle et ségrégation scolaire*, Raisons d'agir, coll. "Cours et travaux", EAN: 9782912107435.

Smith B. (2003). Ontology, *Blackwell Guide to the Philosophy of Computing and Information*, Floridi L. (Ed.), Blackwell, Oxford, p. 155-166.

Appréhender le changement des catégories pour l'étude d'une dynamique spatiale sur le temps long

Christine Plumejeaud-Perreau¹, Lucie Nahassia², Julie Gravier²

1. UMR 7266 Littoral Environnement et Sociétés
2 rue Olympe de Gouges, 17000 La Rochelle, France
christine.plumejeaud-perreau@univ-lr.fr

2. UMR 8504 Géographie-cités
Campus Condorcet, 5 cours des Humanités, 93322 Aubervilliers, Cedex France
{[julie.gravier](mailto:julie.gravier@parisgeo.cnrs.fr) ; [lucie.nahassia](mailto:lucie.nahassia@parisgeo.cnrs.fr)}@parisgeo.cnrs.fr

RÉSUMÉ. À travers trois exemples issus d'études des dynamiques de peuplement en France sur le temps long, intra et interurbaines, cet article montre que l'introduction d'ontologies « a-historiques » a facilité la mise en place d'analyses numériques quantifiant les dynamiques spatiales. Cependant il démontre que ces ontologies ne sont pas neutres, qu'elles ne dissolvent pas la spécificité des sources mobilisées, et constituent au contraire des savoirs situés. Mais au-delà de ces critiques, les auteures tendent à argumenter que le processus de construction et de dialogue établi autour d'ontologie a-historique est très bénéfique à la recherche, qu'il est un socle pour une science interdisciplinaire, partagée et pérenne.

ABSTRACT. Through three examples coming from long-term studies of population dynamics in France, intra and inter-urban, this paper shows that the introduction of some "a-historical" ontologies was required to implement quantitative analyses of spatial evolutions. However, we demonstrate that these ontologies are far from being neutral, and that they do not dissolve the specificity of the sources we used, but on the contrary, they constitute situated knowledge. Beyond these criticisms, the authors tend to argue that the process of construction and dialogue established around ontology is very beneficial to research, that it is a foundation for an interdisciplinary, shared and sustainable Science.

MOTS-CLÉS : catégorie a-historique, ontologie, temps long, savoir situé, reproductibilité.

KEYWORDS: A-historic category, Ontology, Long duration, Situated knowledge, Reproducibility.

DOI:10.3166/RIG.31.47-80 © 2022 Lavoisier

1. Introduction

Cet article propose un retour d'expérience concernant la mise en place et l'utilisation d'ontologies dans le cadre d'analyses de dynamiques spatiales avec rétrospectives historiques sur le temps long. Le temps long entendu ici est celui des historiens et des archéologues qui se déroule sur plusieurs siècles, voire plusieurs millénaires, et non celui, plus communément usité par les géographes ou les urbanistes, qui se déploie sur plusieurs années ou décennies. Le choix de l'analyse diachronique suppose deux difficultés importantes à considérer dans la construction et l'utilisation d'ontologies. Premièrement, les sociétés étudiées sont très diverses, tant dans leurs régimes politiques que dans leurs organisations sociales, dans leurs systèmes économiques, mais également leurs idéologies, leurs croyances, leurs connaissances technologiques, etc., autant d'éléments qui s'inscrivent dans l'espace géographique, au sens de l'espace terrestre (Dauphiné, 1984), et le transforment. Secondement, les sources à partir desquelles il est possible d'appréhender ces sociétés sont très variées dans leurs natures et leurs constitutions. Par conséquent, les objets identifiables à partir de ces sources sont divers. L'exploration des dynamiques spatiales dans la longue durée nécessite donc de comparer l'occupation et les pratiques des sociétés qui se succèdent dans le temps à partir de sources multiples, comparaison qui est justement mise en question par la diversité de ces dernières et des sociétés étudiées.

Rien n'étant comparable par ailleurs, la confrontation des sociétés dans une perspective diachronique exige du chercheur qu'il construise des catégories d'objets qui transcendent cette diversité. Pour ce faire, nous partons du postulat que cette construction doit s'appuyer sur les éléments communs aux différents moments des sociétés et aux différentes sources, et que plus les choses à comparer sont diverses et hétérogènes, plus les catégories utilisées doivent être simplificatrices. Les objets constructibles pour l'étude diachronique de l'espace sont dès lors tenus d'être relativement « a-historique », c'est-à-dire génériques et transposables dans leur description du monde vis-à-vis des périodes analysées. À titre d'exemple, si l'on étudie des édifices religieux dans la longue durée en Europe du Nord-Ouest, ces derniers sont divers : sanctuaires romains, églises catholiques, synagogues, temples protestants, mosquées, etc. Ces édifices sont notamment utilisés dans le cadre de différentes religions et pour des formes de culte variées. Cependant, au-delà de ces différences, tous ces édifices sont des équipements construits dans l'espace terrestre pour accueillir une population qui rend hommage à une ou plusieurs divinités. En cela, chacun de ces équipements pourrait par exemple être catégorisé comme un « équipement collectif de culte ». L'utilisation de cette catégorie autorise ensuite d'étudier les dynamiques de ces équipements, en particulier au moyen d'analyses quantitatives de leurs évolutions dans le temps et dans l'espace.

Cet article a pour objectif de traiter en profondeur des questions laissées encore ouvertes par l'évolution des catégories dans le temps, et leur traitement par des représentations de type ontologiques à des fins d'analyse quantitative des évolutions.

Pour ce faire, nous prenons appui sur trois expériences de recherches personnelles. L'expérience du programme ANR GeoPeuple illustre comment la reconstitution de l'histoire et de la motivation des acteurs ayant cartographié la France sous Cassini, pour l'état-major ou l'actuelle BD Topo de l'IGN a changé en profondeur le contenu et le sens des signes présents sur ces objets, et trouble les usages statistiques de catégories ontologiques a-historiques produites *a posteriori* pour les analyses des changements de l'espace cartographié. Les deux autres expériences sont issues de recherches sur l'évolution de la ville sur 2000 ans, à l'échelle intra et interurbaine avec les exemples de Tours et de Noyon, et interrogent la construction de catégories a-historiques à partir de données archéologiques et historiques pour qualifier et analyser les dynamiques spatiales dans la longue durée. Les ontologies a-historiques utilisées dans ces trois expériences sont des ontologies de domaine. Par ailleurs, elles sont relatives aux deux premières étapes de la construction ontologique telle qu'on l'entend en informatique, à savoir la conceptualisation et la formalisation dans un langage, mais non à la troisième étape de la modélisation ontologique pour utiliser des raisonneurs, que sont les ontologies axiomatisées¹.

Dans un premier temps, il est expliqué comment l'analyse sur un temps long de traces d'activités humaines inscrites dans l'espace requiert de mettre en place un cadre symbolique permettant de comparer ce qui n'est, de prime abord, pas nommé de façon identique. Ce besoin est particulièrement prégnant dès lors que l'analyse des évolutions est menée de manière quantitative – par le biais de statistiques et d'analyses spatiales –, comme l'illustrent les exemples mobilisés dans cette partie. Dans un deuxième temps, une analyse critique de ces exploitations et de ces construits met en évidence les possibles failles et lacunes de ces ontologies a-historiques. Trois niveaux de difficultés sont analysés : le premier est lié à la nature des sources exploitées, le deuxième aux *habitus* disciplinaires, et le dernier aux questionnements scientifiques qui ont guidé la construction de ces cadres ontologiques, et les choix conceptuels qui en découlent. Le troisième temps de cette discussion souligne les bénéfices finalement induits par la mise en œuvre d'ontologies, en particulier pour l'interdisciplinarité et pour la capacité à produire une science plus reproductible grâce à l'explicitation formelle des cadres d'analyse.

2. Usage d'ontologies a-historiques pour l'étude quantitative des dynamiques spatiales dans la longue durée

Pour commencer, nous présentons ici deux ontologies a-historiques avec lesquelles les auteurs ont travaillé pour étudier des dynamiques spatiales dans la longue durée. Les contextes de construction de ces deux ontologies sont très différents – élaboration disciplinaire sur plusieurs décennies pour l'une, composition interdisciplinaire dans le cadre d'un projet ANR pour l'autre – tout comme leur

1. Ontologies axiomatisées : formalisées avec des contraintes et un langage interprétable tel que OWL, qui permettent à une machine de raisonner sur les éléments représentés. OWLTime ou GeoSparql sont des exemples d'ontologies axiomatisées.

domaine d'application – l'évolution de l'espace urbain sur 2000 ans, d'une part et les transformations du territoire français depuis le 16^e siècle, d'autre part. Ces deux exemples mis en regard montrent néanmoins comment une conceptualisation a-historique permet d'étudier les dynamiques spatiales dans le temps long en utilisant les méthodes quantitatives, des statistiques classiques à l'analyse spatiale.

2.1. Un cadre conceptuel issu de l'archéologie urbaine pour l'étude des villes et de leurs interactions dans la longue durée

La première ontologie de domaine que nous utilisons dans nos travaux respectifs est relative à la construction d'un cadre conceptuel en archéologie pour l'étude de l'espace des villes. Dès son émergence en France à la fin des années 1970, l'archéologie urbaine s'est saisie de l'espace urbain avec une entrée à la fois diachronique et comparative. Cette double perspective a abouti à une réflexion toujours ouverte, et à laquelle nous contribuons ici, sur la qualification des sources archéologiques et historiques, des ontologies, et des modèles de données avec des bases communes et un même objectif : pouvoir comparer les objets qui font les villes, dans le temps comme dans l'espace. Une de ces ontologies, largement partagée par les archéologues et construite en collaboration au sein du Centre national d'archéologie urbaine (CNAU), propose ainsi une caractérisation générique et a-historique des objets topographiques qui constituent l'espace urbain, à plusieurs échelles. Nous revenons ici sur l'élaboration de ce cadre conceptuel et sur les usages que nous avons pu en faire dans nos travaux pour analyser les interactions qui font la ville selon des méthodes quantitatives, d'abord à l'échelle intra-urbaine puis au niveau interurbain, à Tours et à Noyon.

2.1.1. Processus de construction d'une ontologie a-historique en archéologie

L'étude de l'espace intra-urbain sur le temps long des villes françaises par les archéologues a débuté dans la seconde moitié des années 1970. Parmi leurs objectifs, citons en particulier la volonté de réaliser des synthèses monographiques ayant pour dessein d'étudier l'évolution de la structure urbaine spécifique à chaque ville, pour être en mesure, dans un second temps, de faire des comparaisons. Il s'agissait de « parvenir à substituer à une collection d'histoires de villes, une histoire des villes. Une telle ambition sera réalisée à la condition que les archéologues aient une approche commune du phénomène urbain » (Galinié 1979, p. 7). Deux grands jalons de l'histoire de l'archéologie urbaine française – synthétisée par Lorans *et al.* (2018) – ont permis de mettre en place cette ambition.

Le premier jalon est la création du Centre national d'archéologie urbaine en 1984 (CNAU), organisme central de ce champ de l'archéologie alors rattaché au ministère de la Culture. En particulier, le CNAU fut l'instigateur d'une collection de

monographies urbaines² réalisées par des archéologues locaux en collaboration avec ceux du Centre (ainsi que quelques géographes), qui devient le lieu d'une réflexion partagée sur les objets observés et analysés. Cette collection hérite de deux ouvrages fondateurs de l'archéologie urbaine sur Londres et Tours (Biddle *et al.*, 1973 ; Galinié et Randoin 1979). À l'image de ces derniers, chaque monographie est constituée d'une partie textuelle et d'un portefeuille de plans-calques qui permettent, par superposition, de visualiser les changements et la pérennité de la structure spatiale. Les éléments relatifs à l'occupation de l'espace intra-urbain sont présentés sous forme de synthèse selon un découpage chronologique. Dans les ouvrages de Biddle *et al.* (1973) et de Galinié et Randoin (1979), la chronologie et la catégorisation des objets sont constituées selon l'évolution de la structure fonctionnelle et spatiale intra-urbaine propre à chacune des deux villes. Les monographies du CNAU, parce qu'elles sont pensées en collection et construites grâce à des allers-retours entre les archéologues, aboutissent, elles, à la création d'un cadre conceptuel commun – dans la qualification thématique des objets étudiés. Dans la partie textuelle, la synthèse de chaque période est complétée par une liste de notices, qui renvoient à des unités d'analyse de l'espace intra-urbain, dénommées « repère urbain » dans la collection. La présentation des notices est systématisée car nous retrouvons toujours le nom du repère urbain, son identifiant lié aux plans-calques (si le repère a été cartographié), sa localisation dans la voirie actuelle, une courte description et des références bibliographiques (figure 1) ; et, surtout, elles sont classées par thèmes. Un thésaurus³ mis en place par le CNAU permet de classer des « termes » (ou « descripteurs », établis à partir des annuaires des opérations archéologiques en milieu urbain) en catégories hiérarchisées : les « rubriques » (ou « valeurs d'usages ») et fonctions (« fonctions urbaines »)⁴, comme l'illustre par exemple la figure 2.

Les repères urbains sont ainsi catégorisés d'un point de vue sémantique de manière homogène et harmonisée à l'intérieur de chaque document et entre les documents. De ce fait, la comparaison de différentes périodes pour une même ville et de différentes villes à une même période devient envisageable. En pratique, les catégories ont été testées sur des villes très diverses : petites et grandes, d'origines protohistorique, antique ou médiévale, etc. (figure 3). Elles ont par ailleurs été

2. Intitulées *Documents d'Évaluation du Patrimoine Archéologique des Villes de France - DEPAVF*.

3. Un thésaurus est une série de termes formalisés et reliés par des relations elles-aussi formalisées (hiérarchie, synonymie, associatives, appartenance à un groupe de concepts, ou de facettes) utilisé pour décrire un domaine. Un thésaurus peut constituer en soit une ontologie, en particulier si sa définition suit les règles du standard SKOS (ISO 25964, 2011), ou bien un thésaurus peut être une des composantes d'une ontologie. Une ontologie permet de spécifier des types de relation plus riches que les seules relations définies par un thésaurus, ainsi que des contraintes sur les relations, ce qui autorise des raisonnements plus sophistiqués.

4. La nomenclature ainsi que le contenu de ce thésaurus ont légèrement évolué au cours du temps et des publications (intégrant parfois même un niveau intermédiaire entre rubrique et fonction, les « chapitres »). La figure 2 n'en présente ici qu'une seule instance.

utilisées par de nombreux archéologues, y compris en dehors des documents de synthèse du CNAU, et l'harmonisation a ainsi bénéficié des allers-retours constants entre ceux travaillant sur des villes particulières et ceux travaillant au CNAU (Collectif, 2004).

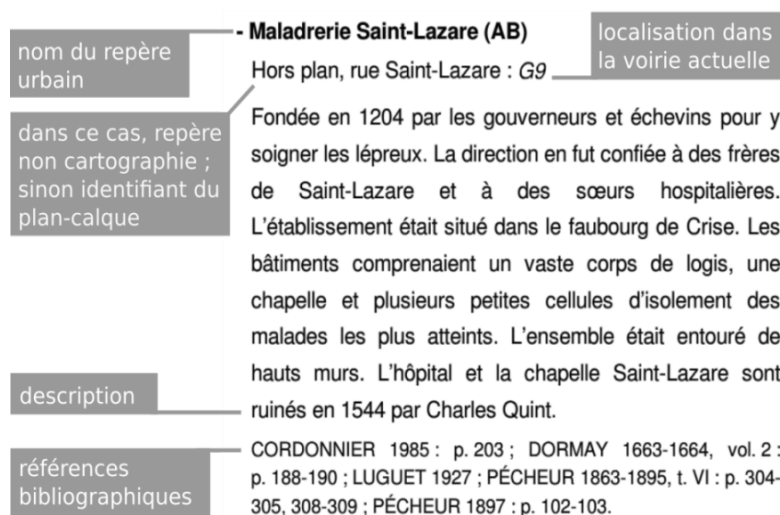


Figure 1. Notice d'un repère urbain de la collection des DEPAVF : Soissons (d'après Roussel et al., 2002, p. 64)

<p>VOIRIE, AMÉNAGEMENTS :</p> <ul style="list-style-type: none"> 1 - VOIES 2 - ESPACES LIBRES 3 - AMÉNAGEMENTS DES BERCES ET VOIES D'EAU 4 - AMÉNAGEMENTS DU RELIEF 5 - FRANCHISSEMENTS 6 - ADDUCTIONS D'EAU 7 - COLLECTEURS, ÉVACUATIONS <p>STRUCTURES DÉFENSIVES ET MILITAIRES :</p> <ul style="list-style-type: none"> 8 - SYSTÈME DÉFENSIF URBAIN 9 - STRUCTURES FORTIFIÉES 10 - GARNISONS, CASERNEMENTS <p>CONSTRUCTIONS CIVILES :</p> <ul style="list-style-type: none"> 11 - ESPACES PUBLICS AMÉNAGÉS 12 - POUVOIR CIVIL, JUSTICE 13 - ÉDUCATION, CULTURE 14 - SANTÉ 15 - SPECTACLE, SPORT 16 - BAINS 17 - COMMERCE, ÉCHANGES 18 - HABITAT PRIVÉ 	<p>ÉDIFICES RELIGIEUX :</p> <ul style="list-style-type: none"> 19 - CULTES PAÏENS 20 - ÉDIFICES CULTUELS CATHOLIQUES 21 - BÂTIMENTS CONVENTUELS OU MONASTIQUES 22 - BÂTIMENTS ECCLÉSIASTIQUES 23 - CULTES AUTRES QUE CATHOLIQUE <p>FUNÉRAIRE :</p> <ul style="list-style-type: none"> 24 - FUNÉRAIRE <p>PRODUCTION :</p> <ul style="list-style-type: none"> 25 - ARTISANAT 26 - AGRICULTURE, ÉLEVAGE 27 - INDUSTRIE 28 - EXTRACTION <p>FORMATIONS NATURELLES : géomorphologie, hydrologie...</p> <ul style="list-style-type: none"> 29 - FORMATIONS NATURELLES
--	---

Figure 2. Rubriques (en police simple) et fonctions (en gras) du thésaurus du CNAU à la fin des années 2000, et tel qu'il est toujours disponible sur le site du ministère de la Culture⁵

5. <https://www.culture.gouv.fr/Sites-thematiques/Archeologie/Archeologie-de-la-ville/Documentation-scientifique-technique>

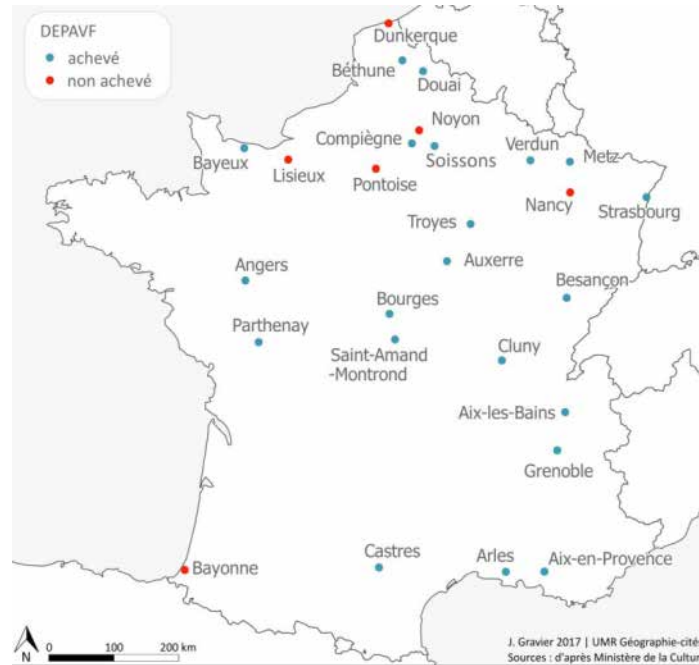


Figure 3. Villes de France métropolitaine ayant fait l'objet d'un DEPAVF (Gravier, 2018, p. 123)

Ce premier jalon a donc permis de mettre en place une ontologie de domaine – relative à la topographie urbaine des villes françaises depuis l'époque gallo-romaine à nos jours – constituée de descripteurs sémantiques a-historiques et hiérarchisés en catégories plus ou moins englobantes. Le second jalon dans la construction de cette ontologie a-historique est celui de son informatisation – par le biais de l'introduction de l'usage des bases de données et des systèmes d'information géographiques (SIG) dans les processus de travail des archéologues. À partir de 1996, le SIG « Topographie de tours préindustrielle » (ToToPI) est mis en place par Xavier Rodier à Tours (Rodier, 2000), qui l'inscrit explicitement dans la perspective de développer des études diachroniques sur la ville : « Il ne s'agissait pas de monter un système de gestion du type "carte archéologique urbaine", mais d'utiliser un SIG comme outil de recherche [...]. Le système devait pouvoir être utilisé par les chercheurs travaillant sur la ville de Tours antique, médiévale et moderne. » (Rodier, 2016, p. 58-59). Du point de vue sémantique, le SIG s'appuie dès l'origine sur le thésaurus du CNAU (Rodier, 2000, p. 6), tout en participant à son évolution et ses ajustements au fil du temps. En revanche, l'intégration des outils informatiques, parce qu'ils induisent une formalisation stricte des objets, est surtout le point de départ de la construction d'une ontologie a-historique plus complète et explicite que

les règles méthodologiques utilisées dans les monographies du CNAU. Cette ontologie développée autour du SIG et des données archéologiques et historiques tourangelles est construite tout au long des années 2000 pour aboutir au milieu de la décennie au modèle « OH_FET », pour « Objet Historique - Fonction Espace Temps » (Rodier et Saligny, 2010). L'ontologie explicite alors la construction des objets étudiés, les « OH » (Objet Historique), en appliquant le modèle conceptuel de Peuquet (2002) (*what, when, where*) à l'espace intra-urbain sur le temps long (Galinié *et al.*, 2004 ; Galinié et Rodier, 2004). Les objets sont ainsi pensés comme le croisement de trois dimensions (Rodier, Saligny 2010) :

- Fonctionnelle : cette dimension sémantique permet de caractériser la nature des objets. Dans la base de données tourangelles elle est constituée de deux niveaux d'analyse emboîtés – les valeurs d'usage et les valeurs urbaines – et reprend peu ou prou le thésaurus du CNAU (figure 2). La nature de ces niveaux est discutée plus en détail dans la section 3.2. Ces catégories permettent de monter en généralité dans la caractérisation des objets étudiés. Par exemple, une église spécifique est définie comme un OH de valeur d'usage « lieu cultuel catholique » et de valeur urbaine « religieuse » et peut, dès lors, être comparée à tous les autres objets pareillement qualifiés de religieux selon les besoins de l'analyse, quels que soient l'espace ou la période considérés.

- Spatiale : cette dimension s'attache à décrire les objets dans l'espace selon leur localisation et leur emprise en des termes génériques. L'emprise est une géométrie (dans le SIG un polygone ou une polyligne, voire un point quand elle n'est pas précisément connue) et la localisation renvoie aux coordonnées géographiques de cette emprise.

- Temporelle : cette dimension est définie pour chaque objet par une date d'apparition et une date de disparition. Elle décrit donc la durée de vie des objets dans un référentiel temporel donné, à savoir le calendrier grégorien.

2.1.2. De l'ontologie à l'analyse : usage dans un cas d'étude intra-urbain

Dans le cadre d'un travail de thèse sur la localisation des activités urbaines dans la longue durée à Tours (Nahassia, 2019), nous sommes reparties des données archéologiques rassemblées dans le SIG ToToPI, que nous avons intégrées à une démarche généralisante et globalisante en considérant toute la durée d'existence de la ville, tous ses quartiers et toutes les activités présentes. Nous avons fait le choix de mobiliser les catégories de valeurs d'usage et de valeurs urbaines pour analyser, de manière quantitative, le changement urbain. La figure 4 ci-après, extraite de ce travail, montre le nombre d'objets, pour chaque valeur urbaine, qui apparaissent ou disparaissent au cours du temps. C'est bien le caractère a-historique des catégories qui donne du sens à ce type de traitement. En effet, du fait que tous les objets, quelle que soit leur époque d'appartenance, sont associés à une caractérisation sémantique commune et générique, ils deviennent comparables les uns aux autres dans la longue durée. Cette analyse permet en particulier de définir des rythmes de changement produits par l'apparition et la disparition des activités dans l'espace urbain.

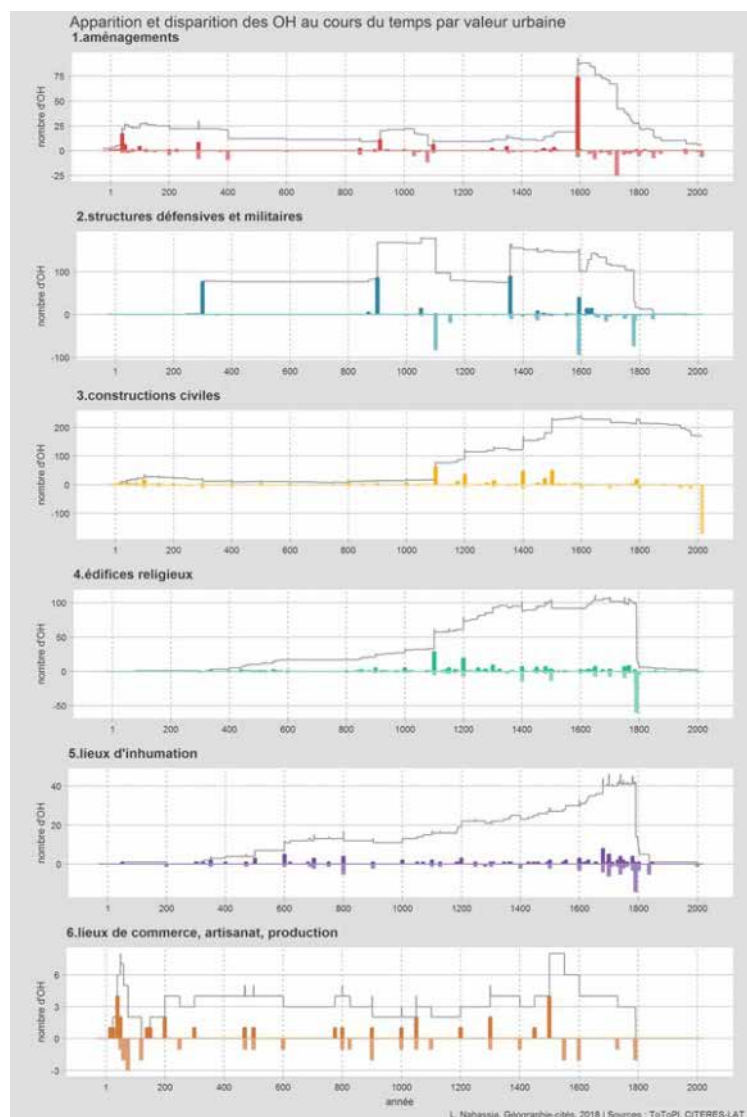


Figure 4. Six catégories génériques pour étudier le rythme du changement urbain à Tours sur 2000 ans (Nahassia, 2019, p. 168) (en couleur : nombre d'apparition et de disparition des objets ; en gris : fréquence cumulée du nombre d'objets)

Les activités militaires par exemple sont caractérisées par un rythme d'apparition de grande amplitude mais à des moments très circonscrits, ce qui correspond aux périodes relativement brèves de construction des enceintes urbaines, objets majoritaires dans cette catégorie et modélisés par de très nombreux OH pour les

différentes parties qui la composent (courtines, tours, portes, etc.). Les édifices religieux (essentiellement chrétiens ici) apparaissent de manière très fréquente à partir de 800 et jusqu'à la fin de la période étudiée. Ce profil temporel est la marque de l'importance que prend et conserve la chrétienté dans les sociétés occidentales pendant de nombreux siècles, et qui prend forme notamment dans la forte présence des activités afférentes dans l'espace urbain : lieux de culte, mais aussi monastères, habitats privés des religieux, etc. De tels traitements statistiques, facilitant des comparaisons dans la très longue durée, sont possibles parce que les observations historiques sont catégorisées de manière générique, et surtout a-historique.

2.1.3. De l'intra à l'interurbain : adaptation de l'ontologie source

Si le modèle OH_FET a d'abord été construit pour l'analyse de l'espace intra-urbain, l'ontologie inhérente au modèle peut être facilement complétée pour aborder la question de l'intra-urbain.

Dans le cadre de notre travail sur Tours (Nahassia, 2019), nous avons par exemple mis en place une nouvelle catégorisation fonctionnelle des objets étudiés, complémentaire des usages et valeurs urbaines, et qui caractérise les objets selon une « portée ». Cette qualification, selon une approche géographique classique, attribue à chaque activité représentée par les objets historiques un niveau de polarisation, qui renvoie à leur influence – économique, culturelle, politique, etc. – qui peut s'exercer au-delà des limites de la ville elle-même. Nous avons qualifié l'ensemble des OH de la BDD ToToPI selon quatre niveaux de portées (courte, moyenne, grande, exceptionnelle) en nous appuyant essentiellement sur la connaissance des chercheurs experts de l'urbain et de Tours (archéologues, historiens, géographes), laquelle a été recueillie à la fois lors de séances de travail conjointes et dans une bibliographie choisie. L'enrichissement de la description des objets est facilité par l'explicitation, en amont, de l'ontologie sous-jacente. Cette nouvelle couche de description, intégrée au modèle de données existant, permet d'interroger par le biais d'analyses quantitatives l'influence de la portée des activités dans l'organisation spatiale des activités, et, dès lors, de saisir la ville dans ses différentes échelles de fonctionnement. Nous avons, par exemple, pu montrer que, toutes époques confondues, plus la portée des activités est grande, plus elles apparaissent dans les zones les plus denses de l'espace urbain tourangeau. Cette nouvelle entrée mène en outre à explorer la ville dans ses relations avec les autres villes, et donc à étudier les conséquences des dynamiques interurbaines sur l'espace urbain : comment évoluent les activités de grande portée (université, cathédrale, etc.), qui placent Tours dans une position dominante par rapport au territoire régional, ou par rapport à d'autres villes ?

Un autre travail d'enrichissement de cette ontologie source à une plus petite échelle peut se trouver dans la thèse de Gravier (2018), qui y construit un cadre conceptuel pour modéliser les relations entre les villes dans la longue durée. Précisons tout d'abord que l'objectif du travail était de proposer une démarche permettant d'étudier ce qui unit l'histoire d'une ville à l'histoire des villes avec

lesquelles elle est en interaction sur 2000 ans, et ce, à partir du cas de Noyon (Hauts-de-France). Pour ce faire, nous avons considéré trois principales échelles d'analyse (figure 5) :

- l'espace intra-urbain, étudié à partir de « briques élémentaires » définies de manière très semblable aux OH ;
- la ville dans son ensemble, envisagée comme la combinaison de ces objets dans l'espace à chaque période de temps ;
- et le système de villes, pensé comme un ensemble de villes interdépendantes.

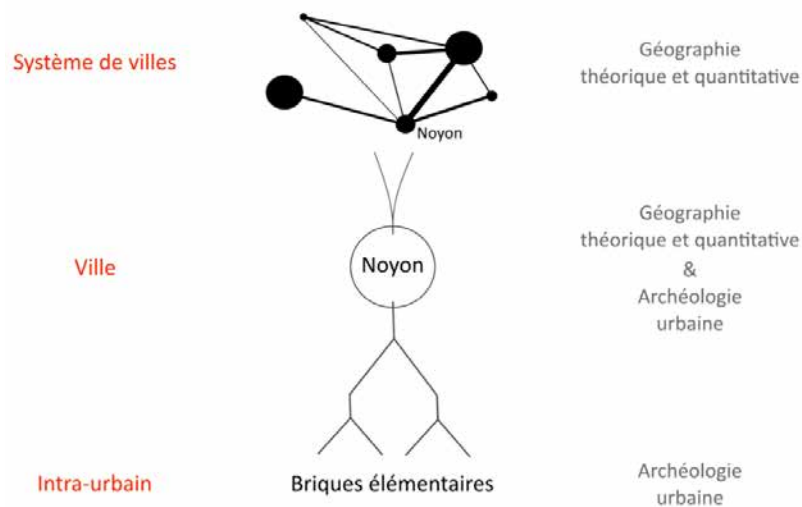


Figure 5. Échelles d'analyse privilégiées et héritages historiographiques dans la thèse de Gravier (2018, p. 36)

À l'échelle interurbaine, pour identifier les systèmes de villes dans lesquels s'inscrit Noyon au fil du temps, nous avons modélisé les relations entre villes par des « relations fonctionnelles » et des « relations potentielles ». Les catégories fonctionnelles pour caractériser les relations entre villes ont été construites à partir du regroupement des catégories fonctionnelles du CNAU (les « valeurs urbaines » évoquées plus haut). Dès le départ, en amont de l'analyse, la conception des relations fonctionnelles devait donc permettre d'articuler les échelles méso et macro-géographiques. À titre d'exemple, la ville de Noyon est une ville secondaire de la cité romaine des Viromanduens durant le 1^{er} siècle après J.-C. dont le chef-lieu est Saint-Quentin. Les deux villes ont donc des relations d'ordre politico-administratif à cette époque. Quant aux relations potentielles, nous avons envisagé un seuil de distance-temps à considérer (une journée aller-retour), c'est-à-dire une catégorie générique et a-historique englobant des réalités multiples selon les époques, en fonction des types de transports et donc des vitesses de déplacement.

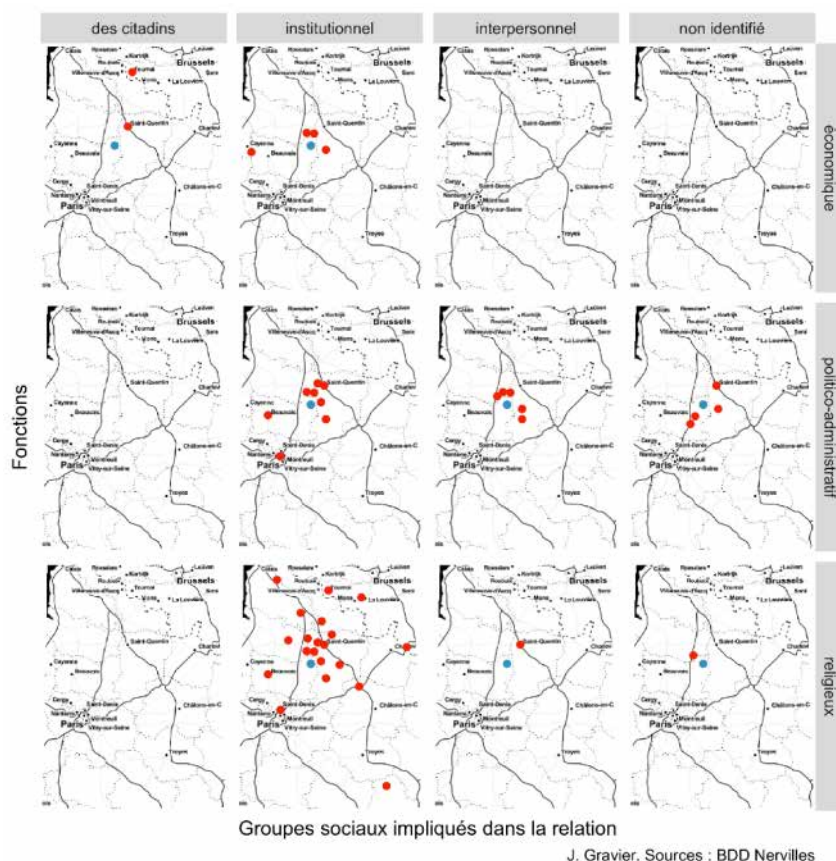


Figure 6. Fonctions et groupes sociaux impliqués dans les relations fonctionnelles entre Noyon – en bleu – et d'autres villes (Gravier, 2018, p. 263)

Nous avons ainsi établi une « ontologie des relations » constituée d'objets a-historiques – les relations fonctionnelles et les relations potentielles –, nous permettant *a priori* d'étudier l'évolution de la ville de Noyon dans les systèmes de villes dont elle fait partie dans la longue durée. Toutefois, une fois ces objets génériques établis, l'enrichissement de cette ontologie s'est révélé nécessaire car les sources disponibles sont très lacunaires. À cette fin, nous avons inclus des éléments de réflexion de la sociologie politique et de l'anthropologie structurale à l'ontologie d'origine (Gravier, 2018, p. 231-247), qui relevait quasi exclusivement du domaine archéologique, afin d'établir une grille analytique retranscrivant aussi les groupes sociaux en jeu dans les relations identifiées, comme présentés dans la figure 6 (les groupes institutionnels, citadins, les relations interpersonnelles, etc.). Nous observons par exemple que les relations institutionnelles d'ordre politico-administratif (sixième cadran) existent principalement à l'échelle micro-régionale :

la plupart des villes sont très proches de Noyon, à l'exception de Beauvais et de Paris plus éloignées. Cette grille permet d'observer que Noyon (en bleu sur la figure) est inscrite dans des réseaux politico-administratifs et économiques d'échelle micro-régionale, mais dans des réseaux macro-régionaux pour ce qui relève du religieux.

2.2. Une ontologie pour étudier l'évolution du territoire français à partir de cartes historiques

La deuxième ontologie a-historique de domaine que nous discutons ici relève d'un contexte totalement différent dans la mesure où elle a été mise en place sur une période courte dans le cadre d'un programme ANR de trois années – programme GeoPeuple⁶ –, et en l'absence de travaux précurseurs. L'ontologie présentée ici hérite donc des objectifs thématiques et méthodologiques de ce projet : étudier l'évolution topographique et fonctionnelle locale du territoire français depuis la Révolution française au regard des évolutions administratives de ses communes, et cela en comparant des cartes historiques – comparaison dont l'automatisation a été explorée. Nous revenons ici sur cette expérience de recherche en explicitant comment ont été construits les objets d'analyse à partir de sources diverses (différentes cartes topographiques historiques et une base de données vectorielle) et comment ils ont été manipulés pour construire un cadre d'analyse pour une étude quantifiée des trajectoires des communes de France sur trois siècles.

2.2.1. Présentation de la question d'analyse et des sources

Depuis 200 ans le territoire français s'est considérablement équipé, densifié au fil de l'accroissement démographique, des progrès technologiques et de l'industrialisation (Lepetit et Royer, 1990, p. 987-1010). L'étude des relations entre l'évolution du paysage français et l'évolution de la répartition de la population sur le territoire était l'objectif du projet GéoPeuple financé par l'Agence nationale pour la recherche (Ruas *et al.*, 2012 p. 6). Le paysage s'apparente dans ce contexte à l'espace topographique (qu'il soit naturel ou anthropique) analysé à travers les traces codifiées que nous offrent les cartes anciennes : cartes de Cassini au 1:86 400 (1750-1789 environ) et cartes d'état-major au 1:80 000 (1866-1895 environ), mais également à travers l'information présente dans le référentiel grande échelle (RGE), en particulier la BD TOPO®, que fournit l'IGN (2011). L'espace administratif porteur d'informations démographiques est matérialisé par le maillage des communes, qui évolue au fil des recompositions communales (Motte et Vouloir, 2008). Le projet s'appuie sur trois cas d'étude, à savoir trois zones où les objets des cartes topographiques anciennes ont été entièrement vectorisés : autour de Reims, Grenoble et Saint-Malo.

6. <http://geopeuple.ign.fr/>

La question était de savoir s’il existe des relations entre l’évolution du paysage selon les sources topographiques et la dynamique démographique et administrative (modification des limites des communes) du territoire français au cours du temps. En effet, les sources topographiques, à l’échelle à laquelle nous les étudions, rendent compte de la distribution des lieux de peuplements et de leur masse au cours du temps (avec la représentation des hameaux comme des villes), mais aussi des objets qui peuvent éventuellement expliquer les évolutions démographiques et administratives : le réseau de transport, le tissu industriel (avec la localisation des usines, leur production) et les facteurs géophysiques (nature du sol, relief, hydrologie). C’est l’étude de ce pan du projet – la caractérisation du paysage topographique des communes du territoire – sur laquelle nous revenons ici.

Dans un premier temps, nous décrivons les résultats de l’analyse, comme si elle pouvait être lue sans s’intéresser dans le détail aux modalités de construction des données qui la constituent, puis nous expliquons ensuite que, justement, ces données sont issues d’un construit qui influence fortement l’analyse (§2.2.2 et §3.1).

*Tableau 1. Indices à évaluer pour chaque source par commune
(dans leurs limites actuelles)*

Indice	Type
La densité de population	Quantitatif
La surface de tissu urbain	Quantitatif
Le nombre d’habitat dispersé	Quantitatif
Le nombre de kilomètre de routes	Quantitatif
Le nombre de routes afférentes	Quantitatif
Le nombre de kilomètre de cours d’eau artificiels	Quantitatif
Le nombre de ponts	Quantitatif
Le nombre de bâtiments industriels	Quantitatif
Le nombre de constructions militaires	Quantitatif
Le nombre de postes	Quantitatif
Le nombre bâtiments religieux de toutes sortes	Quantitatif
Le nombre d’églises	Quantitatif
Le statut de chef-lieu de canton, d’arrondissement ou de département	Qualitatif

Le parti pris de l’analyse a été, dans le cadre de cette étude (Ruas *et al.*, 2013), de prendre le maillage communal actuel comme référence et de calculer à l’intérieur de ce maillage des indices (tableau 1) pour caractériser les trois zones d’étude aux

trois époques d'étude (Cassini, état-major et contemporaine), puis d'étudier leurs changements sous la forme de transition de classe. Les indices possibles qualifient surtout des caractères plus ou moins urbains (postes, routes, tissu urbain, bâtiments industriels, etc.), tandis que les aspects ruraux des territoires sont peu représentés, surtout en termes d'activité agricole. Il n'y a pas d'indice exprimant, par exemple, la surface cultivée, la surface moyenne des parcelles ou le nombre d'exploitations. Nous nous attendons donc à ce que les territoires soient caractérisés dans l'analyse avant tout par le fait urbain – les espaces ruraux, eux, seront décrits en fonction de l'activité industrielle ou en pendant négatif aux espaces urbains (pas d'habitat groupé, peu de services, etc.).

La figure 7 ci-dessous présente un exemple de résultat de l'exploitation de ces indices avec une classification ascendante hiérarchique après le calcul d'une analyse factorielle en composantes multiples.

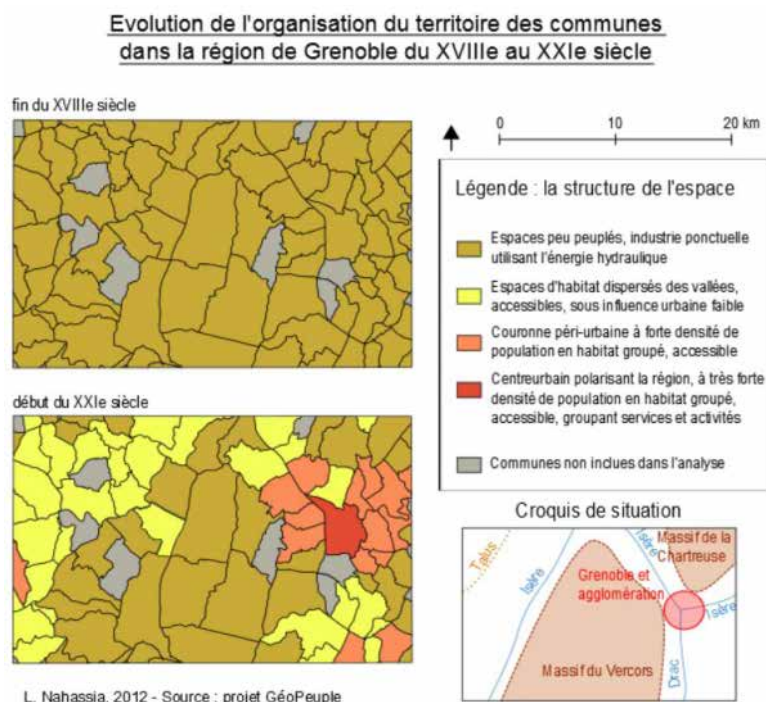


Figure 7. Transitions de classes observées entre environ 1750 (Cassini) et 2011 (BD TOPO©) sur la région de Grenoble (Nahassia et Plumejeaud, 2012, p. 38)

À la fin du XVIII^e siècle, toutes les communes appartiennent à la même classe (la classe grise correspondant à des unités qui n'ont pas été incluses dans l'analyse) : le territoire est peu densément habité, de caractère plutôt rural, mais il est surtout

uniforme. En prenant en considération les trois premiers axes de l'analyse, il n'y a pas de forte différenciation fonctionnelle des communes. Le premier axe de différenciation par l'analyse en composantes principales (ACP) permet d'identifier des communes telles que Grenoble ou La Tronche qui ressortent avec un caractère plus urbain que les autres mais cette spécificité a été amoindrie sous l'influence des deux autres axes factoriels. À l'inverse, en 2011, il existe une très forte polarisation de l'espace autour de la ville de Grenoble, véritable capitale régionale, et dans les vallées du Drac et de l'Isère.

2.2.2. Mise en œuvre d'une ontologie pour dépasser l'hétérogénéité des sources

Les résultats numériques qui précèdent découlent de la fabrique des indices mobilisés, et procèdent de la mise en œuvre d'une ontologie a-historique. Cette ontologie a été conçue lors du projet *ad-hoc* à partir de l'analyse des légendes des cartes, et suivant une conception de l'aménagement du territoire très contemporaine. En effet, chacune des cartes est construite selon des objectifs spécifiques, en suivant des modes de sélection de représentation du réel qui diffèrent, chacune nommant diversement les objets. Ainsi, il a préalablement été nécessaire d'identifier les pictogrammes présents sur les cartes de Cassini et de l'état-major et, dans un premier effort « ontologique », de les regrouper en des catégories qui faisaient sens pour l'analyse et à partir desquelles les indices pouvaient être construits. L'exemple suivant détaille par exemple comment l'indice « nombre d'habitat dispersé » a été fabriqué à la lecture des cartes de Cassini et d'état-major.

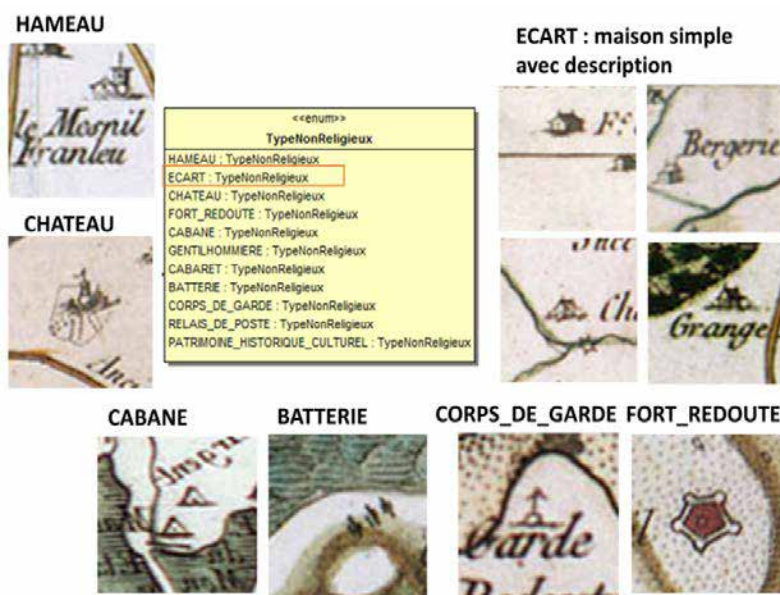


Figure 8. Pictogrammes permettant d'apprécier la densité d'habitat sur Cassini (ici le non-religieux)

La carte de Cassini répond à une volonté d'une meilleure maîtrise du territoire pour améliorer le rendement de l'impôt (Pelletier, 1990, p. 263) et pour accompagner la stratégie militaire : quelles sont les voies d'accès pour les troupes et quelle est leur largeur ? Où sont les villages et les fermes pour s'approvisionner ? Quels sont les marqueurs dans le paysage qui permettent de se repérer sur le terrain ? (Cénat, 2008, p. 28-30). Ainsi, les lieux d'habitat densifié (villages et petites villes) sont signifiés par une église, ce qui justifie la création d'un indice spécifique « nombre d'églises ». Cependant ce sont d'autres pictogrammes, regroupés en base de données dans une classe « Non-religieux » qui rendent compte de lieux d'habitats épars : des châteaux, des granges, des bergeries, etc. (figure 8), et c'est leur décompte qui a permis de renseigner l'indice « habitat dispersé ».

La carte d'état-major quant à elle est un exemple extrême d'un effort d'harmonisation interne à la source, ce qui simplifie en partie le travail de définition d'une ontologie utile pour questionner ici la concentration et la densification de l'habitat. À partir du relevé et de la traduction systématique de tous les termes contenus dans la légende de la carte d'état-major, un glossaire de 400 termes a été constitué. Un grand nombre de ces termes sont en réalité synonymes. Nous trouvons par exemple de nombreux termes relatifs à l'habitat agricole et/ou saisonnier, propres à chaque région : une bergerie se dit cortal dans les Pyrénées-Orientales, orry (orri, horry, orrhy) dans les Pyrénées, en Gascogne et en Ariège (Pégorier 2006) ; une cabane d'estive est une vacherie (Bouyssou, 1987) ou encore un buron en Auvergne, en particulier dans l'Aubrac (Peiffer, 1877). Ce qui est donc remarquable, c'est la volonté d'utiliser le langage des terroirs, et ceci toujours dans l'idée que sur le terrain, le militaire en campagne puisse demander son chemin et comprendre la réponse donnée (Berthaut, 1898). Ainsi, les termes usités sont signifiants dans la région et l'époque où ils ont été relevés, mais peuvent ensuite être tombés en désuétude ou avoir changé de signification. Ils représentent cependant tous à peu près la même réalité, rendant comparables trois régions très éloignées comme le Y grenoblois, la campagne rémoise et la baie du Mont Saint-Michel, et peuvent donc être ordonnés dans la même catégorie. Cette analyse a permis de regrouper ce type de symboles et de dénominations dans une catégorie « Bâti-civil », déclinée en sous-catégories plus spécifiques (figure 9) adaptées à nos questionnements, tels les bâtiments dédiés à la santé ou à l'hébergement (sous-entendu collectif, comme les hospices ou les auberges). En particulier, les sous-catégories « Habitat », « Habitat-agricole » et « Habitat-saisonnier » regroupant diverses dénominations de lieux dans des ensembles sont celles qui permettent de renseigner l'indice « Habitat dispersé ». Sous « Habitat » sont regroupés par exemple les bâtiments que l'on trouve désignés sous les termes Château (Ch^{au}), Manoir, Gentilhommière, Bastide (B^{ide}), Maison (M^{on}), Pavillon (P^{on}), et dans « Habitat saisonnier », par exemple les orrys, cortal, vacherie et buron.

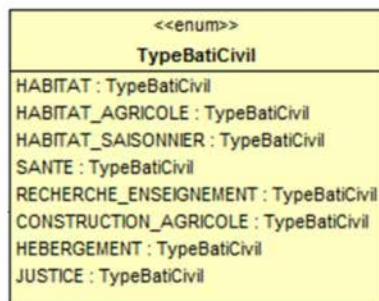


Figure 9. Regroupement et classification des objets de type « bâti civil » en différentes catégories, à partir de la carte d'état-major

Ces représentations diffèrent encore des objets de la BD TOPO© qui ne donnent aucune indication de destination pour l'habitat : il est représenté de façon exhaustive par son emprise au sol, il faut alors trancher par des critères spatiaux entre habitat dispersé ou urbain, et il est impossible d'analyser plus finement les évolutions de sous-catégories d'habitat (agricole, saisonnier ou d'hébergement). Ces divergences entre sources ont donc nécessité une réflexion ontologique préalable pour proposer des indices pertinents pour l'analyse diachronique, impliquant de regrouper des items en catégories a-historiques.

3. Critique de ces ontologies a-historiques

La construction des ontologies a-historiques permet une formalisation stricte des objets et, partant, leur comparabilité dans des systèmes informatiques ou statistiques. Les outils de la géographie quantitative et de l'analyse spatiale peuvent ainsi être mobilisés dans un cadre diachronique, comme illustré par les exemples précédents. Mais l'usage de ces méthodes, parce qu'elles mettent de manière explicite les objets étudiés en confrontation, révèlent en même temps les biais des ontologies mobilisées, qui sont, avant tout, des construits des analystes. Elles dépendent des modèles de pensées inscrits dans le temps de la recherche, dans le contexte de production du savoir et celui de l'usage auquel elles sont destinées, trois biais que nous discutons dans cette partie.

3.1. Conséquences liées à la nature même des sources

Dans les cas d'étude présentés dans la première partie, nous avons construit des catégories a-historiques à partir d'objets qui n'étaient pas initialement destinés à l'usage analytique qu'on en a fait (les cartes par exemple) et, par construction, chaque base de données est peuplée de sur ou sous-représentations de chaque

catégorie à un instant donné. Par ailleurs, ces catégories sont elles-mêmes le fruit de nos propres représentations mentales, ce qui nuance les résultats statistiques de nos analyses.

Par exemple, dans le cas de l'étude des activités à Tours, le résultat illustré dans la figure 4 (« Six catégories génériques pour étudier le rythme du changement urbain à Tours sur 2000 ans ») révèle à la fois des faits thématiques et des biais dus aux conditions de collecte des sources et au contexte de construction de la base de données. Dans le cas des édifices religieux, la disparition à la fin du XVIII^e siècle de très nombreux objets renvoie ainsi à une réalité historique – leur confiscation et leur revente pendant la Première République. Cependant, cette « chute » très visible dans le graphique n'est pas contrebalancée par un regain de cette catégorie quelques années plus tard, qui illustrerait le retour partiel des usages religieux des bâtiments confisqués ; la raison en est que, pour la majorité de ces objets, les chercheurs ont arrêté la datation en 1800.

Dans le cadre du projet Géopeuple, les résultats d'analyse peuvent aussi être discutés de ce point de vue. Par exemple, il reste étonnant que les villes se détachent si peu au XVIII^e siècle dans les résultats d'analyse, en particulier à Grenoble (voir figure 7). Les résultats des classifications signifient que, dans l'ensemble, le profil urbain du XVIII^e siècle est plus proche du profil « rural » du XVIII^e et du profil « très peu peuplé » du XXI^e siècle, que du profil des zones urbaines contemporaines. Ceci peut être compris du fait de la densité de population : même les communes urbaines étaient alors très peu peuplées proportionnellement aux communes urbaines actuelles. Cependant, nous attendions que d'autres éléments, tels que la surface urbaine, fassent ressortir les communes urbaines du XVIII^e siècle.

Or tel n'est pas le cas. Ceci s'explique aussi par la nature de l'information mobilisée. En effet, les données contenues sur les cartes de Cassini sont moins nombreuses que celles de la BD TOPO© : beaucoup moins d'objets y sont représentés. Cela ne veut pas dire qu'il existait effectivement moins de bâtiments au XVIII^e siècle qu'aujourd'hui (dans une certaine mesure c'est le cas pour les villes), mais que le report des objets sur la carte est moins systématique. Par exemple, une ferme en écart sur la carte de Cassini est représentée par un simple point, même si elle est constituée de plusieurs bâtiments ; en revanche dans la BD TOPO© et la carte d'état-major, chacun de ces bâtiments est figuré.

Ce biais a une incidence forte sur les valeurs que prennent les indices, et ceci malgré le fait qu'ils soient normalisés en les rapportant à la même surface d'étude (celle de la commune actuelle). Ainsi la comparaison de leur distribution pour chaque zone d'étude entre chaque époque, comme par exemple celle de l'indice « nombre d'église », entre la carte de Cassini (temps t_1) et la BD TOPO© (temps t_2) sur la figure 10, montre que les indices Cassini ont une distribution moins étendue que les indices de la BD TOPO© et que les valeurs à t_2 sont plus dispersées autour de la moyenne (étendue qui vaut 0,92 en 2011 contre 0,67 sur Cassini) avec un maximum des valeurs systématiquement plus grand – 50 par communes en 2011

contre un maximum de 27 églises par communes suivant Cassini. Ce cas d'espèce concernant la non-comparabilité des distributions des églises est d'autant plus marquant et démonstratif que les cartes de Cassini étaient construites sur une triangulation de la France utilisant les clochers des églises, et que ceux-ci avaient le plus de chance d'être nombreux.

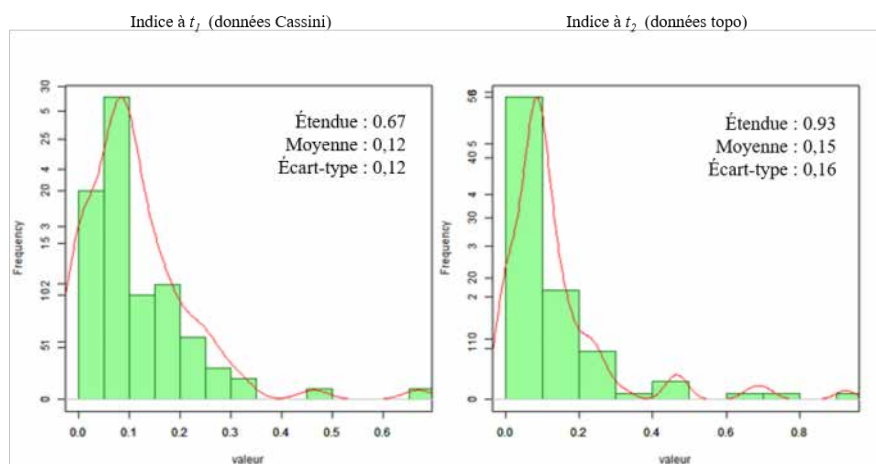


Figure 10. Distribution des valeurs de l'indice « nombre d'églises » pour la zone de Reims (Nahassia et Plumejeaud, 2012, p. 34)

De ce fait, et mécaniquement, lorsque ces indices sont étudiés et analysés conjointement, les communes décrites par les indices Cassini apparaissent comme uniformes. Ce biais a pu être évité pour certains indices (notamment pour l'indice « industrie »), mais pas pour l'ensemble des indices. Il s'agit finalement bien là du résultat d'une étude de la représentation du territoire et non de sa réalité. Tout est question de contexte : si nous analysons les territoires époque par époque, les indices Cassini seront discriminants, et le territoire sera différencié. L'analyse rassemblant les indices des deux époques dilue cette faculté discriminante de l'information topographique issue des cartes de Cassini face au pouvoir de différenciation beaucoup plus important des indices issus de la BD TOPO©.

3.2. Conséquences liées à la pratique des chercheurs

Outre des biais liés aux sources mobilisées dans les analyses, l'usage des ontologies dans le cadre de traitements statistiques révèle également les pratiques des chercheurs qui en sont à l'origine. En particulier, nous pouvons ici discuter des catégories sémantiques du thésaurus du CNAU et de celles qui en découlent, que ce soit celles utilisées pour la base ToToPI ou celles mobilisées pour Noyon, dont

certaines ont en réalité un caractère nettement historique. Autrement dit, alors que ces ontologies se veulent le plus génériques possible, en l'occurrence a-historiques, les pratiques des chercheurs introduisent cependant des catégories qui sont intimement liées à leur discipline ou à une époque.

Ainsi dans ToToPI (figure 11), les objets sont caractérisés par une « valeur d'usage » qui décrit des pratiques sociales et des usages de l'espace (l'habitation, les thermes, la fortification, etc.), qui sont regroupées dans des « valeurs urbaines » de plus haut niveau. La composition de ce thésaurus est finalement le miroir d'une pratique disciplinaire, celle de l'archéologie. En effet, *a posteriori*, nous y voyons d'abord une catégorisation du réel en plusieurs dimensions héritées d'une pensée wébérienne de la ville (Weber, 2014 [1921]) – l'économique, le politique religieux et civil, le militaire – auxquelles ont été ajoutées d'une part, une dimension qui renvoie à un sujet de prédilection des archéologues – le funéraire –, et d'autre part, une dimension pour classer tout ce qui ne rentrait pas ailleurs – les « Aménagements » –, catégorie que d'autres disciplines auraient peut-être pensée autrement. Par ailleurs, ces deux niveaux de description renvoient, à l'origine, à deux niveaux d'analyse de l'espace urbain par les archéologues : l'échelle locale, de la rue au quartier qui est l'échelle des fouilles, et les échelles méso et macro, du quartier à la ville dans son ensemble, qui sont celles de l'analyse.

Ensuite, pour certaines de ces catégories, l'historicité est évidente. Par exemple, la valeur d'usage « Cimetière paroissial » ne peut exister, par définition, qu'à partir du moment où les paroisses sont mises en place dans l'Occident chrétien. Certains traitements statistiques font ressortir l'historicité d'autres classes qui sont pourtant, *a priori*, génériques. En particulier, une analyse factorielle exploratoire a été menée sur Tours pour caractériser des périodes temporelles (périodes régulières de 25 ans) au regard des valeurs d'usage et des valeurs urbaines existant au cours du temps (Nahassia, 2019, p. 179-232). Que ce soit sur les axes factoriels ou dans l'exploration du dendrogramme des classifications ascendantes hiérarchiques subséquentes, nous observons une moins grande variabilité du profil fonctionnel des périodes de 25 ans lorsque celles-ci sont caractérisées par les valeurs d'usages, alors même que ces dernières sont plus nombreuses que les valeurs urbaines. En d'autres termes, alors que les valeurs urbaines participent fortement (en positif ou en négatif) à la caractérisation de multiples groupes de périodes de 25 ans, chaque valeur d'usage a, elle, tendance à caractériser un groupe de périodes et un seul. C'est également ce qui a pu être observé pour le cas de la ville de Noyon sur des périodes de 50 ans. D'ailleurs, les deux premiers axes factoriels révèlent un fort effet Guttman pour les valeurs d'usages contrairement aux valeurs urbaines, typique d'un effet de sériation chronologique (Djindjian, 1991). À l'épreuve des données, les valeurs urbaines restent donc très génériques, et donc a-historiques, tandis que les valeurs d'usage sont plus situées historiquement. En outre, la plus grande stabilité du profil fonctionnel de l'espace urbain tourangeau concerne en particulier un long Moyen Âge, les périodes entre 300 et 1100, caractérisées en particulier par les catégories du religieux (lieux de cultes, couvents et monastères, etc.).



Figure 11. Le thésaurus hiérarchisé en valeurs d'usage et valeurs urbaines de la base de données ToToPI (Nahassia, 2019, incipit)

Ce double motif – les valeurs d'usage sont moins a-historiques que les valeurs urbaines, et les périodes médiévales sont particulièrement stables au regard de la variabilité des valeurs d'usage dans le temps – découle en partie des pratiques de recherche sous-jacentes à la construction de ce thésaurus. En effet, l'ontologie qui caractérise les objets archéologiques en milieu urbain découle d'une construction collective et elle reflète à la fois les besoins et les aspirations des chercheurs qui l'ont élaborée au cours du temps, à savoir des archéologues spécialistes de l'espace urbain. Les catégories ont été constituées en partant du terrain, puisque ce sont d'abord les termes désignant les observations archéologiques qui ont été assemblées en rubriques (les valeurs d'usages), elles-mêmes, par la suite, regroupées en grandes fonctions (les valeurs urbaines) (§2.1.1). Or, l'archéologie urbaine française est fortement constituée par des archéologues médiévistes. D'une part, ce sont donc naturellement les caractéristiques sémantiques centrales dans les études de la ville médiévale qui dominent le thésaurus des valeurs d'usages (en particulier les activités liées à la chrétienté, facteur important de l'organisation spatiale urbaine au Moyen Âge). D'autre part, la ville de Tours est particulièrement bien connue pour cette période et dans cette thématique, et les OH correspondants sont donc particulièrement nombreux au regard des autres catégories – ce qui influence les résultats des traitements statistiques. Il est d'autant plus intéressant que ce phénomène ressorte avec la base ToToPI car elle a été enrichie au cours du temps par de nombreuses études et chercheurs. C'est donc bien une pratique collective – et

non la tendance d'un ou de deux spécialistes – qui émerge ici de l'ontologie, une fois celle-ci mise en pratique par des traitements statistiques.

Enfin, parmi les pratiques des chercheurs ayant des conséquences sur les ontologies réalisées, il nous semble important de relever que les acteurs de la collecte dans le domaine de la recherche veulent généralement conserver une forme de contrôle étroit sur leurs données et l'usage qui s'en fera, tout comme leur autonomie (et ils ont certainement raison dans un sens). Ils préfèrent donc choisir des outils et des modèles dont ils ont la pleine maîtrise, ce qui freine le développement de modèles de bases de données et d'ontologies plus innovants. Cette pratique, plus ou moins conscientisée par les acteurs eux-mêmes, a également des conséquences sur les approches conceptuelles mises en place dans les ontologies. Ceci peut expliquer la faible utilisation d'ontologies axiomatisées proposées par le domaine de l'ingénierie des connaissances, parce que leur emploi requiert des compétences hors champ des chercheurs, et les rend dépendants « du bon vouloir » d'acteurs informaticiens.

3.3. Conséquences liées aux approches conceptuelles

L'ontologie est une construction qui reflète les choix conceptuels de chercheurs, qui sont dictés de notre point de vue par trois facteurs qui agissent conjointement : la nature des sources mobilisées, les *habitus* disciplinaires et le questionnement scientifique qui guide l'analyse. Le poids de ces facteurs peut être illustré par exemple par une question centrale dans la réflexion sur la construction des objets dans la longue durée : le fait d'accorder ou non une identité intrinsèque aux objets. Dans le premier cas, les objets sont dits « endurants » (objets SNAP) car ils « existent dans leur entièreté [...], ont une étendue dans l'espace [...] et forment un tout, observable en chaque instant » (Mathian et Sanders, 2014, p. 32). Dans le second cas, les objets sont « perdurants » (objets SPAN). Ils sont avant tout caractérisés par une durée dans le temps et de ce fait, si l'on observe un objet de ce type à un moment *t*, « on n'a pas accès à son entièreté » (*idem*, p. 32). Ces deux modèles permettent des types d'analyses différents. Le modèle qui sert de support à l'étude des évolutions communales dans GeoPeuple relève du SNAP, tandis que celui qui sous-tend OH_FET se rapproche du SPAN, ils sont donc diamétralement opposés concernant cet aspect d'identité.

En effet, le parti pris de GeoPeuple (Plumejeaud *et al.*, 2015) est d'accorder un identifiant aux communes dont il est possible de suivre les mouvements de fusion, scission, redécoupage au cours du temps, même si leur emprise, leur démographie, leur nom ou d'autres attributs changent. L'usage d'identité dans les communes de GeoPeuple découle d'un *habitus* disciplinaire que nous avons pu observer à plusieurs reprises en histoire moderne. Par exemple, les navires étudiés dans le programme ANR Navigocorpus (Dedieu *et al.*, 2012) se voient attribuer un identifiant unique dans la base de données, suite à une opération de comparaison manuelle, même s'ils apparaissent dans des sources différentes et dans des ports

différents. Cette identité facilite l'étude de trajectoire d'objets (navires ou communes), qui est une question centrale dans ces études. C'est aussi une pratique courante en géographie dès lors que le sujet d'étude touche aux déplacements d'objets dans le temps – par exemple, dans le cas des études des mobilités courtes (déplacements domicile-travail, mouvement d'une tempête) ou longues (études des migrations). Cette approche est possible dès lors que la collecte des informations se fait sur des objets contemporains et/ou aisément accessibles : individus, animaux, phénomènes naturels sont alors conceptualisés et enregistrés dans le temps de l'étude, par questionnaires ou entretiens, *tracking*, observation en temps réel, etc.

Cependant, dans le cas d'études diachroniques qui remontent encore plus loin dans le passé, avec l'utilisation de sources historiques et archéologiques moins complètes, ce type d'information n'est pas nécessairement accessible, ce qui rend la construction d'objets durables difficile. L'intégration d'éléments d'interprétation portant sur l'identité dans la construction d'une base de données historique, au moment de la collecte de l'information, peut s'avérer périlleuse. En effet, cela peut introduire des erreurs qui seront ensuite difficiles à corriger. Il s'agit d'un choix interprétatif fort, qui peut parfois relever de la pure spéculation. C'est une des raisons qui peut expliquer qu'en archéologie, et au vu parfois de la difficulté de « recoller » les morceaux, l'enregistrement strate par strate des micro-éléments et de leur description, même s'ils font partie d'un ensemble plus vaste, se fait sans positionner *a priori* une identité d'appartenance aux fragments collectés. La possibilité reste ainsi ouverte au chercheur, sans être imposée, d'attribuer en aval un héritage, une continuité entre les objets, ou non, au moment où il se saisit de ces objets pour l'analyse. Le parti pris d'OH_FET est donc de ne pas accorder d'identité aux traces archéologiques et de procéder plutôt à une multiplication d'objets « perdurants » sans faire bénéficier les traces plus récentes de propriétés (comme la fonction) de traces plus anciennes qui se superposeraient spatialement. Dans la base de données ToToPI, un objet est une unité fonctionnelle et spatiale homogène et tout changement dans ces deux dimensions aboutit à la création d'un nouvel objet.

Dans l'analyse des activités à Tours à partir de cette base de données, le choix a été fait de ne pas créer de lien d'identité temporelle entre les OH pour les transformer en objets durables – pour la plupart d'entre eux un tel lien est d'ailleurs impossible à établir. Le revers de cette décision est qu'il existe bel et bien dans la base de données des objets qui représentent une même réalité à différentes époques, laquelle est modélisée en différentes unités du fait de changements de forme ou de fonctions, et qui peut être facilement identifiable et reconstituée *a posteriori*. C'est le cas par exemple de la cathédrale de Tours, ou encore de certains cimetières. Un certain nombre d'entre eux changent d'emprise au cours du temps (extensions ou rétraction), ce qui implique dans le modèle OH_FET des changements d'objets, comme l'illustre la figure 12 (le cimetière de Saint-Clément connaît trois emprises entre 1000 et 1800). Ainsi, le décompte de l'apparition des cimetières au cours du temps (tel qu'il est utilisé par exemple dans le résultat d'analyse illustré par la figure 4, §2.2.1) est partiellement biaisé : pour certains d'entre eux il ne s'agit pas de

la création d'une activité, mais de sa réapparition sous une forme différente. Néanmoins, dans le cadre de l'analyse, le changement de forme peut être lu comme une réactivation de l'activité dans le lieu considéré (on agrandit un cimetière plutôt que d'en créer un nouveau ailleurs), ce qui révèle la résilience de celle-ci, et ce qui interroge sur les conditions de cette réactivation (dans quel environnement urbain ? à quelle fréquence ? etc.).

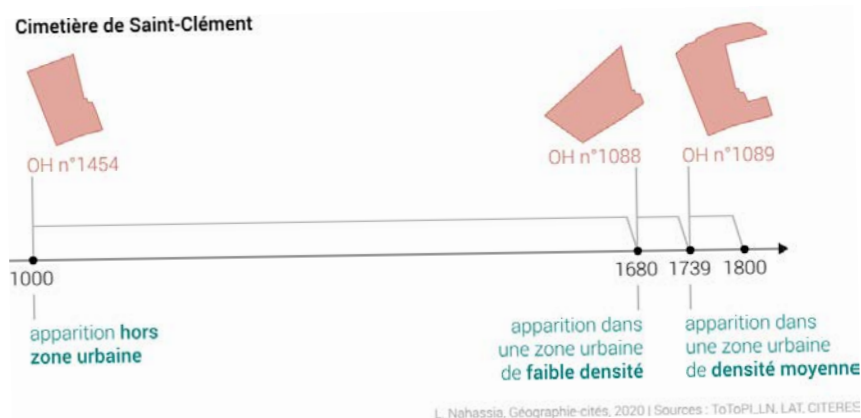


Figure 12. Évolution de la forme du cimetière de Saint-Clément dans la base ToToPI : modélisation en trois OH et interprétation

Une extension du modèle OH_FET propose même d'aller plus loin dans la décomposition de l'identité des objets, en découpant tous les objets d'un même corpus selon leurs superpositions dans les dimensions spatiales et temporelles (Rodier et Saligny, 2010). L'intersection de toutes les géométries et de toutes les durées d'existence des objets aboutit à des morceaux considérés comme « neutres » d'espace et de temps (les « entités spatiales » – ES et les « entités temporelles » – ET). Par exemple, sur la figure 13, l'objet historique OH2 est composé des ES2 et ES3, et de l'ET1. Chaque morceau et chaque dimension peuvent alors être étudiés indépendamment les uns des autres, en se demandant par exemple quels morceaux d'espace sont le plus fréquemment mobilisés par l'occupation urbaine – une approche très utile dans le cadre d'analyses morphologiques de l'espace (voir par exemple Simon, 2015 sur Vendôme). Mais ce découpage fait perdre toute valeur thématique aux objets finalement manipulés, qui ne prennent sens au regard de la réalité que lorsqu'ils sont recombinaison les uns aux autres.

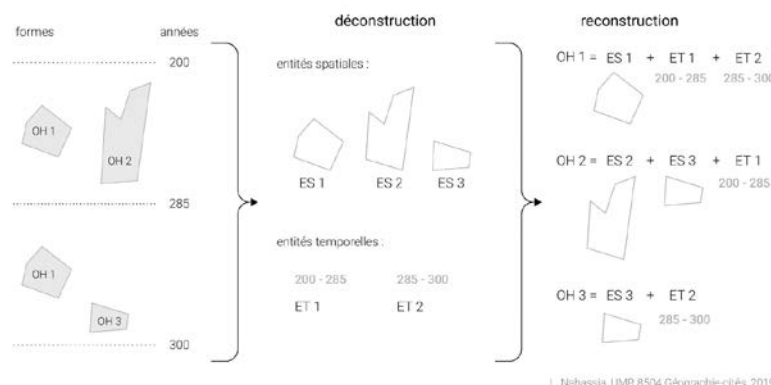


Figure 13. Principe de déconstruction des OH en entités spatiales et temporelles (Nahassia, 2019, p. 91)

Cela rejoint une pratique courante en informatique, qui serait de collecter sans donner d'identité, et de ne procéder à l'identification qu'*a posteriori*. Il existe un pan de recherche en informatique dédié à ces problématiques de conflation/fusion/appariement de miettes d'objets, ancien dans le domaine de la recherche d'information (Lennon, 1981), et renouvelé par la gestion informatisée des cartes (Ruiz *et al.*, 2011) et les bases de données spatiotemporelles (Wang, 2007 ; Recchia et Louwerse, 2013). Ces approches donnent lieu à des développements appliqués à l'Histoire, avec par exemple la thèse de Benoît Costes (Costes, 2016) qui fait suite à ses travaux sur l'appariement de moulins entre les cartes de Cassini et d'état-major (Costes *et al.*, 2014). Mais ce choix peut aussi conduire à des bases de données très redondantes et volumineuses, comme discuté par Plumejeaud *et al.* (2015), et surtout difficilement gérables par les outils que savent manipuler les acteurs de la collecte. Si ce type d'ontologie et de modèle qui ne donnent pas d'identité endurante aux objets peuvent être utilisés en archéologie urbaine, c'est finalement parce que les connaissances y sont, somme toute, relativement limitées en volume, et que les bases de données correspondantes contiennent peu d'objets, même une fois ceux-ci découpés en briques les plus élémentaires.

4. Produire des connaissances reproductibles malgré tout

Si les ontologies sont des construits par des chercheurs et, de ce fait, relèvent de savoirs situés, la construction de ces ontologies a-historiques permet une formalisation stricte des objets et, partant, leur comparabilité dans des systèmes informatiques ou statistiques, mais aussi dans un cadre plus général de construction

des connaissances. En cela, elles facilitent la production de connaissances reproductibles.

4.1. Réutilisation potentielle des données optimisée

Nos différentes expériences de manipulation de bases de données historiques mettent en exergue la nécessité de documenter non pas seulement les données, sous la forme d'un dictionnaire, mais aussi le processus de leur construction, le comment et le pourquoi. De la production d'une ontologie découle normalement la rédaction d'une documentation spécifiant clairement les objectifs des catégories retenues et les modalités de leur construction : les rapports de GeoPeuple répondent à ce besoin par exemple (Nahassia et Plumejeaud, 2012 ; Plumejeaud, 2012a, 2012b, 2012c et 2012d). Ce sont des métadonnées essentielles à la compréhension fine d'un corpus. L'absence de ces métadonnées fait tomber les utilisateurs qui n'ont pas produit les données dans des pièges évidents, et surtout ne rend pas suffisamment compte du fait que les valeurs numériques sont le fruit avant tout d'un processus interprétatif :

« Surtout, la conception d'un life cycle qui part de l'océrisation des sources pour en tirer des textes numériques qui seront annotés et enrichis afin d'en extraire des informations en vue de produire de nouvelles connaissances – à chaque étape de ce processus on rencontre des "données" de nature différente – correspond certes à la méthode informatique mais ne tient pas suffisamment compte d'un acquis essentiel de la méthode historique selon laquelle toute "donnée" historique, tout "fait" historique même élémentaire, résulte d'une construction, c'est-à-dire d'une opération mentale qui prend son origine dans un questionnement [Bizière, Vayssière]. Même les historiens de la fin du XIX^e siècle, théoriciens et praticiens de l'historiographie méthodique et positiviste dont la conception de "fait historique" se rapproche peut-être le plus d'une conception objectiviste de la donnée, étaient conscients de l'importance du questionnement et de la *dimension interprétative de la discipline historique*. » (Beretta, 2017, p. 2-3).

Si ce besoin existe pour toute base de données, il a l'avantage d'apparaître de manière très nette dès lors que les objets étudiés s'inscrivent dans la longue durée, car derrière chaque catégorie se déploient des significations, des méthodes de travail, des appréhensions du monde, etc. potentiellement multiples. Dans le cas de corpus de données plus contemporains, dont l'usage est répandu en géographie, il est parfois moins facile d'identifier ce besoin, les objets pouvant (faussement) paraître plus évidents, car contemporains des chercheurs qui les manipulent. Nous avons par exemple pu nous en apercevoir pleinement lors du post-doctorat effectué par Gravier au sein du LabEx DynamiTe et de l'UMR Géographie-cités, portant en particulier sur l'analyse longitudinale de services publics à partir de la base permanente des équipements (BPE) de l'INSEE (Gravier, 2020). La BPE est produite chaque année depuis 2007. Toutefois, l'INSEE est le compilateur de différentes bases de données et non l'unique producteur. Dès lors, les catégories changent selon les années, du fait de la réorganisation interne de certains services ou de la fusion de plusieurs services

publics entre eux. Il est ainsi nécessaire de se demander systématiquement s'il existe une pérennité des objets étudiés dans le temps et, si non, s'il est possible de regrouper des catégories afin de construire un nouvel objet cohérent pour que l'on puisse étudier sa trajectoire. Pour cela, il faut étudier attentivement l'ensemble des métadonnées élaborées par l'INSEE pour chaque base annuelle, afin de reconstituer le fil des évolutions des catégories pour chaque service étudié, comme l'illustre le cas de Pôle emploi du tableau 2. Il faut également discuter les équivalences d'objets construits dans les BPE. À titre d'exemple, l'ANPE et Pôle emploi, tous deux identifiés en A103 dans les bases 2008 à 2011, sont considérés comme identiques tandis que Pôle emploi résulte de la fusion de l'ANPE et de l'Assédic – ce dernier n'étant pas comptabilisé dans les BPE 2008 et 2009.

Tableau 2. Reconstitution de l'évolution des catégories relatives à Pôle emploi dans les BPE entre 2008 et 2018

2008-2009	2010-2011	2012	2013	2014-2018
Agence Nationale Pour l'Emploi A103	Pôle Emploi A103	Agence de proximité Pôle Emploi A110	Agence de proximité A110	Réseau de proximité Pôle Emploi A122
		Relais Pôle Emploi A111	Relais A111	
		Permanence Pôle Emploi A112	Permanence A112	
		Agence de services spécialisés A113	Agence spécialisée A115	Réseau Spécialisé A115
		Agence thématique A114		
			Relais et maison de service public A116	Réseau Partenarial Pôle Emploi A123
			Point d'Information et de Médiation Multiservices A117	
			Espace Public Numérique et autres A118	

C'est justement un des intérêts de l'utilisation des ontologies que de devoir forcer l'explicitation de la construction des données et, inversement, aider à la critique – productive – de cette construction. Cela est d'autant plus essentiel que les données sont de plus en plus facilement accessibles et nombreuses. D'ailleurs, l'ouverture des données de la recherche qui encourage à la réutilisation des corpus dans des modèles numériques par des personnes hors du champ disciplinaire « producteur » de la donnée ne peut que bénéficier de cette documentation. Ainsi, lorsqu'elles existent, les ontologies fluidifient la réutilisation des données.

C'est aujourd'hui ce que démontre l'ontologie mise en place par les archéologues et relative aux catégories d'analyse de l'espace intra-urbain sur le temps long et aux OH de la base ToToPI. Le modèle de données qui a été construit à partir du cas de Tours est aujourd'hui relativement partagé par ceux qui travaillent sur l'espace intra-urbain aux périodes anciennes – sur la ville de Metz (Gama, *et al.*, 2006 ; Desjardin *et al.*, 2011) ou l'oppidum de Bibracte (Meunier, 2012) par

exemple. Au sein du Laboratoire archéologie et territoire, le modèle a même été enrichi ponctuellement d'un nouveau niveau de catégories fonctionnelles pour aborder des objets archéologiques encore plus fins, à l'échelle des éléments constituant une parcelle ou un bâtiment (cours, logis, jardin, etc.) à partir des cas de Tours et de Vendôme (Lefebvre, 2008 ; Simon, 2015), et d'un niveau plus générique qualifiant leurs portées, comme nous l'avons vu plus tôt (§2.1.2). Dans le cadre de la thèse de Nahassia sur la ville de Tours (2019), le modèle était ainsi suffisamment explicite pour que les données puissent être réutilisables dans un travail de géographie. Bien entendu, pour que cette réutilisation des données puisse être pleinement fructueuse, il a fallu du temps pour être en mesure de s'approprier pleinement la conception archéologique des objets, ainsi que des discussions avec les créateurs et spécialistes de l'espace urbain tourangeau. Cependant cette recherche n'aurait pas pu avoir lieu sans tout le travail de formalisation préalablement effectué par ces derniers. En outre, cette expérience témoigne aussi du fait que l'ontologie permet à un corpus d'être plus pérenne dans le temps. En effet, plus d'une génération s'est écoulée entre le moment du début de la construction de ToToPI et celui de la réutilisation de cette base par Nahassia.

4.2. L'ouverture à la discussion interdisciplinaire fructueuse

L'existence de données décrites dans des ontologies axiomatisées avec des vocabulaires contrôlés est censée favoriser leur réutilisation par les machines, et le croisement d'informations entre disciplines. Notre propos va ici au-delà du discours général sur la « FAIRisation » des données et l'usage des ontologies pour favoriser l'interopérabilité et l'exploitation par des machines. Bien sûr, concernant l'interopérabilité des ontologies, les efforts de la communauté mondiale pour verser des ontologies qui se relient entre elles sont bien réels, conséquents et visibles dans le *Linked Open Data*, espace de référencement et de nommage partagé des ontologies et des données associées sérialisées sous la forme de triplets RDF accessibles sur le Web 3.0, le Web sémantique. Ces efforts visent à favoriser la réutilisation des données. Cependant, nous insistons sur le fait qu'ontologie et Web sémantique ne riment pas forcément avec interopérabilité. D'une part, toutes les ontologies ne sont pas forcément elles-mêmes FAIR, *i.e.* réutilisables en dehors du contexte de leur production (Poveda-Villalón *et al.*, 2020), parce que ces modèles ne sont pas toujours bien méta-décrits eux-mêmes. D'autre part, comme nous l'avons montré dans les paragraphes précédents, les ontologies sont des projections des conceptualisations mentales propres à une époque, une discipline, un type de question de recherche ou d'application. Donc pour réutiliser des données produites suivant deux ontologies touchant au même domaine, comme ce peut être le cas avec des données d'observation en environnement mobilisant Observation & Measurement (Cox, 2016) ou OBOE (Madin *et al.*, 2007) par exemple, il est nécessaire de procéder à des alignements (à une réconciliation) des concepts ; à ce sujet les recherches en informatique ont développé des mécanismes d'alignement automatique (Vargas-Vera et Nagy, 2013), mais sont-ils toujours opérationnels et

mobilisables par des utilisateurs non informaticiens ? Par ailleurs, les outils d'ingénierie des connaissances et les modèles informatiques actuels d'ontologie sont peut-être encore limités pour rendre compte de toute la nuance des connaissances – c'est un problème d'expressivité de ces langages (Claramunt, 2020).

Toutefois, l'existence d'ontologies documentées est le fruit d'un travail de formalisation disciplinaire qui a une valeur très forte et qui, intrinsèquement, contribue également à l'interdisciplinarité. D'abord parce que de nouveaux traitements statistiques, tels que ceux présentés en section 2 sont rendus possibles, car les catégories a-historiques facilitent les comparaisons dans la très longue durée. Mais aussi parce que ces analyses peuvent être interprétées au regard de la réalité, nécessairement plus complexe, des usages de l'espace urbain parce que l'ontologie sous-jacente à ce modèle de données a été et est largement discutée, explicitée et formalisée par les chercheurs qui l'ont composée et qui l'utilisent.

En effet, en structurant et expliquant les choix et les interprétations qui guident la production de ses corpus, le spécialiste d'une discipline de sciences humaines et sociales se voit obligé d'entrer dans la communication avec d'autres spécialistes, de son domaine ou d'autres domaines. L'explication donnée doit en effet être comprise par les interlocuteurs, sinon elle n'a pas de valeur. Le producteur comme le consommateur de l'ontologie doivent donc s'accorder sur le sens des mots utilisés, et comprendre la pensée de l'autre. Sans ce travail de formalisation ontologique, il subsiste souvent des implicites et des non-dits autour d'un corpus. S'il n'est pas toujours facile ni agréable de justifier chacun de ces choix, cette démarche ouvre la porte à la mise en débat de ces derniers, et donc à la comparaison des différentes approches et opinions des chercheurs. L'ontologie facilite donc l'aller-retour entre les différentes positions épistémologiques qui conditionnent les analyses et leurs résultats, un cheminement qui nourrit le questionnement scientifique, et dont le premier bénéficiaire est au moins l'auteur de l'ontologie, suivi de l'ensemble de la communauté scientifique, quelle que soit sa discipline.

Bibliographie

- Beretta F. (2017). L'interopérabilité des données historiques et la question du modèle : l'ontologie du projet SyMoGIH. *Enjeux numériques pour les médiations scientifiques et culturelles du passé*, B. Juanals et J.-L. Minel (dir.), Presses universitaires de Paris Nanterre, cf. HalSHS.
- Berthaut H. (1898). *La Carte de France, 1750-1898. Étude historique*. Imprimerie du Service géographique de l'Armée, Paris, tome 1.
- Biddle M., Hudson D. M., Heighway C. M. (1973). *The Future of London's Past: A Survey of the Archaeological Implications of Planning and Development in the Nation's Capital*, vol. 4, Worcester, Rescue.
- Bizière J-M. et Vayssière P. (1995). *Histoire et historiens : Antiquité, Moyen Âge, France moderne et contemporaine*, Paris, Hachette.

- Bouyssou L. (1987). Le buron, *Volcan cantalien*, Chamina, Clermont, 2e édition.
- Cénat J.-P. (2008). Stratégie, logistique et propagande : l'usage des cartes militaires par Chamlay. *Le monde des cartes*, n° 195, p. 27-37.
- Chareille P., Rodier X. et Zadora-Rio E. (2004). Analyse des transformations du maillage paroissial et communal en Touraine à l'aide d'un SIG. *Histoire et mesure*, vol. 3/4(2004), p. 314-344.
- Claramunt C. (2020). Ontologies for geospatial information: Progress and challenges ahead. *Journal of Spatial Information Science*, 20, p. 35-41.
- Collectif (2004). *Informatisation des Documents d'Évaluation du Patrimoine Archéologique des Villes de France: bilan d'étape 2002-2003*, Tours, ministère de la Culture et de la Communication, La Simarre.
- Costes B., Grosso E., Plumejeaud C. (2014). Du passé au présent. Intégration de données topographiques extraites des cartes de Cassini. *Revue Internationale de Géomatique*, Lavoisier, vol. 24, n° 2, p. 211-230, doi : 10.3166/ri.24.211-230.
- Costes B. (2016). *Vers la construction d'un référentiel géographique ancien : un modèle de graphe agrégé pour intégrer, qualifier et analyser des réseaux géohistoriques*. Université Paris-Est
- Cox S. J. D. (2016). Ontology for observations and sampling features, with alignments to existing models. *Semantic Web*, vol. 8, n° 3, p. 453-470.
- Dauphiné A. (1984). Espace terrestre et espace géographique. *Concepts de la géographie humaine*, S. Bailly (dir.), Paris, Masson, p. 33-43.
- Dedieu J.-P., Marzagalli S., Pourchasse P., Scheltjens W. (2012). Navigocorpus at work. A brief overview of the potentialities of a database. *International Journal of Maritime History*, vol. XXIV, n° 1, p. 331-359.
- Desjardin E., Runz C. De, Pargny D. (2011). Conception et analyses d'informations archéologiques dans un SIG. Retour des projets SIGRem et ArchéoChamp. *SAGEO'11. Rencontres internationales Géomatique et Territoire*, Avignon.
- Galinié H. (1979). Droit de cité pour l'archéologie urbaine. *Nouvelles de l'archéologie*, dossier « L'archéologie urbaine », p. 6-10.
- Galinié H. et Randoin B. (1979). *Les Archives du sol à Tours : survie et avenir de l'archéologie de la ville*, Tours, Laboratoire d'Archéologie Urbaine.
- Galinié H. et Rodier X. (2004). La formalisation des données de la topographie historique. coll. « Informatisation des documents d'évaluation du patrimoine archéologique des villes de France : Bilan d'étape 2002-2003 », p. 21-34, Tours, ministère de la Culture et de la Communication, La Simarre.
- Galinié H., Rodier X., Saligny L. (2004). Entités fonctionnelles, entités spatiales et dynamique urbaine dans la longue durée. *Histoire & mesure*, vol. XIX, n° 3, p. 223-242. Retrieved from <http://histoiremesure.revues.org/761>.
- Gama F., Bourada L., Lansival R., Pernot P. (2006). PCR « Étude documentaire sur le POTentiel ARchéologique Urbain de la ville de Metz », Metz - Moselle, Metz.

- Gravier J. (2018). *Deux mille ans d'une ville en système. Proposition d'une démarche appliquée au cas de Noyon*. Université Paris 1 Panthéon-Sorbonne. Retrieved from <http://tel.archives-ouvertes.fr/tel-02367262>.
- Gravier J. (2020). Reconfigurations territoriales des services publics dans les villes françaises (2009-2018). *Proceedings of CIST'20 - Population, temps, territoires*, CNRS, Ined, Université Paris 1, Nov., Paris-Aubervilliers, France, p. 263-268. Retrieved from <http://hal.archives-ouvertes.fr/hal-03114137>.
- ISO/TC 46 (2011). ISO 25964, Information et documentation - Thésaurus et interopérabilité avec d'autres vocabulaires - Partie I: thésaurus pour la recherche documentaire. <https://www.iso.org/fr/standard/53657.html>
- Lefebvre B. (2008). *La formation d'un tissu urbain dans la Cité de Tours : du site de l'amphithéâtre antique au quartier canonial*. Université François - Rabelais de Tours.
- Lennon M., Peirce D. S., Tarry B. D., Willett P. (1981). An evaluation of some conflation algorithms for information retrieval. *Journal of Information Science*, vol. 3, n° 4, p. 177-183, doi: 10.1177/016555158100300403.
- Lorans E., Rodier X., Jouquand A-M. (2018). Tours: Origins of urban archaeology, new approaches and new questions. *European Journal of Post-Classical Archaeologies*, 8, p. 19-42.
- Madin J., Bowers S., Schildhauer M., Krivov S., Pennington D. et Villa F. (2007). An ontology for describing and synthesizing ecological observation data. *Ecological Informatics*, vol. 2, n° 3, p. 279-296.
- Mathian H. et Sanders L. (2014). *Objets géographiques et processus de changement : approches spatiotemporelles*, Londres, ISTE Éditions.
- Meunier A. (2012). Gestion et exploitation des données spatiales à l'aide d'un système d'information archéologique. Le SIA UrBiS : Urbanisme de Bibracte et Spatialité, *Bibracte, programme de recherche sur le Mont Beuvray*. Rapport annuel 2012, Glux-en-Glenne, Bibracte, Établissement Public de Coopération Culturelle, p. 437-449.
- Motte C., Vouloir M-C. (2008). Frontières administratives et identités communales. Le cas de la France, XVIIIe-XXe siècles. *The Historical Review*, Athènes, vol. V.
- Nahassia L. (2019). *Formes spatiales et temporelles du changement urbain. Analyser la localisation des activités à Tours sur 2000 ans*. Université Paris 1 Panthéon-Sorbonne.
- Nahassia L. et Plumejeaud C. (2012). *Analyse des évolutions des données topographiques et démographiques de la carte de Cassini à nos jours* (Rapport No. L4.1). COGIT/IGN.
- Pégorier A. (2006). Les noms de lieux en France, dictionnaire de termes dialectaux, <http://www.ign.fr/adminV3/display/000/526/725/5267257.pdf>
- Peiffer C. (1877). *Légende territoriale de la France pour servir à la lecture des cartes topographiques*, Paris, Ed. C. Delagrave, éditeur de la Société de Géographie de Paris.
- Pelletier M. (1990). *La carte de Cassini. L'extraordinaire aventure de la carte de France*, Paris, Presses de l'École nationale des Ponts-et-Chaussées.
- Peuquet D. (2002). *Representations of time and space*, New York, NY, USA, Guildford Press.

- Plumejeaud C. (2012a). *L2.0-2 Présentation des données*. Rapport du projet GéoPeuple.
- Plumejeaud C. et Vouloir M-C. (2012b). *L2.1-2 Contenu de la carte de Cassini et sa modélisation en vue de sa vectorisation*. Rapport du projet GéoPeuple.
- Plumejeaud C. (2012c). *L2.1-3 Contenu de la carte d'état-major et sa modélisation en vue de sa vectorisation*. Rapport du projet GéoPeuple.
- Plumejeaud C. (2012d). *L2-3.4 Intégration des données démographiques*. Rapport du projet GéoPeuple.
- Plumejeaud-Perreau C., Cristofoli P., Motte C. (2015). De l'étude des nomenclatures territoriales à la modélisation des dynamiques des territoires administratifs en France. *Revue internationale de géomatique*, Lavoisier, Modéliser les dynamiques spatiales, vol. 25, n° 3, p.355-392
- Poveda-Villalón M., Espinoza-Arias P., Garijo, D. et Corcho O. (2020). Coming to Terms with FAIR Ontologies. *Knowledge Engineering and Knowledge Management*, Cham, p. 255-270.
- Recchia G. et Louwerse M. (2013). A Comparison of String Similarity Measures for Toponym Matching. *Proceedings of The First ACM SIGSPATIAL International Workshop on Computational Models of Place – COMP'13*, Orlando FL, USA, p. 54-61, doi: 10.1145/2534848.2534850.
- Rodier X. (2000). Le système d'information géographique TOTOPI : Topographie de TOurs Pré-Industriel. *Les petits cahiers d'Anatole*, n° 4.
- Rodier X. (2016). *Espace-temps des sociétés du passé*. Université de Franche-Comté.
- Rodier X., et Saligny L. (2010). Modélisation des objets historiques selon la fonction, l'espace et le temps pour l'étude des dynamiques urbaines dans la longue durée. *Cybergeog : European Journal of Geography*. <https://doi.org/10.4000/cybergeog.23175>
- Roussel D., Defente D., De Mecquenem C., Gueugnon Y., Hacquet J., Jagielski K. (coll.) (2002). *Soissons*. Document d'Évaluation du Patrimoine Archéologique des Villes de France, Paris, Ed. du patrimoine.
- Ruas A., Plumejeaud C., Nahassia L., Grosso E., Olteanu A-M. *et al.* (2013). Géopeople : Création et analyse de données topographiques sur 200 ans. *Cartes & géomatique*, Comité français de cartographie.
- Ruiz J.J, Ariza F. J., Ureña M. A., Blázquez E. B. (2011). Digital map conflation: A review of the process and a proposal for classification. *International Journal of Geographical Information Science*, vol. 25, n° 9, p. 1439-1466, doi: 10.1080/13658816.2010.519707.
- Simon G. (2015). *Espace et société à Vendôme du 11e au début du 19e s. : fonctionnement et fabrique d'une ville intermédiaire sur le temps long*, Université François - Rabelais.
- Vargas-Vera M. et Nagy M. (2013). Challenges in ontology alignment and solution to the contradictory evidence problem. *Advances in Artificial Intelligence and Its Applications*, F. Castro, A. Gelbukh et M. González (Ed.), vol. 8265, Springer Berlin Heidelberg, Berlin, Heidelberg, p. 199-216.

Wang H., Tang X., Shi S. (2007). A solution for change detection in spatio-temporal database. *Geoinformatics 2007*, Nanjing, China, juin, p. 675303, doi: 10.1117/12.760449.

Weber M. (2014 [1921]). *La ville*, Paris, La Découverte, coll. « Politique et sociétés ».

Explorer les processus de mobilité passée

Raisonnement ontologique fondé sur la connaissance des pratiques socioculturelles et des vestiges archéologiques

**Laure Nuninger¹, Thérèse Libourel², Xavier Rodier³,
Rachel Opitz⁴, Philip Verhagen⁵, Catherine Fruchart¹,
Clément Laplaige³, Samuel Leturcq³, Nathanael Levoguer³**

1. CNRS/University of Bourgogne Franche-Comté – Chrono-Environnement /
MSHE C.N. Ledoux, 25030 Besançon, France
{laure.nuninger ; catherine.fruchart} @univ-fcomte.fr

2. Université de Montpellier - Espace-Developpement, F-34093 Montpellier
therese.libourel@umontpellier.fr

3. CNRS/Université de Tours - CITERES/MSH Val de Loire, F-37020 Tours
{xavier.rodier ; clement.laplaige ; samuel.leturcq ; nathanael.levoguer}@
univ-tours.fr

4. Départ. d'archéologie, University of Glasgow, Glasgow G12 8QQ, Royaume-Uni
rachel.Opitz@glasgow.ac.uk

5. Faculté des Sciences Humaines/CLUE+, Vrije Universiteit Amsterdam,
1081 HV Amsterdam, Pays-Bas
j.w.h.p.verhagen@vu.nl

RÉSUMÉ. Dans cet article, nous proposons une nouvelle piste de raisonnement élaborée en interdisciplinarité avec des archéologues, conçue pour des études liées à la notion de mobilité. À l'aide d'études de cas, nous analysons comment les chercheurs ont identifié, conceptualisé et relié les traces matérielles décrivant divers processus de mobilité dans une région donnée. Nous élaborons ensuite la construction d'ontologies qui nous permettent de relier explicitement les éléments matériels, identifiés dans le paysage observé, aux connaissances ou à la théorie qui explique leur rôle et leurs relations au sein du processus de mobilité. En combinant les systèmes de parcours formels, dits « Path framework systems » et les systèmes de mouvement informels, dits « Pathway systems », nous montrons, à travers ces études de cas, que les deux types de systèmes ne sont pas intégrés hiérarchiquement, mais plutôt entrelacés. Nous introduisons un nouvel outil heuristique, que nous avons nommé le « track graph », dont l'intérêt est d'enregistrer les éléments matériels observés sous une forme neutre d'un point de vue thématique afin de pouvoir les utiliser pour reconstruire des modèles des trajectoires qui suivent différentes logiques de mobilité.

ABSTRACT. In this article we propose an original approach designed by archaeologists and a computer scientist for knowledge creation in studies of movement. Using case studies, we analyze how researchers have identified, conceptualized, and linked the material traces describing various movement processes in a given region. Then, we explain how we construct ontologies that enable us to explicitly relate material elements, identified in the observed landscape, to the knowledge or theory that explains their role and relationships within the movement process. Combining formal pathway systems, so-called « Path framework systems » and informal movement systems, so-called « Pathway systems », through these case studies, we argue that these systems are not hierarchically integrated, but rather intertwined. We introduce a new heuristic tool, named the “track graph”, that records observed material features in a thematically neutral form. The « track graph » can then be employed to reconstruct the trajectories of journeys which follow different movement logics.

MOTS-CLÉS : archéologie, processus de mobilité, réseau, maillage, ontologie, modélisation, pathway system, path framework system, track graph.

KEYWORDS: Archaeology, Movement process, Network, Meshwork, Ontology, Modeling, Pathway system, Path framework system, Track graph.

DOI:10.3166/RIG.31.81-110 © 2022 Lavoisier

1. Introduction

Dans cet article, nous proposons d’aborder la création de connaissances à partir de multiples dimensions en s’attachant au problème du processus de mobilité passée et à sa matérialisation. La mobilité discutée ici correspond au mouvement déployé par des actants¹ (agents actifs) pour se déplacer, influencer un déplacement ou réaliser une activité. La question posée est : comment, à partir des vestiges matériels observés dans les paysages actuels, pouvons-nous tenter de reconstruire les logiques de mobilité des actants humains et non humains qui ont contribué à modeler l’espace dans le temps long ?

Comme le souligne Hurard *et al.* (2014) « Le champ d’expertise de l’archéologie, la mise en discours de ses données ne se limitent pas aux dimensions techniques et technologiques. Elle est une science de l’analyse spatiale et sociale où production, exploitation, distribution et consommation des ressources et des biens forment une seule chaîne opératoire. La notion de culture matérielle, partagée avec les anthropologues et les historiens, n’enferme pas les archéologues dans la matérialité et la dimension technique de l’objet produit. Elle leur offre au contraire les moyens de penser l’objet, l’espace et le milieu comme des productions sociales aux significations plus ou moins manifestes, des marqueurs privilégiés de l’adaptation de l’homme à un milieu donné. ». Par ailleurs, la modélisation

1. Le terme « actant » fait référence aux travaux de la sociologie de la traduction, dite aussi théorie de l’acteur réseau. Cette notion permet de décrire des interactions au sein de réseaux hétérogènes qui associent des objets ou techniques et des humains, chacun ayant un rôle (Akrich, 2013). Nous l’utilisons pour éviter le déterminisme humain et intégrer l’influence des animaux et des objets dans le processus de mobilité.

systemique et la simulation numérique permet d'envisager un ensemble d'interactions impliquant des dimensions matérielles, cognitives, sociales, spatiales et temporelles. L'intérêt de ce type d'approche est qu'elle nous force à « identifier et caractériser les processus qui sous-tendent le fonctionnement systémique de nos sociétés, actuelles et passées » (Djindjian, 2010).

Les vestiges révélant par définition une vision partielle et biaisée de la réalité passée, notre proposition articule un ensemble de données archéologiques et numériques², et des connaissances théoriques et empiriques pour concevoir les scénarios de reconstructions possibles de ces logiques de mobilités passées, qui contribuent à une meilleure analyse et interprétation de la dynamique des paysages et des systèmes de peuplement.

À l'aide d'études de cas, qui nous servent d'exemples, nous analysons comment les chercheurs archéologues ont identifié, conceptualisé et relié les traces matérielles décrivant possiblement divers processus de mobilité dans une région donnée. La démarche consiste à élaborer des conceptualisations prémices d'ontologies qui nous permettent de relier explicitement les éléments matériels, identifiés dans le monde observé, à la connaissance ou à la théorie qui explique leur relation dans le processus de mobilité.

Bien que les études de cas soient différentes, leur analyse montre qu'elles combinent toutes deux systèmes de mobilité que l'on distingue ici. La mobilité que l'on dit formelle est caractérisée par une infrastructure consacrée à permettre le déplacement et il y a une intentionnalité de déplacement d'un lieu vers un autre. Dans l'autre mobilité, qualifiée ici d'informelle, il n'y a pas d'intentionnalité de déplacement d'un point à un autre, ce dernier est lié au mouvement réalisé par des pratiques socio-économiques variées (cueillettes, travail de la terre par exemple). Ce type de mobilité implique la cooptation de divers éléments et/ou structures³ pour permettre le mouvement et les déplacements au sein d'un espace.

Comme ces deux systèmes ne sont pas intégrés hiérarchiquement mais fonctionnent plutôt de manière entrelacée, nous proposons d'introduire un nouveau

2. On entend par données archéologiques : toutes les données de recherche produites dans le cadre d'une activité de terrain (fouilles, sondages, prospections). Les données numériques correspondant ici à toutes les données produites à partir de sources numériques (imagerie, modèles numériques de terrains, etc.) soit par télédétection, soit par modélisation (par exemple des chemins de moindre coût).

3. La notion d'éléments ou de structures fait ici référence à des objets identifiables dans le paysage par l'archéologue, en altimétrie et/ou dans une image aérienne ou satellite. Par exemple, on considérera comme éléments ou structures : une série de marches taillées dans la roche, l'empilement de pierres qui forme un cairn ou encore une zone avec un état de surface particulier par rapport au voisinage. Une structure étant *a priori* composée de plusieurs éléments fonctionnant ensemble, l'ensemble étant identifié comme un même objet.

concept opérationnel, le « track graph »⁴, qui vise à enregistrer les éléments matériels observés sous la forme la plus neutre possible d'un point de vue thématique. De nature purement analytique, ce concept abstrait de « track graph » ambitionne à terme de devenir un outil de simulation utile à la reconstruction de trajectoires spatiales qui intègrent différents types de logique de mobilité.

Dans cet article, nous présentons la démarche qui nous a permis de construire le concept abstrait de « track graph » en partant d'une réalité concrète. Nous avons développé une approche ontologique que nous précisons en section 3 et que nous illustrons avec deux études de cas détaillées en section 4. La section 5 présente la phase d'abstraction et le concept de « track graph » qui sera la base du raisonnement permettant de reconstruire les schémas de mobilité à partir des diverses données invoquées. Enfin, dans la section 6 nous synthétisons notre démarche et précisons les perspectives envisagées. Mais il convient dans un premier temps (section 2) de présenter les principaux cadres théoriques intégrant la notion de mobilité et de mouvement pour poser la problématique spécifique de la prise en compte des traces matérielles produites par le mouvement.

2. De la mobilité en général et des traces en particulier

2.1. De quelle mobilité parle-t-on ?

Occupons-nous l'espace ? Habitons-nous l'espace ? Cette double question est vraiment significative lorsque nous étudions les systèmes de peuplement et comment le mouvement des actants humains et non humains contribue à sa structuration. Comme précédemment énoncé, à partir de vestiges matériels observés dans les paysages d'aujourd'hui, l'archéologie tente de reconstruire la logique du mouvement des individus et des choses qui ont contribué à façonner l'espace.

Si nous nous référons à une définition générale (issue du CNTRL⁵) le mouvement est la faculté de se mouvoir, de se déplacer, et l'usage que l'on fait de cette faculté permet de le caractériser. Le mouvement peut considérer des individus ou des collectifs et concerne des déplacements effectués sous l'effet de causes diverses. Déplions rapidement les enjeux d'une telle formule. Lorsqu'ils se déplacent, les individus, humains ou non humains, dessinent leur propre itinéraire soit en le définissant (trajet, navigation), soit en suivant un itinéraire existant bien reconnu (transport d'un point de départ à un point d'arrivée). Dans les deux cas, des

4. « Track graph », « Path framework system » et « Pathway system » sont des néologismes que nous avons volontairement conservés sous leur vocable anglophone car ils ont été définis par un groupe de recherche international dans le cadre du projet Movescape (IEA CNRS, resp. L. Nuninger et Ph. Verhagen, 2018-2020). Il n'y a donc pas de traduction *ad hoc* pour ces termes actuellement et comme notre recherche en est encore à ses débuts nous préférons limiter les ambiguïtés.

5. Centre national de ressources textuelles et lexicales : <https://www.cnrtl.fr/>

« lignes » sont tracées dans le paysage, même si leur empreinte physique peut être plus ou moins évidente⁶.

Selon Tim Ingold (Ingold, 2011), la mobilité repose sur un principe simple :

« Le cheminement itinérant (*wayfaring*) est le mode fondamental par lequel les êtres vivants habitent la Terre. Chacun de ces êtres doit ainsi être imaginé comme la ligne de son mouvement ou, plus réalistement, comme un ensemble de lignes. »

« Le monde habité est un maillage réticulaire de ces pistes qui, tant que la vie suit son cours, continuent à se tisser. Le transport, au contraire, est relié à des lignes spécifiques. Chaque déplacement, orienté vers une destination précise, a pour fonction de relocaliser des personnes et leurs effets. Le passager qui part d'un endroit pour arriver à un autre endroit n'est nulle part entre les deux. En les réunissant, les lignes de transport forment un réseau de connexions point par point. »

Ainsi, Ingold se référant à Lefebvre (2000) adopte le terme de « maillage » (*meshwork*) pour distinguer le dessin réticulaire des pistes tracées par le cheminement des actants humain ou non humains guidés par leurs activités, de ce que l'on appelle le « réseau » qui conceptuellement implique l'idée d'un ensemble de lieux interconnectés.

2.2. Comment analyser la mobilité à partir de vestiges inertes ?

Comment ces deux conceptions de la mobilité et du modèle spatial (réseau vs. maillage) qu'elles impliquent est-il perçu par l'archéologue ? Pour tenter de répondre à cette question, il faut d'abord préciser que les vestiges révèlent par définition une vision partielle et biaisée de la réalité passée. Les traces recherchées dans le monde tangible réel sont celles de l'empreinte que les processus divers des déplacements effectués dans le temps ont laissée (traces matérielles issues d'observations diverses, traces « culturelles » : manuscrits, cartes, etc.).

Face à la myriade de possibilités qui s'offrent à eux pour identifier et interpréter ces vestiges en termes de système de mobilité, les archéologues doivent construire des modèles logiques et intelligibles articulant les données matérielles observées sur le terrain ou dans des jeux de données numériques (par exemple un modèle numérique de terrain) avec des connaissances empiriques et théoriques afin de proposer des hypothèses plausibles de mobilité passée et d'interpréter les paysages analysés.

Cette double question initiale focalise notre attention sur les modalités d'articulation entre données et connaissances en ce sens que, suivant Ingold (2011), dans le verbe « occuper » l'acteur est inactif dans un espace donné, même s'il se

6. L'empreinte physique, par impression dans le modelé du terrain, peut-être plus ou moins marqué, d'un simple tassement du sol continu au creusement d'ornières par exemple. L'empreinte peut également être laissée en surimpression par l'accumulation de terre, ou l'installation d'objets (bois, tas de pierres, etc.) voire une construction spécifique (chaussée).

déplace, il suit « passivement » un tracé d'un point défini à un autre, tandis que dans le verbe « habiter » il est actif dans l'espace car il contribue à le définir par la mobilité liée à son activité. Ce dernier point de vue est radicalement différent et il modifie la façon dont les entités (données, connaissances) peuvent être mobilisées et décrites en fonction de propriétés nouvelles et surtout de relations qui peuvent nous amener à repenser ou relire nos interprétations archéologiques des paysages (Aldred, 2016, 2020).

L'interprétation des traces laissées par les activités associées au mouvement dans le paysage actuel est, de fait, soumise à un certain nombre de biais : méthodologiques, observationnels, contextuels et sémantiques.

Biais méthodologiques : dans le cas d'observations, la méthode d'observation choisie détermine le protocole et ce qui peut être détecté. Cela signifie que le filtre méthodologique appliqué et la posture adoptée dans le processus d'investigation statique ou mobile (Aldred, 2016) détermineront la probabilité de détection de certains éléments, selon des critères d'observation du matériel spécifiques au protocole suivi. Le contexte de réalisation et la subjectivité de leur auteur vont également entrer en ligne de compte, les deux influençant la perception du terrain, le contenu des écrits ou des supports.

Biais d'observation : les observateurs ont tendance à observer ce qu'ils ont déjà observé auparavant (Palmer, 2013). La connaissance des éléments et/ou structures détermine dans une large mesure la facilité avec laquelle les observateurs les identifieront (voir, par exemple, Banaszek *et al.*, 2018). Les biais d'observation peuvent donc conduire à des « faux positifs » ainsi qu'à des « faux négatifs ». Si nous ne voyons que ce que nous savons, nous ne découvrirons jamais quelque chose de nouveau.

Biais contextuels : dans la plupart des cas, il est nécessaire de comprendre le contexte des éléments afin de les identifier correctement (Watson *et al.*, 1971) en tenant compte à la fois du contexte environnemental, historique et culturel (Schiffer, 1976). Certaines conditions environnementales vont renforcer ou supprimer l'expression matérielle des éléments et/ou structures, par exemple parce qu'elles sont facilement érodées. D'un point de vue historico-culturel, il est nécessaire de savoir si les éléments peuvent être datés d'une certaine période, et à quoi ils auraient pu servir. Si ces informations contextuelles sont manquantes ou mal interprétées, l'identification des éléments deviendra impossible ou sera erronée (Hodder et Hutson, 2003). En outre, pour un même élément et/ou structure, sa signification peut changer au fil du temps (Case, 1973).

Biais sémantiques : les identifications d'éléments et/ou de structures qui peuvent sembler très évidentes à un observateur, peuvent ne pas avoir la même signification pour un autre observateur, puisque même au sein d'un même référentiel de classement les terminologies appliquées demeurent souvent implicites dans leur signification (Childe, 2015). Ce problème ne fait que s'accroître dès lors qu'il y a plusieurs langages de description, et est à l'origine de nombreux problèmes

liés au croisement de plusieurs jeux de données. Enfin, la signification de chaque terme peut aussi changer avec le temps, car elle est intimement liée au contexte (Hodder et Hutson, 2003).

Pour résumer, lorsque nous recherchons des traces matérielles liées à la mobilité, que ce soit sur le terrain réel ou sur un « terrain numérique », tous ces biais vont jouer et influencer notre capacité à détecter et identifier un élément et/ou une structure.

3. De l'usage d'une démarche ontologique

3.1. *Choses, mots et concepts*

Comment qualifier les traces matérielles observées et identifiées comme potentiellement utilisées pour le déplacement, sans se perdre dans une typologie descriptive qui cloisonne le raisonnement en ensembles morphologiquement homogènes⁷ ? Comment traiter la variation des itinéraires fondée sur un ensemble de traces matérielles morphologiquement hétérogènes, tout en leur donnant un sens ? Comment intégrer cette variation dans notre approche d'identification et d'interprétation des données observées afin de ne pas réduire notre champ de connaissance à ce que nous savons déjà reconnaître ? Enfin, comment pouvons-nous démontrer les relations que nous identifions entre l'observation de traces matérielles hétérogènes et l'ensemble de nos connaissances archéologiques, historiques, anthropologiques et théoriques.

Traditionnellement, la diversité des éléments observés associés au mouvement conduit soit à une multiplication des termes descriptifs, soit à une réduction de la terminologie acceptée, *via* une approche classificatoire ou typologique, telle que celle proposée par Timothy Earle (Earle, 2009). Dans le premier cas, l'utilisation de termes trop spécifiques rend difficile l'identification de connexions entre des éléments qui jouent effectivement le même rôle dans des systèmes de mobilité différents, tandis que dans le second cas, des éléments ayant des morphologies similaires peuvent être regroupés, créant ainsi de fausses connexions entre des entités qui, en réalité, sont associées à différents types de mouvements. Dans chacune des approches, il est très difficile de comparer différentes études de cas ou même différents systèmes de mobilité au sein même d'une étude de cas.

Afin d'éviter ces impasses, nous avons privilégié une approche qui permet de dissocier l'objet de sa terminologie. Pour ce faire, nous nous sommes appuyés sur le triangle sémiotique conçu par Ogden et Richards (Ogden et Richards, 1923). Pour

7. On parle ici de la morphologie des éléments archéologiques. Ainsi, par exemple, les éléments présentant une chaussée bordée par des murets ou des fossés se retrouveront dans un type, tandis que les éléments présentant un simple tassement du sol se retrouveront dans un autre type. Lequel doit-on qualifier de route par exemple ? Il est possible que ce l'on doit qualifier de route rassemble des éléments morphologiques des deux types.

ces auteurs, qui ont travaillé sur le langage, le malentendu entre deux personnes qui utilisent le même terme réside dans le fait que l'on confond souvent le nom (*symbole*) d'un objet et cet objet (*réfèrent*) lui-même (la base du triangle de la figure 1).

En partant du principe que le sens d'un mot est déterminé par l'expérience vécue du locuteur, ils rejettent toute idée de sens canonique ou standardisé. C'est pourquoi, afin de clarifier et de rendre plus explicite la relation, généralement implicitement établie entre un mot et un objet, ils définissent ce qu'ils appellent le domaine de référence ou de pensée (au sommet du triangle). Le domaine de référence est ce que nous appelons un concept, c'est-à-dire l'ensemble des connaissances empiriques et théoriques qui nous permettent de concevoir ce qu'est une « chose » – dans notre cas, une trace matérielle liée au mouvement – et de la relier à un mot qui la désigne et la définit comme telle – dans notre cas, la terminologie qui peut être associée aux parcours ou itinéraires en général (chemin, route, etc.). C'est donc la relation avec le concept, ou le domaine de référence, qui donnera un sens explicite à l'identification d'une trace matérielle comme un *chemin*, par exemple.

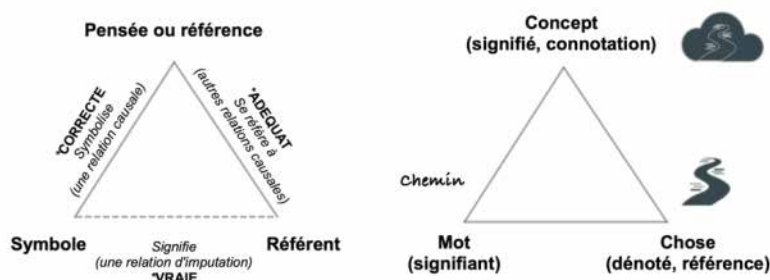


Figure 1. Le triangle sémiotique : le mot (Symbole), la chose (Réfèrent), le concept (Pensée) (d'après Ogden et Richards, 1923)

Fondée sur cette théorie, notre approche consiste donc à analyser la manière dont les éléments matériels identifiés dans le monde observé (terrain, imagerie, MNT LiDAR⁸, etc.) sont mis en relation avec des termes liés au processus de déplacement (sentiers, chemins, routes, pistes, creux, chaussées, canaux, chemins de halage, passages, etc.) dans plusieurs études de cas.

Notre objectif est de développer un cadre conceptuel générique pour décrire et relier des pratiques de mobilité et leur matérialisation (morphologie, nature, etc.) dans des contextes historiques et environnementaux variables, formés de parcours et de trajectoires spatiales enchevêtrés. Ce cadre doit pouvoir intégrer tous les éléments

8. Modèle numérique de terrain (MNT) réalisé à partir d'un relevé altimétrique LiDAR (*Ligh Detection And Ranging* – méthode de télémétrie par onde lumineuse laser) généralement aéroporté.

clés qui permettent d'explicitier les relations entre des éléments hétérogènes afin de reconstruire un ou plusieurs agencements⁹ logiques pouvant constituer un parcours, un sentier, une intersection, voire un réseau ou un maillage.

3.2. Une démarche ontologique pour structurer et formaliser nos descriptions

Afin d'analyser la notion de mobilité dans les études de cas avec une approche interculturelle, nous adoptons une démarche ontologique¹⁰. Cette démarche présente plusieurs avantages : elle fournit une manière structurée et formelle afin de décrire et de mettre en relation les différents termes et bases de connaissances utilisés ; elle permet de clarifier la structure des connaissances disponibles ; et enfin d'envisager les différentes relations entre les termes et concepts utilisés : hiérarchiques, topologiques, temporelles et spatiales.

En pratique, une communauté ou un groupe de recherche décrit la question qu'il a choisie sur la mobilité passée en utilisant des termes et des idées tirés de plusieurs études publiées (figure 2). Les études publiées servent de sources à partir desquelles le groupe de recherche peut constituer un large ensemble de concepts spécifiques à leur contexte¹¹ pour traiter d'une seule idée, plus globale. Ces concepts sont ensuite organisés, dans le but de relier ce qui a été observé, soit des choses ou des noms, avec notre conception des activités qui les produisent, soit des actions ou des verbes dans le texte. Pour créer un modèle opérationnel, nous avons procédé en trois étapes.

Dans un premier temps, le langage utilisé dans les articles pour décrire les éléments du paysage et les actions de mobilité est extrait et analysé. Ensuite, à partir des termes (noms et verbes) et de leurs associations nous décrivons les liens qu'ils

9. Le terme d'agencement est utilisé ici en référence aux travaux de Deleuze et Guattari (1980). « Qu'est-ce qu'un agencement ? C'est une multiplicité qui comporte beaucoup de termes hétérogènes, et qui établit des liaisons, des relations entre eux, à travers des âges, des sexes, des règnes – des natures différentes. Aussi la seule unité de l'agencement est de cofonctionnement » (Deleuze et Parnet, 1996, p. 84). Le concept d'agencement permet de s'extraire de la logique du signifiant de chacun de ces composants, puisque l'agencement forme un tout aux propriétés irréductibles. L'agencement repose sur des relations contingentes, ce n'est donc pas un type fixe mais un espace de possibilités modifiable selon le rôle de chaque composant sur une dimension donnée. On pourrait illustrer ce concept par l'espace multifactoriel en analyse des données.

10. Notre démarche s'appuie ici sur la formalisation d'ontologie informatique dans un objectif final de partage et de réutilisation des connaissances par l'homme ou la machine. Pour ce travail, on adoptera donc la définition suivante « Une ontologie est une spécification formelle, explicite d'une conceptualisation partagée » d'après Studer *et al.* (1998, p. 25).

11. On pourrait parler d'une « bibliothèque de cas » au sens où l'entendent les socio-anthropologues, c'est-à-dire comme un ensemble d'expériences renvoyant à une pluralité de visions sur une situation donnée, que l'on peut analyser ensemble pour donner du sens à un problème particulier (Damien, 1995 ; Tripier, 2009).

entretiennent¹² dans un modèle conceptuel réalisé avec le langage graphique UML - *Unified Modeling Language*¹³.

En rassemblant et en étudiant le langage utilisé dans diverses études de cas, cet exercice vise à produire un modèle conceptuel qui puisse être largement partagé par la communauté des chercheurs intéressés par la question de la mobilité, mais qui ne repose pas sur un modèle prédéfini.

Nous tenons ici à souligner la créativité, les échanges et la dimension heuristique inclus dans la démarche de création d'une ontologie (Nuninger *et al.*, 2020a). Cette démarche qui permet de mieux comprendre et de cartographier les limites et le potentiel de notre domaine de connaissance, nous aide dans le processus d'interprétation des éléments physiques observés. Il est important de noter que l'ontologie qui en résulte fournit un cadre formel qui peut être utilisé pour partager des données, mais reste flexible car d'autres projets peuvent organiser leurs propres données différemment selon le contexte de la recherche.

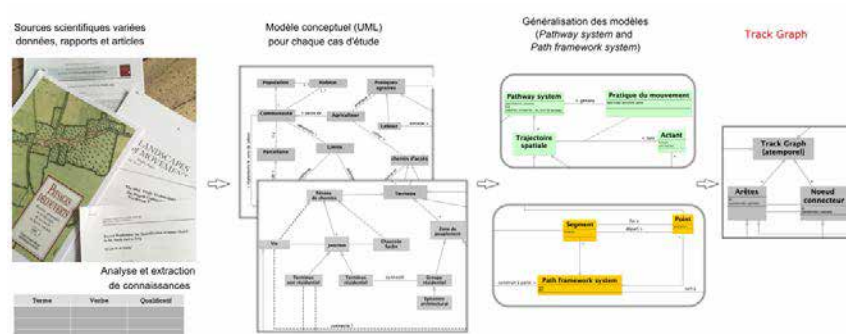


Figure 2. Schéma de la démarche

Pour illustrer notre approche, nous présentons deux études de cas pour lesquelles nous avons modélisé le domaine des concepts de mobilité sous la forme d'un schéma UML représenté par un graphe¹⁴. Il est évident que les définitions des concepts et de leurs relations demeurent critiquables, mais ce sont les éléments de convergence et de divergence de nos différents points de vue sur ces concepts et

12. Par exemple si le texte dit « les routes les plus longues ont servi au voyage et au transport de marchandises », nous distinguerons dans les modèles UML le concept de « route » et son attribut de longueur, relié aux concepts de « voyage », de « transport » par l'action « sert ».

13. Des exemples de ces représentations UML sont présentés en détail dans la partie 4 de cet article. Nous signalons toutefois que nous avons pris quelque liberté avec le langage UML pour exprimer soit l'incertitude de relations parfois hypothétiques, soit des enchaînements d'activités. À ce stade de la réflexion, il s'agit plus d'utiliser la dimension graphique du langage que son côté opérationnel.

14. Réalisé avec un logiciel tel que yEd : <https://www.yworks.com/products/yed>

relations liés à la notion de mobilité qui deviennent la base d'une étude et d'un débat plus approfondis. On évite ainsi de créer une terminologie supplémentaire et trop généralisante de « chemins », chargée d'un bagage conceptuel implicite et inextricable.

4. Les études de cas

L'analyse interculturelle des deux études de cas, utilisées pour illustrer notre approche, est fondée sur trois documents concernant une région de Mésoamérique, Caracol au Belize (Chase et Chase, 2001 ; Chase *et al.*, 2011) et une région du nord-ouest de l'Europe, la Beauce en France (Leturcq, 2008). Ces études de cas ont permis d'enregistrer des traces matérielles liées à des processus de mobilité à l'aide de divers capteurs et protocoles (LiDAR, imagerie, travail de terrain, cartes historiques, documents historiques, et témoignages ethnographiques).

A travers ces études de cas, nous explorons, d'une part, la relation observée, par les auteurs de l'article, entre les éléments qui permettent le déplacement entre des centres de peuplement identifiés créant des connexions à l'échelle régionale, et d'autre part, la relation entre les éléments qui décrivent un fonctionnement de mobilité locale. L'objectif est d'observer dans chacune d'entre elles comment un réseau de communication formel peut s'articuler avec le schéma de mobilité développé par la population *via* soit l'*habitus*¹⁵ (Bourdieu, 1972, 1980), soit des activités quotidiennes répétitives et transmises dans l'ensemble de l'espace habité. Comment les gens se déplacent-ils à travers le paysage pour rejoindre des centres résidentiels et des zones de ressources, ou pour effectuer des pratiques agricoles ? Comment les réseaux formels influencent-ils ces trajectoires quotidiennes ? Comment l'articulation entre ces deux systèmes produit-elle une structure qui imprime sa marque sur le paysage et influence les flux de mouvement ? Rien ne semble évident à cet égard, aucun chemin type ne peut être facilement décrit et reconnu dans les données numériques.

4.1. Région de Caracol (Belize)

Dans la région de Caracol, de nombreuses voies ont été découvertes grâce à l'analyse du MNT dérivé du LiDAR. De nouveaux segments de chaussée s'ajoutent ainsi à l'ensemble de données compilées à partir d'observations précédentes réalisées sur le terrain et à partir d'images satellites (Chase et Chase, 2001). L'étude des données altimétriques LiDAR a révélé une grande partie du système viaire principal reliant l'épicentre principal de Caracol à d'autres sites, appelés *termini* et

15. L'*habitus* est une sorte de matrice à travers laquelle on perçoit le monde et qui influence notre comportement de façon non consciente. Sans déterminisme, mais selon la position sociale et par transmission, il se manifeste par un ensemble cohérent de goûts et de pratiques des individus (Cabin, 2000).

compris comme des centres secondaires, potentiellement résidentiels (Chase *et al.*, 2011).

Le MNT LiDAR montre également qu'il existe une continuité d'habitats et de terrasses entre l'épicentre et ces *termini*. En effet, cette vaste zone de peuplement est caractérisée par de petits groupes de maisons organisés autour d'une petite place, qui sont intégrés dans la zone agraire, selon un modèle reconnu dans d'autres régions mayas. Ces unités de peuplement ne sont pas directement reliées à l'épicentre ou aux centres secondaires par des chaussées. Par ailleurs, il est fort plausible que les éléments composant le système agraire ont permis d'établir des liens solides entre ces habitations et les principaux centres résidentiels (Arnauld et Michelet, 2004). Dans l'étude de cas de Caracol, on a pu noter que les auteurs font référence à des entités qu'ils appellent « *vias* » et caractérisent comme étant des « routes plus informelles et plus courtes » (Chase et Chase, 2001), qui relient des groupes familiaux (ménages) aux principales chaussées, ou qui relient « d'importants groupes résidentiels directement à divers *termini* non résidentiels ». Les auteurs notent qu'à Caracol, « de nombreuses chaussées (...) sont enchevêtrées avec des terrasses agricoles » et que les *vias* peuvent être principalement identifiées grâce à une cartographie détaillée de ces terrasses. Alors que « seulement une demi-douzaine de *vias* ont été formellement identifiées comme rattachées aux longues chaussées [du réseau viaire] de Caracol », ils soulignent qu'« une cartographie détaillée des terrasses permettrait sans aucun doute de trouver d'autres exemples » (Chase et Chase, 2001). Puisque les *vias* identifiées semblent fournir un accès direct aux principales chaussées, les auteurs suggèrent que ces voies étaient en partie utilisées à des fins quotidiennes, pour faciliter les déplacements et la communication des groupes familiaux. Cet exemple illustre comment, derrière le réseau des chaussées, formel et plus lisible dans le paysage, il y a un maillage produit par les pratiques quotidiennes de groupes familiaux allant dans les champs, les centres résidentiels et d'autres lieux du territoire. Pour ce faire, les individus ont produit et utilisé différents types d'éléments et/ou structures tels que ceux observés par l'archéologie (empreintes, sentiers, cairns, etc.). C'est l'agencement de ces éléments dans la chaîne opératoire¹⁶ de leur activité qui permet de définir leurs itinéraires ou trajectoires quotidiennes, comprenant des segments de chaussées formelles, des bords de terrasses agricoles ou des chemins étroits difficilement identifiables.

Nous suggérons, dans un premier temps, de simplifier la description du système de communication maya de Caracol, en s'appuyant sur les articles des archéologues (Chase et Chase, 2001 ; Chase *et al.*, 2011) où l'espace territorial bien délimité appartient à un système de peuplement plus vaste (figure 3).

Ce système de peuplement crée le squelette de l'organisation des populations et des ressources (biens) intégrées dans le territoire par le biais de systèmes politiques

16. On parle de chaîne opératoire au sens défini par Leroi-Gourhan (1998), comme d'un processus décrit par une chaîne d'actions interconnectées impliquant une transformation continue.

et économiques. Le territoire de cette population est composé de ressources agricoles (terrasses agricoles), de ressources en eau (réservoirs), de zones de peuplement (groupes résidentiels et groupes ménages isolés) et de zones vides dites « non résidentielles » qui structurent un paysage particulier. Dans ce paysage, il y a un épiscentre (groupe architectural et résidentiel) et plusieurs endroits appelés *termini* qui peuvent être résidentiels. Chaque composante de ce paysage, décrite ci-dessus, est reliée spatialement par un réseau de chemins composé de chaussées très formelles appelées « saches » et de chemins plus informels appelés « vias ». Les tracés de ces *vias* sont enchevêtrées dans des terrasses agricoles et reliées au réseau viaire formel par des jonctions avec les principales chaussées formelles, les *saches*. Ces chaussées organisées en réseau relient le groupe résidentiel de l'épicentre à d'autres *termini* résidentiels et, dans certains cas, à des *termini* non résidentiels (réservoirs, terrasses agricoles ou autres éléments) par des jonctions. Les groupes familiaux quant à eux n'ont pas de position claire dans cette étude de cas. Il est notamment difficile de savoir comment ils sont reliés à leurs champs (terrasses agricoles), aux sentiers informels ou au réseau viaire formel et, par extension, aux groupes résidentiels (épiscentre et *termini*) et aux lieux non résidentiels (réservoirs ou autres). Bien que les auteurs ne l'affirment pas clairement, on peut supposer que ces groupes familiaux sont le plus souvent reliés au réseau viaire principal des *saches* par les entités nommées *vias* mentionnées dans la première étude et en partie par l'emprunt de segments de chaussées construites (*saches*).

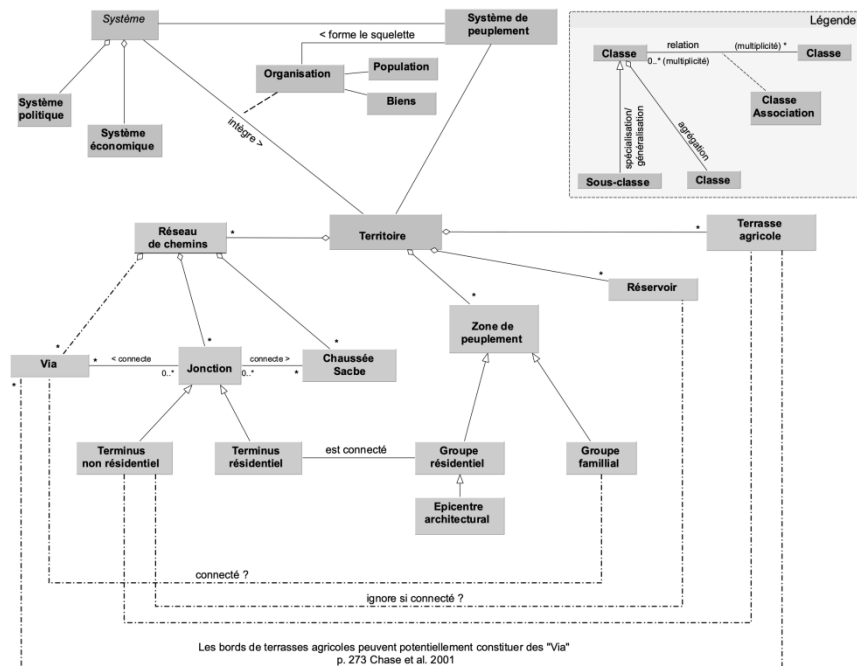


Figure 3. Schéma UML du système de circulation à Caracol, Belize

4.2. La région de la Beauce (France)

Le second exemple traite d'une étude de cas dans le nord-ouest de l'Europe, dans la région de la Beauce (France) au cours de la période médiévale et du début des temps modernes (12^e-17^e siècles après J.-C.). Les principaux concepts et relations seront brièvement présentés pour illustrer l'hétérogénéité des éléments topographiques utilisés pour se déplacer dans le paysage agricole et les circonstances de leur production. Pour une présentation détaillée de cette étude de cas, nous nous référons à un article précédent (Nuninger *et al.*, 2020a). Cette étude est fondée sur les travaux de Samuel Leturcq (Leturcq, 2008), relatifs à l'organisation de paysage de champ ouvert, et en particulier sur la répartition des voies de communication dans ce type de territoires agricoles. Si un réseau routier local dispersé composé de chemins dits formels peut être révélé par les cartes historiques, il n'y a aucune trace évidente de chemins permettant de desservir les petites parcelles enserrées dans les zones agricoles extensives et continues composées d'un grand nombre de petites parcelles en bande contiguës. Dans ces conditions, on peut se demander comment un paysan pouvait atteindre les parcelles de terre qu'il cultivait. Des sources historiques écrites permettent de répondre partiellement en indiquant que les limites de certains champs peuvent être utilisées comme des chemins par les agriculteurs pour accéder à leurs champs (Leturcq, 2008).

Dans cette région, la pratique consistant à utiliser les limites des champs comme chemins est également attestée par le terme « sommière »¹⁷ (Leturcq, 2008). Dans la langue dialectale des paysans de Toury en Beauce, les termes « sommière », « sommier » ou « têtère » désignent une portion surélevée d'un champ » située à la jonction entre deux champs : la crête dite « de labour » (Lachiver, 1997). Cet élément topographique, potentiellement utilisé comme cheminement piétonnier par les agriculteurs, n'a en fait pas été construit pour se déplacer mais résulte des activités agricoles liées à une technique de labour spécifique, ainsi que de règles sociales particulières qui y sont associées, notamment le sens du labour dans la parcelle, imposé par la communauté au sein de chaque quartier de culture. Au fil des ans, l'accumulation de terre en bordure de chaque champ, produite par le retournement de la charrue, s'est transformée en lits surélevés, ou crêtes, que l'on peut encore reconnaître en Beauce, comme dans beaucoup d'autres régions (Beresford et St. Joseph, 1979 ; Juillard, 1953 ; McOmish, 2011 ; Hall, 2014). Avec la réglementation du sens de labour, la crête devient progressivement une limite continue le long de nombreux champs contigus organisés à l'intérieur d'un quartier de parcelles agrégées. Ces limites forment une structure en réseau qui a pu faciliter les mouvements entre les champs, d'une part, et les établissements (villages et fermes isolés), d'autre part.

17. Le terme « sommière » était utilisé dans le langage des paysans de Toury en Beauce pour désigner les crêtes de labours, et il est parfois lié au terme de chemin dans les terriers médiévaux et modernes. Le terrier est un registre des terres appartenant à un seul propriétaire.

C'est en réunissant les données matérielles observées en topographie, grâce au MNT LiDAR par exemple, et la connaissance historique des pratiques agraires dans leur action de mobilité que nous pouvons commencer à cartographier et à visualiser d'autres schémas de déplacement possibles que ceux envisagés initialement. En effet, en combinant le réseau routier formel, cartographié sur des documents historiques, avec celui des crêtes de labour, nous obtenons une image plus complète des possibilités de circulation des communautés rurales. Nous avons synthétisé cela dans un schéma UML où les composantes sont liées par des processus (figure 4). Par exemple, l'agriculteur peut accéder aux parcelles en utilisant des chemins appelés « chemin d'accès ». Il peut s'agir soit de crêtes de labour utilisées comme chemins, soit d'autres chemins plus formalisés tels que des routes ou des itinéraires représentés sur la carte. L'organisation des quartiers dans le paysage tend à structurer la trajectoire des déplacements et à compléter les autres artères de circulation présentes sur le territoire.

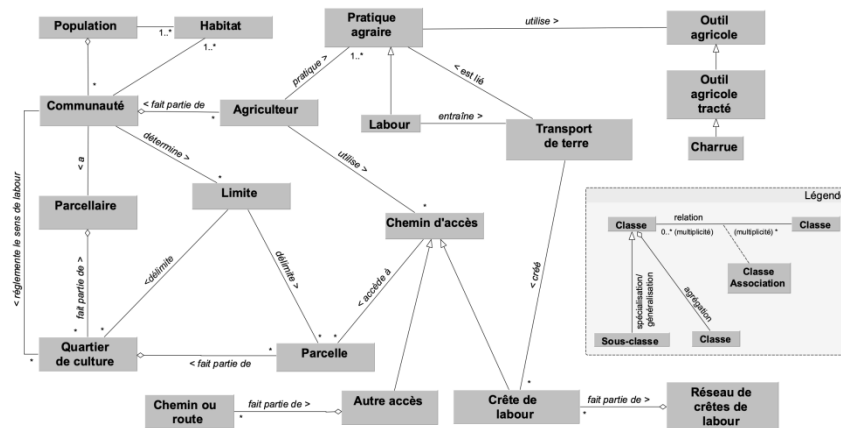


Figure 4. Schéma UML du système de circulation dans la région de la Beauce

Dans cette étude de cas, les connaissances structurées dans le schéma (figure 4) nous permettent de relier des détails importants, tels que le mode d'utilisation de l'instrument aratoire, à une vue plus globale de la structure morphologique du paysage agricole, ce qui est fondamental pour l'étude du système de mobilité de ce territoire. Ce que nous avons appelé « chemin d'accès » est maintenant mieux défini. Nous disposons désormais d'informations sur sa morphologie, son rapport à l'organisation sociale et paysagère, son utilisation et sa temporalité, liées aux règles de la communauté agricole. En se concentrant sur l'élément clé désigné comme « chemin d'accès » on peut inclure des structures similaires qui permettent le mouvement en fonction d'autres types d'objets comme des escaliers, des terrasses, des chaussées ou des canaux, par exemple. Cette entité et les relations qu'elle entretient avec d'autres permettent ainsi de mieux modéliser le réseau de mobilité

local des agriculteurs, qu'il s'appuie sur une structure en terrasse, sur un mode d'organisation en champs ouverts ou encore sur un réseau de canaux (Nuninger *et al.*, 2020b).

4.3. Vers une première étape de généralisation dans la construction des processus de mobilité

Ces deux études de cas illustrent la forte hétérogénéité des traces matérielles liées à la mobilité : chaussées construites, limites de champs ou terrasses agricoles, talus. Tandis qu'il n'y a pas forcément de volonté ou d'opportunité pour construire un chemin, on trouve néanmoins des similitudes dans la manière dont les structures paysagères produites dans le cours d'actions spécifiques (activités, tâches) sont cooptées pour permettre le mouvement tout en étant adaptées ou transformées progressivement. Cette utilisation *ad hoc* des éléments du paysage pour se déplacer dans le paysage crée une sorte de « tissu de cordes entrelacées ou nouées », ou un maillage (Ingold, 2011) de chemins informels. Dès lors que les activités sont normalisées par la communauté, de manière plus ou moins explicite, soit par des règles, soit par des pratiques répétées (*habitus*), elles tendent à structurer progressivement le maillage. Ce maillage a une empreinte dans le monde tangible d'aujourd'hui, caractérisée par des éléments matériels plus ou moins discrets qu'il est nécessaire de réagencer pour qu'ils puissent rendre compte des processus de déplacement d'une population, réalisés pour effectuer une tâche ou pour accéder à un espace particulier dans l'environnement¹⁸. Dans chaque cas, l'agencement de traces tangibles s'appuie sur la coexistence de plusieurs points de vue et usages ou intentions lors des déplacements. C'est cette complexité liée à la richesse des modalités de déplacement qui rend difficile l'analyse des traces archéologiques, par définition discrètes, dans une perspective continue de production et de transformation de l'espace des mobilités. Néanmoins, la généralisation des concepts issus de plusieurs schémas UML (Nuninger *et al.*, 2020) va nous permettre de réduire cette complexité à deux grandes logiques qui sous-tendent le processus de déplacement.

La première est motivée par la nécessité de se déplacer d'un endroit à l'autre *via* un réseau de circulation organisé et socialement reconnu. C'est ce qu'on appelle les « itinéraires formels », qui peuvent être définis « comme des preuves tangibles et physiques d'un itinéraire de déplacement servant de moyen de communication entre des points ou des zones d'activité » (Trombold, 1991). Leur morphologie peut généralement être caractérisée, leur degré de construction est très variable mais souvent identifiable, et l'itinéraire est entretenu grâce à un investissement délibéré de main-d'œuvre (nettoyage et/ou réparation de la chaussée, réparation des

18. L'empreinte laissée dans un même parcours peut être composée d'éléments de limite de champs, de guides ponctuels comme des cairns ou des arbres, de tronçons de routes ou de chemins, de volées d'escaliers ou de passages dans une falaise rocheuse, de bords de terrasse agricole, de parcelles particulières, de venelles, de caniveaux, etc.

bordures). Cet ensemble de routes formelles a généralement laissé son empreinte dans le paysage matériel actuel, mais aussi dans les documents (historiques et cartographiques) et la tradition orale (chansons, récits). Cette logique de mouvement utilisant des itinéraires formels est conceptualisée par ce que nous avons appelé le « *path framework system* » (figure 5). Elle peut être matérialisée par une route, par une série de repères paysagers et/ou par un récit, qui donc ne laisse pas de traces sous la forme de structures archéologiques physiques mais sous la forme d'inscriptions mémorielles dans la pratique du déplacement. Le *path framework system* a donc des propriétés contextuelles et temporelles et il est consciemment conçu par un groupe d'individus, c'est-à-dire par une société, qui le reconnaît.

La seconde logique de mobilité est influencée par l'activité des individus. Elle n'est cependant pas à proprement parler individuelle, puisqu'elle est motivée par un ensemble de pratiques, autres que le simple déplacement¹⁹, répétées, définies et progressivement appropriées par une société dans son ensemble (*habitus*). Elle laisse également une empreinte dans le paysage tangible actuel, mais cette empreinte est composée d'un ensemble hétérogène de traces matérielles, le plus souvent liées aux activités quotidiennes. Elles peuvent varier en fonction du contexte saisonnier, des modes de déplacement ou des outils utilisés dans la pratique d'une activité. Ce type de mouvement est plus difficile à identifier car il n'est pas organisé comme un réseau de lieux connectés mais plutôt comme un maillage de lignes qui se croisent (Lefebvre, 2000 ; Ingold, 2011). Ces lignes peuvent être considérées comme des trajectoires guidées par les activités quotidiennes des individus. Elles sont composées par la mise en relation d'éléments matériels utilisés et produits par des individus en mouvement dans le cadre de leurs activités, mais peuvent aussi être associées à des segments d'itinéraires formels empruntés dans le cours même d'une activité. Tenter d'appréhender ces trajectoires spatiales, nécessite de mobiliser nos connaissances des pratiques, règles, normes et coutumes sociales et économiques. Ces connaissances nous permettent à leur tour de retrouver l'empreinte laissée, en particulier sa forme résiduelle dans le paysage tangible d'aujourd'hui. Elles nous renseignent aussi sur les modalités potentielles d'une articulation avec les composantes du réseau d'itinéraires formels. Nous avons défini cette deuxième logique de mouvement dans un concept appelé « *pathway system* » (figure 5). À la différence du *path framework system*, le *pathway system* est, du point de vue du déplacement, conçu de façon plus ou moins inconsciente, tout au moins sans objectif particulier de relier des lieux, par un groupe d'individus. Il s'agit d'un objet dont les propriétés contextuelles sont essentiellement liées au comportement d'individus ou de groupes dans le cadre de l'*habitus*. Le *pathway system* se rapporte à un temps donné pour une durée donnée, indépendamment de son empreinte dans le monde physique, car il est composé d'un ensemble hétérogène de caractéristiques

19. Par exemple, il peut s'agir d'une activité de cueillette, de production de charbon, de travail de la terre... qui implique un mouvement et un déplacement mais sans intentionnalité directe d'aller d'un point à un autre. C'est le mouvement dans le cours de l'activité elle-même qui définit la trajectoire spatiale.

morphologiques qui sont liées entre elles par une ou plusieurs pratiques spécifiques (*habitus*) à un moment donné.

De manière plus générale, nous avons pu constater une relation complexe entre les différentes traces tangibles observées dès lors que nous les examinons sous l'angle de la mobilité. Archéologiquement, il y a une discontinuité thématique évidente²⁰, que nous devons résoudre par un agencement logique, structuré en fonction des comportements de mobilité (ce qui renvoie à la notion de cofonctionnement, définie en note 9).

À l'échelle d'un espace, d'un paysage, la configuration globale des éléments qui définissent et structurent la mobilité doit donc être comprise dans deux cadres conceptuels radicalement différents mais coexistants : les réseaux et les maillages. Sur le long terme, l'influence de ces deux cadres co-évolue, renforçant la complexité de l'analyse dès lors que l'on tente d'expliquer l'origine des éléments liés à la mobilité observés dans le paysage tangible. Ainsi, par exemple, un segment de route, formellement identifié comme une route dans un réseau bien défini, peut en réalité trouver son origine dans une structure agraire utilisée comme un passage, dont l'usage préfigure le développement postérieur de la route. Inversement, un tronçon de route formellement créé comme route dans un réseau planifié peut par la suite disparaître et être remplacé par une simple limite de champs reconnaissable topographiquement et encore utilisée « officieusement » pour se déplacer.

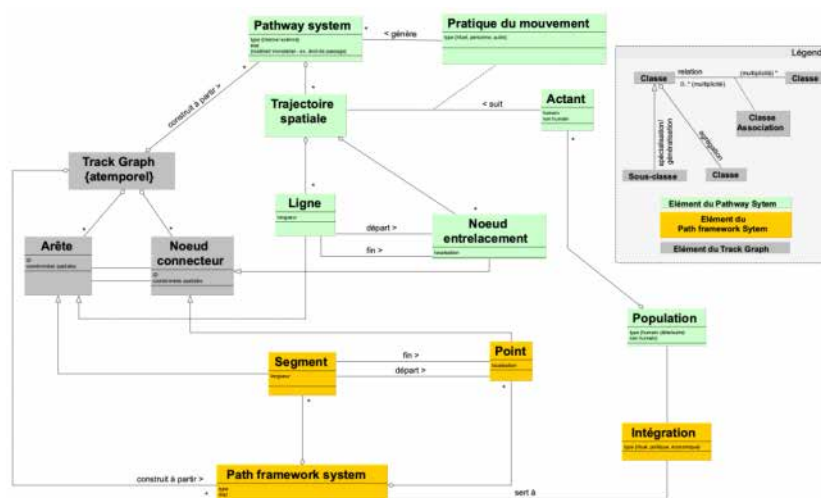


Figure 5. Structure théorique et agencement des trois concepts de pathway system, path framework system et track graph

20. Au sens où des éléments appartenant à des registres différents – éléments d'origine agraire, végétale ou minérale, rituelle... – peuvent composer la trajectoire spatiale d'un déplacement, soit un ensemble cohérent du point de vue de la mobilité.

Si, d'un point de vue théorique, il est possible de présenter cette complexité à partir d'exemples bien décrits et pertinents, il est beaucoup plus difficile d'envisager des méthodes permettant de l'analyser de manière plus globale et systématique en s'appuyant sur des ensembles de données qui sont, par définition, hétérogènes.

5. Vers une abstraction, heuristique du raisonnement : le « track graph »

Pour terminer notre proposition, nous donnons la définition d'un concept analytique strictement abstrait et s'appuyant sur la théorie des graphes, l'objectif étant de le rendre opérationnel dans des analyses spatiotemporelles.

5.1. Motivations et flux de travail conceptuel

L'identification des entités nommées et de leurs relations selon différents points de vue (individuel et collectif, social, économique, politique) nous autorise à définir un objet composite abstrait, appelé « track graph » (figure 5). Celui-ci correspond à une construction qui permet d'interpréter la mobilité selon des éléments variés observés au fil du temps.

L'intérêt de cet objet abstrait réside : 1) dans sa capacité à structurer les données et les connaissances afin d'appréhender le système de circulation d'un espace comme des potentialités issues d'une combinaison d'éléments, et non par la juxtaposition de systèmes *a priori* incompatibles (formels et informels) ; 2) dans la possibilité d'une analyse transculturelle des processus de mobilité conservant le lien logique entre ces processus et les éléments locaux spécifiques interprétés dans chaque cas d'étude.

Ce lien logique est notamment pertinent lorsque l'on considère la manière dont nous générons des connaissances par l'observation répétée et la construction d'hypothèses (voir figure 6). Dans la spirale herméneutique (Ricœur, 1990 ; Hodder et Hutson, 2003 ; Chazan, 2018), l'étape de l'analyse des données est cruciale pour la formulation d'hypothèses et de modèles conceptuels, souvent fortement dépendant de techniques analytiques spécifiques, telles que la télédétection, les SIG, l'analyse de réseaux, les statistiques ou la modélisation, qui supposent une structure formalisée des données observées. Le *track graph* est donc proposé ici comme une structure de données qui devrait permettre un éventail relativement large d'approches analytiques, fondé notamment sur la théorie des graphes. Le *track graph* regroupe (hypothétiquement) toutes les évidences de mobilité observées, présentes et passées, en un ensemble de nœuds et d'arêtes pouvant recevoir un nombre illimité d'attributs²¹. Ces attributs peuvent être associés à chaque élément du

21. Comme nous le verrons ensuite un nœud dans le *track graph* n'implique pas nécessairement qu'on lui attribue une fonction de connecteur dans un réseau, elle peut aussi être comprise comme un point d'entrelacement dans un maillage.

graphe par reconstruction logique *via* des ontologies qui décrivent et structurent l'ensemble des connaissances selon différentes conceptions du monde.

D'un point de vue méthodologique, nous pouvons subdiviser notre chaîne de travail conceptuel en trois phases (figure 6).

La première phase correspond à l'observation de données à partir d'une ou plusieurs sources (terrain, imagerie, MNT LiDAR, cartes, documents historiques, etc.). Les observateurs identifient les caractéristiques archéologiques dans l'ensemble des données disponibles en utilisant un protocole d'observation partagé ou individuel. Ce faisant, ils élaborent des hypothèses et un premier modèle conceptuel pour guider leurs observations et la manière dont ils les enregistreront et les interpréteront.

La deuxième phase consiste à analyser ces ensembles de données observées et la manière dont elles ont été reconnues et interprétées du point de vue du processus de mobilité, afin de proposer un schéma des concepts utilisés et des liens établis par les observateurs pour relier ces différents concepts. C'est précisément ce que nous avons tenté de faire en développant une approche ontologique sur plusieurs études de cas (voir partie 4 et Nuninger *et al.*, 2020b).

Dans la troisième phase, que nous ne développons pas dans cet article, nous pourrions explorer la reconstruction logique de chemins plausibles ou même de schémas de circulation, pour une période donnée, en s'appuyant sur le *track graph* ainsi que sur des connaissances et données complémentaires. Dans notre approche, le *track graph* est un squelette de chemins plausibles qui définit le terrain de jeu dans lequel nous pouvons explorer différents modèles hypothétiques de pratiques de mobilité (modèles induits) à l'aide, par exemple, de modèles de simulation multi-agents. L'objet abstrait représenté par le *track graph* peut donc être utilisé comme support pour reconstruire plusieurs types de structures de circulation (réseaux routiers, parcours). Si ces reconstructions demeurent hypothétiques, elles ont l'avantage de se fonder sur l'ensemble des évidences archéologiques observées. Celles-ci, détachées de leurs attributs d'observation, n'induisent alors aucune interprétation *a priori* dans le processus de reconstruction. Le lien avec les attributs est toutefois maintenu par un identifiant et des coordonnées géographiques (*cf.* 5.2) permettant ainsi de revenir aux données initiales lors d'une phase d'évaluation du modèle de reconstruction.

Cette troisième étape de la chaîne de travail conceptuel est encore en cours. Dans la suite de cet article, nous analysons donc plus en détail le concept de *track graph*, sa construction et son articulation avec deux autres concepts issus de nos ontologies d'observation : le *pathway system* et le *path framework system* (discuté dans la section 4.3).

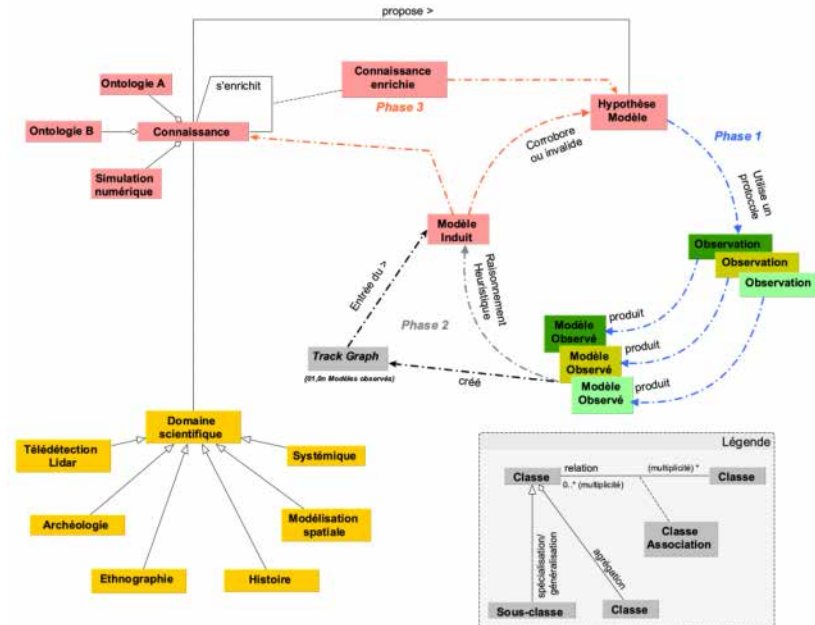


Figure 6. Chaîne de travail conceptuel associée au track graph

5.2. Composition du track graph

Nous définissons le *track graph* (cf. figure 5) comme un objet abstrait, composé d'un ensemble de nœuds et d'arêtes. Il est important de souligner ici que ces nœuds et arêtes sont entièrement abstraits, fournissant une représentation simple, fondée sur un graphe, des connexions entre les entités. Cette approche s'inspire de celle développée dans le contexte de l'archéologie urbaine *via* le modèle OH-FET²² où les entités géométriques représentant des éléments physiques sont abstraites de leurs attributs descriptifs (Rodier et Saligny, 2010) afin de distinguer clairement les propriétés interprétées telles que le type ou la classe, décrites dans les attributs, de la chose abstraite, décrite à travers sa géométrie.

Dans le *track graph* un nœud peut représenter un lieu, par exemple une ville, une place de marché, une habitation individuelle, un repère dans le paysage ou une intersection entre des chemins. Mais une telle intersection pourrait être reconnue comme un carrefour ou une jonction, c'est-à-dire un lieu, ou simplement comme le croisement de deux chemins sans signification particulière, par exemple, l'intersection entre un sentier d'animaux et un chemin de randonnée. La

22. Modèle qui déconstruit l'objet historique (OH) selon ses trois dimensions : fonction, espace, temps (FET).

caractéristique essentielle d'un nœud est que les individus peuvent se déplacer entre eux ou à travers eux. En d'autres termes, la représentation d'un nœud dans le *track graph* n'implique pas nécessairement qu'on lui attribue une fonction de connecteur dans un réseau, elle peut aussi être comprise comme un point d'entrelacement dans un maillage. Dans ce cas, comme nous le verrons plus tard, nous attribuons un statut inactif au nœud du *track graph*.

Le mouvement lui-même a lieu sur les arêtes, représentant des éléments et/ou structures compris comme des chemins. On peut leur attribuer divers attributs, décrivant, par exemple, leurs manifestations matérielles ou leurs fonctions. Au fil du temps, des intersections et des lieux ainsi que des chemins peuvent apparaître, disparaître et réapparaître. Le *track graph* enregistre et accumule chaque trace observée, généralisée cartographiquement sous forme d'éléments géométriques (arêtes et nœuds). Ce graphe est complété au fur et à mesure des observations (figure 7). Une observation permet de créer une arête ou un ensemble d'arêtes et de nœuds, mais une même arête peut correspondre à plusieurs observations (provenant de divers observateurs, et/ou de diverses sources, à différents moments d'observation ou selon différents protocoles).

Les arêtes et les nœuds sont dénués d'attributs liés aux observations faites sur les entités qu'ils représentent, à l'exception d'un identifiant (ID) et de coordonnées spatiales.

Le fait de conserver l'ID et les coordonnées spatiales permet de maintenir un lien avec les observations initiales qui sont décrites par un ensemble d'attributs *via* une ontologie. Tandis que certains attributs sont purement descriptifs, par exemple la taille, la hauteur, la largeur, la longueur, les matériaux, la signature spectrale, les attributs topologiques (croisement, à côté, etc.) et ne changeront donc pas d'un schéma ontologique à l'autre, d'autres attributs en revanche, tels que la classe ou le type, peuvent varier selon les schémas. Ce processus d'abstraction initial lors de la construction de la structure *track graph* permet en retour aux nœuds et aux arêtes de s'enrichir d'une palette d'attributs variés grâce à chaque schéma ontologique et chaque ensemble d'observations. Par exemple, un chercheur interprétant visuellement un MNT LiDAR pourrait identifier et interpréter une anomalie topographique linéaire comme faisant partie d'une ancienne route romaine. Le même élément matériel pourrait être identifié et interprété par un archéologue médiéval comme une limite de champ, ou comme une voie d'accès potentielle pour les agriculteurs. Ce type de controverse sur l'interprétation est à la fois assez courant et source de débats passionnés lorsqu'une équipe de recherche étudie l'évolution du paysage (Williamson, 2016). De même, lorsque l'on combine les données produites par plusieurs équipes, les multiples interprétations créent souvent une complexité en partie artificielle (Rippon *et al.*, 2015). En s'autonomisant par rapport à l'interprétation véhiculée par les attributs, le *track graph* constitue un outil permettant de combiner divers ensembles d'observations. Agissant comme une représentation dynamique des observations, il évolue au fur et à mesure que de nouvelles observations sont faites (un segment peut être partitionné par de nouveaux

nœuds, ou dissocié en deux segments distincts selon l'échelle d'observation). *In fine*, le *track graph* constitue un canevas ou un squelette spatial commun qui peut être utilisé pour recomposer différents systèmes thématiquement significatifs.

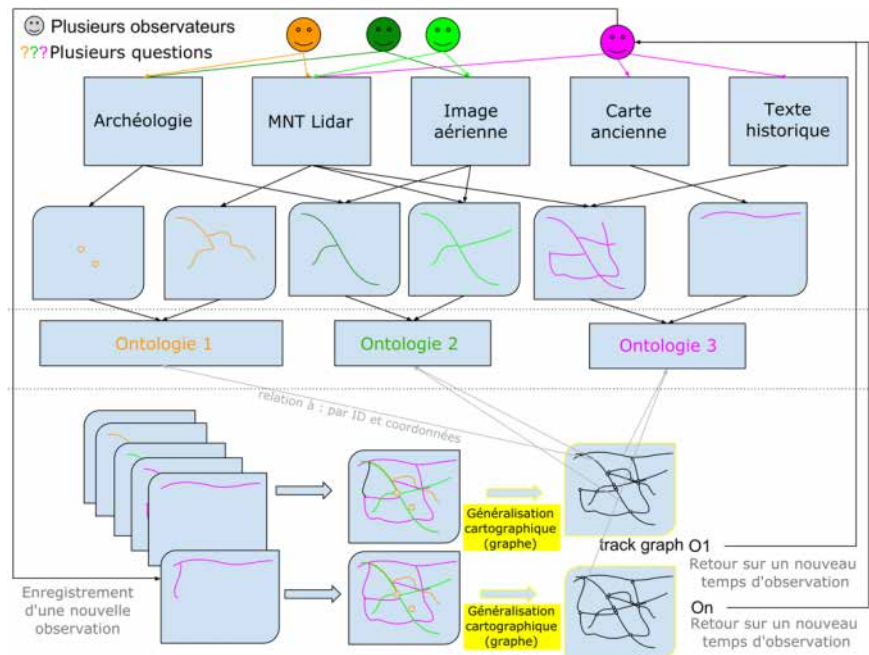


Figure 7. Le processus de création d'un track graph

Compte tenu de notre volonté de considérer le paradigme du maillage dans notre approche de la mobilité, il pourrait être reproché d'utiliser le modèle du réseau comme outil d'analyse. Pour surmonter cette incohérence théorique, nous introduisons la notion pragmatique de statut actif et inactif des nœuds. Pour illustrer notre propos, prenons une observation unique, par exemple une ligne identifiée sur une image aérienne. Elle peut être représentée par une seule arête et deux nœuds à ses extrémités ou par un ensemble d'arêtes reliées par plusieurs nœuds. Dans ce cas, les nœuds n'ont pas le rôle d'origine ou de destination, mais ne feront que relier des arêtes pour représenter géométriquement une entité, tandis que certains de ces nœuds peuvent contribuer à représenter d'autres observations indépendamment de cette entité linéaire. La structure très abstraite du *track graph* permet de traiter les nœuds et les arêtes comme *actifs* ou *inactifs*. Par exemple, prenons un ensemble d'arêtes et de nœuds dans le *track graph* qui représentent une route romaine dans notre modèle. Ceux de ces nœuds qui ne peuvent être reliés à aucune donnée sémantique seront considérés comme inactifs dans le graphe soumis à l'analyse de réseau. Parmi ces nœuds, et selon un autre modèle de mobilité, si l'un d'entre eux

(N) correspond à une jonction qui permet à un agriculteur de rejoindre cette route depuis un passage en bordure de terrasse, alors il sera considéré comme actif. En effet, la jonction entre les deux segments au nœud N est significative pour l'agriculteur, même si ce n'est pas le cas pour d'autres individus qui ne circuleraient que sur l'axe de la route.

Dans le *track graph*, le réseau actif, composé de lieux (nœuds) et de chemins (arêtes) qui sont mobilisés dans un modèle induit, n'est qu'un sous-ensemble de l'ensemble des lieux et des parcours *potentiellement réalisés*. Les arêtes et les nœuds activés par le modèle peuvent alors rendre compte de la structure tangible de la mobilité à des moments spécifiques. Un second intérêt du *track graph* qui découle de sa nature abstraite est la possibilité de l'utiliser comme un cadre analytique qui peut être transposé à plusieurs études de cas relevant de différentes cultures. Cela permet de comparer des modèles de parcours fondés sur la même structure, tout en interprétant chacun d'entre eux selon l'ontologie propre au domaine culturel étudié.

5.3. Résumer et établir des liens : track graph, pathway system et path framework system

Dans cet article, nous avons développé notre approche de deux types de logiques clés qui sous-tendent les processus liés au mouvement : le *path framework system* et le *pathway system*, à partir de plusieurs cas d'études régionaux. Nous avons ensuite exploré comment ces deux conceptions génériques de la mobilité peuvent être exploitées dans l'interprétation de données provenant de sources hétérogènes, et comment les observations qui en découlent peuvent être recombinaisonnées par modélisation à l'aide de l'outil conceptuel *track graph* (figure 5).

Le *path framework system* et le *pathway system* fournissent des modèles de mobilité complémentaires et sont destinés à être utilisés ensemble pour mieux appréhender la structure globale de la circulation dans un paysage donné.

Le *path framework system* est utilisé pour décrire un réseau d'itinéraires conçu par une société, ou au moins reconnu comme tel par un groupe. Cela inclut les chemins formels qui peuvent être utilisés pour le transport ou le voyage d'un point d'origine à une destination. Dans ce cadre conceptuel, le mouvement est essentiellement orienté vers la destination (Wallace, 1993). Dans le cadre de l'élaboration de modèles territoriaux, cette logique pourrait être utilisée pour modéliser la manière dont les systèmes formels de déplacement contribuent à l'intégration sociale, politique ou culturelle de la population.

Le *pathway system* est utilisé pour décrire la trajectoire produite par un ou plusieurs acteurs dans l'exercice de leurs activités. Cette trajectoire est essentiellement informelle, bien qu'elle puisse inclure des éléments du système de parcours formel. Dans le système de parcours, le mouvement est orienté par l'activité plutôt que vers une destination. La logique de mobilité du *pathway system* peut être abordée de manière analogue au processus de parcours tel que développé

par Ingold (Ingold, 2011), mais qui, associée à l'idée d'*habitus* (Bourdieu, 1972, 1980), produit une empreinte dans le paysage créée collectivement sur le long terme.

Le *track graph* est quant à lui un graphe mathématique totalement abstrait mais dans lequel chaque entité (nœud ou arête) représente un élément physique du paysage observé. Le *track graph* relie les deux logiques de mobilité décrites ci-dessus dans un cadre analytique unique. Contrairement au *path framework system* et au *pathway system*, le *track graph* ne contient aucune interprétation de l'utilisation des éléments matériels représentés ni de la date de leur utilisation. Il sert uniquement de support abstrait et géométrique pour les deux cadres conceptuels. Sa structure, composée d'arêtes et de nœuds, est une simple représentation conventionnelle destinée à faciliter l'analyse syntaxique de l'espace.

Dans l'articulation des trois concepts, la gestion du temps mérite une attention particulière. Comme expliqué, le *track graph* est un ensemble de nœuds et d'arêtes qui couvre tous les chemins et lieux potentiels observés et déduits, indépendamment de leur ordre chronologique. En ce sens, le *track graph* lui-même est historiquement et archéologiquement atemporel, car il rassemble des éléments de toutes les périodes. À un moment d'observation donné (O1), la structure du *track graph* enregistre la totalité des chemins réalisés ou potentiels et peut être utilisée pour explorer des chemins plausibles, en utilisant une ontologie spécifique décrivant les comportements de mouvement et les connaissances associées sur un espace-temps donné (figure 7). De nouvelles observations peuvent ensuite être ajoutées au *track graph*, marquées avec leur moment d'observation (O2... On), capturant le développement de la compréhension du paysage et de la mobilité à travers celui-ci *via* un processus itératif de modélisation, d'observation et d'interprétation. Les phases de construction du *track graph* sont liées au moment où les observations sont faites, soit au moment où nous sommes arrivés à un état de connaissance particulier.

Contrairement au *track graph*, le *pathway system* et le *path framework system* sont *historiquement et archéologiquement temporels*. Ils sont délimités chronologiquement et leur dynamique reflète une dynamique historiquement significative. En archéologie, nous nous intéressons à l'analyse des modèles et des dynamiques dans un cadre temporel (T soit de t_0 à t_n) avec un intervalle de temps choisi. Nous pouvons donc nous référer à différentes reconstructions de modèles de mobilité (modèles induits) comme étant valables pour un cadre temporel particulier, en définissant les nœuds et les arêtes qui étaient présents à t_0 , et en définissant ceux qui deviennent actifs ou inactifs à la date de fin t_n . Ainsi, dans notre cadre temporel T, nous pouvons avoir des reconstructions R_1, \dots, R_n .

6. Conclusion

Alors que l'attention des archéologues et des géographes se focalise souvent sur les traces de routes construites ou de sentiers reconnus officiellement *via* la cartographie ancienne, nous démontrons que la compréhension d'un « système » de

mobilité dans le passé repose en réalité sur le lien que nous pouvons établir entre l'observations d'objets hétérogènes et la connaissance de différentes pratiques de mobilité et de différents usages, de différentes activités. Pour reconsidérer notre approche des mobilités passées, nous devons considérer, non seulement le rôle du biais des observateurs en termes de méthodologies employées, de sources de données disponibles ou d'expertise individuelle, mais également celui plus implicitement ancré dans les paradigmes théoriques sous-jacents et la façon dont la connaissance est exprimée dans des domaines de connaissance spécifiques à des contextes culturels et linguistiques extrêmement variés.

Nous avons tenté de démontrer qu'en créant une ontologie des pratiques de mobilité fondée sur une analyse textuelle, nous sommes capables de démêler, de structurer et de clarifier certains biais sémantiques et surtout de porter notre attention sur la nature des relations entre des éléments archéologiques plutôt que sur les éléments eux-mêmes, en particulier lorsqu'il s'agit d'analyser la matérialité d'un processus tel que le mouvement. Le développement d'ontologies constitue un exercice heuristique utile, pour mieux appréhender les raisonnements imbriqués dans des cadres de référence et des chaînes d'inférence souvent largement implicites. La décomposition de références conceptuelles implicites en chaînes de raisonnement explicites et logiques qui décrivent des entités de base et leurs relations permet d'identifier les éléments génériques constitutifs d'un système que l'on peut utiliser pour reconstruire, analyser et comparer les pratiques de mobilités dans un cadre unifié mais avec une approche *bottom-up*.

Notre approche, largement fondée sur une dimension symbolique, celle du mouvement passé lui-même qui n'est pas enregistré, s'appuie néanmoins sur l'empreinte matérielle laissée par le processus de mobilité réalisé dans une chaîne opératoire. L'intelligibilité de cette empreinte passe par l'articulation d'éléments hétérogènes dans des schémas narratifs renouvelés. Ces schémas nécessitent d'être consolidés par des expérimentations (modèle de simulation multi-agents par exemple) et un retour sur les données d'observations afin d'ouvrir et de parfaire notre lecture et notre interprétation des éléments et/ou structures archéologiques dans un paradigme de corps en mouvement. En s'appuyant sur la structure du *track graph* définie ci-dessus, nous proposons un outil conceptuel pour intégrer cette dimension symbolique de la mobilité passée dans des techniques de modélisation qui peuvent être utilisées pour simuler la mobilité et reconstruire des itinéraires potentiels utiles à notre compréhension des modes d'habiter passés.

Nous avons introduit le concept de « track graph » comme un outil analytique potentiel d'exploration et de comparaison des systèmes de parcours en relation avec diverses pratiques de mobilités, et ce, à de multiples échelles spatiales et temporelles. Le « track graph » est défini comme un ensemble de nœuds et d'arêtes représentant tous les éléments observés dans une région d'étude pouvant être liés au mouvement. Fondé sur la théorie des graphes, cet outil rend possible l'application d'approches analytiques bien établies telles que l'analyse de réseau et la modélisation multi-agents. En outre, il permet une description et une compréhension

enrichie des éléments et/ou structures archéologiques observés liés au mouvement grâce à la conceptualisation, la modélisation et la connexion aux systèmes de voies formels (réseaux) et/ou de parcours informels s'appuyant sur l'idée de maillage.

Bien que l'application pratique de cette approche à de nouveaux ensembles de données soit encore dans le domaine de la perspective, nous avons essayé d'illustrer le potentiel et la nécessité de synthétiser des données provenant de diverses sources en utilisant une approche formalisée, mais non standardisée, fondée sur un raisonnement ontologique et des concepts théoriques *a priori* antagonistes mais unifiés dans un système de graphes. Nous espérons que ce travail de clarification conceptuel contribuera au renouvellement de notre approche de la mobilité passée en s'appuyant sur le caractère complexe des empreintes laissées dans le monde tangible actuel, et plus largement, à la question de l'habiter utile à notre compréhension du peuplement et des paysages envisagés dans la longue durée.

Financement

Cette recherche a été financée par le Centre national de la recherche scientifique (International Emerging Action (IEA) PICS07899-MOVESCAPE), VU University Amsterdam (CLUE+ Activity Fund 0055), ZRC SAZU Ljubljana, université de Franche-Comté, université de Tours et université de Glasgow.

Remerciements

La recherche présentée dans cet article a été rendue possible en partie grâce au projet Movescape, dir. Laure Nuninger (CNRS) et Philip Verhagen (Vrije Universiteit Amsterdam) et grâce au ZRC-SAZU pour l'organisation de TRAIL2018 (Postojna, Slovénie) sous la responsabilité de à Žiga Kokalj et Benjamin Stular. Elle a également été soutenue par des postes de professeur invité d'un mois accordés à Philip Verhagen et Rachel Opitz par l'université de Franche-Comté. Des missions ont également été prises en charge par l'université de Glasgow et par l'université de Tours (grâce au projet SOLiDAR, co-organisateur). Des ateliers Movescape ont été organisés à Besançon (MSHE C.N. Ledoux, CNRS), à Tours (MSH Val de Loire et CITERES, université de Tours-CNRS), à Amsterdam (Vrije Universiteit Amsterdam), à Glasgow (université de Glasgow), et ouverts à un public plus large à Atlanta (CAA2017) et à Postojna (TRAIL2018). Enfin, elle a bénéficié du support et des échanges réalisés au sein de l'Action prospective ontologie et dynamiques spatiales du GDR Magis en particulier pour cette publication.

Bibliographie

- Akrich M. (2013). La description des objets techniques. *Sociologie de la traduction : Textes fondateurs*, M. Callon et B. Latour (Éd.), Presses des Mines, Paris, p. 159-178.
- Aldred O. (2016). Past movements, tomorrow's anchors. On the relational entanglements between archaeological mobilities. *Past Mobilities Archaeological Approaches to Movement and Mobility*, J. Leary (Ed.), Routledge, New York, p. 21-47.

- Aldred O. (2020). *The Archaeology of Movement*, Routledge, New York.
- Arnauld M.-C. et Michelet D. (2004). Nature et dynamique des cités mayas. *Annales. Histoire, Sciences Sociales*, vol. 59, p. 73-108.
- Banaszek Ł., Cowley D. et Middleton M. (2018). Towards national archaeological mapping. assessing source data and methodology. A Case study from Scotland. *Geosciences*, vol. 8, n° 272.
- Beresford M. W. et St. Joseph J. K. S. (1979). *Medieval England: An Aerial Survey*, Cambridge University Press, Cambridge.
- Bourdieu P. (1972). *Esquisse d'une théorie de la pratique. Précédé de « Trois études d'ethnologie kabyle »*, Librairie Droz, Genève.
- Bourdieu P. (1980). *Le sens pratique*, Les Éditions du Minuit, Paris.
- Cabin P. (2000). Dans les coulisses de la domination : Le monde selon Bourdieu. *Sciences humaines* (Auxerre), n° 105, p. 24-28.
- Case H. J. (1973). Illusion and meaning. *The Explanation of Cultural Change: Models in Prehistory*, C. Renfrew (Ed.), Duckworth, London, p. 35-46.
- Chase A. F. et Chase D. Z. (2001). Ancient Maya causeways and site organization at Caracol, Belize. *Ancient Mesoamerica*, vol. 12, n° 2, p. 273-281.
- Chase A. F., Chase D. Z., Weishampel J. F. et al. (2011). Airborne LiDAR, archaeology, and the ancient Maya landscape at Caracol, Belize. *Journal of Archaeological Science*, vol. 38, n° 2, p. 387-39.
- Chazan M. (2018). *World Prehistory and Archaeology: Pathways Through Time*, Routledge, New York.
- Childe V. G. (2015). *Piecing Together the Past: The Interpretation of Archaeological Data*, Routledge, London, New York.
- Damien R. (1995). *Bibliothèque et État*, Presses universitaires de France, Paris.
- Deleuze G. et Guattari F. (1980). *Mille plateaux*, éd. de Minuit, Paris.
- Deleuze G. et Parnet C. (1996). *Dialogues*, Paris.
- Djindjian F. (2010). Le rôle de l'archéologue dans la société contemporaine, *Diogenes*, vol. 229-230, n° 1-2, p. 78-90.
- Earle T. (2009). Routes through the landscape. A comparative approach. *Landscapes of movement. Trails, Paths, and Roads in Anthropological Perspective*, J. E. Snead, C. L. Erickson et A. J. Darling (Ed.), University of Pennsylvania Museum of Archaeology and Anthropology, Philadelphia, p. 253-270.
- Hall D. (2014). *The Open Fields of England*, OUP, Oxford.
- Hodder I. et Hutson S. (2003). *Reading the Past. Current Approaches to Interpretation in Archaeology*, Cambridge University Press, Cambridge.
- Hurard S., Roumegoux Y. et Chaoui-Derieux D. (2014). L'archéologie à l'épreuve de la modernité. De l'opportunisme à la maturité, *Les nouvelles de l'archéologie*, n° 137, p. 3-9.
- Ingold T. (2011). *Being Alive: Essays on Movement, Knowledge and Description*, Routledge.

- Juillard É. (1953). Formes de structure parcellaire dans la plaine d'Alsace. *Bulletin de l'Association de géographes français*, vol. 30, n° 232, p. 72-77.
- Lachiver M. (1997). *Dictionnaire du monde rural. Les mots du passé*, Fayard, Paris.
- Lefebvre H. (2000). *La production de l'espace*, Éditions Anthropos, Paris.
- Leroi-Gourhan A. (1998). *Le geste et la parole. Tome II : la mémoire et les rythmes*, Albin Michel, Paris.
- Leturcq S. (2008). Fonction et devenir d'un réseau invisible : les crêtes de labours dans les terroirs beaucerons (XIV^e-XX^e siècles). *Marqueurs des paysages et systèmes socio-économiques*, R. Compantangelo-Soussignan, J.-R. Bertrand, J. Chapman et P.-Y. Laffont (éd.), Presses universitaires de Rennes, Rennes, p. 163-174.
- McOmish D. (2011). Introductions to Heritage Assets: Field Systems. *English Heritage*, p. 1-5.
- Nuninger L., Opitz R., Verhagen P., Libourel T. *et al.* (2020a). Developing FAIR ontological pathways: Linking evidence of movement in lidar to models of human behaviour. *Journal of Computer Applications in Archaeology*, vol. 3, n° 1, p. 63-75.
- Nuninger L., Verhagen P., Libourel T. *et al.* (2020b). Linking theories, past practices, and archaeological remains of movement through ontological reasoning. *Information*, vol. 11, n° 6, p. 338.
- Ogden C. K. et Richards I. A. (1923). *The Meaning of Meaning: A Study of the Influence of Language upon Thought and of the Science of Symbolism*, Routledge et Kegan Paul Ltd., London.
- Palmer R. (2013). Reading aerial images. *Interpreting Archaeological Topography: Airborne Laser Scanning, 3D Data, and Ground Observation*, R. S. Opitz et D. C. Cowley (Ed.). Oxbow Books, Oxford, UK, p. 76-87.
- Ricœur P. (1990). *Soi-même comme un autre*, Seuil, Paris.
- Rippon S., Smart C. et Pears B. (2015). *The fields of Britannia: Continuity and change in the late Roman and early Medieval landscape*, Oxford University Press, USA.
- Rodier X. et Saligny L. (2010). Modélisation des objets historiques selon la fonction, l'espace et le temps pour l'étude des dynamiques urbaines dans la longue durée. *Cybergeog: European Journal of Geography*.
- Schiffer M. B. (1976). *Behavioral Archaeology*, Academic Press, New York.
- Studer R., Benjamins V. R. et Fensel D. (1998). Knowledge engineering: Principles and methods. *Data et Knowledge Engineering*, vol. 25, n° 1-2, p. 161-197.
- Tripier P. (2009). Sociologie et management éclairés par les arts de la guerre ? Les grammaires d'action. *Je connais la situation mais comment agir ? Savoir, savoir agir et agir*, F. Coninck (éd.), L'Harmattan, Paris, p. 27-51.
- Trombold C. D. (1991). An introduction to the study of ancient New World road networks. *Ancient Road Networks and Settlement Hierarchies in the New World*, C.D. Trombold (Ed.), Cambridge University Press, Cambridge, p. 1-9.
- Wallace A. D. (1993). *Walking, Literature, and English Culture: The Origins and Uses of Peripatetic in the Nineteenth Century*, Clarendon Press, Oxford.

Watson P. J., Leblanc S. A. et Redman C. L. (1971). *Explanation in Archaeology: An Explicitly Scientific Approach*, Columbia University Press, New York.

Williamson T. (2016). The ancient origins of medieval fields: A reassessment. *Archaeological Journal*, vol. 173, n° 2, p. 264-287.

Analyse des empreintes guidée par un modèle de connaissances pour la compréhension des dynamiques de navigation maritime côtière

Wissame Laddada¹, Éric Saux²

1. Laboratoire d'informatique, du traitement de l'information et des systèmes
INSA Rouen, av. de l'université, F-76800 Saint-Etienne-du-Rouvray
wissame.laddada@univ-rouen.fr

2. Institut de recherche de l'École navale, École navale, BCRM Brest,
Lanveoc-Poulmic, CC 600, F-29240 Brest Cedex 9
eric.saux@ecole-navale.fr

RÉSUMÉ. Les empreintes spatiales, historiques ou acquises en temps réel, aident à la compréhension des dynamiques de leurs porteurs. Dans un contexte de navigation maritime côtière, la recherche présentée montre la démarche de mise en place d'un modèle de connaissances visant à une analyse sémantique des empreintes de navires. Ce modèle est défini à partir des connaissances recueillies auprès d'experts du domaine complétées par des descriptions textuelles et graphiques. Après avoir décrit les principes de navigation, nous présentons les étapes menant à un tel modèle de connaissances. La recherche est complétée par la formalisation de règles portant sur la visibilité des points de repère. Enfin, un raisonnement conduit à inférer des connaissances supplémentaires de navigation sur lesquelles sont confrontées les empreintes de navigation afin d'identifier les routes ayant une sémantique pour le modèle et celles dénuées de sémantique.

ABSTRACT. The spatial footprints, historical or acquired in real time, help to understand the dynamics of their carriers. In a context of coastal maritime navigation, the research presented details the approach of setting up a knowledge model aimed at a semantic analysis of ship footprints. This model is defined on the basis of knowledge gathered from experts in this field, supplemented by textual and graphic descriptions. After having described these navigation principles, we present the stages leading to such a knowledge model. The research is completed by the formalization of rules relating to the visibility of landmarks. Finally, a reasoning leads to infer additional navigation knowledge on which the navigation footprints are confronted in order to identify the routes having semantics for the model and those devoid of semantics.

MOTS-CLÉS : logiques de description, navigation maritime côtière, ontologie, raisonnement.

KEYWORDS: Description logics, Coastal maritime navigation, Ontology, Reasoning.

DOI:10.3166/RIG.31.111-133 © 2022 Lavoisier

1. Introduction

Étudier et analyser les dynamiques d'entités, de processus ou d'êtres vivants à partir de l'analyse des leurs empreintes spatiales, pour en extraire des informations ou des connaissances, a été de tout temps d'un grand intérêt pour la compréhension des phénomènes naturels ou non naturels (cyclones, désertification, déplacements migratoires, urbanisation, etc.) et des liens qu'ils entretiennent avec leur environnement. Les domaines d'intérêt et les applications sont multiples. Dans le cadre de cet article, nous nous limitons à l'analyse des dynamiques spatiales à partir de modèle de connaissances dans un contexte de navigation maritime côtière pour une application de sécurité maritime.

En préambule, il est important de rappeler la terminologie utilisée dans cet article et usitée en représentation des connaissances. Une **donnée** est une description élémentaire d'une réalité : par exemple une observation ou une mesure effectuée. Deux catégories de données existent : les données brutes et les données (ou jeu de données ou data). Les données brutes n'ont été soumises à aucun traitement ou toute autre manipulation, alors que les données sont généralement le résultat d'un travail préalable sur les données brutes qui permettra de leur donner un sens. Les **informations** font références aux données porteuses de sens. Elles sont construites en fonction de règles et par interprétation des données d'un environnement précis. Enfin, une **connaissance** correspond à l'appropriation et l'interprétation des informations par des êtres humains. Elle peut être formalisée pour être utilisable à des fins opératoires. Le concept de connaissance fait appel aux questions de sens portées notamment par les notions telles que le langage, la sémantique, les croyances, la conscience. Pour passer des données à de l'information et enfin à une connaissance, il faut effectuer au préalable un traitement cognitif de ces données. Définir un **modèle de connaissances** revient donc à stocker les connaissances de sorte que des programmes puissent la traiter. Ceci sous-entend de i) définir des modèles formels permettant de représenter les connaissances, ii) traiter des informations au niveau sémantique et iii) simuler le raisonnement humain par le biais de mécanismes d'inférence. Les objectifs d'un modèle de connaissances sont donc de partager du savoir et de raisonner sur du savoir et ainsi permettre des analyses. C'est dans ce cadre que se situe cette recherche.

En modélisation des connaissances, deux approches peuvent être distinguées et être considérées comme complémentaires. La première approche de modélisation place l'humain au cœur du modèle (« *human driven* »). Ainsi, ses connaissances contenues dans des descriptions graphiques et linguistiques (textuelles et orales), issues d'une extériorisation de ses représentations mentales, elles-mêmes provenant d'une perception des environnements, sont à la source du modèle de connaissances. Cette approche est fortement cognitive et place l'humain à l'origine des connaissances du modèle. En conséquence, ce modèle est plus « proche » des attendus car il modélise des situations et des faits environnementaux qui sont connus et issus de l'homme. La seconde approche consiste à analyser et interpréter les données ou les empreintes (cf. article Fargette et *al.* de ce numéro thématique) afin d'en extraire une sémantique à modéliser. Cette seconde modélisation fait donc émerger une sémantique à partir des données

(« *data driven* ») et prend en considération les différents motifs qui émergent dans ces données. Ces deux modélisations sont complémentaires dans la mesure où la seconde permet de découvrir des inattendus, qui restent à expliquer. Dans cet article, nous montrons comment un modèle de connaissances, construit à partir des connaissances détenues par des experts et contenues dans des ouvrages, peut aider à la compréhension des dynamiques spatiales dans un contexte de navigation maritime côtière.

Dans cet article, nous abordons les différentes étapes entrant dans la définition d'un modèle de connaissances pour la navigation maritime côtière. L'objectif est ici de retranscrire la connaissance des experts en matière de définition de trajectoire maritime côtière au sein d'un système d'information pour que ce dernier puisse inférer des routes de navigation à partir des faits qui sont présents dans la base de connaissances. Deux stratégies d'exploitation des routes inférées peuvent être envisagées. L'une concerne leur exploitation au sein d'un système de routage pour proposer au navigateur des routes de navigation à suivre, l'autre concerne leur exploitation à des fins d'analyse des empreintes de positionnement de bateaux afin de capturer les portions ayant une sémantique pour le modèle et celles dépourvues de sémantique et sur lesquelles une analyse doit être approfondie. Nous optons pour cette seconde stratégie d'exploitation du modèle de connaissances.

Dans la section 2 de cet article, nous nous intéressons à l'extraction des connaissances de l'humain et à leur transcription dans un langage formel. Nous abordons l'étape de conceptualisation du domaine à partir de descriptions graphiques et textuelles réalisées par des experts et à leur formalisation en logique de description. Bien plus que de définir les concepts, nous montrons que ces descriptions peuvent aussi contribuer au peuplement de la base de connaissances. Dans la section 3, nous présentons sa modélisation ontologique en nous fondant sur le langage de représentation des connaissances OWL (*Web Ontology Language*) et le standard GeoSPARQL pour la modélisation de la dimension spatiale. Nous en déclinons les axiomes de notre ontologie à la base du raisonnement axiomatique. Afin d'enrichir le raisonnement, en prenant notamment en compte le processus de visibilité des repères liés à la navigation maritime, nous introduisons dans la partie 4 un raisonnement complémentaire fondé sur des règles SWRL. Enfin, la section 5 est consacrée à un premier niveau d'analyse sémantique pouvant être mené à partir du modèle de connaissances proposé. Nous concluons ensuite cet article et énonçons les perspectives envisagées et la complémentarité des approches numériques pour l'analyse sémantique.

2. De l'humain à la formalisation de ses connaissances

2.1. Extraction de la sémantique à partir de descriptions linguistiques et graphiques de navigation

Les recherches en théorie de l'information spatiale s'attachent depuis une cinquantaine d'années à l'étude des processus de navigation dans des environnements différents (urbain vs. naturel, fortement structuré vs. peu structuré) et à leur modélisation au sein de systèmes spatialisés. Modéliser, analyser et simuler des comportements

humains sont en effet des fonctions essentielles pour une meilleure compréhension de nos environnements et de nos interactions avec ceux-ci. Les descriptions linguistiques (écrites ou orales), graphiques ou mixtes (Przytula-Machrouh, 2004) constituent des données essentielles afin d'observer puis reproduire le plus fidèlement possible ces connaissances au sein d'un système d'information. Chacune de ces formes présente des particularités et des pouvoirs d'expression qui lui sont propres. Le mode linguistique est bien adapté à la description d'itinéraire ou d'environnement dans le sens où la linéarité et la chronologie d'un itinéraire se retrouvent dans le texte ou le discours. Dans ce genre de descriptions, les relations spatiales sont fortement usitées car elles permettent une meilleure appréhension de la spatialité par le positionnement des entités entre elles ou en lien avec l'itinéraire décrit. Les descriptions textuelles sont généralement plus riches et contiennent plus de descriptions d'entités que le mode oral plus simple. Les constructions linguistiques sont naturellement plus élaborées dans les textes que dans le langage oral. Le mode graphique a, quant à lui, le pouvoir de mieux inférer des relations spatiales sur les entités présentes et d'exprimer les différents déplacements à effectuer par rapport à ces entités. L'ensemble des entités statiques décrites est composé uniquement des entités les plus remarquables ou saillantes de l'environnement. Notons que la représentation graphique d'une dynamique de déplacement prend des formes relativement moins variées que celles des entités car elle se fait généralement à l'aide de flèches. Une simple flèche permet de définir le sens, l'orientation et parfois la taille (durée, distance) d'une action de déplacement.

La démarche adoptée pour aborder la problématique de modélisation des connaissances de la navigation maritime côtière a été d'analyser les descriptions textuelles et graphiques de tels déplacements. Ce type de navigation est appelé « navigation à vue » et est essentiellement fondée sur la notion de point de repère visuel ou « *landmark* », et est de ce fait en adéquation avec la problématique de « *wayfinding* »¹. La figure 1 illustre quelques exemples de descriptions graphiques. À gauche, le schéma présente les différents itinéraires pouvant être empruntés dans cette zone géographique côtière et est accompagné de descriptions textuelles (figure 2) complétant ce schéma. À droite, le schéma illustre un exemple théorique de préparation de route de navigation relative aux directives générales pour la conduite nautique (Marine nationale, 2005). Concernant les descriptions textuelles, les ouvrages du SHOM intitulés « Instructions nautiques » (IN) (SHOM, 2018) ou les guides de navigation tels que les « Pilotes côtiers » (Voiles et voiliers, 2006) (figure 2) contiennent des informations d'une grande richesse sur l'environnement maritime et côtier, souvent complémentaires aux descriptions graphiques ou cartographiques, pour répondre aux besoins des navigateurs professionnels, tout en fournissant aux plaisanciers l'information qui leur est nécessaire. Les descriptions des Pilotes côtiers, plus en lien à la navigation à vue, ont l'avantage – par rapport aux IN – de décrire la chronologie d'un itinéraire dans différents contextes (jour/nuit, bonne/mauvaise visibilité ou conditions météorologiques)

1. « Le wayfinding décrit le processus pour effectuer une trajectoire englobant optimisation et raisonnement spatio-temporels tout en prenant en considération notre interaction avec l'environnement qui nous entoure et influence notre itinéraire. », extrait de (Laddada, 2018).

et de décrire les relations spatiales entre différentes entités ce qui en fait notre source de connaissances prioritaire. Moins contemporains, les nombreux ouvrages littéraires anciens sont aussi une source d'information. Certains relatent les principes de navigation à l'aide de connaissances scientifiques et quantitatives d'Eratosthène ou selon la géographie de Strabon dans laquelle les descriptions à la fois cartographique et textuelle (cf. les anciennes cartes nautiques dites « portulans ») servent de support aux marins grecs qui corrélaient la perception visuelle des repères terrestres aux descriptions objectives et intègres de la carte afin de confirmer – ou non – leur positionnement géographique et la validité de leur itinéraire (Kowalski, 2009).

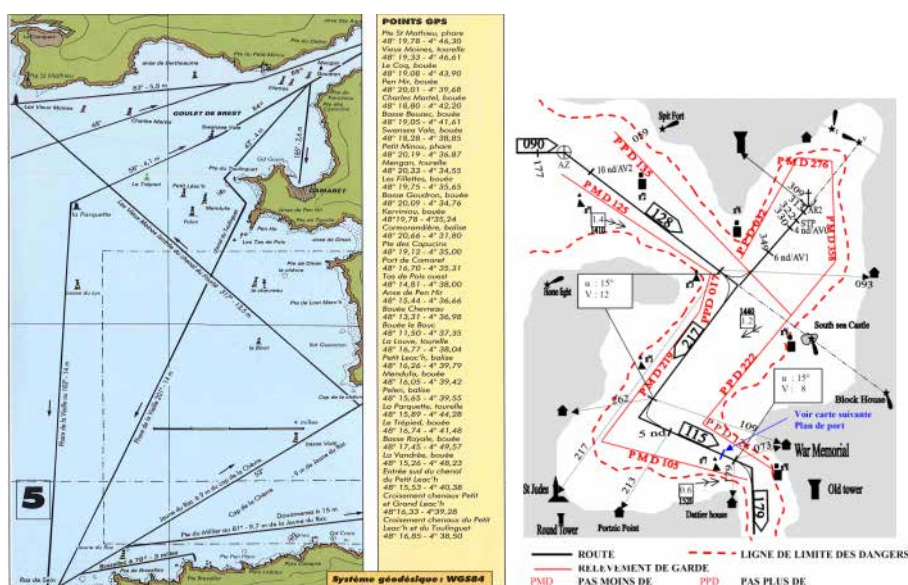


Figure 1. Exemples de routes de navigation : cartographie extraite des *Pilotes côtiers* (à gauche), préparation de routes de navigation suivant les directives générales pour la conduite nautique (à droite)

En venant du chenal du Four, on peut virer la pointe St-Mathieu au plus près en laissant sur bâbord le feu rouge (éclat 4 sec.) visible de 280° à 113° de la tourelle des **Vieux Moines**, puis le feu rouge (4 éclats 15 sec.) de la bouée à sifflet **Charles Martel** que rase l'alignement lumineux à 68°.

En se rapprochant du feu du Petit Minou, on s'écartera de cet alignement à 68° pour laisser à petite distance sur tribord le feu blanc scintillant rapide (9) 10 sec. de la bouée à sifflet du plateau des **Fillettes** et le feu également blanc (3 éclats 12 sec.) de la tourelle de **Mengon**. On notera que ce dernier feu couvre d'un étroit secteur rouge (34° à 54°) la bouée de la **basse Goudron** qui montre un feu rouge scintillant.

En passant à la hauteur du feu de **Portzic**, on peut faire route directement vers les feux scintillants rouge et vert balisant la passe d'entrée sud de la rade abri du port de Brest.

Figure 2. Texte extrait des *Pilotes côtiers* et décrivant l'entrée de nuit dans le goulet de Brest. Ce texte est complété par la description graphique de la figure 1 (à gauche)

L'analyse des descriptions linguistiques textuelles et graphiques nous a permis d'identifier les entités impliquées dans la recherche d'itinéraire dans le contexte de navigation maritime côtière ainsi que les dépendances spatiales et temporelles existantes. Dans l'étape de conceptualisation, nous définissons les concepts de Trajectoire, Route, Amer, WayPoint, PointDeDépart, PointD'Arrivée ainsi que les relations entre deux concepts ou entre un concept et une valeur littérale. Dans ce contexte de navigation, nous montrons l'influence de la dimension temporelle (de jour, de nuit) ainsi que l'influence de la dimension spatiale par l'utilisation des primitives géométriques de base (point, ligne, polygone, ensemble de points, de lignes ou de polygones) et des relations topologiques, d'orientation et de distance. Dans cette analyse, nous montrons aussi l'importance de la modélisation sémantique de la visibilité – d'un ou des amer(s)² –, cette dernière étant soumise à la fois aux influences spatiales (problème d'inter-visibilité) et aux influences temporelles (présence d'un feu sur l'amer de nuit). Suite à cette première analyse, il en résulte l'identification des concepts et relations nécessaires à une modélisation d'itinéraire. Cette étude a été complétée par une étape de formalisation (cf. partie 2.2) visant à raisonner sur ces connaissances afin d'offrir au système des capacités à inférer de nouvelles connaissances (cf. partie 3).

Cette première étape d'analyse des descriptions linguistiques textuelles et graphiques a été, par la suite, complétée par une étape d'extraction des entités spatiales nommées (ESN) à partir d'outils de traitement automatique des langues (TAL). L'objectif est ici d'extraire automatiquement les instances des concepts identifiés lors de l'étape de conceptualisation, ceux-ci étant nécessaires pour le peuplement de l'ontologie résultante. En lien avec l'étude proposée par Lamotte (Lamotte *et al.*, 2019) et qui montre que la particularité des textes des IN (SHOM, 2018) mettent à défaut les approches usuelles en TAL telles que les approches proposées par Perdido³ et CasEN⁴, Lamotte met en place sa propre chaîne d'extraction des ESN à partir de l'outil Unitex⁵. Elle reprend le principe proposé par les auteurs et fondé sur les deux cascades de transducteurs « analyse » et « synthèse » telles que développées dans (Maurel *et al.*, 2011), mais elle définit des graphes d'analyse spécifiques pour annoter les ESN de l'environnement maritime et côtier nécessaires à la navigation en vue de terre. Un exemple résultat est proposé dans la figure 3, dans laquelle sont annotés les entités maritimes et côtières (typées et nommées), les marques de navigation (typées, nommées et géoréférencées) ainsi que les alignements.

2. Objet prééminent à un emplacement fixe qui peut être utilisé pour déterminer un emplacement ou une direction (définition extraite de la norme S-57 de l'Organisation Hydrographique Internationale).

3. Service Web développé par L. Moncla lors de sa thèse (Moncla, 2015). <http://erig.univ-pau.fr/PERDIDO/api.jsp>

4. Outil développé au Laboratoire d'Informatique Fondamentale et Appliquée de Tours (LIFAT). <https://tln.lifat.univ-tours.fr/version-francaise/ressources/casen>

5. Outil proposé par le Laboratoire d'Informatique Gaspard-Monge. <https://unitexgramlab.org/fr>

La ligne joignant la `<maritimeEntity><name><type>pointe</type>` du Toulinguet`</name></maritimeEntity>` au sud à la `<maritimeEntity><name><type>bouée</type>` Charles Martel`</name></maritimeEntity>`, en lisière de l'`<leadingLine type="avant" subtype="name">alignement à Mesure type="orientation">68</Mesure>` du `<name>grand chenal</name></leadingLine>` à 3,8 milles avant le `<maritimeEntity><name><type>phare</type>` du Petit Minou`</name></maritimeEntity>`, marque le début de la zone de pilotage où les navires de plus de 50 m ont une totale priorité de manœuvre. Une veille doit en outre être assurée sur canal 16 en VHF.

À l'approche du goulet, le `<maritimeEntity><name><type>phare</type>` de la `<maritimeEntity><name><type>pointe</type>` du Petit Minou`</name></maritimeEntity></name></maritimeEntity>` accolé à une tour blanche donne avec le `<maritimeEntity><name><type>phare</type>` de Portzic`</name></maritimeEntity>`, une tour élanée sur la pointe à gauche de l'entrée dans la rade, l'`<leadingLine type="avant" subtype="name">alignement à Mesure type="orientation">68</Mesure>` du `<name>grand chenal</name></leadingLine>`. De jour, il n'est pas nécessaire de suivre exactement cet alignement car il ne manque pas d'eau à courir pour un voilier.

POINTS GPS :

```
<navigationMark type="bouée"><name>Charles Martel</name>, bouée<location><lat>48
18,80</lat> <lon>4° 42,20</lon></location></navigationMark>
<navigationMark type="phare"><name>Petit Minou</name>, phare<location><lat>48° 20,19</lat>
<lon>4° 36,87</lon></location></navigationMark>
```

Figure 3. Annotation automatique des entités nécessaires à la navigation maritime côtière : entités maritimes et côtières (en gris clair), marques de navigation géoréférencées (en gris foncé) et alignements (en gris moyen)

L'analyse des descriptions linguistiques textuelles a permis également d'extraire la sémantique d'un concept liée à son affordance⁶ (Tsatcha *et al.*, 2012). La technique proposée permettant d'extraire – à partir de données textuelles nautiques – des connaissances liées à l'affordance d'un concept est fondée sur le principe du vecteur conceptuel d'un mot. Le vecteur conceptuel d'un mot s'obtient à partir des synsets candidats liés à ce mot. Ce vecteur conceptuel de mot a permis de construire le vecteur conceptuel d'une phrase exprimé comme une somme directe des vecteurs conceptuels des différents mots qui la compose. Les phrases utilisées sont celles qui décrivent le concept dans le corpus. Une fois les vecteurs conceptuels de ces phrases déterminés, ceux-ci sont projetés dans un espace de décision formé des vecteurs conceptuels des affordances, par exemple les affordances « manœuvrer, stop, continue, back » – dans les cas traités lors de l'étude – afin de savoir si un bateau doit manœuvrer, stopper, continuer ou faire machine arrière face un objet de l'environnement maritime. Ces affordances sont couramment utilisées par les marins pour modifier la trajectoire du navire. L'espace de décision fait alors émerger une des affordances. Cette possibilité de déterminer l'affordance procure au modèle la capacité de pouvoir fabriquer ou de déduire lui-même sa propre connaissance et ainsi de prendre des décisions raisonnables en toute autonomie. L'affordance ainsi produite est désignée par une propriété (hasAffordance) qui fait partie de l'une des propriétés des états d'une entité définie dans l'ontologie.

6. L'affordance, néologisme proposé par le psychologue américain James Jerome Gibson, traduit fidèlement cette faculté de l'homme, et de l'animal en général, à guider ses comportements en percevant ce que l'environnement lui offre en termes de potentialités d'actions.

2.2. *De la conceptualisation des connaissances à leur formalisation en logique de description*

Nous avons évoqué dans les parties précédentes, l'importance d'analyser des descriptions représentant le processus de navigation dans un environnement donné pour en extraire de la connaissance. Dans ce contexte, le modèle résulte de différentes étapes d'analyse pour une conceptualisation des entités de navigation. Une fois la conceptualisation réalisée, une formalisation des concepts est proposée afin d'offrir au modèle une capacité de raisonnement.

2.2.1. *Conceptualisation des entités de navigation*

Dans l'optique de planifier une trajectoire de navigation en milieu maritime, il faut tenir compte des concepts sur lesquels se fonde la navigation dans cet environnement. Nous nous sommes concentrés sur la navigation côtière (c.-à-d. en vue de terre) fondée sur les repères ou amers. Ces derniers sont exploités pour que les marins puissent suivre un chemin sécurisé jusqu'à leur destination finale. Afin d'explicitier les notions adoptées dans une navigation, nous nous sommes fondés sur les « Instructions nautiques » du SHOM ou les guides de navigation tels que « Pilote côtier » (cf. partie 2.1) mais aussi sur des exemples de descriptions textuelles de marins expérimentés qui représentent une préparation de traversée, dont un extrait est présenté ci-après :

Description 1 : Départ du port de Brest. Route au 254 pour 3.13 Nq. La route est sur un relèvement constant par l'avant sur le phare du Petit Minou.

Description 2 : Route au 191 pour 2.2 Nq. sur l'alignement du phare du Petit Minou par le rocher en forme de chaise de la pointe du Toulanguet.

Etc.

Dans le but d'atteindre une destination, le navigateur doit suivre une trajectoire qui est constituée d'une séquence de routes maritimes. Chacune de ces routes suit un cap tracé sur un repère donné. Ce dernier doit ainsi être visible à partir du bateau. L'objectif premier consiste donc à identifier les principes et les termes employés – dans les descriptions textuelles (cf. partie 2.1) – afin de les intégrer dans une conceptualisation telle que présentée dans ce qui suit.

Une **trajectoire (Trajectory)** est constituée d'une séquence de **routes** successives.

Une **route (Route)** correspond à un **segment** orienté sur un **cap**. Celui-ci est à 254° dans la *Description 1* et à 191° dans la *Description 2*. Le bateau suit ce cap durant une période de temps correspondant à la durée de la traversée sur cette route.

Un **WayPoint** représente un point clé de la trajectoire. Il peut décrire un **point de départ**, un **point d'arrivée** ou un **point de décision**. Chaque **point de décision** de la trajectoire marque le début de la prochaine route et la fin de la route courante. Un point de décision marque le point où un navigateur décide de changer de **cap** et donc de changer de route. Le cap désigne donc l'orientation du navire exprimée en degré. C'est notamment à partir d'un point de décision que les notions de visibilité et de perception des **repères** de l'environnement (amers) jouent un rôle important.

Un **Amer (NavigationMark)** représente un **point** de repère et d'orientation pour fixer un **cap** et définir ainsi une **route**. Un repère a certaines propriétés dues à sa **saillance visuelle** qui peut le distinguer des autres éléments de l'environnement. Comme exemple d'amers, on peut citer les phares, les balises, les roches, etc. Dans les descriptions précédentes, les amers correspondent au « *phare du Petit Minou* » et au « *rocher en forme de chaise de la pointe du Toulinguet* ».

Un **relèvement (Bearing)** correspond à une orientation sur un **amer**, ce qui définit un **angle** formé à partir du bateau entre cet amer et la direction du Nord géographique. Dans la *Description 1*, la **route** suit un **relèvement** à 254° (par l'avant du bateau) sur le phare du Petit Minou.

Un **alignement (LeadingLine)**, comme son nom l'indique, est un concept qui aligne au moins deux amers. Ce concept est représenté par une **ligne** qui forme un angle avec le Nord géographique. Dans la *Description 2*, la **route** suit un **alignement** de 191° formé par le phare du Petit Minou et le rocher en forme de chaise de la pointe du Toulinguet.

Cette première étape, nous permet d'identifier les entités et les dépendances existantes entre elles afin de faciliter la modélisation ontologique qui fait l'objet de la deuxième étape (cf. partie 3). Entre ces deux étapes, une formalisation est nécessaire afin de déterminer l'émergence des dépendances. Suite à une analyse graphique des concepts, la dimension spatiale est incluse dans la formalisation ainsi que la notion sémantique liée à la visibilité. Ceci offre la possibilité d'avoir une modélisation formelle qualitative qui favorisera le mécanisme de raisonnement.

2.2.2. Formalisation des concepts

Nous avons choisi de formaliser les concepts et relations à partir des logiques de description (DL). Les raisons sont multiples. Tout d'abord, et d'après Baader (Baader *et al.*, 2010), les systèmes fondés sur les logiques de description offrent un service de raisonnement sur les prédicats définis (concepts et relations). Ainsi dans l'étape de modélisation ontologique (cf. partie 3), nous pourrions interpréter nos expressions sémantiques avec le langage ontologique OWL (*Ontology Web Language*) – ce dernier étant lui-même fondé sur DL – et ainsi mettre en œuvre un raisonnement axiomatique. L'intérêt d'une formalisation en DL est aussi de pouvoir déployer un modèle avec deux niveaux d'inférence : i) un raisonnement axiomatique, et ii) un raisonnement à base de règles. En effet, il est possible d'ajouter des règles à notre système afin d'enrichir le raisonnement et la sémantique sur les instances de notre base de connaissances. Ces règles seront formalisées en SWRL (langage de règles pour le web sémantique). Nous présentons dans le tableau 1 (Laddada, 2018), les différents concepts formalisés constituant les entités du raisonnement axiomatique. Le raisonnement à base de règles sera quant à lui détaillé dans la partie 4.2.

Dans cette formalisation, les concepts atomiques désignent les concepts qui n'ont pas de définition axiomatique complexe mais une simple définition hiérarchique. A ce titre, le concept *NavigationMark* représente un amer et est désigné par un point

Tableau 1. Formalisation des concepts

Concepts atomiques	Concepts axiomatiques en DL
$OriginPoint \sqsubseteq Point$	$WayPoint \equiv Point \sqcap (DeparturePoint \sqcup DecisionPoint \sqcup DestinationPoint)$
$DeparturePoint \sqsubseteq OriginPoint$	$VisibilityLine \equiv$ $Line \sqcap \exists contains. OriginPoint \sqcap$ $\exists Contains. (NavigationMark \sqcap$ $\exists visibleFrom. OriginPoint) \sqcap$ $\exists hasBAngle. Angle$
$DestinationPoint \sqsubseteq OriginPoint$	
$DecisionPoint \sqsubseteq OriginPoint$	$Bearing \equiv Line \sqcap \exists equals. (VisibilityLine \sqcap$ $= 1 contains. NavigationMark)$
$NavigationMark \sqsubseteq Point$	$LeadingLine \sqsubseteq Line \sqcap$ $\geq 2 contains. NavigationMark$
$Horizon \sqsubseteq Circle$	$DecisionPoint \equiv Point \sqcap$ $\exists within. ((VisibilityLine \sqcap LeadingLine) \sqcup$ $(LeadingLine \sqcap LeadingLine) \sqcup$ $(VisibilityLine \sqcap VisibilityLine))$
$VisibilityLightDistance \sqsubseteq Circle$	$Route \equiv Line \sqcap$ $\geq 2 contains. WayPoint \sqcap$ $\exists within. (Bearing \sqcup LeadingLine)$
$VisibilityWeather \sqsubseteq Circle$	$Trajectory \equiv LineString \sqcap$ $\exists Contains. DeparturePoint \sqcap$ $\exists Contains. DestinationPoint \sqcap$ $\geq 1 Contains. Route$

dans un espace 2D grâce aux coordonnées géographiques (latitude, longitude). Un amer est un point de repère fixe et identifiable sans ambiguïté pour la navigation maritime. Ainsi, pour chaque amer, lui est associé une zone de visibilité dans l'espace : visibilité à l'horizon représentée par le concept *Horizon*, visibilité de son feu – s'il y a – définie par le concept *VisibilityLightDistance* et visibilité selon les conditions météorologiques définie par le concept *VisibilityWeather*. Chacun de ces concepts est représenté géométriquement par un cercle. La vérification de la visibilité d'un amer est réalisée à partir d'un point d'observation défini dans notre modèle par le concept *OriginPoint*. Ce dernier inclut le point de départ (*DeparturePoint*), le point d'arrivée (*DestinationPoint*) et les points de décision (*DecisionPoint*) qui constituent les points de réorientation dans la trajectoire.

3. Modélisation ontologique pour le raisonnement

Le principe de l'approche ontologique, sur lequel nous nous appuyons, est fondé sur le fait que chaque concept impliqué dans la représentation d'une trajectoire est associé à une primitive géométrique comme défini dans la formalisation ci-avant. Ainsi, l'ontologie doit prendre en considération la dimension spatiale et les concepts définis pour une trajectoire. Pour ce faire, nous interprétons en OWL tous les concepts listés précédemment pour la définition formelle d'une trajectoire de navigation maritime côtière et de ses composants. Dans la suite de cette partie, nous présentons la modélisation ontologique que nous proposons en nous fondant sur des ontologies publiées, notamment pour la modélisation de la spatialité. La modélisation globale est donc

constituée de deux ontologies principales ; l'une est définie par nos soins et s'appuie sur une autre modélisation spatiale existante fondée sur le standard de GeoSPARQL. Nous complétons ensuite cette modélisation initiale afin de prendre en compte la modélisation ontologique du concept de visibilité. Par le biais de cette modélisation ontologique, nous abordons les questions liées à l'inférence et au fait que le système puisse déduire de nouvelles connaissances à partir de connaissances initiales. Cependant un raisonnement axiomatique, à lui seul, n'est pas suffisant pour inférer l'ensemble des concepts et notamment les concepts de visibilité. C'est la raison pour laquelle nous enrichissons la sémantique initiale avec des règles SWRL (*Semantic Web Rule Language*).

3.1. Modélisation ontologique de la dimension spatiale

Afin de construire notre ontologie, nous reprenons un modèle ontologique existant pour la représentation de la dimension spatiale à savoir, l'ontologie fondée sur le standard GeoSPARQL⁷. C'est un standard de représentation et d'interrogation des données géospatiales dans le contexte du web sémantique et des données liées. Ce modèle est organisé en trois classes principales : *Geometry*, *Feature* et *SpatialObject* (figure 4). La classe *SpatialObject* est décrite par les deux classes *Geometry* et *Feature*. La classe *Feature* représente une entité spatiale dans un environnement donné, par exemple un phare. La classe *Geometry* recense toutes les formes géométriques définies par l'*Open Geospatial Consortium* (OGC), dont les points, les polygones et les polygones.

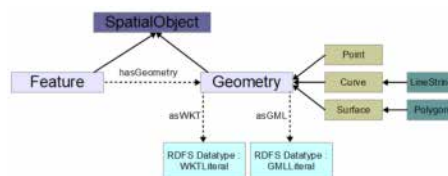
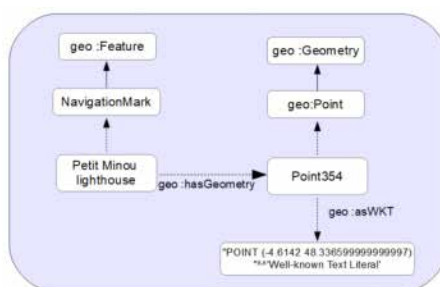


Figure 4. Modèle ontologique du standard GeoSPARQL

Les relations dans ce modèle comprennent principalement le vocabulaire des relations topologiques définies dans les modèles 9-intersections (Egenhofer, Herring, 1990) ou RCC8 (Randell *et al.*, 1992) ainsi que d'autres relations telles que *hasGeometry*, qui relie la classe *Feature* à la classe *Geometry*, ou la relation des données *hasSerealization*. Cette dernière réunit les deux sous-relations de données développées par l'OGC : *asWKT* (*Well-Known Text*) et *asGML* (*Geography Markup Language*). La première sous-relation *asWKT* est relative au format standard utilisé afin de décrire les objets géométriques vectoriels et les systèmes de coordonnées. La seconde sous-relation *asGML* permet de présenter et de manipuler des informations géographiques dans un langage XML. Dans la perspective d'obtenir une modélisation onto-

7. <http://www.opengis.net/ont/geosparql#>

En navigation maritime côtière, une trajectoire de navigation suit une séquence de repères définis comme étant des instances de *Point* dans la formalisation. Selon le modèle de GeoSPARQL, un repère (*NavigationMark*) est une entité spatiale qui doit être une sous-classe de *Feature*. Ainsi, nous interprétons le fait qu'un repère soit un *Point* par la relation *hasGeometry* qui lie une entité spatiale à un point. Plus précisément, un objet spatial dans le monde réel est associé à une géométrie dont les coordonnées géographiques sont décrites par le format WKT. Comme le décrit la figure 5, le « *Phare du Petit Minou* » est de type *NavigationMark* qui est une sous-classe de *Feature*. Ce phare a un point comme forme géométrique qui est décrite par l'instance *Point354*. Ainsi, le « *Phare Petit Minou* » et la géométrie *Point354* sont deux individus reliés par la relation *hasGeometry*. L'individu *Point354* a ses coordonnées géographiques définies par un littéral qui est l'objet de la relation *asWKT*.



Le même principe de modélisation est appliqué à chaque concept d'une trajectoire afin de lui associer une dimension spatiale (figure 6) (Laddada, 2018). Ainsi :

- Afin de modéliser la notion d'orientation pour représenter un cap ou un angle de relèvement, la classe *Angle* est définie. Un angle est lié à la ligne décrivant un relèvement

(*Bearing*) par la relation *hasBAngle*. Une propriété de données nommée *hasAngleValue* de type *Float* est introduite pour représenter la valeur de l'angle d'orientation.

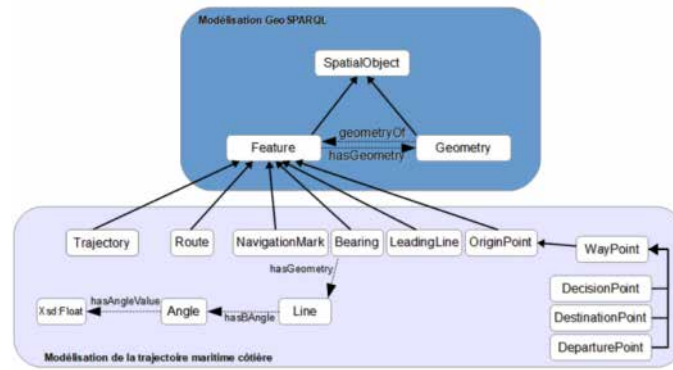


Figure 6. Vue partielle du modèle ontologique de navigation maritime côtière et de ses composants

3.3. Raisonnement axiomatique

Afin de peupler l'ontologie, nous avons extrait automatiquement (cf. figure 3) et manuellement des connaissances à partir des données quantitatives des cartes maritimes du SHOM au format S-57⁸ et des données qualitatives textuelles (Instructions nautiques du SHOM, Pilotes côtiers, etc.). Dans notre ontologie, nous avons fait le choix d'instancier manuellement les alignements et d'instancier automatiquement les *DecisionPoint* à partir d'un algorithme de subdivision de l'espace maritime. Un exemple d'instanciation du modèle à partir de la *Description 1* « Départ du port de Brest. Route au 254 pour 3.13 Nq. La route est sur un relèvement constant par l'avant sur le phare du Petit Minou. » peut être la suivante. Un *Feature*, *Feature1*, est défini dans l'ontologie avec une géométrie, *Line1*. Cette ligne est associée à *Feature1* grâce à la relation *hasGeometry*. Cette ligne est formée par deux points, *PointA* et *Point562*, représentant respectivement les géométries du *Port de Brest* (de type *DeparturePoint*) et du premier point de décision, *DecisionPoint1*. De plus, la relation topologique *contains* associe *Line1* à ces deux points. L'amer, *Petit minou*, de type *NavigationMark* est également défini avec sa géométrie *Point364*. Par ailleurs, une autre ligne, *Line6*, est déterminée décrivant la ligne de visibilité, *VisibleLine1* (*VisibleLine1 hasGeometry Line6*). Cette dernière ligne est formée par le point de départ (*PointA*) et le point *Point364*. La relation spatiale entre la ligne et les deux points *PointA* et *Point364* est introduite dans l'ontologie par la relation topologique *contains*. Ensuite, nous identifions la relation topologique *within* entre les lignes *Line1* et *Line6*.

8. <http://www.s-57.com/>

Le model formel, que nous avons proposé dans le tableau 1, est interprété en s’aidant de la syntaxe Manchester OWL (Horridge *et al.*, 2006). Nous décrivons dans le tableau 2 les axiomes que nous définissons dans l’ontologie grâce à l’éditeur d’ontologie Protégé. En tenant compte des formalisations listées dans le tableau 2, nous remarquons que le concept *LeadingLine* a une formalisation que nous n’adaptions pas dans notre application ontologique (pas de définition en Manchester OWL). Il en est de même pour le concept *DecisionPoint*. La non-application de ces deux formalisations est relative au raisonnement. Les individus instanciant le concept **LeadingLine** ne sont pas inférés mais peuplés manuellement, alors que l’instanciation du concept *DecisionPoint* est automatique à partir d’un processus de subdivision de l’espace.

Au-delà de l’intérêt d’utiliser des ontologies pour la structuration et le partage de données en utilisant des vocabulaires communs, les ontologies informatiques répondent aussi à des besoins de raisonnement. Le modèle ainsi proposé et interprété dans le tableau 2 répond alors au besoin de proposer des routes de navigation maritime qui respectent la sémantique du modèle. Il est par conséquent apte à inférer des *waypoints*, des relèvements, des routes et des trajectoires.

Tableau 2. Axiomes de l’ontologie en syntaxe Manchester OWL

<i>VisibleLine EquivalentTo Feature and (hasGeometry some (Line and (contains some (Point and (geometryOf some NavigationMark)))) and (contains some (Point and (geometryOf some OriginPoint)))) and (hasBAngle some Angle)</i>
<i>Bearing EquivalentTo VisibilityLine and (hasGeometry some (Line and (contains max 1 (Point and (geometryOf only NavigationMark)))) and (visibilityStm value true)</i>
<i>Route EquivalentTo Feature and (hasGeometry some (Line and (within some (Line and (geometryOf some (Bearing or LeadingLine)))) and (contains min 2 (Point and (geometryOf some WayPoint))))</i>
<i>Trajectory EquivalentTo Feature and (hasGeometry some('Line String' and ((contains some (Line and (geometryOf min 1 Route))) and (contains some (Point and (geometryOf some DeparturePoint))) and (contains some (Point and (geometryOf some DestinationPoint))))))</i>

Notons que dans la formalisation (cf. tableau 1), les éléments d’une trajectoire sont des sous-concepts d’une géométrie. En théorie, il suffit donc d’appliquer les relations topologiques directement sur ces éléments. En revanche, la modélisation ontologique est fondée sur le standard de GeoSPARQL dans lequel les éléments d’une trajectoire ne sont pas des sous-classes d’une géométrie mais sont reliés à une géométrie par la relation *hasGeometry*. Afin d’appliquer les relations topologiques sur les éléments, nous devons donc nous référer à leur description géométrique grâce à la propriété *hasGeometry*. Le raisonnement s’appliquant explicitement sur la forme géométrique, il est impératif pour le mécanisme de raisonnement de savoir à quel élément de la trajectoire la géométrie fait référence. Il en résulte l’importance de la définition de la relation *geometryOf*, inverse de la relation *hasGeometry*. Cette relation est définie pour faciliter l’annotation de la formalisation vers l’ontologie, mais il est tout à fait possible d’utiliser aussi la relation inverse *hasGeometry*⁻¹.

4. Enrichissement sémantique relatif à la visibilité des repères

Dans cette partie, nous abordons la problématique relative à la sélection des amers grâce à des critères de visibilité et non pas grâce à des critères de saillances visuelles. En d'autres termes, la perspective ici est de sélectionner les repères visibles à partir d'un point de vue (celui du bateau ou du marin) et qui contribue à la définition d'une route sans danger et ainsi à la génération d'une trajectoire à suivre. En revanche, les questions liées à la saillance des repères – choix d'un repère ayant une forme ou une couleur différente de l'environnement dans lequel il se trouve – ne sont pas abordées ici. Dans notre étude, les repères (*NavigationMark*) sont connus grâce à la couche thématique des « *Landmarks* » – ayant pour acronyme LNDMRK – présente dans les données spatiales au format S-57. L'utilisateur peut compléter cette liste par ses propres repères (avec leur coordonnées géographiques) issus de son expérience. Afin que le système puisse sélectionner des repères, il est nécessaire d'étendre le modèle ontologique défini précédemment afin de prendre en compte la notion de visibilité d'un repère à partir d'un point d'observation.

4.1. Modélisation ontologique de la visibilité des repères

L'idée élaborée est d'inférer des zones de visibilité pour des amers à partir d'un point d'observation en prenant en compte les paramètres influençant la visibilité, c'est-à-dire la rotondité de la terre, les conditions météorologiques (bonne ou mauvaise visibilité) et les différents moments de la journée (visualisation de jour ou de nuit). Ces zones de visibilité sont décrites comme des cercles dans la formalisation (cf. tableau 1) par les concepts *Horizon*, *VisibilityWeather* et *VisibilityLightDistance* (respectivement *VHorizon*, *VWeather* et *VLight* dans le modèle ontologique). Chaque cercle, dont son centre est associé à un amer, est délimité par un rayon définissant la distance à laquelle le repère est visible.

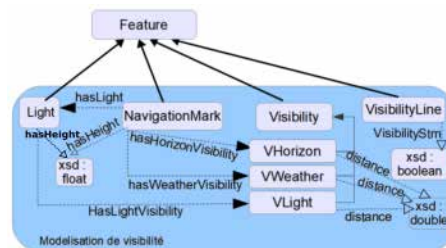


Figure 7. Enrichissement de l'ontologie pour la prise en compte de la notion de visibilité

Chacun de ces trois concepts est indirectement une sous-classe de l'entité *Feature* de GeoSPARQL. Plus précisément, une classe nommée *Visibility* est définie comme une sous-classe directe de *Feature* et inclut les trois concepts de visibilité. La figure 7 (Laddada, 2018) illustre cet enrichissement conceptuel. La hiérarchie développée est une conséquence du fait que la visibilité est modélisée par un cercle ou un polygone

(dans le cas d'un secteur circulaire représentant la portée d'un feu). Nous respectons ainsi le principe du modèle GeoSPARQL dans lequel les *Features* sont assimilés à des formes géométriques. D'autres éléments sont ajoutés au modèle afin de le compléter. Chaque *NavigationMark* est associé au concept *VHorizon* par la relation *hasHorizonVisibility* et au concept *VWeather* par la relation *hasWeatherVisibility*. Le concept *Vlight* est lié par la relation *hasLightVisibility* au concept *Light*. Ce dernier décrit la notion de feu dans l'ontologie et est aussi une sous-classe de *Feature*. Certains amers peuvent être reliés au concept *light* par la relation *hasLight*. D'autre part, en vue de déterminer la distance horizon, la distance de la portée du feu ainsi que celle de la visibilité relative aux conditions météorologiques, nous associons une propriété de type de données *distance* aux concepts *VHorizon*, *VLight* et *VWeather*. Nous attribuons aussi le paramètre de hauteur aux concepts *NavigationMark* et *Light* par la propriété de données *hasHeight*. Cette hauteur est employée comme paramètre dans le calcul et l'inférence de la zone de visibilité relative à l'horizon.

Spécifier qu'un repère est visible dans un environnement naturel, revient tout d'abord à instancier les concepts *VHorizon*, *VWeather* et *VLight* et ceci via des inférences. Le modèle adopte le principe de la ligne de mire (associé au concept *VisibilityLine*) afin de vérifier qu'il n'y pas d'obstacle naturel entre l'élément à percevoir et le point de référence (c.-à-d. le concept *OriginPoint*). Ainsi, le concept *VisibilityLine* est peuplé à son tour par une ligne formée entre chaque *OriginPoint* et un *NavigationMark*. Au moment où un individu de type *VisibilityLine* est créé, il doit être complété par l'information qui assigne la valeur de la visibilité. Ceci nous mène à définir un état (vrai ou faux) – un amer est visible ou non – pour chaque instanciation de type *VisibilityLine*. Dans ce cas, nous affectons à ce dernier concept, la propriété de type de données *VisibilityStm* (type *Booléen*). Théoriquement, l'expressivité du langage OWL qui intervient dans la description de notre modèle ontologique ne permet d'inférer ni de nouvelles informations comme les données numériques et booléennes, ni de nouveaux faits comme la déduction d'une relation entre des individus, alors que ceci est nécessaire pour l'inférence des routes de navigation. Plus précisément, d'une part les coordonnées géographiques des géométries (cercles ou polygones) relatives aux zones de visibilité doivent être inférées. Ces coordonnées sont décrites au format WKT et accessibles dans le domaine par la propriété de type de données *asWKT*. D'autre part, si un objet est visible à partir d'un point d'observation, ce fait doit être inféré entre l'individu de type *OriginPoint* et l'individu de type *NavigationMark*. Ainsi, une ligne de visibilité est créée entre ces deux points dont l'état est inféré à la valeur vraie. Pour combler les lacunes de raisonnement du langage OWL, nous faisons appel à la technologie SWRL qui permet de faire des inférences par des règles. Cette technologie nous permet de compléter notre modélisation avec une nouvelle expressivité sémantique apte à traiter et vérifier le principe de visibilité des amers.

4.2. Raisonnement fondé sur des règles SWRL

Les inférences avec le langage SWRL suivent un processus de raisonnement déductif. Ceci signifie que la vérité des prémisses assure totalement la vérité de la

conclusion, ou qu'à partir d'un fait résultant d'une autre règle on arrive à la vérité de la conclusion de la règle courante. Les règles SWRL que nous établissons sont construites à partir des différents concepts et des relations de l'ontologie existante avec pour objectif final d'inférer si un amer est visible ou non à partir d'un point d'observation (Règle 4). Afin d'arriver à un tel résultat, différentes règles et *built-ins* contribuent à calculer les coordonnées géographiques de chaque zone et à les inférer comme étant de format WKT.

Règle 1 : Visibilité du repère à l'horizon

$NavigationMark(?n) \wedge hasGeometry(?n, ?p) \wedge asWKT(?p, ?k)$
 $\wedge hasHeight(?n, ?h) \wedge hasHorizonVisibility(?n, ?l) \wedge hasGeometry(?l, ?c)$
 $\wedge HorizonArea(?w, ?h, ?k) \Rightarrow asWKT(?c, ?w)$

Dans la Règle 1, le corps de la règle est constitué du concept *NavigationMark*(?n) avec sa géométrie « ?p » définie par la relation *hasGeometry*(?n, ?p). La variable « ?p » a des coordonnées géographiques « ?k » au format WKT décrites par la propriété de données *asWKT*(?p, ?k). L'amer « ?n » est associé au prédicat *hasHorizonVisibility*(?n, ?l) où « ?l » est la zone de visibilité. La valeur « ?l » a une géométrie identifiée par la variable « ?c » dans la relation *hasGeometry*(?l, ?c). Afin de compléter la définition et inférer les coordonnées géographiques « ?w » au format WKT (liées à la géométrie « ?c » de la zone de visibilité à l'horizon « ?l » pour l'amer « ?n »), nous ajoutons un *built-in* *HorizonArea*(?w, ?h, ?k) où « ?h » est la hauteur du repère « ?n ». Techniquement, cette fonction *built-in*, codée en Java, retourne la géométrie d'un cercle « ?w » au format WKT pour laquelle le centre de ce cercle est la position « ?k » de l'amer et pour laquelle le rayon de ce cercle est calculé à partir du rayon terrestre « R » et de la hauteur « ?h » de l'amer, par la formule $Rayon = \sqrt{(2 \times R \times h)}$. Cette règle infère donc les coordonnées géographiques « ?w » au format WKT de la zone de visibilité horizon ayant pour géométrie « ?c » via la propriété de type de données *asWKT*. Cette règle est appliquée sur chaque repère constituant l'environnement.

Règle 2 : Visibilité du bateau à l'horizon

$OriginPoint(?o) \wedge hasGeometry(?o, ?p) \wedge asWKT(?p, ?k)$
 $\wedge hasHeight(?o, ?h) \wedge hasHorizonVisibility(?o, ?l) \wedge hasGeometry(?l, ?c)$
 $\wedge HorizonArea(?w, ?h, ?k) \Rightarrow asWKT(?c, ?w)$

Afin de définir une intervisibilité positive entre un bateau et un amer, la zone de visibilité à l'horizon du bateau doit topologiquement intersecter la zone de visibilité à l'horizon du repère. Ainsi, cette Règle 2 est conçue pour définir les coordonnées géographiques au format WKT de la zone de visibilité à l'horizon pour un bateau. Cette règle est fondée sur le même principe que celui de la Règle 1 en remplaçant l'amer *NavigationMark*(?n) par la position du bateau *OriginPoint*(?o) ayant une hauteur « ?h ».

Règle 3 : Visibilité du repère en fonction des conditions météorologiques

$NavigationMark(?n) \wedge hasGeometry(?n, ?p) \wedge asWKT(?p, ?k)$
 $\wedge hasWeatherVisibility(?n, ?l) \wedge hasGeometry(?l, ?c) \wedge distance(?l, ?d)$
 $\wedge WeatherVisibilityDistance(?w, ?d, ?k) \Rightarrow asWKT(?c, ?w)$

Les conditions météorologiques sont caractérisées par leur changement dynamique. Ainsi, il existe des prévisions sur la visibilité en mer qui fournissent les informations sur la distance de visibilité à laquelle un objet peut être clairement perçu. Dans cette mesure, la *Règle 3* aboutit à l'inférence des coordonnées géographiques « ?w » en format WKT pour le cercle de géométrie « ?c ». Ce dernier désigne la géométrie de la zone de visibilité « ?l » relative aux conditions météorologiques et associée au repère « ?n ». Tout comme pour la définition de la *Règle 1*, nous ajoutons dans la *Règle 3* un *built-in* identifié par *WeatherVisibilityDistance*(?w, ?d, ?k) où l'argument « ?d » précise la distance de visibilité fournie par le biais de la propriété de type de données *distance*(?l, ?d). La règle associe en conséquence la valeur du format WKT « ?w » au cercle de géométrie « ?c » dont le centre est défini par la position du repère « ?n ». Cette règle est également appliquée pour chaque amer constituant l'environnement.

Règle 4 : Intervisibilité entre un bateau (point d'observation) et un amer

$NavigationMark(?n) \wedge hasGeometry(?n, ?p) \wedge asWKT(?p, ?r)$
 $\wedge hasHeight(?n, ?h1) \wedge hasHorizonVisibility(?n, ?l) \wedge hasGeometry(?l, ?b)$
 $\wedge asWKT(?b, ?c) \wedge hasWeatherVisibility(?n, ?w) \wedge hasGeometry(?w, ?a)$
 $\wedge asWKT(?a, ?f) \wedge OriginPoint(?o) \wedge hasGeometry(?o, ?g) \wedge asWKT(?g, ?u)$
 $\wedge hasHeight(?o, ?h2) \wedge hasHorizonVisibility(?o, ?t) \wedge hasGeometry(?t, ?y)$
 $\wedge asWKT(?y, ?x) \wedge VisibilityLine(?s) \wedge hasGeometry(?s, ?k)$
 $\wedge sfContains(?k, ?g) \wedge sfContains(?k, ?p)$
 $\wedge IsVisibleLine(?i, ?r, ?c, ?f, ?u, ?x) \Rightarrow VisibilityStm(?s, ?i)$

La *Règle 4* stipule qu'une fois que les trois zones de visibilité sont inférées (zones de visibilité relatives à l'horizon pour le bateau et pour le repère et zone de visibilité liée aux conditions météorologiques), nous pouvons vérifier si l'amer est visible à partir du bateau par le biais du principe d'intervisibilité. Plus précisément, le *built-in* *IsVisibleLine*(?i, ?r, ?c, ?f, ?u, ?x) est à vrai si le repère « ?n » – de hauteur « ?h1 » – ayant les coordonnées géographiques « ?r », avec une zone de visibilité à l'horizon « ?c » et une zone de visibilité définie par les conditions météorologiques « ?f », est visible à partir du bateau – de hauteur « ?h2 » – ayant une position géographique « ?u » et une zone de visibilité à l'horizon « ?x ». Les sous-conditions *sfContains*(?k, ?g) et *sfContains*(?k, ?p) sont intégrées afin de vérifier que la ligne de visibilité « ?s » – ayant pour géométrie « ?k » – passe par la position du bateau – ayant pour géométrie « ?g » – et la position de l'amer en question – ayant pour géométrie « ?p ». En conséquence, cette règle assigne, via la propriété de type de données *VisibilityStm*(?s, ?i), un état de visibilité à l'individu « ?s » de type *VisibilityLine*. Cet état est de valeur booléenne qui est à vrai si le repère est visible à partir du bateau et à faux dans le cas contraire. La valeur booléenne est stockée dans le paramètre « ?i ». Le *built-in* *IsVisibleLine*(?i, ?r, ?c, ?f, ?u, ?x) retourne la valeur « true » si les conditions ci-dessous sont satisfaites dans l'ordre :

- Les deux zones de visibilité à l'horizon « ?x » et « ?c » pour le bateau et pour l'amer doivent être en intersection topologique ;
- La localisation du bateau « ?u » doit être topologiquement dans la zone de visibilité définie par les conditions météorologiques « ?f » ;
- En considérant la ligne de visibilité, aucun obstacle présent sur la ligne n'a une élévation qui excède l'élévation de la ligne de mire.

5. Discussions et perspectives

Les systèmes de positionnement (GPS, AIS) sont omniprésents dans de nombreux secteurs. Dans le domaine maritime, les systèmes de positionnement par GPS et AIS (« Automatic Identification System ») servent à un suivi de route mais aussi à une exploitation de données ayant trait à divers phénomènes (pollution, migration d'animaux marins, etc.). Dans leur analyse, les opérateurs ayant en charge la surveillance du trafic maritime se doivent de fédérer et traiter les informations provenant des moyens de détections terrestres (radars, systèmes électro-optiques, etc.) et des moyens embarqués (AIS, etc.). En complément des systèmes de surveillance à vocation professionnelle, les données d'empreintes de navires peuvent aussi être accessibles pour des utilisateurs lambda par des serveurs cartographiques⁹. Si les systèmes actuels intègrent couramment des systèmes d'alarme fondés par exemple sur la réglementation maritime comme le sens de circulation dans les dispositifs de séparation de trafic (DST) ou sur la présence d'un navire dans une zone de faible profondeur en dessous d'une isobathe de sécurité fixée, les opérateurs doivent aussi déterminer si l'empreinte d'un navire a du sens ou non. En effet, un non-sens pourrait révéler un phénomène inattendu et une situation d'insécurité (p. ex. avarie de barre, trafic de marchandises non autorisées, etc.). Un des enjeux actuels des systèmes de navigation est d'intégrer de la sémantique au sein des systèmes afin de faciliter l'interprétation des empreintes de navires. Cette recherche a pour objectif applicatif final d'apporter un premier niveau d'analyse sémantique de trajectoires de navigation maritime côtière en lien avec une modélisation des connaissances du domaine. Ainsi, deux approches peuvent être envisagées.

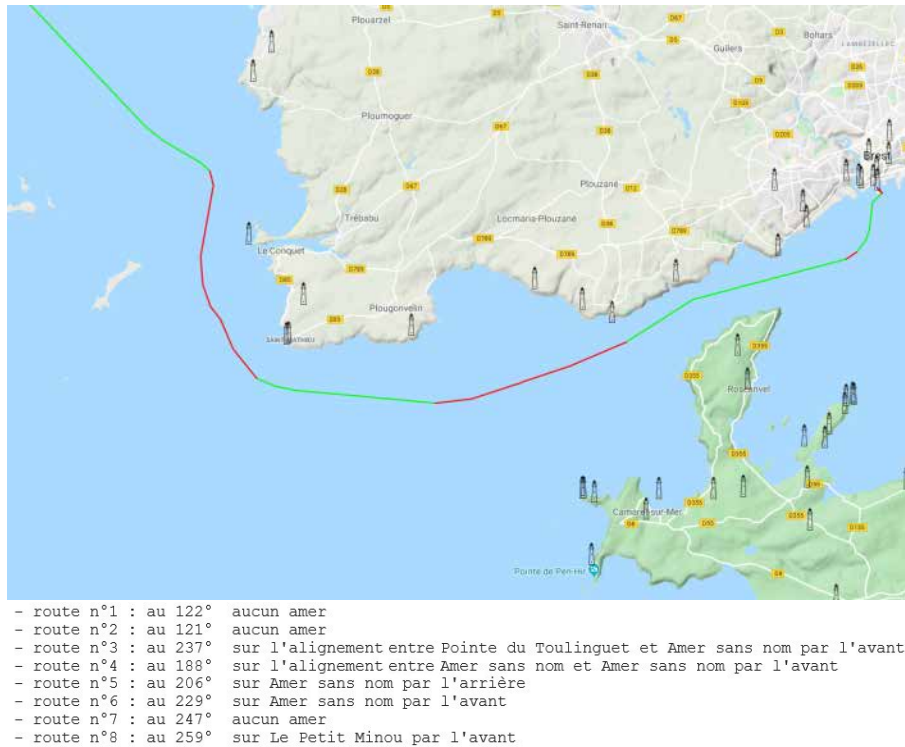
Une fois le modèle de connaissances ainsi défini, ce dernier peut être peuplé par des données réelles contenues dans les ENC (*Electronic Navigational Chart*) au format S-57 ou dans des descriptions textuelles telles que celles des « Instructions nautiques » du SHOM (cf. section 2.1). Une première stratégie d'analyse sémantique consiste à confronter les routes inférées à partir de l'ontologie instanciée, routes dites « théoriques », aux routes dites « réelles » des bateaux obtenues par filtrage et simplification de leurs données de positionnement. Dans le cadre de notre étude, les routes réelles sont obtenues à partir d'empreintes AIS de navires.

Plus précisément sur le mécanisme d'inférence, la modélisation présentée dans cet article est fondée sur deux raisonnements. Le premier est axiomatique résultant des

9. par exemple, le serveur cartographique <https://www.marinetraffic.com>.

définitions des axiomes (cf. tableau 1), le second est fondé sur des règles SWRL. Pour maintenir la décidabilité et la consistance de notre ontologie les deux mécanismes de raisonnement sont séparés et ne sont pas exécutés en parallèle. Ainsi, l'ontologie principale est séparée en deux ontologies. D'une part, nous avons une base de connaissances avec toutes les définitions axiomatiques complexes des concepts mais sans les règles SWRL et d'autre part, nous avons une ontologie comprenant les règles SWRL et tous les concepts de l'ontologie mais sans les définitions axiomatiques complexes. Une fois les ontologies définies, un algorithme de subdivision de l'espace génère des points de vue (*OriginPoint*) fictifs, ce qui permet d'exploiter l'ontologie avec les règles SWRL (sans les axiomes) afin de générer les points de décision et lignes d'intervisibilité à partir de ces points de vue. Ce premier raisonnement représente donc un processus de peuplement de l'autre ontologie, qui est enrichie avec les axiomes (sans les règles SWRL), afin d'en déduire les segments portant sur un relèvement ou un alignement. En effet, l'assertion de ces lignes d'intervisibilité et des points de décision contribue à l'inférence des connaissances définies par les axiomes complexes et, de ce fait, générées par le biais du processus de raisonnement axiomatique. Plus précisément, ce dernier peuplera l'ontologie avec des *WayPoints* et ainsi avec des segments décrivant les routes théoriques qui suivent un ou plusieurs *landmark*. Le mécanisme de raisonnement décrit ici est élaboré dans un mode ouvert (*Open World Assumption*). Ceci signifie que si une information n'est pas connue cela n'affirme pas que cette information soit fausse. Cette supposition nous a incités à ajouter des définitions restrictives sur les instances en attribuant à chaque définition une exclusivité. Par exemple, si une ligne est définie par deux points elle ne le sera que par ces deux points. Ou encore, si deux lignes ont une intersection il ne peut y avoir qu'une seule intersection définie par un seul point.

La figure 8 illustre l'analyse d'une empreinte de navigation maritime côtière dans laquelle les sections de couleur verte indiquent les portions ayant une sémantique connue pour notre modèle, c'est-à-dire les portions à proximité (à une distance paramètre du traitement) d'une route théorique fondée sur un alignement ou un relèvement sur un amer. À l'inverse, les portions dénuées de sens sont indiquées en rouge. Ces portions en rouge n'indiquent pas forcément une situation d'insécurité mais soulignent que ce comportement n'est pas pris en compte dans le modèle de connaissances qui est défini et qu'une attention plus particulière doit être apportée. La figure 8 est complétée (en bas) par la description de la sémantique des routes présentes dans la trajectoire. Afin d'approfondir l'analyse, un projet complémentaire a été initié pour prendre en compte la sémantique des objets de type « Bouée » dans l'analyse des empreintes (Fournier, Tyschenko, 2019). Cependant, et en lien avec les propos introductifs, cette approche symbolique permet de faire émerger les situations qui sont modélisées dans le modèle de connaissances et d'en fournir une sémantique. En revanche, elle est incapable d'expliquer les situations inattendues. C'est la raison pour laquelle les approches numériques fondées sur l'analyse des données sont complémentaires à la nôtre et peuvent renforcer l'analyse sémantique des traces de navigation maritime côtière (Camossi *et al.*, 2013 ; Vandecasteele *et al.*, 2014).



*Figure 8. Analyse sémantique d'une trajectoire de navigation maritime côtière.
En vert, les routes ayant une sémantique pour le modèle de connaissances et
en rouge, celles dénuées de sémantique. En bas, et pour une trajectoire orientée
de droite à gauche, la sémantique associée à chaque route*

Une autre approche d'analyse sémantique consiste à instancier tout d'abord le concept *Route* de l'ontologie par des routes réelles et d'inférer ensuite les éléments (repères, alignements, relèvements) contribuant potentiellement à définir la sémantique des routes réelles. Cette approche différente nécessite de compléter la formalisation axiomatique ou à base de règles de l'ontologie existante afin d'en déduire les éléments à associer aux routes de navigation. Elle fait donc l'objet de recherches ultérieures permettant de compléter la présente recherche. Notons cependant que cette dernière approche est moins adaptée aux analyses sémantiques « temps quasi réel », puisque dans la première approche les routes théoriques peuvent être précalculées et être ainsi directement comparées aux données réelles, réduisant alors les temps de traitement majoritairement imputables au processus d'inférence.

Bibliographie

- Baader F., Calvanese D., McGuinness D. L., Nardi D., Patel-Schneider P. F. (2010). *The description logic handbook: Theory, implementation and applications* (2nd éd.). USA, Cambridge University Press.
- Camossi E., Villa P., Mazzola L. (2013). Semantic-based anomalous pattern discovery in moving object trajectories. *Computing Research Repository (CoRR)*.
- Egenhofer M., Herring J. (1990). *Categorizing binary topological relationships between regions, lines, and points in geographic databases*. Rapport technique. Department of Surveying Engineering, University of Maine.
- Fournier T., Tyschenko R. (2019). *Analyse de trajectoire de navigation maritime fondée sur la sémantique*. (Rapport de projet informatique 2e année, Ecole navale)
- Horridge M., Drummond N., Goodwin J., Rector A., Stevens R., Wang H. (2006). The manchester OWL syntax. In *Proceedings of the 2006 OWL experiences and directions workshop (OWL-ED2006)*.
- Kowalski J.-M. (2009). *Thalassographie : mesure, représentation et description des espaces maritimes en grèce ancienne*. Thèse de doctorat non publiée, Université de Nice, France.
- Laddada W. (2018). *Vers une émergence des systèmes d'information géographique maritime fondés sur la connaissance : application aux systèmes d'aide à la navigation*. Thèse de doctorat non publiée, Université de Bretagne Occidentale, France.
- Lamotte L., Abadie N., Saux E., Kergosien E. (2019). Extraction de connaissances pour la description de l'environnement maritime côtier à partir de textes d'aide à la navigation. In *Proceedings de la conférence extraction et gestion des connaissances (EGC), Bruxelles*, p. 341-348.
- Marine nationale. (2005). *Directives générales pour la conduite nautique, force d'action navale, édition 2005*. Marine-nationale.
- Maurel D., Friburger N., Antoine J.-Y., Eshkol I., Nouvel D. (2011). Cascades de transducteurs autour de la reconnaissance des entités nommées. *Traitement Automatique des Langues*, vol. 52, n° 1, p. 69-96.
- Moncla L. (2015). *Automatic reconstruction of itineraries from descriptive texts*. Thèse de doctorat non publiée, Université de Pau, France.
- Przytula-Machrouh E. (2004). *Information verbale et information graphique pour la description d'itinéraires : une approche pluridisciplinaire*. Thèse de doctorat non publiée, Université René Descartes, Boulogne-Billancourt, France.
- Randell D. A., Cui Z., Cohn A. G. (1992). A spatial logic based on regions and connection. In *Kr'92: Proceedings of the third international conference on principles of knowledge representation and reasoning*, p. 165–176.
- SHOM. (2018). *Instructions nautiques, France (côtes nord et ouest), du cap de la hague à la pointe de penmarc'h (version à jour au 19 décembre 2018)*. SHOM.
- Tsatcha D., Saux E., Claramunt C. (2012). A modeling approach for the extraction of semantic information from a maritime corpus. In S. Timpf, P. Laube (Eds.), *Proceedings of the 15th international symposium on advances in spatial data handling (SDH)*, p. 175–191. Springer-Verlag Berlin Heidelberg.

Vandecasteele A., Devillers R., Napoli A. (2014). From movement data to objects behavior using semantic trajectory and semantic events. *Marine Geodesy: Special Issue on Coastal and Marine Geographic Information Systems*, vol. 37, n° 2, May, p. 126-144.

Voiles et voiliers. (2006). *Pilote côtier. De Saint-Malo à Brest. 9e édition.*

La genèse systémique d’empreinte pour une maîtrise de l’observation de la Terre

Mireille Fargette¹, Maud Loireau², Najet Raouani³,
Thérèse Libourel⁴

1. Institut de Recherche pour le Développement - IRD, UMR 228 Espace-Dev,
500 rue Jean-François Breton, 38093 Montpellier Cedex 05, France
mireille.fargette@ird.fr

2. Institut de Recherche pour le Développement - IRD, UMR 228 Espace-Dev,
Univ. de Perpignan-UPVD, 52 rue Paul Alduy, F-66860 Perpignan Cedex 09,
maud.loireau@ird.fr

3. Institut supérieur agronomique Chott-Mariem, UR 13AGR 04,
Université de Sousse, Tunisie
nraouani@yahoo.com

4. Université de Montpellier, UMR 228 Espace-Dev,
500 rue Jean-François Breton, 38093 Montpellier Cedex 05, France
therese.libourel@umontpellier.fr

RÉSUMÉ. Ce travail s'intéresse à l'observation, à la connaissance scientifique construite à partir de ce qui est perçu (Lien de Sens) d'un monde systémique complexe. La démarche conduit à proposer le concept d'empreinte dans le cadre scientifique interdisciplinaire « Système - Réalité - Monde perçu - Modèle », à le mettre à l'épreuve des données, puis à proposer la démarche d'ontologie systémique. Celle-ci permet de déployer le Lien de Forme du domaine systémique au monde perçu, d'analyser et décrire la part pertinente de la donnée et de montrer en quoi l'ensemble de ce travail essentiellement symbolique peut contribuer à la maîtrise sémantique, technologique et numérique de la collecte et analyse des données et de l'observation de la Terre. Les illustrations portent principalement sur l'oasis. Ce travail discute ce qu'il apporte dans la construction d'une science de l'observation.

ABSTRACT. This work is interested in observation, in scientific knowledge acquired from what is perceived (Link making Sense) from a complex systemic world. The approach leads to proposing the concept of imprint within the interdisciplinary framework "System – Reality – World as perceived – Model" and testing it against data, then to proposing systemic ontology as an approach. This makes it possible to deploy the Link making Shape from the systemic domain to the world as perceived, to analyze and describe the relevant part in the data and to show how the whole of this mostly symbolic work can contribute, with respect to semantic, technological and digital aspects, to better control data collection and

analysis and Earth observation as a whole. Illustrations mainly focus on oasis. This work discusses its contribution to constructing an observation science.

MOTS-CLÉS : modélisation conceptuelle, ontologie systémique, science de l'observation, passages Système - Monde perçu, existence géographique, contenu de donnée, reconnaissance du signal, analyse de physionomie, interdisciplinarité, oasis, Lien de Sens, Lien de Forme.

KEYWORDS: Conceptual modeling, Systemic ontology, Observation science, Paths linking Systems to the world as perceived, Geographical existence, Data content, Signal recognition, Physiognomy analysis, Interdisciplinarity, oasis, Link making Sense, Link making Shape.

DOI:10.3166/RIG.31.135-197 © 2022 Lavoisier

1. Introduction

Le travail qui suit s'intéresse à l'observation de la Terre et à son suivi. Toutefois, on ne peut aborder cette question en termes scientifiques sans s'interroger préalablement sur celle de la relation de l'homme au monde qui, de façon empirique, passe par l'expérience d'observation, celle-ci prise dans son acception la plus large. Nous abordons les deux points, observation du monde et relation au monde, et les liens qui les unissent dans le travail présenté ici afin d'asseoir une démarche et profiler une science de l'observation.

L'aventure technologique est indissociable de l'aventure humaine (Leroi - Gourhan, 1964, 1965) ; elle occupe incontestablement une place grandissante dans la pratique de gestion du quotidien, mais aussi dans la façon de questionner le monde, de le voir et le concevoir, de le connaître et l'interpréter. Ainsi, par exemple, les avancées des dernières décennies en matière d'observation de la Terre par voie satellitaire restituent de toujours plus imposants jeux de données acquis par des capteurs toujours nouveaux pour ce qui est de leurs caractéristiques technologiques (gammes et registres des canaux exploités, capacité à affiner la résolution, récurrence temporelle). De tels jeux de données intéressent l'activité scientifique *s.s.* (acquisition de connaissance) ou l'activité gestionnaire (suivi, gestion relative à ce monde, prise de décision relative à ce temps). Ils proposent de nouvelles caractéristiques pour renseigner et dessiner le monde, critères qui s'ajoutent ou se confrontent à ceux fournis par les capteurs physiologiques ou leurs substituts technologiques directs. De même, le déploiement d'une puissance de calcul toujours croissante et celui de l'intelligence artificielle s'immiscent entre le monde et l'homme (Ganascia, 2020).

L'aventure humaine s'ancre dans sa relation au monde. Les travaux conduits en anthropologie, sociologie, psychologie, psychanalyse et philosophie ont montré la richesse, en complexité comme en diversité, des schèmes de représentation que l'être humain construit seul ou collectivement et vit individuellement ou en partage au sein de groupes d'individus plus ou moins organisés. Ce constat porte en particulier sur ce qui le lie au monde qui l'entoure, qu'il s'agisse d'une conception de rang individuel ou collectif (Descola, 2010, 2011). Les déploiements de

mythologies, de cosmogonies, de spiritualités, de pensées d’ordre philosophique ou scientifique fondent autant de paradigmes qui assurent la relation au monde selon leurs ontologies propres, explicites ou implicites. On constate l’extrême relativité de telles représentations et leur diversité quant à leur degré d’explicitation.

1.1. Enjeux, problématique et objectifs

Dans le monde toujours plus ouvert qui est le nôtre (monde hyperconnecté, donnant la parole à des millions de voix, la distance n’entravant plus la communication), l’information de toute sorte abonde, se partage, se contredit. Nous nous confrontons à la profusion, à la diversité, parfois la contradiction et même l’incompatibilité des points de vue et à la « relativité des choses » qui en découle. L’académie d’agriculture, parmi d’autres, s’interroge sur ce relativisme et sur la place des sciences dans un tel contexte (Chevallier-Le Guyader, 2019). Dans le même temps et paradoxalement, l’urgence de la communication bien comprise, celles de la compréhension des choses et de la compréhension mutuelle sont toujours plus grandes, le besoin de justesse de notre relation au monde aussi, clé du discernement et, en conséquence, du bien fondé de nos décisions. En effet, les enjeux planétaires appellent une méthode qui permette un dialogue constructif et l’élaboration de solutions coordonnées. En l’état, et sans prôner pour autant l’unicité du monde et de ses représentations, ce qui serait délétère et totalitaire, est-il possible de poser la question : existerait-il une représentation « correcte (ou la moins fausse) » du monde ? Comment la figurer, l’apprécier ? Selon quelle qualité, quels critères ? Une représentation adéquate du monde (approchant au plus près possible la réalité) serait un premier gage de succès des actes ensuite posés sur ce monde.

Étant considéré l’enjeu et admise la diversité aux traits rapidement brossés ci-dessus, étant aussi reconnue notre filiation ancrée dans la culture occidentale qui a vu naître et s’élaborer la démarche scientifique¹, elle-même nécessairement (Bachelard, 1934 : « La science crée de la philosophie » ; Wolff, 2010) adossée à la pensée philosophique, nous nous reconnaissons dans le même temps en exploration des interfaces de larges pans de connaissance et de larges pans d’inconnu. Par voie de conséquence, nous nous reconnaissons aussi non spécialistes de si larges « espaces » auxquels le monde et le questionnement contemporains nous confrontent cependant. Dans ce contexte pourtant plus restreint, la complexité des enjeux et la diversité des connaissances scientifiques à impliquer nous met à nouveau en face de la diversité des points de vue, ces derniers ayant pour origine cette fois les diverses disciplines, des sciences de l’homme et des sociétés aux sciences de la nature, de la vie, de la terre et de l’univers. Pour être partagé de façon satisfaisante entre spécialistes de disciplines parfois très éloignées les unes des autres, ce savoir, comme toute représentation, doit être explicité entre disciplines. Or le domaine

1. sans la résumer en un positivisme trop sommaire, comme nous en met en garde Bouleau (2017).

scientifique possède les méthodes pour raisonner, analyser, décrypter, représenter en partage, les connaissances et, *in fine*, les rendre explicites, accessibles, ouvertes.

Comment se situer et vivre dans le monde ? Comment situer les nouveaux modes d'observation et voies d'acquisition de données évoqués ci-dessus dans le cadre de diversité précédemment évoqué ? A partir d'observations, construit-on des mondes parallèles (sans communication ni relation de sens entre eux), des mondes concurrents ou des mondes complémentaires ? Des facettes d'une même réalité ? Suit-on le même schème qui nous confirme (et risque éventuellement de nous enfermer) dans nos propres représentations et une certaine routine ?

Les questions de correspondance de jeu de données à jeu de données mais aussi de communication entre individus se trouvent donc posées à propos de ce qui est observé puis représenté.

De façon plus générale, pour habiter notre monde « avec intelligence et sagesse » (Bouleau, 2021) tout en tirant parti d'un déploiement technologique, il est important de mettre en relation, d'une part une interprétation plurielle et partagée de ce monde, fondée sur la connaissance scientifique issue de la contribution combinée et interdisciplinaire des domaines scientifiques concernés (l'écologie, la biologie, la géologie, la géographie, l'agronomie, les sciences humaines et sociales, etc.) avec, d'autre part, les propositions et les nouveaux acquis de l'observation de la Terre et de son suivi (observation diachronique) menés par exemple en observatoire, en plein essor depuis quelques décennies.

Dans ce contexte, notre réflexion a donc pour origine le questionnement croisé sur, d'une part, les notions d'observation, de suivi environnemental et d'observatoire scientifique sociétés - milieux (Loireau *et al.*, 2017), d'autre part, le renouvellement constant et la démultiplication de la capacité d'observation² de la Terre par satellite (Cazenave, 2013), et s'intéresse aux voies possibles d'un travail en interdisciplinarité. Elle pose la question scientifique : comment observer la Terre pour traiter de l'interprétation de la réalité ?

De façon pragmatique, pour comprendre le monde qui nous entoure (le découvrir, lui donner sens, s'y situer, s'y repérer, identifier les formes qui s'y trouvent), **l'important est, pour commencer, de bien observer**. La figure 1 et l'encadré 1 ci-après posent la problématique de l'expérience d'observation avant toute normalisation par la démarche scientifique. Trois moments particuliers de cette aventure y sont pointés :

- trouver la **signification** d'un contenu observé, donc créer de la connaissance ;

2. capacité d'observation et de suivi par télédétection ; l'ampleur des espaces couverts, la fréquence temporelle toujours augmentée des collectes de données, la variété de la gamme spectrale et de la taille pixellaire démultiplient le potentiel et la rapidité de réponse à nombre de questions d'ordre environnemental (Cazenave, 2013).

- par extension, faire se correspondre (et donc donner du sens par l’établissement d’une **correspondance**) des contenus obtenus « en parallèle », observés au cours d’expériences d’observation indépendantes sur une même fenêtre d’observation ;
- savoir **reconnaître** l’observé à l’aide d’un savoir préalable, un champ de connaissance plus vaste et déjà établi.

Ces trois démarches pourtant distinctes sont souvent regroupées sous le seul terme d’interprétation.

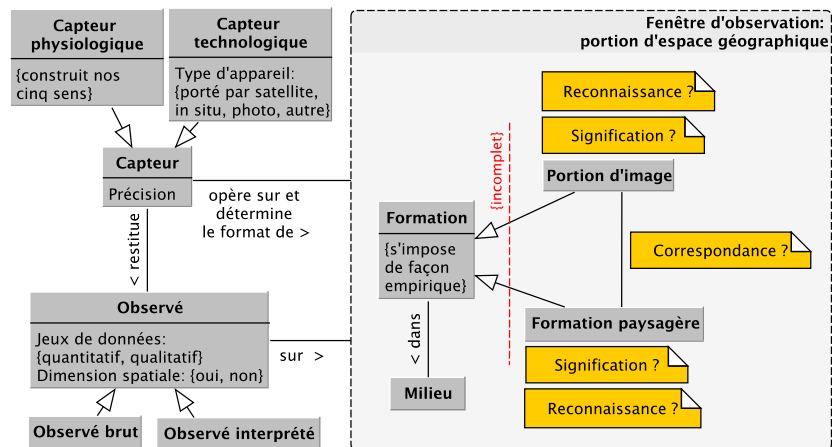


Figure 1. Expérience d’observation - Problématique de départ, avant toute normalisation par la démarche scientifique (voir encadré 1)

L’intention de ce travail est d’asseoir une démarche d’observation en précisant ce qui associe observation et connaissance dans la relation au monde que nous entretenons. Elle est aussi d’élargir l’assise, l’enracinement interdisciplinaire de la démarche. Nous gageons que comprendre et expliquer les relations mises en jeu au moyen d’une ontologie systémique y contribuera en même temps qu’elle épaulera l’exercice de l’interdisciplinarité, lui aussi requis par l’objectif.

1.2. Premières difficultés et grands traits de la démarche

Toutefois, des difficultés apparaissent dès la lecture de l’encadré 1. Une forte subjectivité imprègne en effet la sémantique des termes tels que par exemple observation, perception et monde perçu, interprétation, ainsi que, en bout de course, le résultat du processus d’observation (pris dans son ensemble) vers une représentation du monde. La relation au monde, propre à chacun, met encore l’accent sur la forte relativité des propos sur le monde et des interprétations que l’on en tire. Enfin, l’absence d’explicitation de la démarche augmente la confusion. Que

met on par exemple, pour ce qui relève de la démarche, derrière les termes signification, correspondance, reconnaissance, interprétation, posés en figure 1 ?

Encadré 1 - Expérience d'observation - Problématique de départ, avant toute normalisation par la démarche scientifique (voir figure 1)

De façon très générale, le processus d'observation dépend avant tout de la présence d'un capteur (qu'il soit physiologique, technologique, etc.), celui-ci déterminant le format de la fenêtre d'observation qui lui est associée. Ce capteur [ou un ensemble de capteurs, de même type ou pas, repartis dans la fenêtre] . Les capteurs physiologiques portés par des organismes biologiques rapportent, sur les caractéristiques de milieu ou d'un élément d'intérêt l l l'origine par l'homme, qui mettent en jeu des propriétés et mécanismes physico-chimiques complexes. Le développement technologique propose d'autres modes de capture [à l'image ou voisins de ceux détenus par des organismes biologiques e.g - l l chauve-souris pour détecter les obstacles, ou un système d'orientation des oiseaux sensible au magnétisme terrestre], d'autres capteurs l 'intérêt] qui renseignent sur d'autres caractéristiques du milieu tout aussi complexes [l'opacité et la couleur au-delà du visible, la sonorité au-delà de l'audible, la dureté, la rugosité au-delà du palpable, le magnétisme]. Par exemple, l'observation de la T ll ll enregistrées les radiations reçues de la Terre, dans et au-delà du registre du visible. Les capteurs construits de main d'homme captent des données qui se distinguent par leur nature l l l'origine percevoir ; le monde reconnu comme tangible s'étend du fait de l'évolution technologique des capteurs. Le monde perçu est donc capteur-dépendant et certaines caractéristiques du monde nous sont transparentes, faute du capteur *ad hoc*.

De façon générale, une observation est une expérience de la réalité rendue tangible par la présence de capteur *ad hoc* opérant dans la fenêtre d'observation (il faut donc qu' l

l l'origine du jeu de données (l') est dépendante des caractéristiques propres du capteur, de la sensibilité aux conditions extérieures, etc. et de son mode opératoire. L' l l l réalité vers laquelle le c .

La fenêtre d' 'espace géographique, dont l'emprise l l , voire des contingences de lieu [e.g. une barre d'immeuble ou une montagne qui barre la vue ou la prise de vue photographique] et des caractéristiques du capteur [e.g. la taille de la fauchée d'un capteur satellite]. Dans l'espace délimité peuvent émerger des formations qui « s'imposent » aux sens, par leur limite et leur contenu qui les distinguent de ce qui les entoure [elles viennent spontanément, qu'il s'agisse d'une montagne ou d'un lac, d'individus représentant des êtres vivants, de « constructions » d'origine animale ou humaine, etc.], constat plus ou moins empirique produit par la perception, mais aussi phénomène vécu, « expérimenté dans son corps » l'origine de l'expérience phénoménologique : cf. Husserl in Bakewell (2018) ; voir aussi Pitte (2010) dans une perspective plus géographique]. De telles formations peuvent alors polariser, organiser l'observation, par la structuration de l'espace qu'elles suggèrent.

L’expérience d’observation peut (facultativement) tenir compte de ces formations dans son protocole d’observation et/ou d’analyse. Sinon la fenêtre d’espace unique sur lequel l’observation est menée. L’expérience d’observation commence par être brute ; elle s’élabore ensuite, lorsque le jeu acquis est progressif

’autres, interprété. L’expérience d’observation brute est par exemple un jeu de données sous sa forme de réception, ou bien une formation (un sous-ensemble, empiriquement délimité du jeu de données) dont la présence et la physionomie (contenu souvent synthétique et qualitatif) sont constatées. Chez les organismes conscients, percevoir, distinguer des formations dans le milieu et leur attribuer du sens (signification), ou bien les reconnaître (reconnaissance) relèvent de l’interprétation au sens large. De plus, il peut y avoir télescopage de l’étape « capture de données par un capteur » avec l’étape « interprétation de l’ », des filtres s’interposant et/ou l’interprétation étant plus ou moins indissociable de la réception - sensation reçue du capteur.

Non raisonnée, l’interprétation est arbitraire ; conditionnée, programmée, elle est contrainte [l’apprentissage diffère dans sa genèse, selon qu’une intelligence s’active, ou qu’une « intelligence artificielle » est en construction (apprentissage par machine). Le conditionnement (chez les êtres vivants) peut être vu comme un cas limite qui ne nécessite pas ce que l’on entend en général par intelligence]. Raisonnée, elle repose plus librement sur une étape analytique et sur un savoir. L’interprétation appelle l’intelligence ; elle est tout d’abord empirique, intuitive ; elle peut devenir plus explicitement fondée sur l’

Quand plusieurs expériences d’observation sont menées sur la même fenêtre d’observation, par différents capteurs ou par renouvellement de l’expérience, la question de leur correspondance se pose.

Tout ceci montre la difficulté de préciser **une méthode pour asseoir l’observation** alors que l’expérience d’observation, l’expérience du monde ne sont jamais les mêmes, d’un individu à un autre, chacun les vivant personnellement, par et dans son corps et selon sa propre histoire individuelle et collective. Le courant de pensée de la phénoménologie (*cf.* Husserl à l’origine puis l’école plus largement, *in* Bakewell, 2018) a amplement documenté cette position mais aussi des romanciers qui évoquent ce vécu au travers du corps (Yourcenar, 1951). Au flou contenu dans la relation de l’individu au monde s’ajoute la complexité de la relation collective d’une société au monde. Il s’ensuit à nouveau un relativisme qui accentue encore la difficulté et semble rendre vaine toute tentative de « normalisation ». Dans ce contexte, le mouvement post-moderne de sociologie des sciences « est une préciosité universitaire totalement incapable de construire une pensée collective » (Bouleau, 2021), qui dénonce la difficulté mais n’apporte aucune proposition constructive pour la circonvier. En fin de compte, une forte **polysémie** existe, qui contribue à la confusion déjà mentionnée.

Avant de poursuivre et, parce que ce travail est interdisciplinaire et s’adresse à des lecteurs , il est important d’admettre que la polysémie des termes est inévitable. Or la communication non ambiguë nécessite, au moins pour le temps de l’échange, un **vocabulaire partagé**. Nous proposons, en annexe, une liste de termes que nous définissons ; ce sont des concepts-clés dans le cadre de ce

travail : empreinte et genèse d'empreinte, entité et épure, existence et existence géographique, formation, milieu, observé et monde perçu, ontologie systémique, forme et *Lien de Forme*, physionomie, réalité, référentiel, sens et *Lien de Sens*, système, modèle et compartiment, tangible. Pour chaque terme mentionné, il ne s'agit pas d'imposer de façon définitive une définition à une large communauté au sein de laquelle un consensus n'existe pas. Il est cependant nécessaire que le lecteur assume cette polysémie et accepte, pour le temps de sa lecture, le sens que les auteurs attribuent à ces termes. Cette liste, à entrées multiples et concertantes, est plus riche qu'un glossaire puisqu'elle montre aussi les relations que les termes entretiennent entre eux (dans la cohérence de notre travail) et/ou en quoi ils se distinguent les uns des autres.

Nous illustrons ci-dessous, à partir d'éléments en figure 1, trois points qui nous semblent cruciaux pour dégager ensuite une piste pour l'étude de la question posée.

– **Point 1** : L'observé résulte de l'observation qui est une action, un processus. On constate cependant beaucoup de flou à propos de la notion d'expérience d'observation. Certains ont déjà en tête l'expérience scientifique qui se distingue en particulier par son souci d'objectivité et son intention de prendre (du mieux possible) la distance nécessaire à l'impartialité. Plus précisément, l'observation consiste dans ce cas en l'acquisition des données selon une méthode, un protocole qu'elle veut rigoureux, vérifiable, reproductible et partageable avec tous et donc tout d'abord explicite (l'ontologie formelle du concept d'observation OBOE³ et les principes directeurs pour la gestion et l'administration des données scientifiques FAIR⁴ s'y emploient). Nombreux sont ceux qui voient cependant dans l'expérience d'observation (et c'est le parti de la figure 1 et de l'énoncé de son titre) le processus dans son entier, de l'observation jusqu'à la « leçon » que l'on en tire, jusqu'à l'interprétation ; c'est bien le cas, vécu, de tout homme dit « d'expérience ». La **subjectivité** et le **relativisme** s'engouffrent par cette brèche et sont source de nombre de malentendus non seulement dans le résultat énoncé à partir de ladite expérience mais aussi sur la notion de processus opérant qui mène au résultat : un mode quasi-réflexe ou conditionné, le mode le plus sommaire ; un mode pensé, mais encore représenté par pure intuition (non explicité), ou bien un mode pensé, empruntant des cheminements que l'on précise pour mener une science de la compréhension ou bien une science de l'explication, par exemple selon des raisonnements de nature probabiliste ou logique (de type inductif⁵, déductif⁶, hypothético-déductif, abductif⁷).

3. The Extensible Observation Ontology: <https://github.com/NCEAS/oboe>

4. Findable, Accessible, Interoperable, Reusable : <https://www.ouvrirlascience.fr/fair-principles/>

5 'induction (George, 1997) est une forme d'inférence qui généralise ce qui est connu de quelques cas à l'ensemble des cas de la même espèce ; plus largement, elle englobe toute inférence comportant une « projection » du connu vers d'autres cas non connus. Sa conclusion ne serait qu'une conjecture dont la vérité ne peut être garantie (Hume, in George, 1997). Il est parfois possible de réduire un raisonnement inductif à un raisonnement déductif.

– **Point 2** : La relation que nous avons au monde (figure 2 et encadré 2), fondée sur un bipôle monde tangible - représentation symbolique, est ambivalente : cette relation fonctionne tantôt dans un sens (acquisition issue d'observation qui va s'insérer dans l'abstraction), tantôt dans l'autre (lecture du monde tangible avec le concours de représentations déjà acquises) ; elle est à l'origine de processus itératifs pas toujours conscients ni contrôlés. Ces processus, non seulement souvent implicites mais aussi propres à chaque individu, sont la source de flou dans le mode d'acquisition de connaissance sur le monde, sur la qualité des acquis ou la pertinence des décisions qui s'ensuivent. **Subjectivité, imprécision** dans le message, voire **incommunicabilité** représentent de nouvelles sources de difficulté.

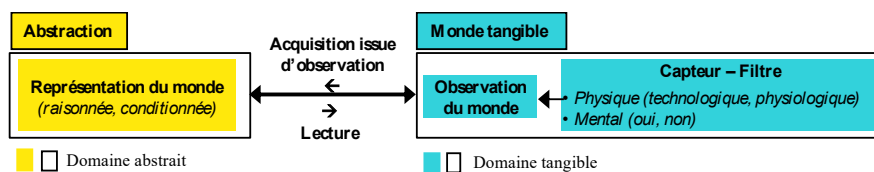


Figure 2. A l'origine de la relation au monde (voir encadré 2)

Encadré 2. A l'origine de la relation au monde (voir figure 2)

La relation que nous avons au monde qui nous entoure consiste dès l'origine en la conjugaison d'observations et de représentations. Elle peut être schématisée par une relation complexe qui confronte et relie dans un va et vient continuels deux domaines, le domaine du tangible qu'il est possible de percevoir avec le concours de capteurs, et le domaine de l'abstraction : 1) du monde tangible vers l'abstraction, une relation d'acquisition d'observations menées sur le monde, une démarche composite mais non explicitée, allant de l'observation jusqu'à la symbolisation des représentations et des connaissances, 2) du domaine de l'abstraction vers le monde tangible, une mise en relation de ces deux mondes par la lecture des produits de l'observation selon les schèmes abstraits préalablement acquis. La démarche n'en est pas explicitée non plus. La lecture se fonde donc sur les représentations intériorisées du monde, qu'elles aient été transmises par l'éducation ou acquises par auto-apprentissage. L'apprentissage est une phase de construction, l'auto-apprentissage est une phase de construction autonome ; induite à la suite d'une observation par une remise en cause, une absence de réponse selon les schèmes préalablement connus, elle entraîne la révision des représentations en cours.

6. dont la voie royale en science est (et a longtemps été) le raisonnement hypothético-déductif.

7. 'tion (George, 1997) consiste à expliquer u

desquelles on pourrait inférer le phénomène à expliquer.

La relation au monde tangible consiste donc en ces liens de deux types, souvent impliqués
 processus itératifs, 1) où capteurs physiologiques ou technologiques interviennent en tant que
 filtres auxquels peuvent se surimposer des filtres mentaux liés aux représentations préalables,
 pour aboutir à des observations plus ou moins objectives ; 2) où les représentations du monde
 ou même de quasi-conditionnements où la réflexion intervient peu ou pas.

Une telle ambivalence de la relation est classique jusque dans le domaine
 scientifique où toutefois le souci de clarté suppose qu'elle soit prise en compte,
 explicitée et maîtrisée. Par exemple les observatoires OSAGE (Loireau *et al.*, 2015,
 2017) distinguent dans leur fonctionnement, le mode recherche (qui permet
ad hoc.
 la figure 1 : la
 connaissance
 (sup

– **Point 3 :** En conséquence, se conjuguent assez spontanément l'intuition et
l'empirisme⁸ dont les sources de pertinence se perdent aux origines mêmes de notre
 relation au monde et dont on ne connaîtra jamais les tâtonnements initiaux ni la
 première itération. Goodman (1954, *Fact, Fiction and Forecast*, in Lecourt, 2018)
 s'intéresse à ces commencements où « l'esprit est en marche dès le départ, émettant
 des projections spontanées tous azimuts. Peu à peu, le mécanisme corrige et canalise
 son processus projectif ». Par exemple, le concept de formation mentionné en
 figure 1 témoigne de ce moment où l'on se heurte au fonds de représentation et de
 connaissance, venant de la nuit des temps et fruit de cette relation au monde. Il
 charrie sa part d'induction mais il est aussi, souvent, le support premier
 d'observations plus méthodiques qui enrichissent le corpus de connaissance.

D'emblée, pour tenir compte des difficultés soulevées, notre démarche doit :

- rechercher **l'objectivité** (tout en sachant qu'elle ne peut-être totale) et donc se
 donner un cadre qui nous le permette, autant que possible ;
- distinguer et mentionner de façon claire tout au long du travail quelle
 orientation est utilisée de la relation au monde ; nous les appelons *Lien de Sens*⁹ (du
 monde tangible à sa représentation symbolique) et *Lien de Forme* (de la
 représentation abstraite vers le monde tangible) ;

8. où intervient plus particulièrement le raisonnement de type inductif.

9. Pour montrer qu'il s'agit d'un groupe nominal doté du sens précis que nous lui donnons ici,
 l'expression

- admettre que la connaissance ne peut embrasser le monde dans son entier ni que la lumière soit faite sur l’intégralité de notre relation au monde ; procéder par partie, par point de vue, par modèle partiel/incomplet et imparfait ;
- mais admettre toutefois que l’ensemble de ces difficultés ne nous empêche pas de poser un référentiel existant au-delà de notre seule relation (subjective, limitée, partielle) au monde et de la représentation (incomplète, imparfaite) que l’on s’en fait ;
- avoir conscience de la polysémie de nombreux termes et, pour y remédier, partager, le temps de la lecture, un même vocabulaire.

Cette vigilance ayant été mise en avant, le travail commence par poser le cadre conceptuel (2) où sera proposée l’hypothèse heuristique¹⁰ idoine qui structure la suite du développement. Il illustre en partie 3 l’ensemble des différents liens qui animent la relation au monde (*Lien de Sens* et *Lien de Forme*). Le cheminement du *Lien de Forme*, du monde systémique au monde perçu, est explicité (4) pour déboucher vers une maîtrise de l’observation (5) avant de discuter (6) ce qu’apporte notre hypothèse heuristique, tout particulièrement l’hypothèse de genèse d’empreinte dans la construction d’une science de l’observation, puis l’utilité de sa mise en pratique en observatoires société-milieu.

Enfin, nous posons l’intention de nous situer dès que possible dans le monde de la géographie et, cela va de soi, des sciences qui sont en relation avec elle. En effet, l’intention ultime de la géographie embrasse, de façon très large, l’étude de la relation de l’homme à son milieu.

2. Cadre conceptuel

2.1. Démarche scientifique

La démarche scientifique est une méthode qui élabore une représentation raisonnée du monde. Les figures 3 et 4 en précisent les processus abstraits et les concepts :

- la démarche dans son ensemble améliore autant que possible l’indépendance de l’objet observé du sujet observant, démarche vers l’objectivité ; le *cogito* et l’objectivation sont à l’avènement de la pensée et du raisonnement scientifique (Descartes *in* Moreau, 2016) ; au 17^e siècle, le chimiste et physicien Robert Boyle

¹⁰ A l’image de la carte heuristique, qui est une représentation schématisée et simplifiée de la réalité, permettant de visualiser les concepts et les relations entre ces concepts. Plus précisément ici : « Partie de la science qui a pour objet les procédures de recherche et de découverte » (*Le Robert*), qui analyse la découverte de faits. En psychologie, on parle de carte heuristique pour désigner une représentation schématisée et simplifiée de la réalité, permettant de visualiser les concepts et les relations entre ces concepts (*Mind Mapping*). Hypothèse heuristique : hypothèse adoptée provisoirement comme idée directrice indépendamment de sa vérité absolue (CNRTL).

précise qu'il faut s'effacer devant les faits en les rapportant de façon honnête, en se détournant de tout excès interprétatif ou spéculatif (Shapin et Schaffer, 2011) ;

– la part expérimentale de la démarche (Bernard, 1865, 1877) et la réfutabilité (Popper *in* Juignet, 2015) permettent le test d'une idée formulée sous forme d'hypothèse et de gagner en certitude, dans le cadre d'une théorie, de façon rationnellement justifiée ;

– l'ordonnement par la causalité et la théorisation (Krivine, 2018) élaborent une représentation raisonnée du monde et construisent la connaissance scientifique.

Par le travail sur soi (au sens individuel ou collectif au sein d'une société), ces processus cognitifs permettent l'élaboration progressive puis l'enrichissement de la connaissance scientifique par raisonnement-réflexion sur l'acquis, par son organisation en une mémoire ordonnée et partagée, en un savoir transmis par l'éducation. C'est la mise en situation (par confrontation au monde) de la connaissance, représentée à cet effet par un modèle (ou plus simplement par un énoncé hypothétique) lié à un questionnement, qui permet de progresser (en gagnant en certitude).

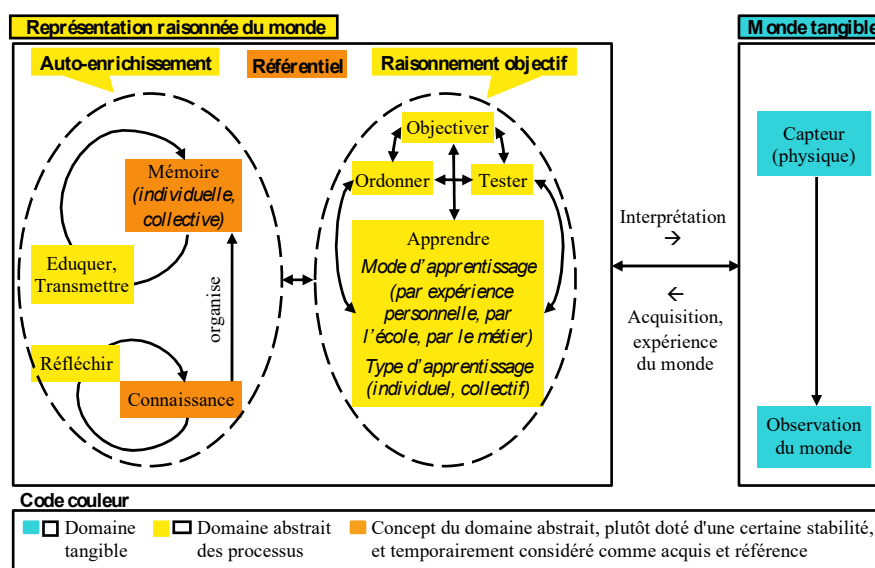


Figure 3. Raisonner et réfléchir - Vers une représentation raisonnée du monde

La démarche scientifique (figure 4) revient à une mise à l'épreuve volontaire d'un modèle et donc de la connaissance qui l'étaie. Mis à l'épreuve, confronté à des données lors d'un protocole dûment construit, le modèle, posé comme hypothèse, se

voit (on non) corroboré et la certitude attribuée à l’énoncé ou au modèle, augmentée ou dénoncée.

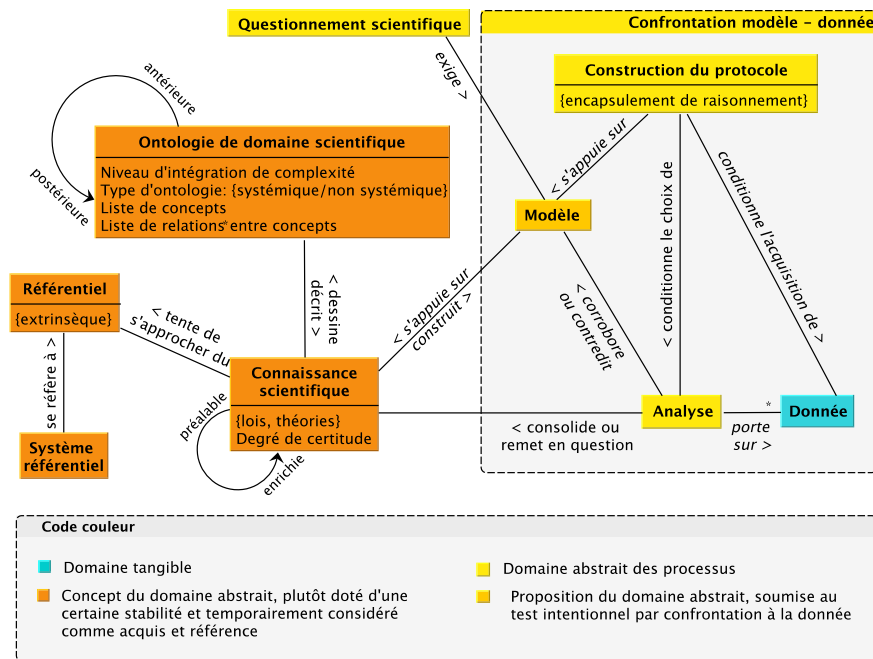


Figure 4. La connaissance et la démarche scientifiques

La démarche scientifique organise à la fois l’observation du monde (protocole d’observation, protocole d’analyse) et les acquis qu’elle en tire ; elle admet l’existence d’un **référentiel** (abstrait, conceptualisé), duquel la connaissance tente de s’approcher, progressivement au fil des études. La manifestation perçue du monde est liée à l’expérience d’observation que l’on en a, inféodée à nos sens et capteurs ; elle est donc manifestation filtrée, partielle voire déformée par nos sens et nos interprétations ou représentations. Nos expériences sont avant tout « incarnées » (cf. Merleau-Ponty *in* Bakewell, 2018). Sartre (*in* Bakewell, 2018), tout en participant lui aussi au courant phénoménologique¹¹, ne remet cependant pas en cause l’existence d’une vérité extrinsèque, à laquelle nos expériences sont confrontées sans pouvoir l’atteindre pleinement. Avec un positionnement ontologique différent, le mythe de la caverne de Platon témoigne aussi (par l’imperfection des formes projetées sur le mur de la caverne) de la quête d’un référentiel externe ; même si nous ne pouvons nous extraire de la caverne, la limite

11. Courant initié par Husserl au début du 20^e siècle (*in* Bakewell, 2018) puis largement discuté ensuite en particulier en France. Courant d’analyse de l’expérience individuelle en tant que reçue, ressentie, vécue et donc relative à chacun, subjective.

de l'entreprise est assumée, comme l'admet simplement Bouleau (2021). Ainsi l'approche scientifique ne signifie pas pour autant l'absence de présupposés ou de questionnements métaphysiques associés qui guident le travail (Bachelard, 1940), la formulation d'hypothèses à tester ensuite (*cf.* Popper *in* Lecourt, 2018), et peut-être même tout le contraire : « Si l'esprit peut changer de métaphysique, il ne peut se passer de métaphysique » (*cf.* Simondon *in* Lecourt, 2018). Dans la démarche scientifique, cette notion extrinsèque, indépendante et existant en dehors de nous, devient la « référence ». Cette démarche tente de dépouiller toute expérience de sa part subjective, et donc d'objectiver. Une cohérence s'échafaude, par raisonnement et réflexion, en modèle¹² pour décrire un système ; cohérence construite à partir des faits constatés ou étudiés ; elle ne relate pas des événements mais construit des relations, restitue des enchaînements (Rousseau, 1989 [1755]) et, par itération, elle est toujours remise à l'épreuve de nouveaux faits. Les faits corroborent (et non pas parfaitement vérifient) ou bien réfutent l'état des connaissances ou l'hypothèse testée (*cf.* Popper, *in* Juignet, 2015). Popper parle de « vérisimilitude », car on n'atteint pas la vérité absolue.

E ê ' ll l l
 mémoire en l' l l'
 l' l
 l' l - ,
 l ,
 (l' l l l ()
 ordonne des concepts ou des idées, mais aussi infère, propose des hypothèses, construit des modèles, fournit des interprétations) ; la connaissance scientifique se , l -dessus,
 l l' , l'abstraction (figure 3) ;
 celle-ci construit lois, théories¹³ et dessine ontologies de domaines scientifiques plus ou moins complexes (figure 4). Nous proposons de nommer ontologie systémique une ontologie fondée sur la connaissance de systèmes complexes, en particulier celle de leur structure-fonctionnement¹⁴ et de leur dynamique. Cette approche ontologique systémique, en tant qu'abstraction construite pour définir le référentiel extrinsèque,

12. Un modèle ou un ensemble de modèles, chacun, plus simple que le tout, ne parlant que , - chacun des mille points lumineux (étoiles émettant ou joyaux réfléchissant la lumière), dit quelque chose de la lumière (Serres, 2014).

13. Une « bonne » théorie non seulement explique, propose avec une vraisemblance l l' ll l' l : inconnu, et anticipe sur son observation ultérieure ; , l : e.g. l' l' l l N par Le Verrier, en utilisant la théorie de la gravitation de Newton ; hypothèse confirmée en 1846 par la première observation (*a posteriori*), de cette planète, non encore connue (Krivine, 2018).

14. Structure et fonctionnement sont les deux faces d'une même pièce, au moment de l'observation. On aborde le l' l' ... l l'on est par exemple morphologiste ou physiologiste, géographe ou écologue.

pourrait être rapprochée de la position de Leibniz (Triclot, 2005) qui recherche le « calcul parfait et ainsi toucher la vérité » ou à celle de Gonsseth (*in* Bachelard, 1950) qui réfléchit à l’exactitude en mathématiques.

Nous approchons et appelons **système** un référentiel dont on reconnaît, au plan des idées, l’existence objective par la « cohérence » qu’il instaure au niveau fonctionnel et évolutif. Il s’agit d’une organisation qui donne sens, dont l’existence (idéelle/abstraite) n’est liée à la contingence matérielle et à la contingence spatiale que dans un second temps (à partir du moment où le système est implanté, voir plus loin la figure 5).

Le concept système appelle un certain nombre d’autres concepts explicités par Fargette *et al.* (2019a) : le fonctionnement dont découle le concept de processus ; la structure dont découle le concept d’organisation ; la dynamique à l’origine du concept de changement et donc d’histoire ; le compartiment¹⁵, modèle extrait par tel point de vue sur le système, qui interprète, organise, décrit le système¹⁶. Ces concepts ont été explorés en les illustrant dans le cas des sciences biologiques et écologiques (Fargette *et al.*, 2019a) et, plus largement des sciences de la Terre et de la vie (Fargette *et al.*, 2018).

Le **modèle** (figure 4) permet de tester une hypothèse ou de vérifier ce que l’on connaît du système référentiel dont il décrit et représente de manière symbolique les caractéristiques et propriétés mises en perspective par un point de vue ; c’est un construit (de tête d’homme¹⁷) relatif à un point de vue et à la qualité de la connaissance sur laquelle il se fonde ; il représente, à un moment donné, une organisation d’un pan de savoir correspondant à une question posée ou au point de vue qui lui est associé. Descartes (*in* Moreau, 2016) aurait utilisé le terme « synthèse » à propos de la démarche qui le construit ; on peut aussi parler d’interprétation. Il est plus ou moins consolidé ou spéculatif selon le besoin et peut même être hypothétique ; à la base d’une question scientifique posée, il est alors la formulation d’une hypothèse scientifique. Sa manipulation permet de travailler une question selon un point de vue spécifique (Minsky, 1995). Les points de vue modélisateurs et les modélisations qui en résultent sont divers dans les outils et symboles qu’ils utilisent et les démarches qu’ils mènent ; même pour des

15. Le compartiment approche le système ; le compartiment « parfait » est système (Fargette *et al.*, 2019a).

16. Souvent, particulièrement dans le cas de systèmes complexes et malgré la conscience de la nécessité d’une vision globale, on se contente d’une vision partielle, d’une vision d’une seule facette et propriétés en une seule vision ; la notion de point de vue est incontournable, à laquelle correspond le concept de compartiment ; le système est représenté, de ce point de vue-là, par un modèle.

17. Le relativisme culturel apporte une complexité supplémentaire à la notion de point de vue et la représentation qui en découle. Sans occulter de tels écarts de points de vue, liés à la diversité intra-inter-culturelle, il faut reconnaître que la notion de point de vue est relative à la partie vs totalité » de ce relativisme et à la diversité des perspectives induites par celle des points de vue.

questionnements apparentés, ils montrent leur diversité par le domaine de connaissance que chacun traite ou le focus qu'il adopte. Un concept peut lui-même être décrit par un modèle¹⁸ ayant opté pour un focus particulier (point de vue). Une ontologie est un modèle qui ordonne des concepts selon certains critères, par exemple ' systémique.

La connaissance scientifique dont la forme de savoir est légitimée par la démarche scientifique, construit un monde intelligible et une interprétation (une proposition d'interprétation), un ensemble de théories ; exprimées sous forme de modèles qui décrivent/définissent/« disent » (approche par modélisation, en l'état de la connaissance) le système référentiel ou bien énoncent des hypothèses à son propos. Il existe une **dialectique entre modèle et système** et il y a parfois redondance ou au contraire hiatus entre eux selon l'état de la connaissance, selon les points de vue, toujours parcellaires. L'identité exacte système - modèle verrait l'objectif atteint, la réponse trouvée, sans erreur, le modèle s'identifiant alors à la limite vers laquelle il tend.

C'est sur le relief qu'apporte, telle la vision binoculaire, cette dialectique que se fonde la (l'acquisition de) connaissance et sur ses aspérités (mise en évidence de contradictions ou de vacuités) que naît le questionnement scientifique, que s'élaborent interprétation et hypothèses à tester. La démarche scientifique n'est jamais à l'arrêt. Ce travail laborieux pour améliorer le résultat est peut-être à rapprocher des notions défendues par Rousseau (1989 [1755]) de conjecture (conjecturer n'est pas imaginer mais « tirer de la nature des choses ce qui est le plus probable ») et de perfectibilité (en tant qu'elle permet à l'homme de s'améliorer, ici dans son entreprise de connaissance).

La connaissance scientifique ne peut être certitude incontestable. Par contre, selon Descartes, Leibniz et nombre de scientifiques, elle ne peut que viser la vérité, qu'avoir pour intention d'approcher la/une vérité¹⁹ ; admettre « une » et non « la » vérité revient à adopter « un » modèle pour en parler (à partir « d'un » certain point de vue) et à actualiser ce modèle, à renouveler si nécessaire le point de vue, voire changer de paradigme. La connaissance scientifique, une fois « acquise » (testable, testée et re-testable), est transmise et devient collective : elle correspond à « un » nouvel état de connaissance orienté, avec un certain de degré de certitude, vers « le » référentiel et, en l'état, corpus de connaissance partagée, non dogmatique, toujours possiblement remise en cause, quand cela s'avère nécessaire.

En résumé, la démarche scientifique (figure 4) se construit à partir d'une question de recherche plus ou moins large. Elle exige un modèle formel²⁰, construit

18. par exemple, l'espèce, la cellule ou la division cellulaire, l'écologie, le travail, le concept d'épure, modèle géographique (*cf.* partie 4).

19. Est-ce que le concept de vrai, ne serait-ce que en tant que limite ?

20. Un modèle formel peut être conceptuel, logique, mathématique, etc.

sur la base de la connaissance scientifique disponible et formulée sous forme d'ontologies dont le niveau d'intégration reflète la complexité. Le modèle mobilise et ordonne les seuls concepts et relations pensés et posés comme pertinents étant donné le focus (objectif, échelle) pensé et posé comme adapté à la question ; il synthétise une hypothèse complexe. Érigé sur la connaissance préalablement acquise, il conduit aux protocoles adéquats d'observation et d'analyse (modes d'échantillonnage, de collecte et d'analyse des données) qui concrétisent le raisonnement et explicitent l'action (le raisonnement est encapsulé dans le protocole). L'analyse porte sur des données observées. Les résultats de l'analyse corroborent ou contredisent ce modèle. En conséquence, la connaissance scientifique qui lui est liée s'en voit consolidée, enrichie ou bien au contraire réfutée, invitée à la révision.

Ainsi le modèle, par les choix qu'il fait, pose une hypothèse sur l'organisation ou le fonctionnement, propose une interprétation dont on accepte cependant la remise en cause lors d'un test qui le confronte à une (nouvelle) expérience du tangible, à un (nouveau) jeu de données.

L'ensemble est bien sous-tendu par une représentation raisonnée du monde et l'abstraction qui l'accompagne : au monde complexe que nous observons correspondraient des systèmes (complexes²¹ et des modèles complexes proposant des points de vue, des interprétations, des hypothèses) tandis que la dialectique modèle-système permettrait de poser le test. Les propos de Krivine (2018) « trouver comment approcher au mieux une fonction idéale définie dans des espaces de très grande dimension », de Jean Perrin (1913, *Les Atomes*, in Krivine, 2018) « La science remplace du visible compliqué par de l'invisible simple », sont autant de témoignages de cette démarche, chacun usant d'un vocabulaire, d'une image, etc. pour parler de « l'âme du fait » défendue par Poincaré (in Bouleau, 2017) car, au-delà du constat du simple fait, c'est comme si, dans un moment de fascination, et avec le vocabulaire utilisé ici, le système en soi et complètement compris d'une part et la compréhension complète du système d'autre part se croisaient dans la réflexion pour, un temps, ne faire qu'un [« On les croyait des "buts" » ils deviennent des moyens. », cf. Bachelard, 1950]. Et n'est-ce pas cela que nous recherchons ? Le « système complètement compris » s'approcherait alors du système référentiel [comme le calcul parfait des mathématiciens, l'approche holiste des systémiciens, la vérité des philosophes, l'exactitude²² discursive de Gonseth, cf. in Bachelard, 1950]. Même si nos modèles scientifiques imparfaits reflètent nos interprétations et représentations, ils tendent vers ce référentiel alors même que nous ne le saisissons jamais complètement [« une éternelle reprise des sources du vrai », cf. Bachelard, 1950].

21. agrégations et/ou emboîtements de systèmes élémentaires ou complexes en interactions, tout.

22 d'exactitude (Bachelard, 1950).

2.2. Existence géographique

En s'appuyant sur le cadre formel système - modèle de la démarche scientifique et en déclinant ci-dessous la réflexion dans le cas de la géographie, il est posé le concept d'existence²³, en tant que concept partagé par toute science quand elle déploie ses lois, ses raisonnements, ses outils, et le concept de réalité²⁴ dans le cas de l'existence géographique. La figure 5 et l'encadré 5 proposent un modèle général où sont introduits et précisés le concept d'existence, l'existence géographique et la réalité de l'instance planète Terre.

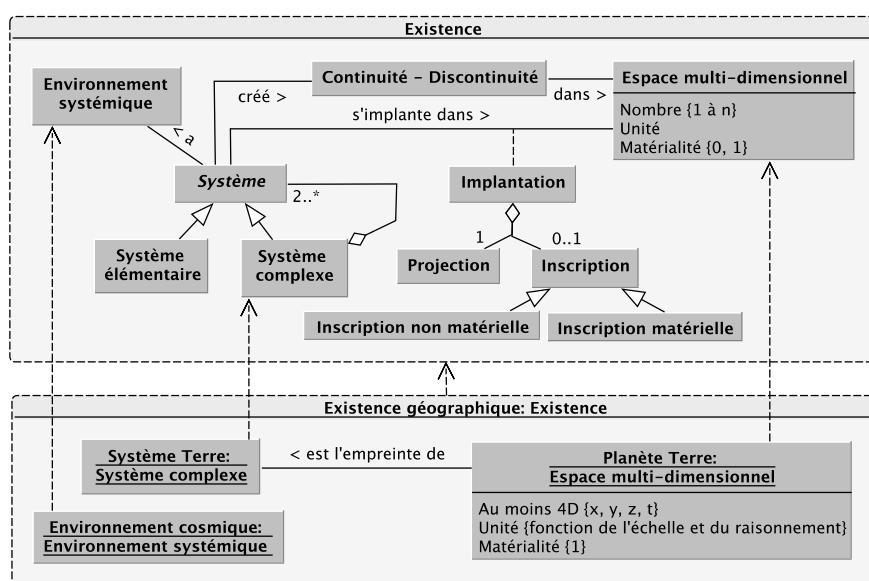


Figure 5. Existence géographique et réalité de la planète Terre (voir encadré 5)

Dans le cadre ainsi défini, système Terre est une instance de système complexe, fonctionnant et évoluant (dynamique) dans son environnement cosmique, lui-même instance d'environnement systémique. Une fois le système Terre implanté, avec toute sa complexité dynamique et fonctionnelle, en planète Terre, les processus systémiques se déploient et, par voie de projection et inscription, la planète Terre devient le théâtre des phénomènes énergétiques et géographiques leur correspondant et qui s'y inscrivent. La matérialité serait donc le principe à l'origine de la réalité,

23 ' ' 1 concret » ; les symboles, les idées existent.

24 1 1 ' II apteur, une forme « en soi », non encore perçue ni pensée par quiconque. De composante matérielle, la réalité est référentiel concret (cf. partie 2.4).

inscrite dans un espace appréhendé à l'échelle planétaire ou aux échelles infra-planétaires et possédant quatre dimensions (x, y, z, t) dites spatiotemporelles.

Encadre 5. Existence géographique et réalité de la planète Terre (voir figure 5)

Un système, en tant qu'organisation abstraite, est en capacité de fonctionner dans son environnement systémique, du fait de sa qualité de cohérence (viabilité) et selon les principes qui le constituent. Par son implantation dans un espace multidimensionnel, sa structure et son fonctionnement sont rendus effectifs et attestent de son existence dans cet espace. L'existence tient à l'implantation, que celle-ci consiste en projection seulement ou, facultativement en plus, en inscription ; la projection relève de l'instantané tandis que l'inscription, matérielle ou non, rend possible la mémorisation (*i.e.* au-delà du labile).

Ce faisant (*i.e.* dès qu'il est implanté), le système créé des continuités et des discontinuités dans l'espace et les limites où il s'inscrit.

La connaissance scientifique conçoit progressivement et affine le modèle de (chaque) système par l'étude des caractéristiques et des continuités - discontinuités de l'espace lui correspondant (domaine de définition, nombre de dimensions et limites ; selon quels registres et unités ? Parmi celles-ci, et pour ce qui nous intéresse ici, la dimension temporelle intervient-elle ? ' ainsi conçu fait-il appel à une matérialité ou bien reste-t-il dans le monde de l'abstrait ?).

La géographie constate et étudie un cas particulier d'existence (l'existence géographique), dont le sens est ontologiquement attaché au système Terre, instance de système, et à sa logique, imprimée dans la réalité de la planète Terre. Les points de vue qui y sont portés sont divers, portant sur son fonctionnement (selon différentes composantes/points de vue : mécaniques, énergétiques, etc.), sa matérialité, son histoire et ce à différentes échelles et selon divers focus.

Nous positionnons le concept d'empreinte :

- la planète Terre, inscrite dans un espace quadridimensionnel, est l'empreinte du système Terre ;
- l'empreinte est associée à la fois au concept d'inscription matérielle et à la notion de mémoire qui l'accompagne (Viers, 1967 ; Leroy-Gourhan, 1965).

Une fois le système Terre implanté, projeté et (donc) inscrit dans la matérialité, l'empreinte résulte du fonctionnement du système et de ses changements de fonctionnement au cours du temps. L'empreinte est non seulement reflet de l'instantané²⁵ mais elle peut aussi enregistrer, retenir²⁶ (ou effacer) les événements et ainsi révéler un passé, une histoire.

25. Empreinte instantanée : *e.g.* vue par satellite des formations cycloniques (nébulosité, température, pression, vent).

26. Empreinte avec mémorisation : *e.g.* les formes géologiques, les biomes, etc., et leur répartition sur la planète.

Les sciences géographiques, adossées à d'autres sciences de la Terre, de la vie, des sociétés, s'intéressent à ce qu'un espace de ce monde révèle du système Terre.

2.3. Ontologie de domaine : cas de la classification pour ordonner la connaissance et pour aider à reconnaître

La méthode de classification présentée en figure 6 (et encadré 6 ci-après) peut être utilisée pour proposer un ordonnancement de la connaissance et ainsi aider à en définir le contenu ; elle consiste à regrouper des instances autour de concepts, selon une organisation hiérarchique, celle-ci fondée sur le partage organisé (ici emboîté) de leurs attributs ou propriétés. Les classes ainsi constituées et organisées correspondent à un point de vue, en l'occurrence le choix des critères retenus et de leur poids, qui détermine leur place dans la hiérarchie. Pour procéder, elle se fonde sur l'analyse et la comparaison des instances que l'on veut classer et sur les choix préalablement faits (*a priori*), eux-mêmes fondés sur de la connaissance plus ou moins assise et une théorie plus ou moins solide.

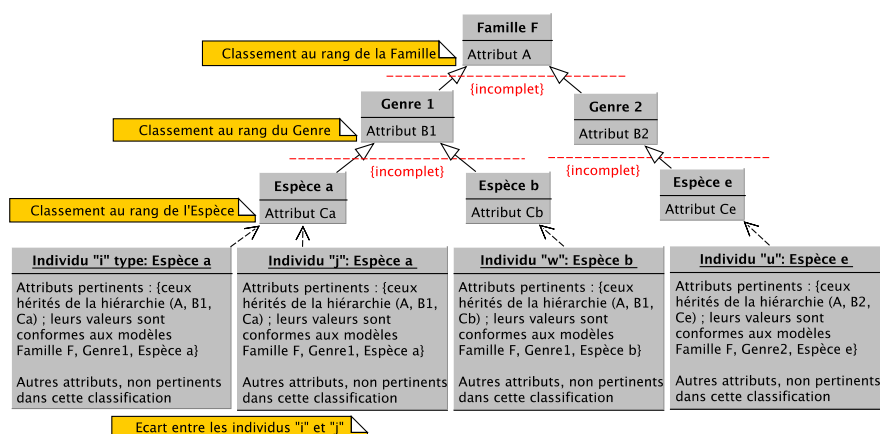


Figure 6. De l'ordonnancement de la connaissance à la reconnaissance - Classification et identification (voir encadré 6)

Par exemple, dans les sciences biologiques, sur le fonds d'étude et de description d'individus, la méthode de classification commence par l'observation, la comparaison d'individus et leur regroupement guidé par leur ressemblance²⁷ : caractères partagés (approche structurale) et/ou leur héritage commun (approche dynamique, évolutive). Il s'agit de la phase de construction d'une classification où les individus ainsi regroupés appartiennent par exemple à une même espèce (dont ils

27. où s'inscrit une variation « naturelle » (figure 6).

sont des instances) et donc à une lignée ontologique en remontant les rangs de la classification. La classification des organismes, ensuite l’arbre évolutif du vivant, ont été et sont encore en permanente construction et révision ; à un moment donné, chacun est le reflet de l’état des connaissances selon ce point de vue. Une classification à la fois bien menée (méthodologie) et menée sur un objet d’intérêt et selon un point de vue pertinents, présente un contenu doté de sens et reflète un arbre du vivant référent, approché par un certain état de connaissance. En particulier une connaissance de type systémique (en l’occurrence ici celle relevant de la biologie et de l’évolution) infuse ses acquis de connaissance dans les classifications de cet exemple. Réciproquement, sur la base de la connaissance contenue et ordonnée dans une telle ontologie et en reconnaissant le sens attaché aux concepts biologiques d’espèce et d’évolution, on peut assigner à une classe de rang espèce, et par là identifier et nommer, des individus observés selon le protocole *ad hoc*.

**Encadré 6. De l’ordonnancement de la connaissance à la reconnaissance -
Classification et identification (voir figure 6)**

L’ensemble des attributs d’une classe se range :

- ou bien dans le « α » partagé par toutes les classes appartenant à une même classe supérieure dans la hiérarchie ;
- ou bien en « corolle », correspondant à certains critères (de type caractère ou modalité de caractère) qui au contraire apportent des éléments de distinction entre ces mêmes classes.

A chaque rang de la classification, les classes s’organisent et s’emboîtent selon une hiérarchie fondée sur le partage des caractéristiques- α

En descendant dans la classification 1 α (l’ensemble des attributs partagés par l’ensemble des membres de la classe), est toujours plus grand. En représentation des connaissances, la hiérarchisation opère par spécialisation de concepts. En biologie, par exemple, on parle de taxons de rang espèce, genre, famille, ordre, etc. en remontant les rangs de la classification.

La mesure et la comparaison des attributs entre individus d’une même espèce dans le monde réel (instances) révèlent un écart (statistique) donc une diversité, une dispersion possible autour de la valeur constatée chez l’individu de référence appelé type. La nature et l’origine de la variation sont importantes à connaître et vont être discutées en partie 4 (*cf.* figures 11 et 14).

Réciproquement, à partir de donnée(s) observée(s)/mesurée(s) sur un individu, l’assignation de celui-ci à une classe-Espèce (c’est-à-dire l’identification dans le contexte de la classification en jeu) concentre l’observation sur les caractères pertinents (les attributs de classe correspondante (Il (α ’une classe de rang supérieur. Dans le même temps, elle prend en compte l’écart au type au moment de l’assigner, ou pas, à telle classe et donc d’en faire une de ses instances. Une telle décision est assistée par modèles statistiques ou probabilistes.

D’autre part, selon le rang de classement choisi et donc l’ensemble des attributs dits « de α » qui le définit (et dont dépend le protocole d’observation), deux instances seront assignées à la même classe ou au contraire, seront dans des classes différentes. Par exemple,

l'observation est celui de l'Espèce a , soit « A, B1, Ca », les individus i et j seront identifiés en tant qu'instances de l'Espèce a , tandis que les individus w l'observation est celui du Genre 1, soit « A, B1 », alors les individus i , j et w relèveront de la même classe [taxon] (Genre 1) tandis que l'individu u restera distinct.

Où P est l'ensemble des critères de discrimination et, partant de là, les caractéristiques à observer et le résultat de la démarche d'identification : l'observé. Nous reviendrons sur ce point dans la suite du travail.

Une classification serait exacte d'une part, si elle représentait au plus près le référentiel ; d'autre part, elle éviterait plus d'ambiguïté dans l'identification (*i.e.* l'assignation d'un nom à un individu) ; les requis en termes de clarté et distinction seraient alors remplis (Descartes *in* Moreau, 2016).

Le recul dans le domaine des sciences biologiques (des sciences de la systématique à celles de l'évolution) donne un relief particulier à la démarche de classification, notamment par l'attention portée au cadre conceptuel, aux types de caractères retenus (Barriel et Bourgoïn, 2000), et sur lesquels portent les « observations », et aux soubassements conceptuels, mathématiques ou statistiques (Page et Holmes, 2004 ; Baele *et al.*, 2018) pour construire les classifications et les « arbres ». Ce recul met en lumière comment, au fil du temps et du questionnement scientifique, non seulement l'amélioration de la capacité d'observation et d'analyse, mais aussi les ajustements, voire les révolutions d'interprétation qui définissent le cadre théorique, façonnent et affinent le modèle approchant le système ; le concept systématique d'évolution et la théorie de l'évolution du vivant sont devenus décisifs, en tant que paradigme pour décrire, interpréter, comprendre le vivant (Lecointre et Le Guyader, 2001), pour donner les bonnes clés d'organisation des taxons (élaboration d'une ontologie issue de la systématique) et réciproquement d'identification d'individus (assignation à une classe, appartenance à un taxon).

En observation de la Terre, domaine scientifique et technologique encore récent (Cazenave, 2013 ; Kugler *et al.*, 2019), la terminologie de classification est aussi utilisée, dont la méthode reflète au moins en partie cette approche vieille de plusieurs siècles dans le domaine de la systématique en biologie (Linné, 1758 ; Fischer et Tassy, 2014), de quelques décennies dans celui de l'évolution du vivant (Lecointre et Le Guyader, 2001) et de quelques millénaires dans la pensée aristotélicienne (*cf.* Aristote *in* Dhondt, 1961). En télédétection par exemple, on observe des portions d'espace géographique. Par l'analyse de données spatialisées (leur contenu, leur organisation et leur dynamique spatiale) dans une image satellitaire par exemple, on souhaite, *in fine*, distinguer, qualifier sur des polygones des témoins de structures [*e.g.* *landcover*, forêt, ville, etc. *cf.* Lucas *et al.*, 2019], ou de matière (Lane *et al.*, 2015), de processus [*e.g.* production agricole et ses

modalités, cf. Fargette *et al.*, 2017] ou de dynamique temporelle [e.g. déforestation, cf. Village Daza *et al.*, 2020].

Il est intéressant de noter aussi que Rousseau (1989 [1755]) évoque les premiers tâtonnements des hommes qui ont découvert la possibilité de catégoriser le réel. Il envisage la problématique de la « bonne généralisation » dont dépendent, selon lui, les réussites ou les échecs de la science comme de la philosophie.

2.4. Hypothèse heuristique et démarche d'ensemble

L'expérience du monde (cf. figure 1) et la connaissance que l'on en acquiert (cf. figures 2 et 3) relient et confrontent dans des processus d'aller-retour complexes et itératifs un bipôle : le monde tangible tel que reçu de notre expérience d'observation et le monde représenté, ordonné selon notre découverte et notre compréhension. Pour comprendre le monde et notre relation au monde par une approche systémique, pour tenir compte aussi des difficultés de la problématique d'observation citées en introduction et des pistes mentionnées pour y remédier (cf. point 3 en partie 1.2) et convenir au cadre conceptuel présenté ci-dessus, nous proposons :

- de remplacer le bipôle initial (monde tangible-monde représenté) par un quadripôle [modèle - système - réalité - monde perçu] dont nous détaillons ci-dessous chacun des pôles et leurs associations ;
- de découpler pour un temps le processus d'aller - retour présenté dans ce bipôle pour considérer, séparément et l'un après l'autre, le *Lien*²⁸ de Sens et le *Lien* de Forme.

Ainsi, le temps de ce travail, l'hypothèse idoine (cf. l'idonéisme²⁹ de Gonseth) est proposée (figure 7), qui constitue notre hypothèse heuristique (cf. note 10) entre tout et partie, entre abstrait et concret³⁰ (elle est organisée selon les deux axes de la boussole posée sur la figure 7) : 1) verticalement : le tout (version holiste : système et réalité, registre du haut) *vs* la partie (version point de vue : modèle et monde perçu, registre du bas) et 2) horizontalement : l'abstrait (système et modèle, registre de gauche) *vs* le concret (réalité et monde perçu, registre de droite).

Les concepts de système, modèle, système référentiel ont été introduits en partie 2.1 ; réalité et empreinte, en partie 2.2. La distinction entre le concept de monde perçu (mentionné en partie 1) et celui de réalité permet de rendre compte de

28. « Lien » est défini dans le sens du Centre national de ressources textuelles et lexicales (<https://www.cnrtl.fr/definition/lien>): « un élément qui réunit, rattache deux choses entre elles, assure leur relation, les met en rapport ; le rapport lui-même ».

29. « L'idonéisme se situe dans une « IV 47 in Bachelard (1950) » [Gonseth,

30. Il est intéressant de noter que Bachelard (1938) considère comme « scientifique la relation concret-abstrait ».

- 1 **l'observation** : « le monde perçu témoigne de la réalité » et réciproquement « la réalité se manifeste dans le perçu (*i.e.* les données

observées) ». Toutefois le monde perçu restitue une partie seulement, incomplète et biaisée (filtrée, déformée) de la réalité. En prenant conscience des contraintes « incontournables », *a priori* (les préjugés selon le

l D l , l'
en repoussant le référentiel au- l l' l
maîtriser l l , l l adéquats. Cette maîtrise
l l' ;

– entre modèle et monde perçu, la **confrontation** : « ce qui est perçu du monde est formalisé³¹ en un modèle qui explicite une organisation » (modèle systémique conçu à partir de données analysées) et réciproquement « un modèle prédit, décrit (et éventuellement recherche dans les données) ce que l'on pourra observer » (entre une connaissance modélisée voire même théorisée et le monde perçu, modèle de préfiguration de la donnée « attendue », qui peut assister la construction de protocole d'observation – cf. parties 4 et 5) ;

– entre modèle et système, la **compréhension par le raisonnement** (cf. système compris) passe par l'enrichissement progressif de la connaissance et par la théorisation ; l'intégration vers le tout gagne en complexité et tend vers le système référentiel : « le modèle décrit une proposition d'organisation/propose une interprétation selon une vision partielle du système (*i.e.* un point de vue) » ; « il enrichit et s'intègre dans une vision plus large » ; et réciproquement « le système plus complexe est à l'origine de la proposition de modèle en tant qu'hypothèse à tester, à confronter au monde perçu et, par-là, à la réalité » ;

– entre système et réalité, nous proposons le processus de « genèse d'empreinte » (concept explicité dans ce travail mais voir aussi Fargette *et al.*, 2018, pour sa première mention) qui tient à l'implantation et au fonctionnement du système référentiel : « Le système se projette et s'inscrit dans la réalité » (ce faisant il génère une empreinte) et réciproquement « l'empreinte révèle l'activité systémique ». Wittgenstein (1921, *Tractatus*, in Lecourt, 2018) pose une

« "ressemblance" de nature logique – un isomorphisme – entre le "tableau" que constitue la proposition » [ici, le système référentiel] « et la réalité. Ainsi la logique est le "miroir du monde". »

En conséquence, les concepts d'empreinte et de passage peuvent être définis comme suit.

– **Empreinte** : Précédant la diversité des points de vue mentionnée plus haut, ce qui se donne³² à voir a pour origine un système (complexe), celui-ci étant générateur d'empreinte dans la réalité du monde. La genèse d'empreinte procède du fait du système, indépendamment du sujet³³ qui ensuite observe, du capteur qui capte. L'empreinte préexiste à toute observation. L'empreinte est donc inscription dans le

31 S l V l' l l' l (MOOC « Renaissance en Val de Loire », université de Tours).

32 , l , de tangible.

33 Q ' l l l

référentiel concret (réalité) et découle d'une logique systémique relevant du référentiel abstrait (système référentiel). La réalité peut être vue comme un agrégat d'empreintes, chacune relevant d'un système à un niveau d'intégration donné.

La relation système - empreinte proposée ici est à rapprocher de celle entre système et « physionomie du lieu » défendue dans le domaine de la géographie (Claval, 1974 ; Brunet, 1974 ; Brossard et Wieber, 1984) comme dans celui de disciplines qui lui sont liées, tout en mettant un accent sur l'aspect plutôt systémique, comme par exemple l'agronomie (Deffontaines, 2004 ; Michelin, 2008). En conséquence, suivre l'évolution d' , ou de façon plus générale l'évolution de tout système socio - bio - physique d'une part, et le suivi du changement de la physionomie des espaces concernés d'autre part, sont deux points de vue sur la même relation.

– **Passage :** Le « passage » tel que posé en figure 7 est donc un raccourci du monde perçu vers la logique systémique qui lui est liée. Cette simplicité apparente ne doit en rien faire oublier la complexité à laquelle il s'adosse. Le passage, thème central de l'heuristique, entre les mondes abstrait et concret peut être parcouru dans un sens ou dans l'autre (*Lien de Sens* et *Lien de Forme*). Lorsqu'il est explicité en modèle, il propose une interprétation qu'il est possible de tester. Sinon, il demeure simple concordance (corrélation, covariation ou autres associations selon le mode de « calcul » adopté).

Nous constatons que l'hypothèse heuristique proposée en figure 7 mobilise la confrontation scientifique et la compréhension progressive vers une connaissance intégrée du monde (discutées en partie 2), qu'elle introduit le concept de genèse d'empreinte et celui d'empreinte qui en découle, et qu'elle positionne la science de l'observation pour maîtriser l'observation.

Nous proposons d'illustrer, en partie 3, la genèse d'empreinte de notre hypothèse heuristique à partir d'exemples en situation oasienne. Ce faisant nous détaillons un exemple de *Lien de Sens* en partie 3.1.1 (construit ici selon un raisonnement inductif) ; nous examinons le bien fondé du concept d'empreinte en partie 3.1.2 (selon un raisonnement hypothético-déductif et la confrontation modèle - monde perçu) ; une fois l'hypothèse d'empreinte corroborée en partie 3.1.2, nous introduisons en partie 3.2 le modèle géographique « Épure » et par là l'illustration du *Lien de Forme*.

Le *Lien de Forme* sera plus méthodiquement exploré en partie 4 (selon une démarche discursive³⁴) pour synthétiser en partie 5 certains aspects du contexte systémique auxquels se confronte l'observation et qui conduisent, en partie 6, à une science de l'observation.

34. Qui procède par raisonnements successifs (opposé à intuitif) : cf. Bachelard (1950) commentant Ferdinand Gonseth.

3. Illustrations en oasis

Les illustrations s'appuient sur un travail réalisé sur l'oasis de Nefta (sud de la Tunisie) précédemment publié (Fargette *et al.*, 2017, 2019b).

3.1. Du monde perçu à l'abstraction : le Lien de Sens et le concept d'empreinte

L'intention n'est pas de revenir ici sur le détail des protocoles d'acquisition et d'analyse des données, ni sur celui du contenu précis de la typologie des systèmes de culture, de la description des types ou sur les chiffres de la performance d'assignation (voir pour cela le contenu de Fargette *et al.*, 2017). Nous présentons plutôt la démarche suivie (le raisonnement et les opérations), conformément au cadre formel que nous nous avons donné et en quoi elle rencontre l'hypothèse heuristique (*cf.* figure 7).

3.1.1. Illustration d'un Lien de Sens en oasis

Pour répondre à la question « Quels types d'agriculture en oasis ? », une expérience d'observation et son analyse ont été menées par un agronome, en tant qu'expert de ce domaine scientifique, pour aboutir à l'établissement d'une typologie de systèmes de culture (Fargette *et al.*, 2017). En suivant les éléments de démarche scientifique présentés en figures 4 et 6, la figure 8 illustre le raisonnement et les opérations effectuées.

Il y a une diversité de types de systèmes de culture en oasis, qui sont liés à la diversité des types de fonctionnement, comme autant de spécialisations du système de culture. Chaque type résume un ensemble cohérent de traits en lien avec les équipements, les savoirs et les actions déployés dans les parcelles agricoles de l'oasis.

Raisonnement

– La question « Quelles agricultures en oasis ? » (figure 8) est abordée par le concept de système de culture ; le point de vue choisi sur le système est celui de la « typologie » ;

fonctionnels, diversité systémique organisée qui conjugue les traits en un fonctionnement effectif sur chaque parcelle agricole. La typologie rend compte de la diversité des types de fonctionnement, comme autant de spécialisations du système de culture. Chaque type résume un ensemble cohérent de traits en lien avec les équipements, les savoirs et les actions déployés dans les parcelles agricoles de l'oasis.

– Cette proposition s'appuie sur la connaissance agronomique qui, lors d'une approche systémique en exploitation agricole, croise, dans une ontologie du domaine (*cf.* figure 4) agronomique, des aspects biologiques, environnementaux et socio-

(liées). Toute exploitation agricole présente un ensemble de traits fonctionnels qui sont autant de propriétés plus ou moins diversifiées, spécifiques et

cohérentes pour assurer son fonctionnement et sa longévité. Une exploitation agricole, par son « profil fonctionnel » se rattache à un type de cette classification.

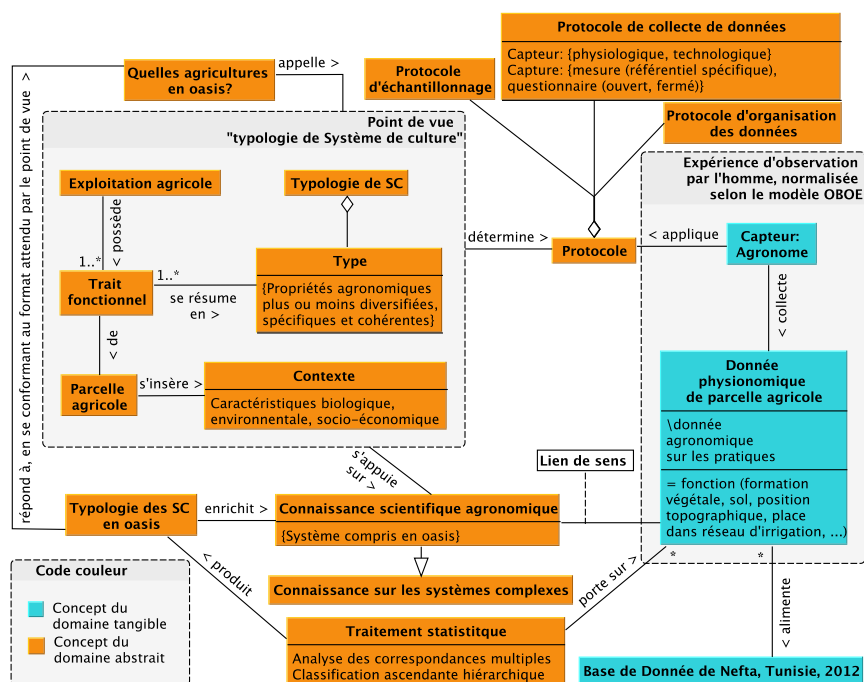


Figure 8. Acquisition d'un Lien de Sens en oasis

– Les protocoles d'échantillonnage et de collecte des données tiennent compte de la connaissance disponible sur les systèmes en jeu et le contexte où ils se déploient, ce qui implique aussi temps et espace. Façonnés par la connaissance experte que l'agronome a des facteurs et des interactions pertinents pour le système de culture en jeu, les protocoles suivent son point de vue pour expliciter les jeux de données à acquérir ; le protocole d'organisation des données peut aussi prévoir la structure d'une base de données et l'explicitation des métadonnées, ici encore à partir de la connaissance experte de l'agronome. Ceci est un préalable à la collecte-même des données qui prend ensuite une forme adaptée au contexte (fiche d'observation directe sur le terrain, grille d'entretien avec les acteurs).

Opérations

– L'expérience d'observation (figure 8) : la collecte des données (jeu de données D1) a eu lieu en l'oasis de Nefta, de mars à juin 2012 ; elle a été guidée par les protocoles et menée par le capteur, en l'occurrence un homme, praticien formé à cet exercice ; elle porte sur la physionomie des parcelles agricoles et repose aussi,

fondamentalement, sur la connaissance et sa modélisation (l'ontologie du domaine agronomique, contenue dans l'expérience de l'expert). La question modélisée selon le point de vue « typologie de système de culture » est à l'origine de la collecte des données.

L'ensemble est décisif pour structurer la base de données. La donnée collectée respecte aussi le format de l'ontologie de l'observation OBOE³⁵ (Madin *et al.*, 2007).

– Le traitement des données : l'analyse multivariée des données distingue, organise la diversité que déploie dans l' , en l'occurrence l' , dans des parcelles agricoles en classant les parcelles selon le bouquet de pratiques³⁶ qui y est constaté. Sont ainsi dégagés des types, par classification. L'organisation fonctionnelle de chacun des types est synthétisée, à partir des traits fonctionnels décrits tels qu'en témoigne l'observation des pratiques dans les parcelles et regroupés sous un même type. Ce résultat enrichit la connaissance agronomique en fournissant, en oasis, un ordonnancement dont on pourra ensuite vérifier la pertinence. Il est assigné un type à chacune des parcelles où a été menée la collecte de données.

La typologie des systèmes de culture oasiens ainsi obtenue est le résultat d'une expérience cognitive qui répond à la question initialement posée. Il s'agit d'une classification faite de spécialisations, qui s'inscrit dans celle, plus large, des systèmes de culture en général, une ontologie « simple », c'est-à-dire une hiérarchie arborescente, faite d'emboîtements. En reprenant les termes de l'hypothèse heuristique (*cf.* figure 7) : le monde agricole oasien, perçu par un agronome (qui récolte un jeu de données D1), est formalisé en un modèle systémique qui s'intègre dans une compréhension plus large des agrosystèmes. Entre monde perçu et système, nous venons d'acquérir de la connaissance, de parcourir un *Lien de Sens*.

3.1.2. Le concept d'empreinte : mise à l'épreuve en oasis

En suivant la démarche scientifique présentée en figure 4, la figure 9 illustre en oasis le raisonnement et les opérations effectuées pour confronter l'hypothèse d'empreinte aux données.

S l l l l l
du point de vue « typologie des systèmes de cultures » utilisé en partie 3.1.1 pour construire la typologie en oasis.

Raisonnement

– S1 l' , (cf. figure 7), des jeux de données, indépendants (monde perçu) et portant sur la même empreinte (réalité), devraient

35. Ontologie formelle à un haut niveau de conceptualisation / généralisation explicitant la l' : <https://github.com/NCEAS/oboe>.

36. en tant que témoins de traits fonctionnels.

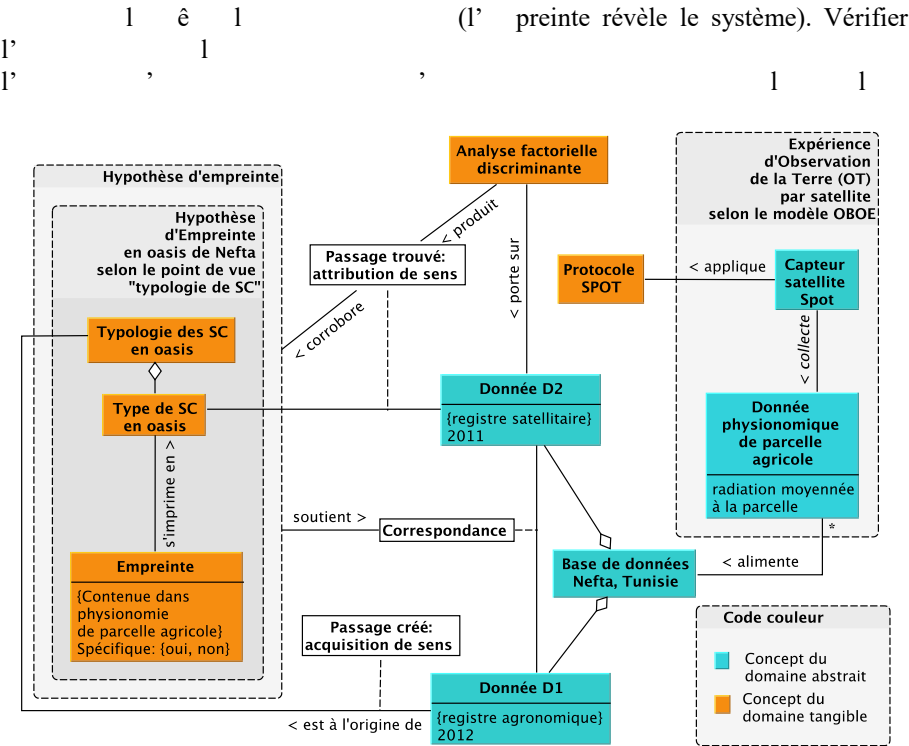


Figure 9. Mise à l'épreuve en oasis du concept d'empreinte

Dans le cas de l'oasis, la logique systémique contenue dans la typologie de systèmes de culture devrait être contenue dans l'empreinte de chacun des types de systèmes de culture, une signature leur correspondre dans la réalité, que l'on devrait retrouver dans la physionomie³⁷ des parcelles agricoles (monde perçu). Dans le cas le plus net, il y aurait spécificité (signature au sens le plus strict³⁸) : à un type de système de culture serait associé, sous un mode d'observation donné, une seule physionomie et *vice versa*.

– De plus, les jeux de données correspondant à deux modes d'observation indépendants des parcelles agricoles portent sur la même empreinte (issue du même système de culture en place) et de ce fait devraient manifester la même logique systémique (*i.e.* ne pas être indépendants l'un de l'autre).

37. cf. encadré 1, cf. la « physionomie des lieux » introduite en partie 2.4, à la suite de la

38 O , cf. vocabulaire partagé.

Plus précisément, nous proposons de confronter la représentation par typologie des systèmes de culture (en l'occurrence la classification en types, obtenue ci-dessus à partir du jeu de données D1, partie 3.1.1) à un deuxième jeu de données (D2) dont la nature et le mode d'acquisition sont indépendants du premier, en l'occurrence des données satellitaires. Ces deux jeux de données portent sur la physionomie des parcelles agricoles, chacun dans son registre d'observation et selon son propre protocole d'observation. Pour rappel, c'est dans la parcelle agricole que s'imprime le système de culture.

– Nous souhaitons vérifier si cet autre jeu de données (D2) est compatible avec la première représentation (référentielle). Pour ce faire :

- Une relation, un passage est recherché ici entre la physionomie de chaque parcelle (en l'occurrence maintenant dans le jeu de données D2) et les types de systèmes de culture précédemment définis à partir de D1 ; cette typologie précédemment acquise sur des bases agronomiques³⁹ sert en quelque sorte de référentiel.

- L'existence d'une telle congruence sera admise selon la qualité (mesurée par le taux de véracité des propositions d'assignation de chaque parcelle à un type systémique de la typologie de référence). Un passage serait ainsi permis entre le jeu de données D2 et la typologie référentielle.

- La vérification de l'existence de cette congruence corroborera l'hypothèse heuristique.

Opérations

– L'expérience d'observation de la Terre (notée OT en figure 9) par capteur satellitaire (oasis de Nefta, juin 2011) est différente et indépendante de la première expérience d'observation par un agronome, quant à la nature de son registre de données : radiations appartenant à des canaux spécifiques de l'imagerie Spot, rapportées à chacune des parcelles (par ailleurs typées dans le référentiel). Les données OT ramenées à chaque parcelle correspondent à sa physionomie, selon cet autre mode d'observation.

– Nous recherchons, par analyse factorielle discriminante, le passage entre la physionomie OT de chaque parcelle (jeu de données D2) et un type de système de culture appartenant à la typologie systémique de référence. La qualité du résultat est vérifiée par le test d'assignation⁴⁰ des données OT de chaque parcelle à un type de système de culture.

– On trouve un taux satisfaisant, de l'ordre de 80 %, de conformité à la typologie systémique (voir Fargette *et al.*, 2017 pour le détail de l'analyse).

39. Les capteurs utilisés en agronomie de terrain et en télédétection sont de natures indépendantes.

40

La diversité de physionomie des parcelles, contenue dans le jeu de donnée D2 OT, est compatible avec la typologie systémique pré-établie à partir de données agronomiques (D1). Si elle ne l'avait pas été, l'hypothèse d'empreinte en oasis aurait été contredite et, par-là, l'hypothèse heuristique. Bien que les méthodes d'observation soient indépendantes, les jeux de données ne le sont pas. L'hypothèse heuristique est corroborée. Les jeux de données D1 et D2 sont liés (correspondance), soutenus par l'empreinte.

3.2. De l'abstraction au monde perçu : le Lien de Forme et l'intuition du modèle géographique « Épure »

Pour répondre à la question « Qu'est-ce qu'une oasis ? », nous empruntons un certain nombre d'éléments à un travail précédemment publié (Fargette *et al.*, 2019b), travail de terrain et d'échanges entre experts du monde oasien. Nous souhaitons une définition générale, qui ne retienne que les caractères suffisamment partagés par toute oasis, pour n'en exclure que le minimum. Nous avons le choix de l'aborder d'un point de vue plutôt descriptif : les structures et les formes que l'on voit se dessiner ; ou bien plutôt explicatif, systémique : les fonctions et les processus constatés, qu'ils relèvent du monde bio-géophysique ou de celui des hommes, d'ailleurs fortement imbriqués la plupart du temps. Or, selon l'hypothèse heuristique proposée en partie 2.4 et figure 7, les deux définitions se répondent, même si chacune parle sa langue propre (système ou géographique) puisque le concept d'empreinte fait le lien entre système abstrait et réalité concrète.

Pour illustrer le *Lien de Forme* (du système vers le monde perçu, cf. figure 7), nous commençons par une description en termes systémiques de l'oasis, pour en « déduire » ensuite (par la traduction utilisant un vocabulaire pictographique) ce que les éléments abstraits de description (en termes de fonctionnement, structure, dynamique) donnent à voir dans le concret. De manière synthétique à partir de Fargette *et al.* (2019b), nous décrivons ainsi (encadré 10) ce que nous appelons l'Entité⁴¹ Oasis (équivalent de compartiment systémique de Fargette *et al.*, 2019a) puis l'Épure Oasis qui profile en paroles (encadré 10) puis en modèle (figure 10) ce qui se donne à voir (équivalent de modèle géographique) (cf. partie 4.1).

L'épure, très simple, dessinée en figure 10, ne retient que le principal, le plus important, le plus partagé, de ce qui a été conçu, pour commencer, dans la compréhension systémique et la description synthétique de l'Entité Oasis. La correspondance systémique - géographique est contenue dans ce modèle de passage, sorte de programme, de préfiguration de la forme. L'entité (système), « donne forme » (genèse d'empreinte), « se donne à pré-voir », « prend forme », selon ce modèle géographique que nous nommons épure, avant même que de se montrer réellement dans l'empreinte.

41. système ou quasi-

1 ' 1 considère trop ouvert pour être système.

Encadré 10. De l'entité oasis (compartiment systémique) à son épure (modèle géographique) (voir figure 10)

Entité Oasis

L'oasis est rendue possible par une ressource en eau rare mais accessible en zones sèches et son réseau de distribution, des cultures végétales et des habitants produisant et coordonnant le tout ; elle communique avec l'ailleurs, proche ou lointain.

Chacune joue le rôle de relai pour des échanges qui la dépassent. Les systèmes de culture sont spécifiques en agroforesterie et combinent de très diverses cultures intercalaires, associant plantes pérennes et cultures annuelles ; les parcelles agricoles sont regroupées pour être irriguées (système intensif). Les agriculteurs et leurs familles s'organisent pour le partage de services (ressource en eau, production agricole, services urbains, etc.) et leur protection (tempêtes de sable, insécurité, etc.).

Leurs habitations sont le plus souvent indépendantes mais proches des parcelles cultivées pour en faciliter l'accès, ne pas entraver les travaux agricoles et optimiser les aménagements collectifs, en particulier d'irrigation ; des flux denses, diversifiés et à double sens existent entre les parcelles cultivées et les lieux d'habitation. L'oasis résulte d'une adaptation de l'homme au milieu en zone aride, de sa capacité à s'organiser en société dans des conditions de milieu contraignantes (conditions climatiques, isolement). Une oasis n'est pas autarcique.

Épure Oasis

L'oasis se situe le long d'une rivière, à proximité d'une source de piémont, de puits artésiens ou de forages, au débouché d'une *foggara*. Elle se caractérise toujours par un périmètre irrigué permanent, qui contraste fortement avec les abords arides de l'oasis et présente une forte densité végétale à plusieurs étages, dont ceux de ligneux pérennes. Il est toujours connecté à une zone d'habitat permanent groupé autour de marqueurs urbains (marché, école, mosquée, etc.).

Outre le réseau de desserte de proximité, l'ensemble possède des voies de communication au long cours (pistes, chemins de terre, routes goudronnées... : cf. cordon ombilical). L'orientation, l'extension, les continuités ou discontinuités spatiales de l'oasis dépendent des conditions locales du milieu.

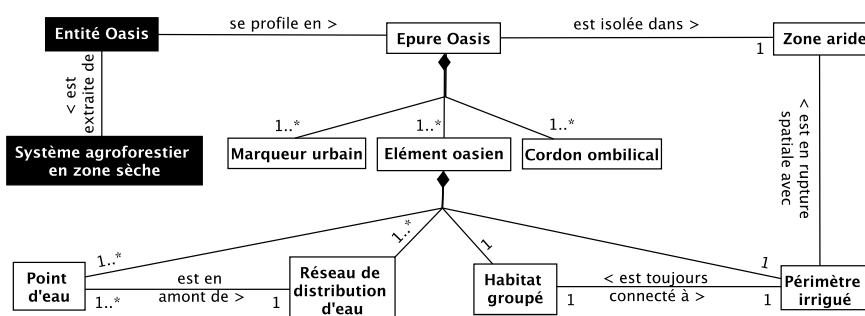


Figure 10. De l'entité oasis (compartiment systémique) à son épure (modèle géographique) (voir encadré 10)

Ainsi, en reprenant à partir de l'hypothèse heuristique posée en figure 7, la connaissance systémique de l'oasis (*cf.* le système compris) permet de proposer un modèle de passage (modèle géographique, épure) qui prédit, préfigure une forme à venir, décrit (langage pictographique) l'oasis, sorte de « pré-image », de « portrait-robot », qui trouve son origine dans l'entité que le point de vue porté sur le système a dégagé. Le cheminement le long du *Lien de Forme*, passage de l'abstrait au concret, est bien là ; il mérite d'être précisé.

Toujours selon l'hypothèse heuristique et de façon générale, aux entités (systèmes ou compartiments extraits de systèmes, parties de système), correspondent des formes projetées, inscrites sur la planète *via* le processus de genèse d'empreinte. Elles seront donc représentées dans une fenêtre d'observation, un paysage⁴² ou une image satellitaire (*cf.* figure 1), etc. Quelle que soit l'échelle planétaire considérée ou la (les) discipline(s) impliquée(s), une entité répond de la forme à venir. Il en est ainsi des entités systémiques de type biologique, biogéophysique⁴³ (Lacoste et Salomon, 1999) ou social⁴⁴ (Ramognino, 2014), qui façonnent aussi la planète. Chacune de ces disciplines a recours au concept d'entité, abstrait, de nature systémique et qui contient en germe (modèle, programme) dans son épure, la forme concrète prise dans la réalité, forme découlant de la cohérence systémique. Vont y figurer les caractéristiques matérielles et spatiotemporelles⁴⁵ (registres conceptuels décrivant la réalité, *cf.* figure 5) relatives à la structure⁴⁶ de l'entité ou aux signes de son fonctionnement⁴⁷ et de ses interactions⁴⁸ (donc traduites du registre systémique), chacune en puissance de s'inscrire dans le concret (par la genèse d'empreinte). Le *Lien de Forme* aussi est donc très général.

La suite du travail propose de parcourir, de façon discursive, le chemin du monde des systèmes jusqu'au monde perçu et donne consistance au raccourci appelé *Lien de Forme*. Il en décrit les étapes et en discute les enseignements.

4. Construction du signe : ontologie systémique de l'empreinte et *Lien de Forme*

L'ontologie de l'empreinte se réfère à l'ascendance systémique de l'empreinte (*cf.* figure 7). Nous parlerons d'ontologie systémique. Nous allons maintenant

42. ' 1 graphes, des agronomes, ou des écologues...

43. qui prendront la forme *e.g.* ' 1 1 ê 1 1 ï

44. 1 1 ' 1 ' ll

45. *e.g.* pour une entité biologique : la traie l'œ l'
papillon, en passant par les différents stades de la larve chenille.

46. *e.g.* pour une entité biologique : structure individuelle : un cône de patelle attaché au fond rocheux aquatique ; structure collective : un récif corallien.

47. *e.g.* un barrage de castors, un paysage agricole...

48. *e.g.* 1 ' (' ; turricules des lombrics, témoins de leur activité et de leur contribution à la fertilité des sols.

parcourir le *Lien de Forme*, du système originaire de l’empreinte à une forme concrète, observable (réalité) qui sera observée (monde perçu).

4.1. Concevoir l’entité, modéliser l’épure

Une entité (monde abstrait) est extraite (par un point de vue) du système ; elle correspond (figure 11) à une portion (de système) dont la logique dépend d’un regroupement cohérent de traits fonctionnels. On peut aussi qualifier l’entité (figure 11) de façon synthétique par sa nature⁴⁹ et sa fonction⁵⁰. Le langage du modèle de l’entité est systémique, celui des lois (structurelles, fonctionnelles, adaptatives) sélectionnées par le point de vue.

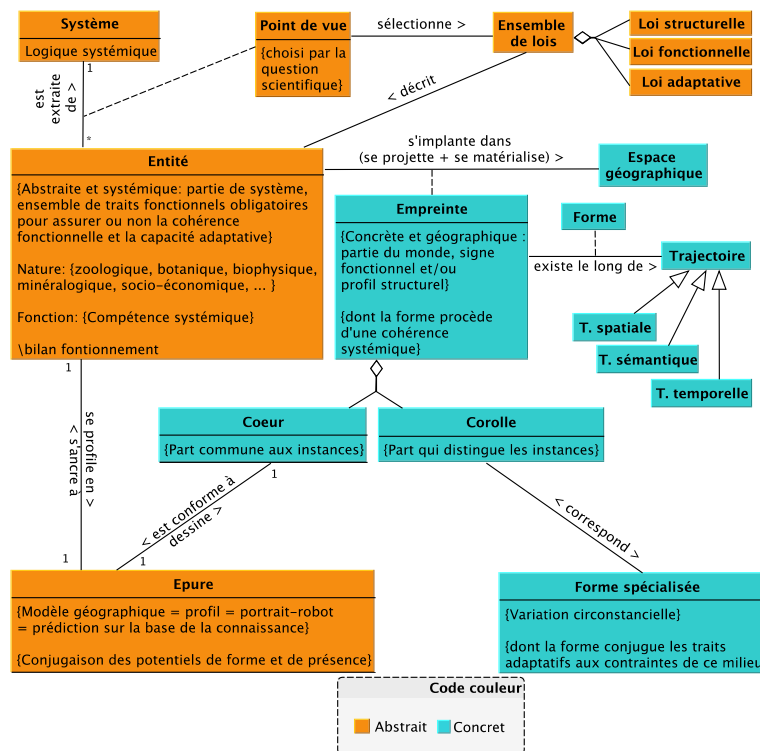
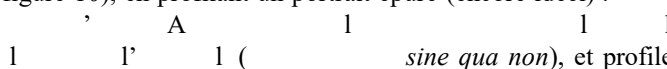
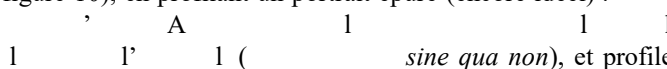


Figure 11. Ontologie systémique de l’empreinte

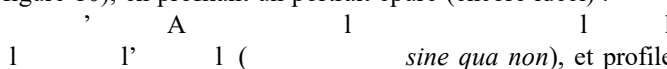
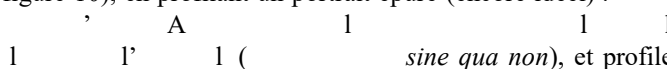
49. attachée à la ou les domaines scientifiques concernés par le point de vue.

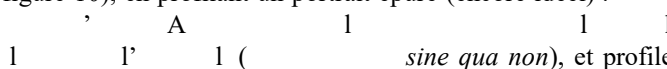
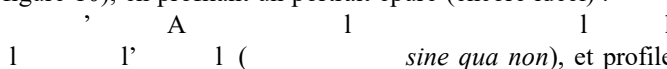
50. indiquée par le bilan (mono- ou multidimensionnel) de son fonctionnement dont se détachent éventuellement une ou plusieurs compétences systémiques notables.

Alternativement, la description de la logique systémique peut emprunter un vocabulaire pictographique, imagé qui anticipe l'empreinte matérielle (cf. encadré 10 et figure 10), en profilant un portrait épuré (encore idéal) :  et profile dans ce « portrait-robot » les traits de forme signifiants (*i.e.* se rapportant aux traits fonctionnels, pré-requis systémiques, et par conséquent obligatoirement présents). Ainsi, une fois le processus d'observation en place, ce que l'on verra de l'empreinte dans toutes ses instances (cf. plus loin, figure 14) est l'empreinte (sémantique) profilée à partir de cette seule entité. De même que sa forme épurée, sa distribution géographique⁵¹ potentielle  (i.e. obligatoire, partagé par toutes les instances).

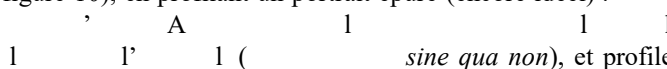
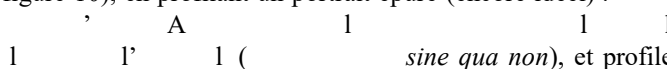
L'épure conjugue donc le potentiel de forme et celui de présence (*i.e.* distribution géographique), qui seront acquis au moment de l'implantation de l'entité, en même temps que se concrétise dans l'espace géographique (projection et matérialisation) l'empreinte⁵² correspondante (figure 11). Le système Terre dans son entièreté se concrétise en planète Terre (elle est son empreinte) ; chacune des entités extraites du système Terre se matérialise, à diverses échelles infra-planétaires, en empreintes.

En résumé, toute entité, système ou (quasi-)système extrait du système Terre et implanté (c'est-à-dire projeté et qui se matérialise) a une empreinte (figure 11).

L'empreinte est concrète et positionnée dans l'espace géographique ; ses instances appartiennent à la réalité. Elle est le résultat de l'inscription dans la réalité de l'activité fonctionnelle ou du profil structurel de l'entité dont elle procède (ontologie systémique) et ce conformément au modèle décrivant son épure. Les traits  « essentielles et invariantes »⁵³ qui existent le long d'une trajectoire sémantique, spatiale et temporelle. Au-delà de la notion de généricité (ce qui relève d'une abstraction)  instances) que nous étudions ici en considérant l'ascendance systémique, le concept d'empreinte contient toutefois une notion de diversité (ce qui relève de la corolle), souvent en relation avec le milieu une fois l'implantation faite. La capacité du modèle géographique « épure » à décrire ce qui se donnera à voir à partir de données systémiques seulement, lui confère un potentiel prédictif intéressant (cf. partie 5).

51. localisation possible des instances  ; elle se lit en « présence ou absence » en tel ou tel lieu, préalable au déploiement de la forme en ce lieu, *e.g.* les conditions minérales ; l'empreinte  dessine telle forme (géomorphologie).

52. Descartes parle de corps en tant que « partie du monde » (cf. Descartes in Belhoste, 2018).

53. *e.g.* modalités sémantiques : l'entité  - chenille - papillon et la somme des formes  ; modalité spatiale : zone de distribution géographique, mobilité de chaque stade ; modalités temporelles : caractère stable ou labile, modifiable, périodique, saisonnalité de chaque stade.

E l ê œ

distingueront les unes des autres par leur corolle. Celle-ci tient aux formes spécialisées, variation d'origine circonstancielle liée aux circonstances de lieu (qualités et contraintes de milieu, interactions de logique systémique en ce lieu). Par exemple, la capacité d'adaptation des êtres vivants, par les lois et mécanismes qui la régissent, est diversifiante ; on en retrouve les effets dans les attributs et propriétés de corolle (correspondant aux formes spécialisées issues de l'adaptation). Le regroupement des caractères en entités et l'ordonnancement des entités selon le partage des caractères se retrouve dans une classification telle que ce qui est présenté en figure 6.

– *Importance du choix de l'ancrage systémique* : Un système peut être à l'origine de plusieurs entités (figure 12). Plus ou moins apparentées mais générées par des points de vue différents (e.g. point de vue « système de culture », point de vue « état sanitaire », point de vue « composant espèce »), les entités obtenues n'entretiennent pas le même lien de parenté avec le système (spécialisation, composition, état) ; elles ne mettent pas l'accent sur les mêmes caractéristiques systémiques (cf. entité en figure 11) ; par voie de conséquence, elles présentent des épures différentes. Pour les cas proposés en figure 12, les épures correspondant aux déclinés du système agroforestier en oasis sont différentes, même si les entités traitent toutes de végétation et de cultures en oasis.

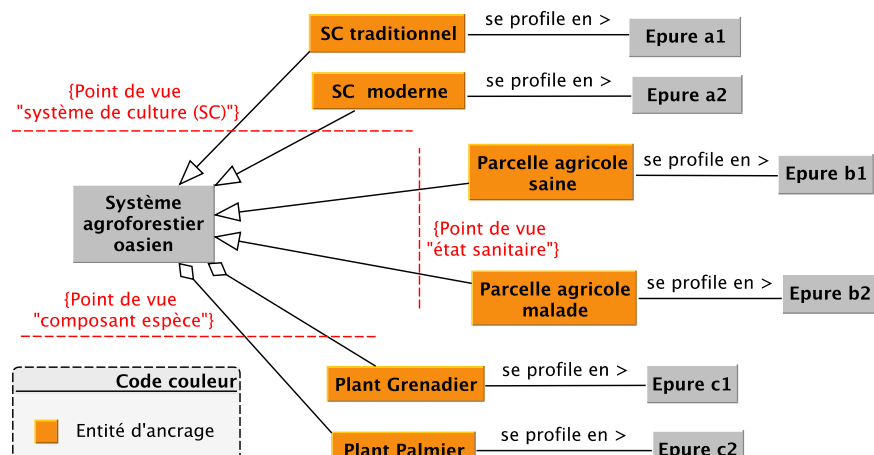


Figure 12. Importance du choix de l'entité d'ancrage - Illustration en oasis

Par exemple, l'interprétation des valeurs de l'indice NDVI (associé à la réflectance des pigments chlorophylliens des plantes), classiquement utilisé pour

à voir ses caractéristiques d'œ Il en une synthèse matérielle, de forme générale et/ou spécialisée, à laquelle peut s'adjoindre, sous une forme concrète aussi, son héritage (systémique) archivé⁵⁶.

E l' ê l
, l l l des critères de localisation
permettant de décrire le voisinage et la mobilité. De sa composition et sa géométrie
résultent par exemple un nombre de propriétés physiques (barycentre, inertie, etc.)
en lien avec les caractéristiques de la matière impliquée, en alliage ou pas (densité,
émetteur ou capteur de radiations, opacité, rugosité, etc. l'
l l',
positionnée l l' ll', l

4.3. Vers l'observation d'instances d'empreintes : les concepts de caractère, de signe et de signal

L'empreinte est un assemblage (agrégat) logique de caractères ; ses instances sont décrites selon ces caractères à partir de données observées (cf. caractéristiques de l'instance ; figure 14).

L'ensemble décrit en figure 13b fournit un premier format (générique de haut niveau) au recensement des caractéristiques de toute empreinte lors de l'expérience d'observation menée sur ses instances. Ces caractéristiques s'organisent de façon générale selon les quatre axes d'un « schème⁵⁷ » (figure 14) : 1) celui de l'espace géographique, de nature concrète et de dimension 4 (x, y, z, t), 2) celui du comportement spatio-temporel et du mouvement dans cet espace, 3) celui de la connectivité (dans l'espace, avec le milieu et avec le voisinage plus ou moins lointain), 4) celui de l'activité dans l'espace et de son résultat (en lien avec le bilan de son fonctionnement systémique, cf. figure 11).

En tant que construit abstrait adossé à une ontologie de domaine, l'entité référente annonce les caractères dont l'empreinte hérite et, parmi ceux-ci, ceux qui contribuent au dessin de son épure. Dans cette représentation, ce qui est « accidentel⁵⁸ » n'a pas statut de caractère ; en conséquence il n'est pas enregistré ; au mieux, il l'est en tant qu'anecdote. Nous avons en effet vu (cf. figure 12) que la

56. e.g. une vallée fluviale inscrite dans une vallée glaciaire, ce qui révèle la modification de p', l l

57. chez Kant, représentation intermédiaire entre les phénomènes perçus par les sens et les l'

58 P l l', ê a) pertinente à noter lors du recensement de soldats blessés (et à ce titre figurer dans les caractères recensés), ou bien b) le', ll (l : œ ll l -

question posée induit le point de vue sur le système et le focus choisi, par là l'entité d'ancrage ; ceci est déterminant dans la définition des caractères de l'empreinte qui correspondent, en matérialité, aux caractères annoncés dans l'épure, demeurant au

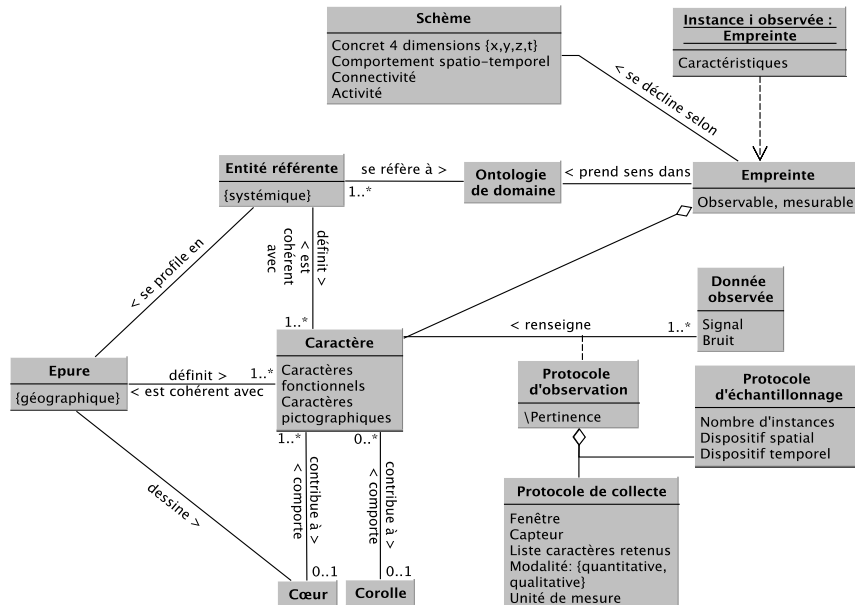


Figure 14. Instance d'empreinte

Le concept d'empreinte, inscription de l'entité dans la réalité (cf. figure 7) s'instancie⁵⁹. Réciproquement et de façon idéale, l'instance observable dans la réalité témoigne de l'empreinte.

Selon cette vision, les données pertinentes sont celles qui renseignent les classes de caractères de l'empreinte et les protocoles d'observation (échantillonnage et collecte) sont construits et adaptés à cet effet. Ainsi, la démarche d'ontologie systémique surplombe le raisonnement et assiste la conception et la structure de la base de données au service de l'expérience d'observation. On retrouve ici le commentaire de Hansen (cf. Lecourt, 2001) : « tout compte rendu d'observation est lesté de théorie » ; et on peut alors dire que la théorie le lesté dès la préparation de sa méthodologie. En partie 3.1.1, le savoir expert de l'agronome préparant sa campagne de terrain (protocoles) en est une illustration : les données collectées sont physiologiques (cf. figure 8) ; tout en décrivant des pratiques agronomiques

59. Inscription en n « copies » () ; par exemple,

61. dans ce cas (*e.g.* l'œl l'ê l
 , w l' l l' , l ê
répétabilité sont pris en charge dans les protocoles (*e.g.* ,
 , ê ll , ll l l ol etc.)

protocoles précis d’observation sur le terrain. Le signal forme le contenu pertinent de la donnée.

Il y a un décalage du signe au signal ; il est toutefois connu et, en remontant le lien, il est théoriquement possible de « sonder » (par la recherche de la présence du signal dans le contenu de la donnée acquise), la présence ou non des signes correspondant à l’épure dans le jeu de données.

Au plus près du terrain, l’adéquation du capteur et celle du protocole d’observation à la problématique peuvent être optimisées (choix d’une meilleure focale, diminution du bruit) par la connaissance préalable des signes contenus dans l’épure et celle des propriétés des capteurs. De nouveaux capteurs, ayant une sensibilité mieux adaptée au registre des signes, peuvent aussi être conçus.

La donnée contient aussi du bruit, lié aux caractéristiques du capteur ainsi qu’aux conditions de capture. Cependant la pertinence ou non du contenu de la donnée ne se limite pas à la limitation technique des capteurs. L’apport de ce travail sur le *Lien de Forme* dans le décryptage (pertinent vs non pertinent) du contenu de la donnée sera plus précisément discuté en partie 5.

Le chemin que nous avons suivi est long, de l’entité (abstraction de départ, fondée sur la connaissance acquise antérieurement) à la donnée (telle qu’on peut la collecter plus ou moins empiriquement, cf. figure 1). De façon idéale, ce chemin pourrait être consolidé par les travaux relevant de nombreuses compétences, de la systémie⁶² à laquelle il est fait appel dans tant de domaines scientifiques, à l’informatique et l’électronique⁶³ (cf. partie 5), pour expliciter, ordonner la connaissance et la formaliser selon le *Lien de Forme*.

5. Du signe à la maîtrise de l’observation

Nous proposons trois perspectives sur le travail effectué, la première synoptique sur les mondes systémique et perçu (figure 16), la deuxième analytique sur la donnée (tableau 1), et la troisième sur la manière dont ce travail s’inscrit dans la recherche de la pertinence de la donnée et de son analyse (figure 17).

Nous croisons en figure 16 les deux approches distinctes pour comprendre le monde, l’une fondée, de prime abord, sur l’approche systémique, l’autre sur

62. “In the context of systems science and systems philosophy, systemics is an initiative to study systems. It is an attempt at developing logical, mathematical, engineering and philosophical paradigms and frameworks in which physical, technological, biological, social, cognitive and metaphysical systems can be studied and modeled. The term “ ” was coined in the 1970s by Mario Bunge and others, as an alternative paradigm for research related to general systems theory and systems science” (Wikipedia).

63

traitement et de l’

1

l’

1

l'organisation de l'espace géographique, toutes deux complémentaires et dialoguantes. La première vise une compréhension holiste ; elle adopte toutefois, pour gérer la complexité, une démarche par compartimentation des systèmes. En raisonnant sur les acquis de connaissance, elle organise des concepts, en crée de nouveaux si nécessaire (en l'occurrence ici l'entité d'ancrage, l'empreinte et son épure). La seconde exploite l'expérience du monde (la part plus particulièrement expérientielle) pour prendre en compte les caractéristiques⁶⁴ de l'espace (ses continuités ou discontinuités ou les gradients) qui conduisent à sa fragmentation en formations et à son organisation spatiale. L'enjeu est de faire converger ces deux approches vers une représentation cohérente, une compréhension commune.

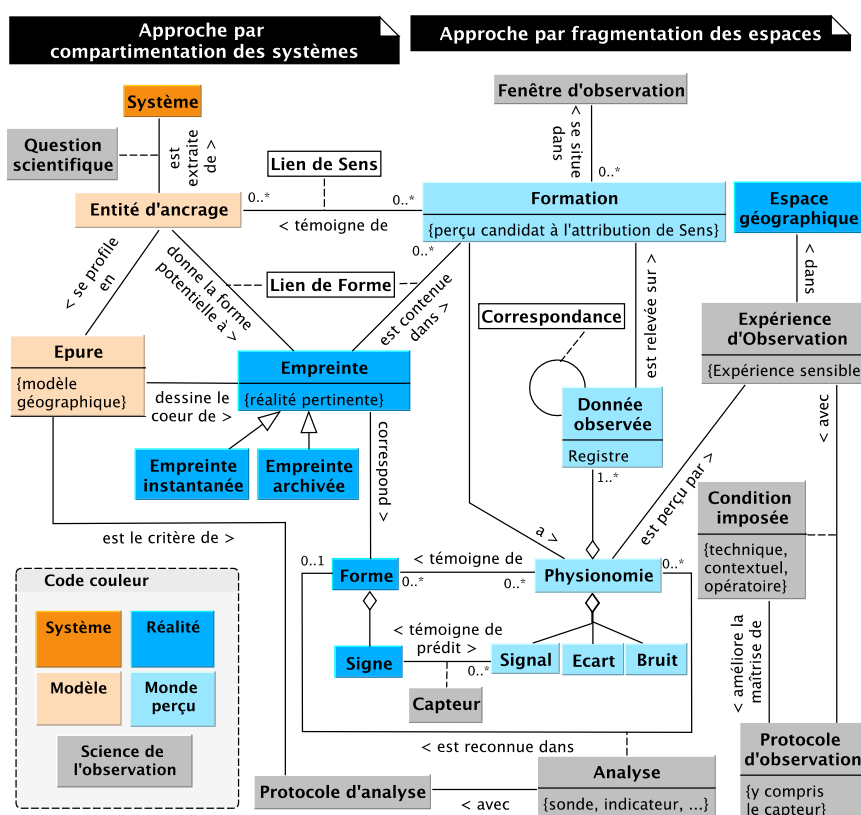


Figure 16. Schéma synoptique système - Monde perçu et science de l'observation

La figure 16 propose une synthèse des concepts clés de ce travail et de leurs relations :

64. obtenues de façon intuitive ou méthodique.

– Le passage entre système et monde perçu se décline à différents niveaux⁶⁵ : entité-formation ; empreinte-formation ; forme-physionomie ; signe-signal ; il ouvre la voie au potentiel prédictif de l'épure (dessin de sonde) ; la sonde recherche dans le monde perçu des signaux directement prédits par le signe (et permet ainsi de reconnaître et donner sens à la donnée) ; toutefois il est aussi possible de leur substituer des indicateurs leur correspondant tout en étant issus d'un jeu de données d'un autre registre ; ceci est parfois préférable, voire même nécessaire, quand la collecte des seconds est techniquement plus simple ou moins onéreuse, par exemple les indicateurs paysagers présentés et argumentés par Loireau et Fargette (2022) ;

– Le *Lien de Forme* et le *Lien de Sens* qui ont participé au déroulé du raisonnement y sont resitués ; en observatoire (cf. partie 6), ce sont eux qui activent les phases de recherche (acquisition de sens) ou de suivi (attribution de sens) ;

$-L'$ l' ∂l l'
 (cf. figure 7) est présentée l' l ;

– Les éléments à maîtriser pour passer d’une expérience d’observation originelle et non formalisée (*cf.* figure 1) à une science de l’observation restent en grisé.

L’empreinte, si elle est localisée là où a lieu l’expérience d’observation, constitue la partie pertinente de la donnée, le signal contenu dans la donnée et en lien avec la question posée ; l’empreinte contribue par sa présence et sa forme, à la physionomie du lieu. Toutefois, entre d’une part la physionomie (*cf.* encadré 1) d’un lieu, d’une fenêtre d’observation, d’une formation et, d’autre part, la forme de l’empreinte, des bruits se superposent dans la donnée physionomique *in situ* (figure 16 et tableau 1). Ils sont contingents⁶⁶ du lieu (en tant que co-occurrence aléatoire) et correspondent à des accidents qui frappent (en dehors de la logique systémique retenue et de son épure) dans sa matérialité la portion d’espace concernée par l’empreinte. De tels accidents ne sont pas pertinents ; leurs sources ne relèvent pas de l’ontologie systémique considérée et brouillent le déchiffrement qui recherche (par voie de sonde) l’empreinte dans la physionomie. Par exemple⁶⁷, sur une surface cultivée (empreinte A, se référant à un système de culture A), se trouve un bombement correspondant à un tumulus (sépulture se référant au système néolithique B). Un tel constat n’a rien à voir avec une activité agricole ; il n’appartient pas au système A et ne figure pas dans l’entité « parcelle agricole », ni dans la description de l’empreinte correspondante. Le système B est indépendant (sans interaction, au sens systémique) du système A. Le bombement lié au tumulus est contingent du lieu et induit une anomalie de ce point de vue, à ne pas prendre en compte dans la description de la parcelle. L’existence d’un nombre indéterminé de composants systémiques s’inscrivant⁶⁸, en se superposant et se combinant, sur un même espace géographique,

65 l l l' l l' l l' ll
perçu.

66. en raison de relations topologiques qui ont été codifiées par ailleurs.

67. Ceci est un exemple de succession chronologique d'évènements indépendants en un lieu.

68. De façon imagée, nous parlons d' « millefeuille systémique ».

chacun responsable de sa propre empreinte, complexifie la physionomie de l'espace concerné.

A l'issue de la démarche proposée, l'élaboration du *Lien de Forme* a permis de spécifier le contenu du signal ; la démarche permet aussi de pointer les sources de difficulté, les facteurs qui brouillent le signal avec du bruit simultanément collecté et contenu dans la donnée observée. Le tableau 1 recense les biais induits par les contraintes de protocole, au bruit né de la capture-même de la donnée.

Tableau 1. Clé de lecture du contenu de la donnée - Séparer le bon grain de l'ivraie

Constituants de la donnée	Sources	
	Nom	Pertinence
–	Signe	–
Signal	Cœ l' (de signes)	1
Écart lié à la variation adaptative	C ll l' (adaptation circonstancielle)	0/1
Cumul : surimpression au cours de l'	Mémoire (empreinte archivée)	0/1
Bruits causés par des événements fortuits : – s , mais indépendants ; – s , indépendamment à différentes époques de l'.	Contingence du lieu de l' l' l' (aléa)	0
Bruits liés : – au biais de la mauvaise adéquation du capteur au signe ; – au biais de la mauvaise représentativité l' ll .	Limites du protocole ,	0
Bruits liés : – à l' ; – aux conditions du milieu au moment de l' ; – au technicien.	Limites dans le déroulé de l'	0

Il est important d'en tenir compte au moment de l'interprétation de l'observé, quelle que soit la méthode d'analyse retenue. Bien qu'un même jeu de données (par exemple le contenu d'une même image satellitaire) puisse être utilisé pour traiter diverses questions, et même si certaines sources d'erreur ou d'imprécision, strictement inféodées aux caractéristiques du capteur, sont donc invariantes quelle que soit la question traitée, il est aussi primordial de noter que la part pertinente de la donnée et donc sa capacité (voire son incapacité) à répondre à la question posée, sera variable car relative avant tout à la question posée. La part pertinente tient en effet à l'ancrage sur l'entité pertinente et en cela à la prise en compte de ce qui est variation adaptative, circonstancielle mais aussi à la part (variable) de l'histoire qui peut cumuler, surimprimer les formes pourtant relatives au même système (empreinte archivée).

Des données sont acquises par observation puis analysées, que la démarche d'observation isole ou non⁶⁹, en préalable à l'analyse, des données relatives à des formations distinctes, candidates à l'attribution de sens. Ainsi, avec ou sans fragmentation préalable (figure 16), une formation⁷⁰, selon que l'analyse de sa physionomie, en lien avec la question (i), y reconnaît la présence de signaux relatifs à l'épure (i), deviendra, avec un indice de confiance à déterminer, témoin de l'entité (i) et endossera (avec ce degré de confiance) le sens contenu dans l'entité (i) (figure 17). Le *Lien de Forme* se « branche » de façon optionnelle sur un processus d'observation et d'interprétation du monde, pour l'épauler et l'enrichir par la connaissance préalablement acquise. Dans l'exemple « oasien » (cf. figure 10), un algorithme construit à partir de l'Épure Oasis interviendrait alors dans l'analyse des données.

Toutefois, constater dans la donnée la présence du signal (i), conforme à l'épure (i) n'assure pas (du moins pas avec un indice de confiance de 100 %) l'association de la donnée à l'entité (i) en tant que donneuse de sens (i) ; il s'agit seulement de corroboration de cette hypothèse (raisonnement poppérien). Par contre, ne pas y reconnaître la présence du signal (i) assure⁷¹ la négation de la proposition (réfutation). Cependant, en tenant compte de cette contrainte d'ordre logique, l'inférence issue du *Lien de Forme* (sur les caractéristiques d'épure, de signe et de signal) pourra ensuite être associée à d'autres raisonnements dans des combinaisons d'analyse plus performantes.

Aussi l'apport majeur de l'analyse selon le *Lien de Forme* consiste dans le concept de genèse systémique d'empreinte (ontologie systémique) et donc dans l'importance du choix pertinent de l'entité d'ancrage (entité référente) sur laquelle se

69 ' 1 ' 1 l' 1 1
 dans la seule démarche ' (1 1
 70 S l' ' l' 1 l' 1 1
 ê ' 1 l' 1 (rait
 71 ' 1 l'

fonde, pour chaque empreinte, la modélisation géographique de son épure. Cet exercice, en précisant le contenu sémantique de chaque empreinte, permet de pointer ce qu'il est pertinent de noter au cours de l'observation et ce qui n'est pas pertinent étant donné la question (point de vue) et le focus (l'entité) choisis. Ensuite, afin de distinguer ou assimiler des épures proches, il pourrait être constitué des bibliothèques de références, elles-mêmes fondées par exemple sur des classifications de formes, de signes ou de signaux. De telles classifications expertes (*cf.* figure 6) enrichissent la mémoire et affinent la recherche de sens de toute connaissance scientifique correspondante.

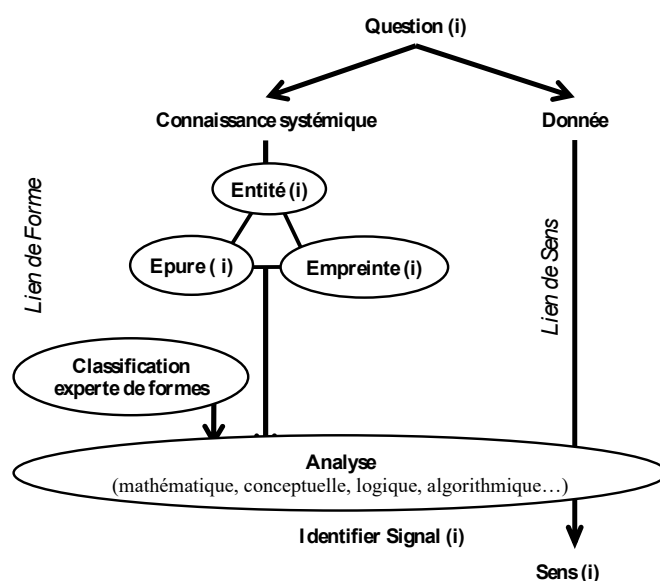


Figure 17. Conjugaison du Lien de Sens et du Lien de Forme

L'enjeu de ce travail est de fait plus large qu'un apparent détour par le *Lien de Forme* ; la mise en jeu de ce dernier rejoint l'un des défis actuels analysé par l'Inria (2016) :

« travailler en intelligence artificielle en combinant des sources de connaissances et des sources de données ; intégrer des modèles et des ontologies dans le processus d'apprentissage. »

Alexandre (2019) pointe le même enjeu entre intelligence artificielle et neurosciences. La genèse systémique d'empreinte est une proposition de source de connaissance pour enrichir les processus d'apprentissage dans le domaine de l'observation de la Terre. En effet, une expérience d'observation taille une fenêtre d'observation de façon (plus ou moins) arbitraire, avec plus ou moins de

considération de l’étendue de chacun des systèmes en jeu sur cette étendue. A l’issue du travail d’observation, les capteurs restituent un corpus de données correspondant à des portions d’empreintes, à la superposition et au cumul d’empreintes plus ou moins partielles (un millefeuille systémique) et ceci sans l’a priori d’un point de vue. Si au contraire un point de vue est assumé lors de l’expérience d’observation (y compris son analyse), le protocole qui la conduit a pour objectif de minimiser le bruit contenu dans l’observé selon ce point de vue. Il se fonde en particulier sur la démarche d’ontologie systémique de l’empreinte. L’algorithme construit à partir du « Modèle géographique Épure » pourrait ainsi saisir et mettre à disposition la connaissance experte contenue dans l’épure pour accompagner et même guider l’observation.

L’expert apporte une connaissance structurée (modélisée, ordonnée, classée, etc.) dans son domaine de compétence, un supplément d’information qui pourra faciliter ou améliorer l’observation et l’analyse :

- raccourcissement du raisonnement, simplification de l’analyse et moindre temps de calcul en précisant la question⁷² ;
- meilleure qualité du résultat en proposant un meilleur focus du protocole d’observation et d’analyse, en enrichissant et modélisant l’hypothèse, en préfigurant certaines sources de difficultés.

Ce raisonnement n’est cependant pas linéaire. Les inférences issues du *Lien de Forme* pourront (facultativement) intervenir dans la logique des raisonnements qui conduisent l’observation⁷³, puis son analyse⁷⁴. Elles pourraient aussi intervenir dans la combinaison de raisonnements pour aboutir à des séquences hautement plus interconnectées. S’ouvrent alors, en psychologie et sciences cognitives, l’investigation des raisonnements humains et, en intelligence artificielle, l’algorithmique pour asseoir les approches par reconnaissance ou par prédiction, le traitement de la donnée et du signal.

6. Discussion : pour une science de l’observation

Notre travail s’intéresse, à partir de l’expérience du monde, à ce construit scientifique, raisonné à partir d’une question (Bachelard, 1938 : « C’est précisément ce sens du problème qui donne la marque du véritable esprit scientifique. ») qui initialise et formalise le point de vue. Le construit scientifique est en constante élaboration entre abstrait et concret, systémique et perçu. Il est le lieu où se développent et s’expriment les questionnements scientifiques, les méthodes,

72. question vague « Q ’ -ce- ’ ? » remplacée par « l’ ’ 1 ’ ’ , oasis ».

73 l ’ ’ 1 ’ ’ Il construction des capteurs.

74 1 l ’ ’ 1 ’ ’ etc.

raisonnements et protocoles, par lesquels il donne corps et sens à l'expérience empirique du monde.

Raisonnement

Notre travail montre que les raisonnements des systèmes au monde perçu sont complexes ; ils conjuguent, ordonnent un ensemble d'inférences plus élémentaires de type déduction, induction, abduction. La psychologie (George, 1997) dit que de telles combinaisons ne sont pas obligatoirement uniques et universelles et montrent souvent un polymorphisme dans l'expression qu'elles peuvent prendre d'un individu à l'autre devant une même question. La diversité des formes d'inférences assure la flexibilité du raisonnement et, par-là, son adaptabilité. Poincaré (1913, *in* Bouleau, 2017) faisait le commentaire qu'il n'y a pas de logique ni d'épistémologie indépendantes de la psychologie. La psychologie distingue cinq temps dans une heuristique⁷⁵, que notre travail illustre : a. *la prise en compte du problème* - cf. partie 1, l'introduction ; b. *l'incubation, recherche de solution, rumination parfois longue* - cf. partie 2, la connaissance et la démarche scientifique, que nous nous devons de respecter ; c. *l'illumination* - cf. figure 7, partie 2.4, l'hypothèse d'empreinte ; d. *l'explicitation qui descend dans les détails* - cf. parties 3, 4 et 5, la démarche exposée ; e. *la validation* - cf. partie 3.1.2, le test d'hypothèse sur le concept d'empreinte.

D'où l'importance du raisonnement pour la démarche scientifique, et de faire la part aussi (cf. partie 2, cadre conceptuel), entre le conditionnement (ni explicite, ni raisonné) qu'il faut éviter, l'intuition (pas nécessairement explicite mais qui peut, par exemple, susciter des hypothèses) et le raisonnement que l'on veut rationnel ; celui-ci est non seulement conscient mais il doit suivre Descartes (*in* Moreau, 2016) qui pose la nécessaire clarté d'explicitation et le besoin de séparation claire entre les items. Il est intéressant de noter que la clarté exigée dans la démarche n'exclut cependant pas l'intrusion du « rêve », source parfois d'inventivité « transgressive » au moment de poser théories, hypothèses ou modèles à tester et d'énoncer les prémisses de la démarche. Ainsi Kekulé en 1890, dans un discours devant le Société allemande de chimie (*in* Krivine, 2018, préface de J.C. Ameisen) :

« Apprenons à rêver, Messieurs, et alors, peut-être, nous découvrirons la vérité. Apprenons à rêver. Mais gardons-nous de révéler nos rêves tant qu'ils n'auront pas été évalués et éprouvés par la raison et l'entendement de nos états de veille. »

Théorie

Krivine (2018) insiste sur le besoin de théorie, et de « bonne » théorie⁷⁶, à tester, sans laquelle l'abondance de données perdrait de son utilité. Il est donc important de construire les bonnes propositions d'interprétation (cf. figure 7), celles qui sont les

75. [https://fr.wikipedia.org/wiki/Heuristique_\(mathématiques\)](https://fr.wikipedia.org/wiki/Heuristique_(mathématiques))

76. « U

pas vrai. » (Krivine, 2018).

meilleures candidates à procurer du sens à l'expérience d'observation du monde, celles qui sont riches de données potentielles à collecter (*i.e.* données paysagères de terrain, contenu d'image satellitaire, *cf.* figure 1) pour les confronter à ce que la théorie avance (confrontation aux données, à l'expérience). Les concepts d'empreinte et d'épure d'empreinte ont été construits dans cette intention.

A l'opposé, la collecte et l'analyse, voire la fouille de données et le big data, dans un cheminement sans *a priori*⁷⁷, sans prémisse (ni intuition explicite ou implicite, ni hypothèse préalable), a pour vocation d'être riche en suggestions, propositions (hypothèses, modèles ou scénarios) qu'il s'agit d'extraire pour ensuite les investiguer plus avant. Une telle démarche de « fouille » est un point possible d'initialisation d'un raisonnement, parfois d'une itération, pour construire de la connaissance. Le travail présenté en partie 3.1.1 en est un exemple qui, à partir de données agronomiques, propose une typologie de systèmes de culture, référente pour la mise à l'épreuve du concept d'empreinte en partie 3.1.2.

De façon plus large, nous théorisons dans notre travail sur le passage entre systèmes et monde perçu, eux-mêmes associés à deux types de théories préexistantes ; l'une de type systémique, sur la base de travaux engagés par Fargette *et al.* (2018, 2019a) ; l'autre de type géographique (George et Verger, 2000 ; Gouyon et Leriche, 2010) au sens large, incontournable dès que l'intérêt de la réflexion porte sur la Terre, son fonctionnement et son histoire géologique, biologique, anthropologique et sociale.

Maîtrise de la démarche

Le caractère scientifique de la démarche tient aussi à l'ordonnancement explicite des concepts, ici dans le cadre formel de l'hypothèse heuristique système - réalité - monde perçu - modèle (*cf.* figure 7), et à la possibilité de confronter la théorie aux données, ici la mise à l'épreuve (test) du concept d'empreinte dans la théorie de passage. Il reste maintenant à mener la mise à l'épreuve sur données, en intelligence artificielle, du concept d'épure (modèle géographique) et de l'utilité des concepts entité d'ancrage systémique, épure d'empreinte, forme et signe.

Les protocoles d'observation précis et documentés (rendus ainsi partageables et répétables), couplés à des raisonnements complexes mais explicités, consolident une connaissance scientifique acquise selon un processus d'observation maîtrisé. Le processus d'observation se distingue du seul processus expérimental (expérimental à distinguer d'expérientiel), très sollicité depuis le 19^e siècle⁷⁸. Avec des *habitus*

77. *i.e.* des protocoles « neutres / génériques », absence de « théorie fondatrice » sur la réalité, absence de « choix » l'« observé » (OBOE)

78. « rapport au monde » (Bouleau, 2017) l'« expérimental » (Belhoste, 2018) l'« expérimentation, pièce majeure de la démarche et des sciences (*i.e.* l'« expérimentation, pièce majeure de la démarche et des sciences »)

sensiblement différents, ils sont complémentaires dans ce qu'ils contribuent à construire (explorer, expliciter, décrire, tester).

Enfin l'étude menée sur les sources d'écart et de bruit dans la donnée (cf. tableau 1) permet d'augmenter la pertinence, la précision et la fiabilité des résultats par la maîtrise de la démarche d'observation : en adaptant au mieux (par le choix de l'entité d'ancrage) les protocoles de collecte et d'analyse de la donnée à la question posée ; en poursuivant l'amélioration du protocole par exemple par le développement de capteurs mieux adaptés à la question [« Les instruments scientifiques doivent être considérés non comme des outils perfectionnés mais comme des "théories matérialisées" », cf. Bachelard, *in* Lecourt, 2018], pour ne pas citer, cela va sans dire, l'amélioration du contrôle du déroulé même de l'observation.

Science de l'observation

La rigueur et l'explicitation du raisonnement, l'utilité d'une théorie rendue testable, la maîtrise et la précision de la démarche de collecte et d'analyse des données, la capacité de contrôle et d'amélioration de ce contrôle font d'une démarche d'observation ainsi construite une science de l'observation.

Observatoire scientifique

Les questionnements contemporains complexes⁷⁹ sur l'environnement et plus largement sur les changements globaux et la mondialisation nécessitent recherche et suivi environnemental et sollicitent les observatoires scientifiques. Par exemple, les observatoires scientifiques en appui aux gestionnaires de territoire (observatoire OSAGE, Loireau *et al.*, 2017) décrivent les dispositifs scientifiques, techniques et organisationnels qui facilitent le déploiement d'une science de l'observation. La théorie du passage système - monde perçu y serait appropriée pour solliciter et intégrer des connaissances scientifiques nombreuses et diverses. La confrontation d'une telle pluralité, voire l'interdisciplinarité qui peut en naître, doit aboutir à l'enrichissement mutuel.

Il existe un parallèle entre l'acquisition de connaissance et l'utilisation de connaissance commentées dans ce travail et les séquences de recherche et de suivi du modèle OSAGE d'observatoire (Loireau *et al.*, 2017), deux types de démarches où la connaissance préalable n'a pas le même statut par rapport à la démarche d'observation ; dans le premier cas, la connaissance est construite (acquisition de sens) à partir de l'analyse de données observées tandis que dans le second, la connaissance acquise sert à attribuer du sens à ce qui est observé. Dans notre travail, la partie 3.1.1 illustre typiquement ce qui peut avoir lieu au cours d'une séquence de

l P l l' vs l' ,

processus écologiques en observatoire.

79. par exemple les « wicked problems » ou « wicked questions » posés dans le domaine de la « Sustainable Science » (Wehrden *et al.*, 2018 ; Oberlack *et al.*, 2019) correspondent à la complexité par excellence.

recherche en observatoire tandis que la partie 3.1.2 prépare le déploiement sur de larges espaces d’une observation dans une séquence de suivi (en se donnant les moyens d’acquérir des indicateurs relevant du registre satellitaire).

Dans le dispositif scientifique de tels observatoires, s’élabore le travail sur la connaissance systémique plurielle en interdisciplinarité (cet aspect se rapprocherait du domaine symbolique débattu dans ce volume) ; dans leur dispositif technique, sont menées la gestion et l’analyse des flux de données (ce qui se rapprocherait aussi bien du domaine numérique que symbolique). Une ontologie fondée sur la connaissance systémique est un gage de bon ajustement et de pérennité non seulement des raisonnements mais aussi des outils conçus et des passerelles en interdisciplinarité. Par ailleurs, le croisement des disciplines impose l’explicitation des non-dits⁸⁰. Ce croisement et l’enrichissement mutuel qui doit en résulter peut être grandement aidé par l’« ingénierie ontologique », en analysant les relations entre disciplines et en articulant leur complémentarité. L’intelligence artificielle, par sa composante ontologique, d’une part assisterait ce rapprochement interdisciplinaire, d’autre part utiliserait les raisonnements systémiques (langage partagé, utilisé dans différents domaines) en tant que source d’algorithmes, constructions associant modèles et inférences et/ou données.

Les approches symbolique et numérique débattues dans ce volume ne s’opposent donc pas dans notre proposition d’une maîtrise de l’observation de la Terre. Coexistantes dans tout processus cognitif, concertantes même et complémentaires, elles interviennent dans l’expérience du monde, le rapport du concret (réalité et monde perçu) à l’abstrait (système et modèle), tantôt la découverte du monde (et sa description), tantôt sa lecture. Synthèse symbolique spontanément opérée dès l’origine puisque le corps et le cerveau sont intrinsèquement liés (Bouleau, 2017 ; mais aussi cf. école phénoménologique, in Bakewell, 2018), l’expérience du monde (fondée sur l’observation, la perception) est fondatrice : philosophie rationaliste de Spinoza, liberté de penser de Kant⁸¹, courant des Lumières dans son ensemble (Belhoste, 2018). De ce cheminement itératif abstrait - concret résulte un processus de construction du rapport au monde. Fondée sur cette itération, une telle construction permet d’ordonner, commenter l’expérience du monde et lui fournit un sens. De l’adoption d’une démarche aux critères scientifiques, il découle que le sens fourni enrichit la connaissance, mais aussi se réfère et se nourrit de la connaissance scientifique acquise.

Marx, dans son *Introduction générale à la critique de l’économie politique*, a écrit (in Krivine, 2018) :

80

l’

l’

non seulement bien sûr dans le « pack de connaissance » mais aussi dans les pratiques du métier, la « façon de fonctionner ».

81. pour qui la liberté de penser ne peut se concevoir sans l’appui de l’approche de la connaissance telle qu’elle est défendue par les Lumières.

« La méthode qui consiste à s'élever de l'abstrait au concret n'est pour la pensée que la manière de s'approprier le concret, de le reproduire en tant que concret pensé. »

En d'autres temps, dans ses lettres, Galilée (*in* Wismann, 2017) qui s'interrogeait sur la place que tiennent les mathématiques à l'intérieur de la physique telle que celle-ci commençait à se développer avec lui et autour de lui à son époque, considérait qu'elles correspondaient à « l'encodage du divin dans la création ». Plutôt que de considérer de telles positions comme caduques, opposées, inconciliables, ... Il exprime, avec les mots propres à sa culture et son siècle, ce besoin partagé de mettre en relation l'abstrait (symbolique) et le concret (numérique) ; ce que commentent Kant (1781, 1787) :

« Notre connaissance procède de deux sources fondamentales de l'esprit, dont la première est le pouvoir de recevoir les représentations (la réceptivité des impressions), la seconde le pouvoir de connaître par l'intermédiaire de ces représentations un objet (spontanéité des concepts) ; par la première nous est donné un objet, par la seconde celui-ci est pensé en relation avec cette représentation (comme simple détermination de l'esprit). Intuition et concepts constituent donc les éléments de toute notre connaissance, si bien que ni des concepts, sans une intuition leur correspondant de quelque manière, ni une intuition sans concepts ne peuvent fournir une connaissance. [...] C'est seulement dans la mesure où ils se combinent que peut se produire de la connaissance. »

et Bachelard (1950) :

« En élevant le débat, si la connaissance est un essentiel et constant ajustement de la nature de l'esprit et de la nature des choses, on voit bien que les dialectiques du concret et de l'abstrait - de l'intuition et de la rigueur - de l'expérimental et du rationnel sont des dialectiques motrices, dès l'instertile l O l synthèse dialectique qui est, comme dit Gonseth, « le rythme même du progrès scientifique ». »

Remerciements

Ce travail a bénéficié d'une aide de l'État gérée par l'Agence nationale de la recherche au titre du programme investissements d'avenir pour le projet EQUIPEX GEOSUD portant la référence ANR-10-EQPX-20.

Bibliographie

- Alexandre F. (2019). *De quelles façons l'intelligence artificielle se sert-elle des neurosciences ?* Évènement « Le procès de l'IA », Univ. de Bordeaux : <https://theconversation.com/de-quelles-facons-lintelligence-artificielle-se-sert-elle-des-neurosciences-124402>.
- Bachelard G. (1934). *Le nouvel esprit scientifique*, Alcan, ISBN 2-13-054249-2.

- Bachelard G. (1938). *La Formation de l’esprit par la destruction et la reconstruction de la connaissance objective*. Paris, Vrin, ISBN 2-7116-1150-7.
- Bachelard, G. (1940). *La philosophie du non*, PUF.
- Bachelard G. (1950). L’exactitude discursive. *Études de philosophie des sciences*. n° 10, *Revue de la philosophie*, N° 1 (Suisse), Éditions du Griffon, p. 7-10.
- Baele G., Dellicour S., Suchard M., Lemey P., Vrancken B. (2018). Recent advances in computational phylogenetics. *Current Opinion in Virology*, 31, p. 24-32.
- Bakewell S. (2018). *Au café existentialiste : la liberté, l’être et le cocktail à l’abricot ?* Albin Michel, EAN : 9782226392732.
- Barriel V. et Bourgoïn Th. (2000). Caractères. *Biosystema*, n° 18. Société française de systématique, Paris, ISBN : 2-906892-18-1.
- Belhoste B. (2018). *La science moderne de la Renaissance aux Lumières*. MOOC, Université Paris 1 Panthéon-Sorbonne-16009 : <https://www.fun-mooc.fr/fr/cours/la-science-moderne-de-la-rennaissance-aux-lumieres/>.
- Bernard C. (1865). *Introduction à l’étude de la médecine expérimentale*. Reed. Champs, Flammarion, Paris, ISBN : 9782081307582.
- Bernard C. (1877). *Leçons sur le diabète et la glycogénèse animale*, B B II.
- Bouleau N. (2017). *Penser l’éventuel ; Faire entrer les craintes dans le travail scientifique*, Ed. Quae, Sciences en questions, eBook : 02556NUM.
- Bouleau N. (2021). *Ce que Nature sait, La révolution combinatoire de la biologie et ses dangers*, PUF et discussion autour de cet ouvrage : vidéo 1 : <https://youtu.be/RwjeBBXRgVY> ; vidéo 2 : <https://youtu.be/JeKvcm7YQ-Y> ; vidéo 3 : <https://youtu.be/XOAM-MQmh0k>.
- Boutin J.-P. (2022). *La ville et sa croissance favorisent-elles les moustiques ? Oui ! Mais, la méthode pour l’évaluer n’est pas encore au top !* Site de la Société francophone de médecine tropicale et de santé internationale. https://societe-mts.fr/la-ville-et-sa-croissance-favorisent-t-elles-les-moustiques-oui-mais-la-methode-pour-levaluer-nest-pas-encore-au-top/?utm_source=sendinblue&utm_campaign=Newsletter_N_6_2022&utm_medium=email.
- Brossard Th., Wieber J.-C. (1984). Le paysage, trois définitions : un mode d’analyse et de cartographie. *Espace géographique*, tome 13, n° 1, p. 5-12. <https://doi.org/10.3406/spgeo.1984.3887>.
- Brunet R. (1974). Analyse des paysages et sémiologie. Éléments pour un débat. *Espace géographique*, tome 3, n° 2, p. 120-126. <https://doi.org/10.3406/spgeo.1974.1460>.
- Cazenave A. (2013). *La terre et l’environnement observés depuis l’espace*. Leçons inaugurales du Collège de France. Collège de France / Fayard.
- Chevallier-Le Guyader M-F. (2019). Éditorial. *Le Mensuel*, n° 46, sept., Académie d’agriculture de France. <http://i5vg.mjt.lu/nl2/i5vg/m6jky.html>.

- Claval P. (1974). Géographie et sémiologie. *Espace géographique*, tome 3, n° 2, p. 113-119. <https://doi.org/10.3406/spgeo.1974.1459>.
- Deffontaines J-P. (2004). L'objet dans l'espace agricole. Le regard d'..... In: *La revue de la Société française de géographie*, vol. 12, n° 3, p. 299-304 ; <http://www.cairn.info/revue-natures-sciences-societes-2004-3-page-299.htm>.
- Descola P. (2010). *Diversité des natures, diversité des cultures*, Bayard Ed. ISBN : 978-2-227-50065-5.
- Descola P. (2011). *L'écologie des autres ; l'anthropologie et la question de nature*, Paris, Ed Quae, coll. « Sciences en questions » : <https://doi.org/10.4000/developpementdurable.9639>.
- Dhondt U. (1961). *Science suprême et ontologie chez Aristote*. Revue philosophique de Louvain. Série 3, tome 59, n°61, p. 5-30. <https://doi.org/10.3406/phlou.1961.5068/>
- Fargette M., Loireau M., Raouani N., Kadri Zara M., Simon H., Libourel T., Sghaier M., Leibovici D. (2017). Modélisation de la diversité des systèmes de culture, de l'observation à la détection : cas de l'..... N (D, Tunisie). *Cah. Agric.* 2017, vol. 26, n° 4, article n° 45009, Doi :10.1051/cagri/2017042 fdi:010070799.
- Fargette M., Loireau M., Ben Kathra N., Kiari H., Libourel T. (2018). Conceptual analysis of Climate Change in the light of Society-Environment relationships; Observatories closer to both systems and societies. *Developing and Communicating Climate Change Information for Decision Making*, Serrao-Neumann S., Coudrain A et Coulter L. (Eds), Springer, Nature, Dordrecht, The Netherlands, p. 29-48. https://link.springer.com/chapter/10.1007/978-3-319-74669-2_3; https://doi.org/10.1007/978-3-319-74669-2_3.
- Fargette M., Loireau M. et Libourel T. (2019a). The relationship between Man and his environment: a systemic approach of the viability of « System Earth ». *Coviability of Social and Ecological Systems : Reconnecting Mankind to the Biosphere in an Era of Global Change. Vol. 1 : The Foundations of a New Paradigm*, O. Barrière et al. (Eds), Dordrecht, Springer Nature, eBook ISBN : 978-3-319-78497-7 ; Doi : 10.1007/978-3-319-78497-7 ; <https://www.springer.com/us/book/9783319784960>.
- Fargette M., Loireau, M., Sghaier, M., Raouani, N. et Libourel, T. (2019b). The future of oases in North Africa through the prism of a systemic approach: towards which type of viability and coviability? *Coviability of Social and Ecological Systems: Reconnecting Mankind to the Biosphere in an Era of Global Change. Vol. 2 : Coviability Questioned by a Diversity of Situations*, O. Barrière et al. (Eds), Dordrecht, Springer Nature, eBook. <https://www.springer.com/us/book/9783319781105#otherversion=9783319781112>.
- Fischer M.S. et Tassy P. (2014). Analyse cladistique : le débat Mayr-Hennig de 1974. *Biosystema*, n° 29. Société française de systématique, Paris.
- Ganascia, J. G. (Ed.). (2020). Communication et connaissance : Supports et médiations à l'âge de l'information, CNRS Éditions via OpenEdition.
- George C. (1997). *Polymorphisme du raisonnement humain : une approche de la flexibilité de l'activité inférentielle*. Presses universitaires de France, Paris, coll. « Psychologie et sciences de la pensée ». <https://www.cairn.info/--.htm>.

- George P. et Verger F. (2000). *Dictionnaire de la géographie*. Terme « géographie », p. 212, 7^e édition, © Presse Universitaire de France, 1970, ISSN : 21330479243.
- Gouyon P.H. et Leriche H. (sous la dir.)(2010). *Aux origines de l'environnement*, Fayard. France info.
- Inria (2016). *Intelligence artificielle le défi actuel le défi d'Inria*, livre blanc Inria. <https://www.inria.fr/fr/intelligence-artificielle-les-defis-actuels-et-laction-dinria>.
- Juignet P. (2015). Karl Popper et les critères de la scientificité. *Philosophie*. <https://philosciences.com/112>.
- Kant E. (1781, 1787). *Critique de la Raison pure, Logique transcendantale, introduction*, AK, III, 75, p. 143. <https://www.les-philosophes.fr/auteur-kant.html>.
- Kolimenakis A., Heinz S., Wilson M.L., Winkler V., Yakob L., Michaelakis A., Papachristos D., Richardson C., Horstick O. (2021). The role of urbanisation in the spread of Aedes mosquitoes and the diseases they transmit. *A Systematic Review*. PLoS Negl Trop Dis. 2021 Sep 9, vol. 15, n° 9: e0009631. doi: 10.1371/journal.pntd.0009631. PMID : 34499653; PMCID: PMC8428665.
- Krivine H. (2018). *Comprendre sans prévoir, prévoir sans comprendre*. Cassini.
- Kugler T. A., Grace K., Wrathhall D. J., de Sherbinin A et al. (2019). *People and Pixels 20 Years Later: The Current Data Landscape and Research Trends Blending Population and Environmental Data*, Springer Nature. <https://doi.org/10.1007/s11111-019-00326-5>.
- Lacoste A et Salamon R. (1999). *Éléments de biogéographie et d'écologie*, Nathan, coll. « fac ».
- Lane M. D., Bishop J.L., Dyar M. D., Hiroi T., Mertzman S.A., Bish D. L., King P. L. et Rogers A. D. (2015). Mid-infrared Emission Spectroscopy and Visible/near-infrared Reflectance Spectroscopy of Fe-Sulfate Minerals. *American Mineralogist*, 100, p. 66-82. <https://ur.booksc.eu/book/37940353/5c699e>.
- Lecointre G. et Le Guyader H. (2001). *Classification Phylogénétique du Vivant*, Éditions Belin.
- Lecourt D. (2018). *La philosophie des sciences*. 7^e édition (1^{ère} édition 2001), PUF, ISBN: 978-2-13-080397-3.
- Leroi-Gourhan A. (1964). *Le geste et la parole. Tome 1. Technique et langage*, Éditions Albin Michel, EAN : 9782226198402.
- Leroi-Gourhan A. (1965). *Le geste et la parole. Tome 2. La mémoire et les rythmes*, Éditions Albin Michel, EAN: 9782226198419.
- Linné C. (1758). *Systema Naturæ*.
- Loireau M., Fargette M., Desconnets J.-C., Libourel T. (2015). Observatoire Scientifique en Appui à la GEstion du territoire (OSAGE) : entre espaces, temps, milieux, sociétés et informatique. multigr. Conférence internationale annuelle SAGEO - *Spatial Analysis and Geomatics*, 2014/11/24-27, Grenoble. http://horizon.documentation.ird.fr/exl-doc/pleins_textes/divers16-02/010064550.pdf

- Loireau M., Fargette M., Desconnets J.C., Khiari H. (2017). Observatoire scientifique en appui aux gestionnaires de territoire, entre abstraction OSAGE et réalité ROSELT/OSS. *Revue Internationale de Géomatique*, vol. 27, n° 3, p. 303-333.
- Loireau M., Fargette M. (2022). Science paysagère au service de l'observatoire scientifique Sociétés-Milieus en appui à la gestion territoriale. jimis:8762 - *Journal of Interdisciplinary Methodologies and Issues in Sciences*, 5 janvier, vol 6 - *Observatoires scientifiques Milieus / Sociétés, nouveaux enjeux* : <https://doi.org/10.46298/jimis.8762> ; HAL: hal-03430947v2.
- Lucas R., Mueller N., Siggins A., Owers C., Clewley D., Bunting P., Kooymans C., Tissott B., Lewis B., Lymburner L. et Metternicht G. (2019). *Land Cover Mapping using Digital Earth Australia*. Data, 4, 143. <https://www.mdpi.com/journal/data>.
- Madin J., Bowers S., Schildhauer M., Krivov S., Pennington D. et Villa F. (2007). *An Ontology for Describing and Synthesizing Ecological Observation Data*. *Ecological Informatics*, vol. 2, n° 3, p. 279-296.
- Michelin Y. (2008). *L'approche sémiologique au service de la mise en évidence du lien produit agricole - paysage : l'exemple de l'AOC Saint-Nectaire*. <http://epublications.unilim.fr/revues/as/3409>.
- Minsky M. (1995). Matter, mind and models. Rev. Version of the essay. *Semantic Information Processing*, M. Minsky (Ed.), Cambridge MA, MIT Press.
- Moreau D. (2016). *La philosophie de Descartes*. Vrin, « Repères philosophiques », Paris. Philosophie 2.0, une introduction à la philosophie : <https://www.les-philosophes.fr>.
- Oberlack C., Breu T., Ciger M., Harari N., Mathez-Stiefel S-L., Messerli P., Moser S., Ott C., Providoli I., Tribaldos T., Zimmermann A. et Schneider F. (2019). *Theories of Change in Sustainability Science. Understanding How Change Happens*. GAIA, vol. 28, n° 2, p. 106-111.
- Page R.D.M. et Holmes E.C. (2004). *Molecular Evolution. A Phylogenetic Approach*, 3rd edition, Blackwell.
- Platon, *La République*, VII.
- Pitte JR. (2010). *Le génie des lieux*, CNRS Éditions, coll. « Débats ».
- Ramognino N. (2014). *Les entités sociales : réflexions ontologiques*. *SociologieS, Grands résumés*. <http://journals.openedition.org/sociologies/4650>
- Livet P. et Nef F. (2009). *Les Êtres sociaux. Processus et virtualité*, Paris, Éditions H M M N R <http://sociologies.revues.org/4647>].
- Rousseau J.-J. (1989 [1755]). *Discours sur l'origine et les fondements de l'inégalité parmi les hommes*, Gallimard, coll. « Folio Essai ».
- Serres M. (2014). *Yeux*, Éditions Le Pommier, Paris, ISBN : 978-2-7465-0779-1.
- Shapin S., Schaffer S. (2011). *Leviathan and the Air-pump: Hobbes, Boyle, and the Experimental Life*, vol. 1, Princeton University Press.

- Triclot M. (2005). « Penser, c’est calculer » : éléments pour une préhistoire de l’informatique. *Journée d’études EDIIS*, EDIIS, Lyon 1, Feb., Lyon, France. Hal-01526653.
- Viers G. (1967). *Éléments de géomorphologie*, Nathan.
- Village Daza D. A., Sanchez Moreno H., Porta L., Portantiolo Manzolli R., Bolivar-Anillo H. J. et Anfuso G. (2020). *Mangrove Forests evolution and Threats in the Caribbean Sea of Colombia*. Water, vol. 12, n° 4, 1113. <https://doi.org/10.3390/w12041113>.
- Wehrden H. von, Guimarães M. H., Bina O., Varanda M. *et al.* (2018). Interdisciplinarity and transdisciplinarity research: finding the common ground of multi-faceted concepts. *Sustainable Science*. <https://doi.org/10.1007/s11625-018-0594-x0123456789>.,-volIV.
- Wismann H. (2017). D’où vient la notion de fonction ? *Sciences de la vie, sciences de l’information*, Gaudin T., Lacroix D., Maurel M.-C., Pomerol J.-C. (dir.), Colloque de Cerisy, p. 176-183.
- Wolff F. (2010). *Notre humanité. D’Aristote aux neurosciences*, Éditions Fayard, Paris.
- Yourcenar M. (1951). *Mémoire d’Hadrien*, Éditions Plon, Paris.

Annexe – Vocabulaire consensuel et à entrées multiples

Voici le sens donné à certains termes à l’issue du dialogue entre les auteurs étant donné leurs points de vue disciplinaires en biologie, écologie, agronomie, géographie et science de la donnée et des modèles (vocabulaire partagé). Ce sont tous des concepts-clés dans le cadre de ce travail.

Cette liste est à entrées multiples et montre les relations que les termes entretiennent entre eux et/ou en quoi ils se distinguent les uns des autres (vocabulaire croisé). A sa façon, ce corpus propose en soi un résumé du contenu sémantique de ces concepts-clés.

Compartiment (*cf.* « système »)

Empreinte, genèse d’empreinte (*cf.* figures. 7, 11, 13, 16)

De nature concrète, l’**empreinte** est un concept charnière entre les domaines systémique et géographique. Elle permet la communication entre eux, fonde le passage et conduit le raisonnement ontologique qui donne aux choses du monde concret leur sens systémique. L’empreinte permet de raconter (orthographe et grammaire) le monde réel (*cf.* « réalité ») qui procède de système(s), de leur structure, de leur fonctionnement, de leur dynamique (*cf.* « système »). C’est par la **genèse d’empreinte** que la réalité prend forme, par l’empreinte, inscription de système(s) dans la matérialité. Par son ascendance, l’empreinte se réfère (ontologie) à une entité systémique unique, l’entité d’ancrage (*i.e.* compartimentation systémique) dont elle matérialise l’existence. L’empreinte est logique ; c’est de son ascendance systémique, via la genèse d’empreinte, qu’elle tient sa cohérence, et par là le sens qu’elle donne à la réalité, ainsi que les grandes lignes de sa forme et de son

étendue. Sans l’empreinte, le monde perçu serait inintelligible. L’empreinte est incontournable. L’empreinte est un concept géographique.

L’empreinte est ontologiquement distincte de la formation et de sa physionomie (cf. « formation » et « physionomie ») ; l’empreinte ne correspond pas (ou pas totalement) à la formation, elle ne retient d’elle que ce qui parle d’une (seule) origine systémique. Elle peut aussi spatialement chevaucher plus d’une formation, ou même s’inscrire au-delà de la formation. C’est par les concepts d’empreinte et d’épure (cf. « épure ») que l’on peut dénouer la confusion que risque de contenir ce que l’on appelle formation, perçue et arbitraire (donc sujette à caution).

Entité et épure (cf. figures 10, 11, 12, 16)

De nature abstraite, l’**entité** est un compartiment (cf. « système ») extrait d’un système par un point de vue. Elle regroupe les propriétés principales, nécessaires à une logique systémique selon ce point de vue disciplinaire, pluri ou inter-disciplinaire. Elle peut être quasi système.

Selon le *Lien de Forme*, l’entité d’ancrage devient référente ; elle se profile en **épure**, concept abstrait où sont recensées les caractéristiques et de forme (cf. « forme ») encore en potentiel de s’inscrire dans la réalité. L’épure est modèle géographique qui dessine l’empreinte, lui donne forme. L’exemple donné avec l’oasis en présente les propriétés (en termes systémiques) et les formes correspondantes (en termes géographiques). Dans la démarche d’ontologie systémique, l’entité d’ancrage est décisive dans le dessin de l’épure qui en découle.

Épure (cf. « entité »)

Existence et existence géographique (cf. figure 5)

Concept très général, l’**existence** tient à l’implantation d’un système dans un espace à n dimensions, et donc au déploiement de ses lois, de sa logique dans cet espace. L’**existence géographique** tient à l’implantation d’un système (et au plus haut niveau d’intégration, du Système Terre) dans l’espace géographique, caractérisé par ses 4 dimensions (spatio-temporelles : x, y, z, t) et sa matérialité.

Extrinsèque (cf. « référentiel »)

Formation (cf. figures 1, 16)

La formation est perçue empiriquement par l’observateur, émerge « spontanément » du milieu qui la contient (par fragmentation de l’espace), « elle saute aux yeux », elle est constatée sans qu’une interprétation fondée sur un raisonnement ou une connaissance préalable ne semblent être à l’origine de son constat. L’enjeu est ensuite de lui attribuer un sens possible selon une logique systémique.

Forme (cf. « ontologie systémique »)

Lien de Forme (cf. « ontologie systémique »)

Lien de Sens (cf. « sens »)

Milieu

Le milieu correspond à l'ensemble des conditions, des facteurs qui caractérisent le contexte environnemental (bio-physico-socio-écologico-écono-systémique) auquel est soumis un élément d'intérêt dans un espace géographique.

Modèle (cf. « système »)

Monde perçu (cf. « observé »)

Observé et monde perçu (cf. figures 1, 7, 15, 16 ; tableau 1)

L'**observé** est ce qui est restitué par des capteurs physiologiques ou technologiques (cf. « tangible ») ; il se présente sous la forme de tableaux de données non agrégées ou bien sous forme synthétique, par exemple la physionomie d'une formation. L'expérience d'observation dont il découle est toutefois une expérience limitée par la nature des capteurs. La perception (quelle qu'elle soit la nature) est une expérience limitante. Le **monde perçu** résulte, dans son ensemble, de cette expérience. Il est variable selon la technologie de capture mise en jeu. Celle-ci est évolutive avec l'amélioration des technologies et la mise en place de nouvelles technologies (qui proposent de nouveaux registres de perception). N'oublions pas (même si ce n'est pas le sujet ici) que des représentations individuelles et collectives, culturelles peuvent se surimposer aux capteurs eux-mêmes, en tant que filtres (manque d'objectivité) ou infléchir l'interprétation. De tels « biais » peuvent de plus se modifier dans le temps. Les protocoles d'observation et la démarche scientifique dans son ensemble ont pour intention de minimiser de tels biais.

Ontologie systémique, *Forme* et *Lien de Forme* (cf. figures 7, 11, 12, 13 14, 15, 16, 17)

Avec le recours à l'organisation systémique et plus particulièrement dans ce travail au processus de genèse d'empreinte, la démarche (telle une volée de marches qui représentent les étapes du passage) recherche a) l'ascendant systémique à partir duquel poser le raisonnement (analyse critique du choix de l'entité posée pour l'ancrage) ; puis b) la forme que cette entité prendra en tant que projet (modèle géographique) [il s'agit de l'épure], puis dans le monde réel [il s'agit de l'empreinte], puis dans le monde perçu [il s'agit d'un bouquet de signaux]. Le **Lien de Forme** en est le raccourci intuitif qui « se descend » de l'organisation systémique vers le monde perçu (données). L'**ontologie systémique** est la construction formelle de ce raisonnement qui s'ancore (en amont) dans une entité systémique et pose (en aval) la forme, la contextualise et la distingue des physionomies apparentes (cf. physionomie). La **forme** est une propriété de l'empreinte : l'empreinte prend forme ; la forme est la synthèse dessinée à partir de l'épure (modèle géographique) et concrétisée dans la réalité ; elle est décrite, répertoriée selon la liste de caractères pertinents que définit l'entité. La forme correspond, dans le monde concret au cœur et à la corolle de l'empreinte, respectivement « programmés » dans l'épure et la variation circonstancielle. Parce qu'elle se réfère à une seule entité d'ancrage par son ontologie systémique, la forme est pure.

Le *Lien de Forme* permet, selon un point de vue, de caractériser la forme, dotée de logique systémique, pour la rechercher dans un observé complexe (cf. « observé »). De façon idéale, les protocoles d'observation doivent alors tenir compte, au plus près, des propriétés de forme de l'empreinte. L'ontologie systémique donne ainsi les clés d'une « lecture à rebours » de l'espace géographique. L'organisation de la connaissance selon cette démarche est utile à la « lecture du monde » (en y reconnaissant des formes et par conséquent des empreintes).

Physionomie (cf. figures 8, 9, 16)

La physionomie relève du monde perçu (cf. « monde perçu »). Lors d'une expérience d'observation, et dans le cas où l'espace est fragmenté, une physionomie est attribuée à une formation. Si l'espace n'est pas fragmenté, la physionomie se rapporte à la fenêtre d'observation ; elle correspond à une synthèse de l'observé sur cet espace.

La physionomie peut contenir un nombre de formes (cf. « forme » et « empreinte »), juxtaposées ou surimposées, correspondant à autant de points de vue différents initiant chacun son ontologie propre, chacun à partir d'une entité d'ancrage pertinente et unique. En effet, en un lieu donné, différents systèmes sont à l'œ 1 1 D 1 1 1 le seul point de vue retenu, ces effets aux origines diverses sont des accidents fortuits et contingents du lieu (sans aucune logique apparente) ; ils figurent cependant aussi dans la physionomie. On serait dans l'erreur de les confondre avec l'empreinte, de les lui attribuer, d'où l'utilité du travail fondé sur la démarche d'ontologie systémique. De telles observations sont hors-sujet et participent au bruit d'une donnée composite, brouillée et, en cela, « dégradée », bruit que la qualité des protocoles et la logique d'échelle d'observation tentent de minimiser.

La physionomie observée contient donc plus que l'empreinte (la forme pure) et, en ce sens, brouille l'empreinte. Ni la formation ainsi « observée », ni sa physionomie ainsi « décrite », ne sont proprement systémiques.

Lire une physionomie (dans toute sa diversité) c'est y décrypter, y distinguer les empreintes à l'aide de leurs épures respectives (leurs modèles géographiques respectifs, utilisés en tant que critères dans les protocoles d'observation et/ou les protocoles d'analyse). Chacune, individuellement et en interaction, participe à la synthèse observée dans la physionomie générale.

Réalité (cf. figures 5, 7, 16)

Elle procède de l'existence géographique (cf. « existence ») et elle est de composante matérielle. Elle est en soi, elle existe, qu'elle soit perçue ou non par un capteur, qu'elle soit pensée ou non par un « agent ». Non encore diminuée, appauvrie, déformée, fragmentée par la perception, elle est entière et référentiel concret. La réalité procède de système(s) implanté(s) et inscrit(s) dans l'espace géographique (elle procède de leur structure, de leur fonctionnement et de leur dynamique). La réalité est une empreinte (la planète Terre, empreinte du système

Terre) ; elle peut aussi être considérée comme un agrégat d'empreintes, celles-ci correspondant aux inscriptions de systèmes de moindre niveau d'intégration (*i.e.* des composants systémiques du système Terre).

Référentiel et extrinsèque (*cf.* figures 4, 7)

Dans le domaine scientifique dont la démarche commence par l'objectivation, le référentiel est extrinsèque c'est-à-dire qu'il ne dépend pas de l'être qui raisonne ou du capteur qui observe. La réalité (*cf.* « réalité ») est référentiel concret. Le système (*cf.* « système ») est référentiel abstrait. Référentiel « système » et référentiel « réalité » sont intimement liés en sciences, en particulier par le concept d'empreinte.

Sens et Lien de Sens (*cf.* figures 7, 8, 16, 17)

Il s'agit de la signification recherchée, du sens systémique trouvé, *i.e.* de l'intelligibilité conférée par une cohérence systémique du comment (comment c'est fait - structure typée ; comment ça fonctionne - fonctionnement expliqué ; comment ça change - évolution suivie, expliquée, etc.). Le **Lien de Sens** « se remonte » du monde perçu (données) vers l'organisation systémique (structurelle, fonctionnelle, évolutive).

Système, modèle et compartiment (*cf.* figures 4, 7)

Ensemble de lois, structuré et fonctionnant, le **système** est le principe organisé (... mais aussi se transformer (évolution) ou disparaître s'il n'est plus viable (extinction) (Fargette *et al.*, 2019a). Le **modèle** (ce que l'on comprend du système) tend vers le système. Le modèle parfait serait celui qui présenterait le système dans son entier (version holiste). Le système est la référence extrinsèque que chaque modèle ne peut représenter que selon un point de vue, par compartimentation systémique. Le modèle est construit et représente un **compartiment** (une partie de système, parfois quasi-système) extrait par un point de vue sur le système. De nature abstraite, le système est en soi (et, une fois implanté, il existe), qu'il soit modélisé ou non, qu'il soit pensé ou non. Non encore diminué, déformé, compartimenté par la raison et ses limites, il est entier et référentiel abstrait.

Tangible (*cf.* figures 15, 16)

Caractéristique de la réalité (le référentiel concret), capable d'être perçue (touchable, visible, etc.), tout autant que soit à disposition le capteur adéquat (*cf.* « observé »). Le fait d'être concret ne suffit pas pour être observé : il faut aussi que les conditions de perception et les propriétés physiologiques ou technologiques du capteur permettent la capture du signal, formule dégradée du signe. Dans ce travail le terme tangible est une extension à tout capteur de la définition du CNRTL : « Que l'on peut connaître en touchant ; fait d'être perceptible par le toucher ».

Une ontologie orientée objet pour la modélisation des socio-écosystèmes

Éric Masson

Laboratoire TVES ULR 4477, Université de Lille, Cité Scientifique,
av. Paul Langevin, 59650 Villeneuve-d'Ascq, France
eric.masson@univ-lille.fr

RÉSUMÉ. Cet article de positionnement théorique est une contribution ontologique pour la modélisation des socio-écosystèmes (SES). C'est également une prise de position épistémique qui est ancrée sur la démarche orientée objet. Nous proposons ainsi un modèle d'organisation des connaissances qui permet d'intégrer plus de complexité dans la déconstruction pluridisciplinaire et l'analyse des SES. Cette proposition ontologique orientée objet s'appuie sur six concepts de haut niveau d'abstraction permettant une portabilité transdisciplinaire autour des structures, fonctions, connexions, phases, échelles et adaptations des SES. Après avoir défini notre positionnement sur la modélisation spatiotemporelle orientée objet, nous présentons les six concepts et nous explicitons l'intérêt d'une ontologie de haut niveau, orientée objet, pour les SES pour conclure sur les perspectives d'organisation des connaissances qui prolongent notre proposition.

ABSTRACT. This article is a positional paper and an ontological contribution to the modelling of socio-ecosystems (SES). It is also an epistemic position that is anchored on the object-oriented approach. We thus propose a model of knowledge organisation that allows more complexity to be integrated into the multidisciplinary deconstruction and analysis of SES. This object-oriented ontological proposal is based on six concepts of high abstraction level allowing a transdisciplinary portability around the structures, functions, connections, phases, scales and adaptations of SES. After defining our position on object-oriented spatiotemporal modelling, we present the six concepts and explain the interest of a top level, object-oriented, ontology for SES to conclude on the perspectives of knowledge organisation that extends our proposal.

MOTS-CLÉS : géomatique, ontologie, objet, système, environnement, société, structure, fonction, connexion, phase, échelle, adaptation.

KEYWORDS: Geomatic, Ontology, Object, System, Environment, Society, Structure, Fonction, Connection, Phase, Scale.

DOI:10.3166/RIG.31.199-230 © 2022 Lavoisier

1. Introduction

Les socio-écosystèmes (ou SES dans la suite de l'article) sont des systèmes réputés complexes (Gunderson et Holling, 2002 ; Allen et Holling, 2008 ; Jorgensen *et al.*, 2007 ; Fustec et Lefeuvre, 2000, Le Moigne 1999 ; Partelow, 2018 ; Baartman *et al.*, 2020 ; Hervé *et al.*, 2016 ; Liu *et al.*, 2007) car ils hybrident et métabolisent des processus et des relations entre les sociétés et leurs environnements écologiques (Loireau *et al.*, 2017 ; Barnaud, 2000). Depuis plusieurs décennies (Kuipers, 1978, 1994 ; Laurini, 1990 ; Saint-Gérard, 2002 ; Tissot et Cuq, 2004 ; Tran *et al.*, 2017), leur modélisation et leur analyse spatiale s'appuient sur des systèmes d'informations géographiques (SIG) à tel point que le développement même des outils et méthodes implémentés dans les SIG peut être considéré comme une réponse au besoin d'intégration de cette complexité (Burrough et Frank, 1996 ; Clementini et Di Felice, 1996 ; Clementini *et al.*, 2021). Cependant certains aspects de la complexité des SES constituent des verrous scientifiques (Le Moigne et Morin, 2007) qu'il est difficile de modéliser sans les simplifier par des approches plus ou moins compliquées. Parmi ceux-ci nous pouvons citer les phénomènes d'apprentissage (Ehresmann et Vanbremeersch, 2007 ; Daniels et Walker, 1996 ; Morin, 1986 ; Le Moigne, 1999), d'auto-organisation (Bak, 1996 ; Morin, 1986), d'émergence (Folke *et al.*, 2003 ; Folke, 2006 ; Jorgensen *et al.*, 2007), de résilience (Gunderson et Holling, 2002 ; Jorgensen *et al.*, 2007), d'évolution adaptative (Steneck, 1986 ; Gunderson et Holling, 2002 ; Jorgensen *et al.*, 2007, Walker *et al.*, 2004 ; Folke *et al.*, 2003), de directionnalité (Jorgensen *et al.*, 2007), d'intentionnalité (Brentano, 1976 ; Johansson, 1992, 2004 ; Smith *et al.*, 1994 ; Husserl, 1900), d'homéostasie (Ehresmann et Vanbremeersch, 2007 ; Meurnier *et al.*, 2014 ; Von Bertalanffy, 1968), de dyschronie (Alter, 2003), etc. Les dimensions spatiotemporelles des SES participent d'une complexité qui s'accommode très mal de modèles réductionnistes ou déterministes (Malczewski et Jankowski, 2020). Il ne suffit pas de positionner les éléments de connaissance d'un SES dans un modèle spatial 3D-temporel pour lever les incertitudes liées aux 1) conditions initiales, car il y a toujours un état de connaissances et de dépendances antérieures du SES qui n'est pas intégré, et 2) conditions aux limites, tels les phénomènes de forçage exercés par des externalités qui, par définition, sont externes donc exclues du modèle. Cette sensibilité aux conditions initiales/aux limites est source d'incertitude majeure dans les modèles systémiques en particulier pour les facteurs sociétaux des SES qui possèdent un comportement rarement déterministe (Allen et Holling, 2008 ; Le Moigne, 1999 ; Bak, 1996).

Un premier défi à relever pour modéliser les SES tient dans la capacité du/des modélisateur(s) à assembler une base de connaissances qui doit nécessairement s'appuyer sur plusieurs corpus disciplinaires des sciences sociales et de l'environnement. La première des complexités s'observe donc dans la formalisation des connaissances (Loireau *et al.*, 2017) où chaque partie prenante 1) projette son domaine de connaissances dans une démarche pluridisciplinaire, 2) intègre son domaine de connaissances dans une démarche interdisciplinaire ou 3) hybride son

domaine de connaissances dans une démarche commune de type transdisciplinaire. Cette première étape de la modélisation revient à compiler et à organiser les connaissances et leurs corpus sémantiques dans une démarche ontologique dont l'objectif est d'identifier les objets étudiés (les êtres) et d'en décrire leurs propriétés (leurs étant). Pour un SES, l'objectif ontologique consiste donc à modéliser les composants (systèmes et sous-systèmes), les propriétés (fonctions et processus) et les connexions (liens et interrelations). Les SES étant des systèmes spatiotemporels, les outils et les méthodes de la géomatique participent également à l'objectif ontologique de modélisation. Il devient alors incontournable d'intégrer les connaissances disponibles sur les objets étudiés dans un modèle systémique « prêt à l'emploi » construit et organisé en fonction de leurs dimensions spatiales, par des entités graphiques (points, lignes, surfaces et pixels, volumes et voxels), dans leur dimension temporelle, par une localisation dans le temps (dynamique temporelle), et dans leurs dimensions analytiques, par leurs propriétés (attributs et relations).

Cet article propose un modèle d'organisation des connaissances s'appuyant sur une démarche ontologique transdisciplinaire qui est généralisable à l'étude des SES. Cette proposition est le résultat d'une recherche théorique menée en parallèle de projets de recherches dont les enjeux de pluridisciplinarité (géologie, écologie, génie civil, informatique, géomatique, géographie, urbanisme, sociologie, économie, histoire, archéologie) concernent différents thèmes (risques naturels, gestion intégrée des ressources en eau, gestion intégrée des ressources environnementales, économie circulaire appliquée à la gestion des sédiments, systèmes participatifs d'aide à la décision, etc.) pour lesquels les outils de la géomatique (SIG et télédétection) étaient systématiquement mobilisés en lieu et place de système d'information (ou SI). L'objectif ontologique recherché ici est de mieux articuler les connaissances pluridisciplinaires afin de mieux intégrer la complexité des SES dans les outils de modélisation spatiotemporelle.

Cette démarche place notre proposition dans la catégorie des ontologies de haut niveau (Sowa, 1995 ; Vogt, 2011 ; Partridge et Stefanova, 2003 ; Garcia, 2020 ; Arp *et al.*, 2015) ou *Top Level Ontology* (TLO dans l'article). Plusieurs ontologies de très haut niveau sont reconnues telles que SUMO (Niles et Pease, 2001), DOLCE (Gangemi *et al.*, 2002), BORO (Partridge et Stefanova, 2003), UFO (Guizzardi, 2005), SWEET (Raskin et Pan, 2005 ; DiGiuseppe *et al.*, 2014), GFO (Herre, 2010) ou encore BFO (Arp *et al.*, 2015) et couramment utilisées comme fondation pour des ontologies de domaine permettant l'instanciation d'ontologies applicatives (Ye *et al.*, 2011 ; Arp *et al.*, 2015). Nous ne développons pas ici les différences entre chacune de ces TLO et nous ne faisons pas non plus de comparaisons entre notre proposition et une ou plusieurs de ces TLO car ce n'est pas l'objectif de cet article.

Notre proposition s'appuie sur le paradigme orienté objet (OO dans l'article) qui a révolutionné la programmation dès les années 1960 et constitue également le fondement de la modélisation des bases de données géographiques. Si l'application de celui-ci à l'analyse d'image à références spatiales (aériennes et satellitaires) semble récente (Blaschke, 2010 ; Blaschke *et al.*, 2014) sous la dénomination

GEOBIA (*GE*ographical *O*bject *B*ased *I*mage *A*nalysis), elle prolonge en fait des travaux plus anciens sur l'imagerie aérienne (Bausch, 1979 ; Fu et Mui, 1980), dans la poursuite de travaux sur l'extraction automatique d'objets dans les images numériques statiques (Prewit, 1971) ou dynamiques (Radig, 1978). Le concept d'objet, pris dans le sens de chose, d'item, d'entité réelle ou abstraite est lui-même consubstantiel de toute ontologie comme le rappelle Barry Smith : « Ontology as a branch of philosophy is the science of what is, of the kinds and structures of the objects, properties and relations in every area of reality. » (Smith, 1999). Il n'est donc pas surprenant qu'une école de pensée philosophique propose d'appliquer le paradigme OO à l'ontologie elle-même dans ce que Graham Harman définit comme l'« Object Oriented Ontology » ou OOO (ontologie orienté objet) (Harman, 2018) et que l'objet soit lui-même considéré comme l'un des rares concept théorique transdisciplinaire (Heller, 1984, 1990 ; Carnap, 1922 ; Casati et Varzi, 1999).

Cette contribution s'articule donc en trois parties. La première s'appuie sur réflexion sur le concept orienté objet comme abstraction élémentaire pour la modélisation spatiotemporelle appliquée aux SES.

La deuxième partie analytique propose de définir un ensemble de six concepts utiles à la modélisation du métabolisme des SES. Nous démontrons la « portabilité » pluridisciplinaire de ces six concepts en élicitant chacun d'entre eux avec des exemples pris dans les domaines géomorphologique, écologique, urbain et culturel qui participent au fonctionnement des SES.

Enfin, la troisième partie intègre les réflexions des deux précédentes dans une OOO qui s'appuie sur les six concepts appliqués aux SES pour construire un modèle à six dimensions. Cette proposition théorique ambitionne d'agréger plus de complexité, à défaut de toute la complexité, pour l'analyse des SES spatiotemporels en substituant un positionnement des connaissances dans un modèle spatiotemporel multidisciplinaire à un positionnement des connaissances dans une OOO 6.0 (ontologie orienté objet de dimension 6) transdisciplinaire. L'importance de la dimension adaptative et la question de la topologie de notre proposition pour la modélisation des SES seront également discutées.

2. Modélisation spatiotemporelle orientée objet pour les socio-écosystèmes

2.1. L'objet, une abstraction élémentaire

Dans notre proposition OOO 6.0, chaque entité, attribut ou relation du SI est un conceptuellement appelé un objet. Notre concept d'objet correspond donc simplement à une abstraction physique ou sémantique. Par abstraction physique nous considérons que notre objet représente toute chose constituée de matière dans l'espace physique du monde réel. Par abstraction sémantique nous considérons que notre objet représente toute chose immatérielle portée ou échangée par des objets physiques localisables dans l'espace du monde réel. L'objet est donc une abstraction

au sens le plus large. Ce concept possède aussi l'avantage d'inclure la notion de délimitation et il intègre également, conceptuellement, la notion d'assemblage. Un objet peut donc être composé des éléments qui le constituent définissant à la fois sa singularité et ses propriétés.

Un objet devient lui-même un élément s'il participe de l'assemblage d'un autre objet de rang supérieur, un système. Nous reconnaissons cependant deux objets d'exception à cette logique élément-assemblage : la particule élémentaire et l'univers. Elles constituent pour les limites infranchissables, internes (limite inférieure pour la particule) et externes (limite supérieure pour l'univers), de la connaissance donc de la modélisation. La première est à la fois élément et objet : l'internalité de l'objet particule élémentaire est donc exclue. La seconde est l'objet de tous les éléments : l'externalité de l'objet univers est aussi exclue. L'objet univers a donc pour dimension une étendue, un intervalle ou une enveloppe d'existence, dont la limite interne (limite inférieure, fondement) est la particule élémentaire et la limite externe (limite supérieure, totalité) est l'univers lui-même. De la particule à l'univers, il existe un ensemble d'assemblages ou de combinaisons possibles, donc d'objets, dont le nombre ou la complexité n'est pas défini ou déterminé, *a priori*. La première des décisions dans la modélisation d'un objet d'étude consiste donc à délimiter, à borner, la part de complexité et le système d'objets qui est lui-même l'assemblage d'éléments à étudier.

Les objets peuvent être matériels (Heller, 1984, 1990) ou immatériels (Arp *et al.*, 2015 ; Smith et Mark, 2003 ; Smith et Varzi, 2000 ; Harman, 2018). Cette distinction ontologique très ancienne remonte à Aristote (Hennig, 2008). S'ils sont matériels, ils disposent d'une réalité physique, qui est observable (perçue par un capteur) et que l'on peut localiser et dimensionner (grandeur physique) dans un segment du système physique espace-temps. L'objet matériel (physique), a pour dimensions une étendue, délimitée par une internalité (limite inférieure) et une externalité (limite supérieure), dans ses quatre dimensions, il existe donc dans un volume 3D-temporel (4D). Cette étendue 4D est une enveloppe qui possède deux surfaces : une surface interne et une surface externe. La différence entre ces deux surfaces est un intervalle qui exclut 1) un volume interne (l'internalité), qui peut alors être considéré comme une boîte noire interne, et 2) un volume externe (l'externalité) qui peut alors être considéré comme une boîte noire externe.

Les objets immatériels, quant à eux, ne disposent pas d'une réalité physique mais disposent d'une réalité cognitive et sémantique que l'on peut également localiser et dimensionner dans l'espace et le temps. Ce sont des abstractions que l'on ne peut pas observer ou dimensionner sans utiliser un modèle conceptuel. Ainsi une connaissance est un type d'objet immatériel qui peut être rattaché à la localisation matérielle d'un individu ou d'un groupe qui la partagent et l'utilisent. Les objets immatériels sont des objets intentionnels (Johansson, 1992 ; Smith et Woodruff-Smith, 1995 ; Smith, 1996) ou sensibles (Harman, 2018) qui sont plus complexes que les objets matériels car ils encapsulent une part plus ou moins grande de subjectivité de l'individu ou du groupe qui le porte.

L'objet immatériel, comme l'objet matériel (physique), a lui aussi pour dimension une étendue, un intervalle, délimitée par une internalité (limite inférieure) et une externalité (limite supérieure). La distinction entre délimitation construite (*fiat boundary*) qui correspond au modèle de la réalité et la délimitation véritable (*bona fide boundary*) qui correspond à l'objet du monde réel (Smith et Varzy, 2000) pose le problème du réductionnisme en modélisation (*fiat boundary*) et de la difficulté à délimiter certains objets (*bona fide boundary*) dans le monde réel (Smith et Mark, 2003). Dans notre approche, un objet est donc une enveloppe dont l'abstraction numérique est conçue pour modéliser un domaine de connaissances des SES. Nous appelons objet géographique un assemblage d'éléments, matériel, immatériel ou hybride, localisable dans un système de coordonnées terrestres. Nous appelons objet spatiotemporel un objet géographique qui est également localisable dans le temps.

2.2. L'objet spatiotemporel

Dans les bases de données construites pour l'analyse des SES la modélisation des informations s'articule sur l'association entité/attribut/relation. Dans les SI spatiotemporels, les entités sont identifiées et décrites géométriquement dans un espace respectant la métrique d'un système de coordonnées géographiques (Laurini, 1990). Les objets du « monde réel » sont modélisés selon des entités discrètes par des vecteurs (figure 1) ou selon des espaces continus échantillonnés par des pixels (Laurini, 1990). Les attributs sont donnés (valeurs initiales) ou construits (valeurs déduites ou observées) pour décrire la (les) valeur(s) sémantique(s) de chaque entité. Les relations entre entités à références spatiales (relations spatiales) et entre attributs (relations attributaires) sont, elles aussi, données ou construites afin de modéliser et d'organiser les traitements numériques qui sont implémentés dans le SI. Attributs et relations sont les propriétés des entités. L'association entre géométries, attributs et relations repose sur le modèle entité-association (Bouillé, 1975, 1977, 1978 ; Saint-Gérand, 2002 ; Pirot et Saint-Gérand, 2005) qui lie chaque entité à une description sémantique normalisée (attributs, relations et métadonnées) par un identifiant.

Les entités sont produites et contraintes par leurs modèles numériques, elles sont dimensionnées selon une résolution spatiale et temporelle définie par l'utilisateur. Elles sont pensées et construites pour résoudre une question thématique elle-même modélisée en fonction de connaissances comprises dans des cadres théoriques et méthodologiques choisis par l'utilisateur. Elles résultent donc d'un processus de réductionnisme nécessaire à tout exercice de modélisation qui dépend fortement de l'utilisateur.

Dans le « monde réel », les objets géographiques présentent une dynamique qui dépend des systèmes territoriaux, c'est-à-dire des contextes spatiotemporels, dans lesquels ils évoluent. L'objet est donc une interface entre ses relations internes (constitutives) et ses externalités (contextes). A certaines échelles d'analyse, l'objet géographique peut-être indéterminé ou non localisé en raison de paramètres

d'observation, de temps ou de métabolisme. Ses limites peuvent apparaître, disparaître, évoluer ou être floues comme dans le cas de la délimitation des zones humides (Barnaud, 2000 ; McMullen et Meacham, 1966) ou d'une montagne (Smith B. et Mark, 2003). L'objet géographique du monde réel peut aussi être actif, inactif, évolutif, latent (en veille) ou sous-jacent (encapsulé). Si les notions de résolutions spatiale (géométrie), temporelle (dynamique), sémantique (description) et de contexte (relations) jouent un rôle fondamental dans l'observation du « monde réel », s'y ajoute également un facteur de réduction (échelle de simplification) dans la modélisation systémique et dynamique des objets étudiés. Cependant, il existe un hiatus entre la complexité du « monde réel » et le réductionnisme nécessaire à tout exercice de modélisation (De Sède-Marceau et Moine, 2011). Ce problème n'est pas nouveau y compris en géomatique.

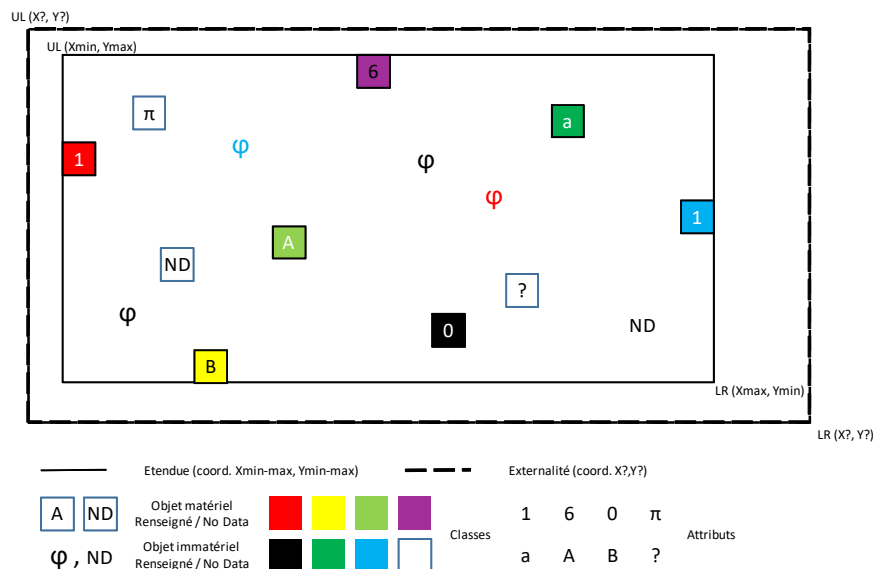


Figure 1. Exemple de représentation conceptuelle des objets vectoriels dans un SI à références spatiales (les objets matériels ou immatériels peuvent appartenir à des classes et possèdent – ou non – des attributs quantitatifs ou qualitatifs)

2.3. Socio-écosystèmes et ontologie orientée objet

Si la conceptualisation OO est largement recevable dans les disciplines relevant des sciences et technologies, elle n'est pas aussi largement acceptée dans celles relevant des sciences du vivant et plus particulièrement des sciences humaines et sociales. La finalité du concept d'objet pose un problème éthique lorsqu'il s'agit d'étudier les êtres vivants en général et les êtres humains en particulier. En effet, l'objet, dans sa définition la plus large, est une chose façonnée, un artefact construit

pour être manipulé en fonction d'un objectif, d'une intention ou d'une finalité (Gruenfeld *et al.*, 2008). Or, l'idée même qu'un être vivant, qu'un humain ou qu'une société soient considérés comme des objets, donc soient manipulables et manipulés par la science et pour la recherche, est une posture scientifique qui n'est pas tenable. Cette insoutenabilité alimente largement les débats sur l'éthique en santé (Halioua, 2017), en environnement (Hess, 2017) ou en sciences sociales (Marx, 1844 ; Nussbaum, 1995 ; Gervais *et al.*, 2013 ; Hess, 2017). Le terme d'objet est dérangeant et incite les sciences humaines et sociales et une bonne partie des sciences du vivant à lui préférer d'autres concepts plus respectueux de l'être tels qu'individu, sujet, famille, groupe, population, société, etc.

Sans poursuivre plus avant sur ce débat éthique qui dépasse la finalité de cet article, il me semble utile de clarifier la place de l'objet dans la modélisation des SES qui est proposée. L'approche OO retenue considère l'objet comme une abstraction de la réalité, matérielle ou immatérielle, des éléments constitutifs d'un SES. Un être vivant n'est donc pas un objet au sens strict mais son abstraction, dans le système d'information, est considérée comme un objet renseigné, une représentation, un élément de connaissance participant à l'étude d'un SES.

Les objets, matériels ou immatériels, parce qu'ils sont construits intentionnellement, sont délimités, ils ont des formes, des longueurs des surfaces, des volumes (3D). Parce qu'ils sont délimités, ils partagent aussi une co-limite, une topologie, avec leur environnement, leur voisinage, d'autres objets qui en deviennent des externalités (cas de forçage) et des relations. Parce qu'ils sont créés ou observés à un instant *t*, ils sont également une temporalité (1D). Enfin, parce qu'ils existent, les objets sont aussi renseignés ou renseignables selon une sémantique qui participe d'une ontologie elle-même développée dans le cadre de la discipline scientifique de référence. Cette approche 4D disciplinaire n'est pas satisfaisante pour l'analyse de systèmes complexes tels que les SES car elle se rapporte à un modèle physique (au sens disciplinaire) de la réalité qui laisse peu de place à l'autodétermination et au métabolisme tels qu'on les rencontre dans les SES (Le Moigne, 1999 ; Holling *et al.*, 2002 ; Peterson, 2008). La proposition théorique ambitionne d'ajouter un niveau ontologique entre le référencement 4D et la description sémantique des éléments et relations des objets d'un SES.

3. Six concepts pour l'analyse des socio-écosystèmes

3.1. Des concepts pluridisciplinaires

L'extraction de connaissances disciplinaires multiples nécessite un effort d'abstraction dont la portée doit être a/trans-disciplinaire. L'objectif alors recherché consisterait à agréger/intégrer/hybrider les savoirs et les méthodes autour de l'objet scientifique commun. Les systèmes environnements-sociétés étant des objets scientifiques complexes et pluridisciplinaires, leur étude repose sur le

développement de modèles théoriques facilitant/autorisant des raisonnements et des résolutions de problèmes complexes.

Dans le foisonnement des possibles pour pallier cette contrainte théorique, nous proposons l'hypothèse d'un modèle d'abstraction combinant six dimensions (figure 2). La première, structurale, est indispensable à une approche OO. Elle implique une topologie discrète, consubstantielle à la notion de système et d'objet. Un système est par définition une organisation d'objets peuplant une structure. La deuxième, fonctionnelle, établit les actions, processus, rôles, etc. qui s'opèrent par et entre les objets. La troisième, relationnelle, caractérise la connexité des interactions, c'est-à-dire la topologie des transactions entre les objets. La quatrième, séquentielle, introduit une temporalité discrète, événementielle, une phase qui s'applique à l'activation des interactions entre les objets. La cinquième, scalaire, renvoie à une résolution, ou granularité, systémique. Elle définit les internalités et externalités qui délimitent, dans le sens des conditions aux limites d'un modèle, la résolution systémique des interactions. La sixième, adaptative, formalise les adaptations qui s'appliquent aux objets en réponse aux rétroactions du système. Elle est indispensable au bouclage des interactions entre les objets et elle permet la formalisation d'un état transitoire du système. Plasticité/rigidité, flou-vague/discret, déterministe/chaotique, prévisible/aléatoire, résilient/résistant, incertain/déterminé, stable/instable, etc. peuvent ainsi être dissociés des fonctions qu'elles sont déterminées.

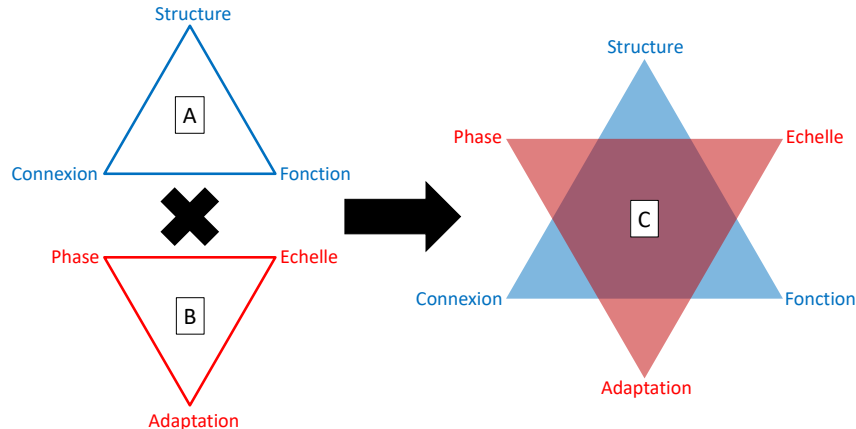


Figure 2. Ouvrir à plus de complexité dans la modélisation des SES avec (A) transposition du modèle Entité-Fonction-Relation à 3 classes et (B) intégration de 3 nouvelles classes pour un modèle de complexité (C) à 6 classes par objet

Dans la figure 2, 2A correspond à la transposition du modèle entité-fonction-relation, lui-même dérivé du modèle ontologique « entité-propriétés » (Armstrong, 1997 ; Johansson, 2004 ; Munn et Smith, 2008 ; Grenon, 2008). 2B correspond à

une transposition d'un modèle de contingence permettant d'inclure la temporalité, le réductionnisme (compris comme un facteur d'échelle) et la dimension évolutive d'un système complexe. Enfin, 2C correspond à un modèle d'objet systémique temporel qui, sans être « multistatique », intègre plus de complexité que le modèle 4D. Les six concepts retenus dans notre proposition sont donc structure, fonction, connexion, phase, échelle et adaptation (figure 2, tableaux 1 à 6). Ces six concepts sont des classes d'abstraction dont le choix sémantique obéit à un objectif de portabilité transdisciplinaire pour une application à la modélisation des SES dans un contexte scientifique pluridisciplinaire.

3.2. *Elicitation des six concepts proposés*

Notre proposition est élicitée à partir d'instances d'objets choisis (tableau 1 à 6) dans des SES relevant de quatre domaines 1) Géomorphologie, 2) Écologie, 3) Urbanisme et 4) Culture. Pour le premier domaine les objets étudiés sont principalement matériels et inertes (abiotique), pour le deuxième les objets sont matériels et mixtes (biotiques et abiotiques), pour le troisième les objets sont matériels - mixtes (biotiques et abiotiques) et immatériels (culturels), pour le quatrième domaine les objets sont mixtes matériels et immatériels (culturels).

3.2.1. *Concept de structure*

Une structure (tableau 1) est un assemblage, une organisation, une distribution, une hiérarchie ou une panarchie (Gunderson et Holling, 2002) entre les objets d'un système qui est considéré comme discret donc susceptible d'être segmenté ou réorganisé.

Tableau 1. Instances du concept de structure

Domaines	Structure
Géomorpho- logique	<p>Une forme de relief est une structure géomorphologique. C'est un assemblage d'objets inertes (abiotiques). Nous excluons volontairement la partie biologique qui contribue à la l'évolution d'une forme de relief.</p> <p>Ex. : une plage de dépôt : un assemblage d'éléments, des particules sédimentaires. C'est aussi un élément d'objet de rang supérieur tel qu'une plage est elle-même élément d'un objet cordon littoral, élément de l'objet littoral, élément de l'objet bordure continentale, élément de l'objet continent, lui-même élément de la planète Terre. Entre l'élément particule sédimentaire et l'objet terre, il existe un emboîtement de structures géomorphologiques.</p>
Écologique	<p>Une forêt est une structure biogéographique. C'est un assemblage d'objets inertes (abiotiques) et biologiques (faune et flore).</p> <p>Ex. : un sol forestier : un assemblage d'éléments minéraux (fraction minérale) et biologiques (fraction organique). C'est aussi un élément</p>

	(substrat) de l'objet sol-plante (incluant différentes classes d'objets plante herbacée, arbustive et arborée), élément de l'objet sol-plante-atmosphère incluant les éléments physiques (gazeux, liquides et solides), élément d'un objet peuplement arboré lui-même élément d'un objet forêt. Entre l'élément fraction organique du sol et l'objet forêt, il existe une structure forestière.
Urbain	<p>Une ville est une structure urbaine. C'est un assemblage d'objets matériels inertes (abiotiques, bâtis, etc.), biologiques (homme, faune et flore) et d'objets immatériels culturels (langue, histoire, arts, religions, etc.).</p> <p>Ex. : un sol urbain : un assemblage d'éléments « naturels » biotiques et abiotiques et artificiels (réseaux techniques). C'est un élément (fondation) pour des objets abiotiques (voiries, bâtiments), et biologiques (populations humaine, animale et végétale) eux-mêmes éléments d'un objet îlot urbain, élément d'un objet quartier urbain, lui-même élément d'un objet ville. Entre l'élément fraction organique du sol et l'objet ville, il existe une structure urbaine.</p>
Culturel	<p>Une mémoire est une structure culturelle. C'est un assemblage d'éléments immatériels (souvenirs) d'une temporalité passée. Nous excluons volontairement le support biologique (individuel) ou technologique (analogique ou numérique) de la mémoire.</p> <p>Ex. : un souvenir (lieu et temporalité donnés) : un assemblage d'éléments de perception humaine. C'est un élément de l'objet mémoire individuelle, elle-même élément de l'objet mémoire d'une population, élément d'un objet mémoire collective de l'Humanité, lui-même élément d'un objet mémoire collective de l'Histoire de l'Humanité. Entre l'élément souvenir et l'objet mémoire collective de l'Histoire de l'Humanité, il existe une structure de la mémoire.</p>

3.2.2. Concept de fonction

Une fonction (tableau 2) est un opérateur qui mobilise un ou plusieurs élément(s) en entrée, pour produire un ou plusieurs objet(s) du système en sortie qui est considéré comme multifonctionnel.

Tableau 2. Instances du concept de fonction

Domaine	Concept de fonction
Géomorpho- logique	<p>L'assemblage des éléments érosion, transport et accumulation est un objet fonctionnel de la morphogenèse. Exemples :</p> <ul style="list-style-type: none"> – L'érosion est une fonction géomorphologique qui consomme une énergie en entrée pour produire un objet altération de matière inerte en sortie (ex. un front d'altération d'un substrat). – Le transport est une fonction géomorphologique qui consomme une énergie en entrée pour produire un objet propagation de matière inerte en

	<p>sortie (exemple un flux solide dans un écoulement).</p> <ul style="list-style-type: none"> – L'accumulation est une fonction géomorphologique qui dissipe de l'énergie en entrée pour produire un objet stockage de matière inerte en sortie (ex. une plage de dépôt).
Écologique	<p>L'assemblage des éléments dégradation organique, anémogamie et cycle du carbone est un objet fonctionnel du métabolisme forestier. Exemples :</p> <ul style="list-style-type: none"> – La dégradation organique est une fonction forestière qui consomme une énergie biochimique en entrée pour produire un objet sol forestier en sortie (ex. l'humification). – L'anémogamie est une fonction forestière qui consomme une énergie éolienne en entrée pour produire un objet transport des pollens entre les organes mâles et femelles d'espèces végétales (ex. reproduction de l'épicéa). – Le cycle du carbone est une fonction forestière qui consomme une énergie solaire en entrée pour produire un objet séquestration du carbone dans la biomasse en sortie (ex. le bois des arbres).
Urbain	<p>L'assemblage des éléments construction, mobilité pendulaire et valorisation est un objet fonctionnel qui participe du métabolisme urbain. Exemples :</p> <ul style="list-style-type: none"> – Construire est une fonction urbaine qui assemble des éléments de ressource minérale en entrée pour produire un objet bâti en sortie (ex. l'artificialisation des sols urbains). – La mobilité est une fonction urbaine qui consomme une énergie pour produire un objet transport des usagers entre différents lieux de la ville en sortie (ex. la mobilité pendulaire). – La valorisation d'un bien immobilier est une fonction urbaine qui utilise des éléments de consentement à payer en entrée pour produire un objet valeur en sortie (ex. le prix d'un appartement).
Culturel	<p>L'assemblage des éléments lire, apprendre et enseigner est un objet fonctionnel qui participe à la production de la mémoire culturelle. Exemples :</p> <ul style="list-style-type: none"> – Lire est une fonction culturelle qui assemble des éléments symboliques en entrée pour produire la restitution orale d'un objet sémantique en sortie (ex. lire une carte). – Apprendre est une fonction culturelle qui assemble des éléments de connaissance en entrée pour produire un objet savoir en sortie (ex. réviser un cours de géographie). – Enseigner est une fonction culturelle qui mobilise des éléments de savoirs en entrée pour produire un objet de diffusion des connaissances en sortie (ex. un cours de cartographie).

3.2.3. Concept de connexion

Une connexion (tableau 3) est une mise en relation, un connecteur qui relie un ou plusieurs éléments à un ou plusieurs objets du système. Un objet peut être connecté (déconnectable ou non), connectable (déconnecté) ou isolé (non connectable et non connecté), il dispose ainsi d'un domaine de connexions qui peut être mis en relation avec son voisinage systémique.

Tableau 3. Instances du concept de connexion

Domaine	Concept de connexion
Géomorphologique	<p>Le réseau hydrographique est un assemblage qui connecte des ruissellements de surface élémentaires à un ruissellement d'ordre supérieur. Il peut être connecté, connectable (déconnecté) ou isolé (non connectable). Exemples :</p> <ul style="list-style-type: none"> – Un réseau exoréique connecte des ruissellements de surface à une mer ou un océan. – Un réseau endoréique connecte des ruissellements de surface à une cuvette topographique. – Un réseau hydrographique aréique déconnecte des ruissellements de surface par percolation/infiltration vers un écoulement et un stockage souterrains.
Écologique	<p>Un réseau écologique est un assemblage qui connecte différents substrats biophysiques qui peuvent être connectés, connectables (déconnectés) ou isolés (non connectables) à des organismes vivants. Exemples :</p> <ul style="list-style-type: none"> – Les corridors écologiques (trames vertes, bleues, brunes, etc.) correspondent à un réseau d'habitats pour une espèce ou un groupe d'espèces et constituent un support de la biodiversité et des interactions entre les écosystèmes. – Les réseaux trophiques sont des chaînes alimentaires reliant différentes ressources (composés organiques et inorganiques) à des producteurs primaires (autotrophes) à des espèces (hétérotrophes) d'herbivores ou de prédateurs. – Une forêt est un écosystème qui connecte différents corridors écologiques avec différentes chaînes trophiques.
Urbain	<p>Un réseau urbain est un assemblage qui connecte différents échanges ou flux urbains qui peuvent être connectés, connectables (déconnectés) ou isolés (non connectables) par des infrastructures urbaines. Exemples :</p> <ul style="list-style-type: none"> – Les réseaux de transport sont des supports de la mobilité entre des usagers et différents lieux et usages de la ville. – Les réseaux techniques sont des infrastructures permettant des échanges de flux d'énergie, de déchets et d'informations entre des producteurs et des consommateurs de la ville. – Les réseaux culturels sont des échanges et des mises en relations qui connectent différents acteurs culturels avec des usagers de la ville.
Domaine culturel	<p>Un réseau social est un assemblage qui connecte plusieurs individus. Exemple :</p> <ul style="list-style-type: none"> – Un réseau d'acteurs institutionnels est une mise en connexion de différents acteurs politiques ou administratifs qui ont un rôle dans un espace social de l'échelle locale à l'échelle globale. – Un réseau d'acteurs culturels est une mise en connexion de producteurs, médiateurs et consommateurs d'une offre culturelle dans un espace culturel de l'échelle locale à l'échelle globale. – Un réseau social numérique est une mise en connexion numérique d'individus ou de groupes d'individus dans un espace numérique privé ou public individuel (familiale, professionnelle, etc.) ou communautaire (communauté d'intérêt) du local au global.

3.2.4. Concept de phase

Une phase (tableau 4) est une étendue temporelle qui décrit une initialisation, une durée et une extinction pour un objet événement du système qui est considéré comme une phase d'états/un état de phases des temporalités pour des événements spatiotemporels (actifs, inactifs, en veille, ouverts, délimités). Une phase correspond donc à une enveloppe de temps durant laquelle le comportement de l'objet reste inchangé en structure, fonction, connexion, échelle et adaptation. Une phase n'est pas synonyme de stabilité du système, car l'instabilité peut relever d'un fonctionnement transitoire entre deux phases de stabilité.

Tableau 4. Instances du concept de phase

Domaine	Concept de phase
Géomorphologique	<p>Un événement géomorphologique est une enveloppe temporelle dans la chronologie (temps linéaire) d'une forme de relief. Exemples :</p> <ul style="list-style-type: none"> – Une plaine alluviale est une accumulation temporelle de phases de submersions/alluvionnements dans la chronologie des flux hydrosédimentaires d'un fond de vallée. – Un cycle morphologique est une succession d'états de phases, altération (érosion), exportation (transport) et accumulation (accrétion), qui participe à la morphogénèse (création/évolution) d'une forme de relief. – Une érosion est une phase temporelle durant laquelle un agent érosif exerce une force supérieure à la capacité de résistance d'une forme de relief.
Écologique	<p>Un événement forestier est une enveloppe temporelle dans la chronologie (temps linéaire) d'une forêt. Exemples :</p> <ul style="list-style-type: none"> – Dans la forêt méditerranéenne les incendies se produisent principalement pendant la phase temporelle saison sèche. – Dans la forêt des latitudes tempérées, le printemps est une phase d'état à forte croissance végétale. Sur le temps long plusieurs phases d'états se succèdent pour un arbre (pousse, maturité, senescence). – Une phase écologique forestière est un intervalle de temps observable correspondant à un écosystème stable (écotope).
Urbain	<p>Un événement urbain est une enveloppe temporelle dans la chronologie (temps linéaire) d'une ville. Exemples :</p> <ul style="list-style-type: none"> – Une inondation urbaine est une phase temporelle durant laquelle l'eau submerge des espaces urbains dont la vocation d'usage est incompatible avec la présence de l'eau. – Une panne sectorielle d'électricité est une phase temporelle durant laquelle plusieurs espaces urbains sont déconnectés du réseau d'approvisionnement électrique. – Une élection municipale est une phase temporelle durant laquelle l'état institutionnel de la ville est modifiable.

Culturel	<p>Un événement culturel est une enveloppe temporelle dans la chronologie (temps linéaire) culturelle. Exemples :</p> <ul style="list-style-type: none"> – Une création artistique est une phase temporelle qui associe un avant (maturation) un pendant (réalisation) et une révélation (exposition) pour une œuvre d’art. – La civilisation égyptienne est une phase temporelle durant laquelle la culture égyptienne existe et se diffuse en Afrique, en Méditerranée et en Asie mineure. – Un orchestre philharmonique est une phase temporelle durant laquelle un assemblage de musiciens classiques est regroupé en un lieu de concert pour réaliser une œuvre musicale.
----------	---

3.2.5. Concept d’échelle

Une échelle (tableau 5) est une enveloppe réductionniste qui définit un domaine de validité qui est borné par des limites exclusives internes (internalités) et externes (externalités) pour un objet du système. Les externalités (adjacences) et les internalités (sous-jacentes) sont des forçages potentiels, au-delà des conditions limites d’un modèle, que le réductionnisme a écarté pour réduire la complexité du système. Il s’agit ici de la valeur relative d’un objet par rapport à son contexte systémique, un ratio de complexité en quelque sorte, afin modéliser les connaissances en tenant plus ou moins compte des internalités et des externalités de l’objet dans le système étudié.

Tableau 5. Instances du concept d’échelle

Domaine	Concept d’échelle
Géomorphologique	<p>Un domaine géomorphologique est une enveloppe réductionniste qui borne un domaine de validité par des internalités et des externalités pour une forme de relief donnée. Exemples :</p> <ul style="list-style-type: none"> – Le littoral est un domaine géomorphologique qui exclut la plus grande part des espaces océaniques et continentaux. – Un effondrement karstique est le résultat d’une dynamique (internalité) qui modifie le fonctionnement interne d’un relief karstique. – Un cratère d’impact est le résultat d’une dynamique (externalité) qui modifie la morphologie de la surface de la croûte terrestre.
Écologique	<p>Un domaine forestier est une enveloppe réductionniste qui borne un domaine de validité par des internalités et des externalités pour une forêt donnée. Exemples :</p> <ul style="list-style-type: none"> – La lisière est un domaine de la forêt qui exclut la plus grande part des espaces forestiers et non forestiers. – Une clairière est le résultat d’une dynamique (internalité) où la strate arborée n’est plus le fonctionnement dominant dans un massif forestier. – Une zone agricole est un domaine non forestier (externalité) qui contribue à la mobilité et au renouvellement des espèces forestières.

Urbain	<p>Un domaine urbain est une enveloppe réductionniste qui borne un domaine de validité par des internalités et des externalités pour un espace urbain donné. Exemples :</p> <ul style="list-style-type: none"> – La frange urbaine est un domaine de la ville qui exclut la plus grande part des espaces urbains et ruraux. – Les corridors écologiques urbains sont des trames (internalité) qui contribuent au maintien d'une biodiversité urbaine. – Une zone de champs captant est un domaine non urbain (externalité) qui participe à la recharge des nappes utilisées pour l'alimentation en eau potable de la ville.
Culturel	<p>Un domaine culturel est une enveloppe réductionniste qui borne un domaine de validité par des internalités et des externalités pour un espace culturel donné. Exemples :</p> <ul style="list-style-type: none"> – La pratique du violon est un domaine musical qui exclut la plus grande part de l'art de la lutherie et de la pratique des autres instruments d'orchestre. – Le tuba est un instrument optionnel (internalité) qui contribue à la sonorité de la section cuivre d'un orchestre symphonique. – La maison de la radio est une institution (externalité) qui met à disposition de l'Orchestre philharmonique de France (OPF) un environnement culturel (infrastructure, organisation, diffusion, etc.) de l'OPF.

3.2.6. Concept d'adaptation

L'adaptation (tableau 6) est nécessaire (voir paragraphe 4.3) sinon tout objet serait donné *a priori* et présenterait une intégrité totale et intemporelle, donc infinie et insensible à toute forme d'évolution y compris la fin de son existence. Cela supposerait une reproduction à l'identique et à l'infini de cycles sans pertes ni gains car ceux-ci constitueraient en soit une adaptation du système existant. Notre proposition considère donc que les SES sont évolutifs et qu'ils s'adaptent/se réadaptent aux conditions/forçages issus de leur contexte systémique.

Tableau 6. Instances du concept d'adaptation

Domaine	Concept d'adaptation
Géomorpho- logique	<p>L'évolution d'un relief est une adaptation de son état initial en réponse à un assemblage de contraintes issues du métabolisme géomorphologique.</p> <p>Exemple :</p> <ul style="list-style-type: none"> – Le recul d'un cordon dunaire est une adaptation du trait de côte en réponse à un assemblage de contraintes d'érosion, de transport, d'accrétion, etc., exercées sur le matériel sédimentaire côtier. – La création d'une chaîne de montagne est une adaptation du relief en réponse à un assemblage de contraintes tectoniques, de géodynamique interne, de lithologie, etc., exercées sur la croûte terrestre. – La construction d'un barrage d'avalanche est une adaptation géomorphologique en réponse à un assemblage de contraintes techniques (ouvrage), anthropiques (populations riveraines), etc., exercées sur le couloir d'avalanche.

Écologique	<p>L'évolution d'une forêt est une adaptation de son état initial en réponse à un assemblage de contraintes issues du métabolisme forestier. Exemples :</p> <ul style="list-style-type: none"> – L'évolution d'un sol forestier est une adaptation d'un complexe édaphique en réponse à un assemblage de contraintes biologiques, chimiques, minérales, climatiques, etc., exercées sur l'interface biosphère-lithosphère. – L'évolution de la biodiversité forestière est une adaptation de la présence d'espèces en réponse à un assemblage de contraintes climatologiques, écologiques, anthropiques, etc., exercées sur l'écosystème forêt. – L'évolution des services écosystémiques forestiers est une adaptation des fonctions et supports écologiques en réponse à un assemblage de contraintes écologiques, économiques, culturelles, etc., exercées sur l'écosystème forêt.
Urbain	<p>L'évolution d'une ville est une adaptation de son état initial en réponse à un assemblage de contraintes issues du métabolisme urbain. Exemples :</p> <ul style="list-style-type: none"> – L'étalement urbain est une adaptation de l'espace périurbain en réponse à un besoin de croissance de la population et des activités urbaines. – La requalification urbaine est une adaptation de la vocation de certains espaces urbains en réponse à une réorganisation et à une optimisation des activités et des usages de la ville. – L'attractivité urbaine est une adaptation des rapports ville-périphérie en réponse à la concentration des pouvoirs et des valeurs culturelles et économiques.
Culturel	<p>L'évolution d'une œuvre d'art est une adaptation de sa création initiale en réponse à un assemblage de ses multiples réceptions issues du contexte culturel. Exemples :</p> <ul style="list-style-type: none"> – La création d'une œuvre d'art est un phénomène créatif, émergeant en réponse à un assemblage de sensibilités, d'expressions, de techniques, etc., exercées sur un support matériel ou immatériel. – L'acceptation d'une œuvre d'art est un phénomène de réception en réponse à un assemblage de sensibilités, de valeurs, de cultures, etc., exercées sur une proposition artistique. – La relégation d'une œuvre est un phénomène de rejet ou d'oubli culturels en réponse à l'évolution d'un assemblage de sensibilités, de valeurs, de contextes culturels, etc., exercés sur une proposition artistique.

4. Proposition d'ontologie orientée objet à six dimensions pour les socio-écosystèmes

Il est possible d'extraire (segmenter) des objets dans des jeux de données spatiotemporels tout en les renseignant avec des valeurs attributaires et des relations de hiérarchie (*supra* 2.3). Il est aussi possible de définir des concepts qui transcendent les disciplines (*supra* 3.1 et 3.2). Nous proposons maintenant de combiner le paradigme OO avec les six concepts proposés.

4.1. Problèmes de modélisation

Plusieurs problèmes sont posés par la modélisation quand il s'agit d'analyser un SES spatiotemporel (Armstrong, 1988 ; Arenas *et al.*, 2018 ; Bittner *et al.*, 2009). Ils résultent principalement du réductionnisme nécessaire pour transposer la complexité du monde réel dans un modèle numérique spatiotemporel. Il faut également tenir compte du fonctionnement déterministe d'un SI qui applique rigoureusement le modèle et seulement les règles du modèle aussi intelligent soit-il (Gruber, 1995). Dans l'étape réductionniste de la création d'un modèle, le choix d'une ontologie préexistante, la création d'une ontologie *ad hoc* et la délimitation (internalités et externalités) du SES étudié renvoient aux disciplines des parties prenantes aux dépens de la transversalité. Or, pour aborder les SES comme des objets complexes, il est nécessaire de dépasser ces limites disciplinaires (Hervé *et al.*, 2016). Dès qu'une partie d'un système dispose de fonctions de cognition-mémorisation (apprentissage) ou d'une capacité d'autodétermination (libre arbitre), alors le comportement prédictif déterministe s'efface au profit d'un comportement auto-adaptatif (Le Moine, 1999). S'attaquer à la modélisation de la complexité des SES dans un environnement numérique 3D temporel revient à considérer que les SES sont des systèmes modélisables avec les outils de la physique. L'argumentation qui consisterait à dire que les SES sont observables dans le « monde réel » qui a une « existence physique réelle » est un raccourci scientifique qui n'est pas acceptable. Il est du même ordre que celui qui postulerait qu'il serait possible de réaliser l'analyse scientifique d'une peinture de maître en se basant sur son cadre en bois comme conditions aux limites et sur la seule position (4D) relative des éléments de matières utilisés par le peintre comme conditions initiales. Dans le meilleur des cas, ce modèle de connaissances et d'analyse basé sur un tel réductionnisme physique ne pourrait que décrire un état de phase physique de la peinture, en imaginant que l'on puisse déconstruire la succession des gestes du peintre à partir de l'assemblage des traces laissées par la technique employée. Mais, en aucun cas, il ne pourrait expliquer ce que le peintre a souhaité représenter ni dans quel contexte culturel le résultat final obtenu aurait atteint les objectifs recherchés par l'artiste. Il ne pourrait pas plus expliquer comment l'œuvre est reçue, cognitivement, par un individu inconnu et portant un regard sur le tableau. Dans cet exemple, qui associe des éléments matériels physiques et des éléments immatériels culturels dans un objet peinture, il est évident qu'un modèle de connaissance uniquement basé sur un modèle physique 4D est inadapté. De même sans ces connaissances physiques, il serait difficile de décrire et de reconstruire à posteriori l'usage des pigments sur une surface de fibre qui a offert à l'artiste un vecteur d'expression sensible et culturelle.

Ce qui est recherché dans notre proposition, ce n'est pas de résoudre le problème de la modélisation de la complexité des SES mais de doter le modélisateur d'une TLO (Guarino, 1997 ; Arp *et al.*, 2015) permettant de construire un modèle systémique sur base de connaissances pluridisciplinaires. A partir des six concepts proposés (structure, fonction, connexion, phase, échelle et adaptation) nous proposons donc, en théorie, de modéliser les SES en dehors de l'espace 4D en les

projetant dans un espace de phases dont les dimensions numériques ne sont plus physiques mais OO et systémiques. Ce basculement de référentiel nous semble ouvrir des perspectives de modélisation dans lesquelles notre proposition de concepts peut être utilisée pour dimensionner et renseigner les objets des SES.

4.2. Une ontologie pour les socio-écosystèmes

Dans notre proposition OOO 6.0, l'objectif recherché est le positionnement des connaissances dans une ontologie qui décrit un modèle d'interactions complexes au sein d'un SES. Passer d'un modèle physique du monde réel (4D) à un espace de phases dimensionné sur la base théorique des six concepts proposés revient à créer un modèle de connaissances qui n'est plus basé sur la position des objets dans le domaine spatiotemporel (le monde réel) mais sur la position relative des objets de connaissance dans l'espace des phases (tableau 7) d'un système de connaissances étudié. Là où un modèle 4D découpe et dimensionne les objets en volumes (surfaces et distance) temporels par leurs positions spatiotemporelles, notre proposition lui substitue des objets dimensionnés par leur position relative dans un espace renseigné par les six concepts proposés (tableau 7). Or, pour dimensionner, il faut pouvoir mesurer en utilisant une métrique adaptée car une dimension est avant tout une étendue mesurable (les distances en mètres, le temps en secondes, etc.). Qu'en est-il dans notre modèle ?

Tableau 7. Matrice conceptuelle intégrant les 6 concepts proposés

Dimen- sions	Structure	Fonction	Connexion	Phase	Échelle	Adaptation
Structure	Meta structure	Fonct. structurale	Connex. structurale	Phase structurale	Échel. structurale	Adapt. structurale
Fonction	Struct. fonctionnelle	Meta fonction	Connex. fonctionnelle	Phase fonctionnelle	Échel. fonctionnelle	Adapt. fonctionnelle
Connexion	Struct. connective	Fonct. connective	Meta connexion	Phase connective	Échel. connective	Adapt. connective
Phase	Struct. phasée	Fonct. phasée	Connex. phasée	Meta phase	Échel. phasée	Adapt. phasée
Échelle	Struct. scalaire	Fonct. scalaire	Connex. scalaire	Phase scalaire	Meta échelle	Adapt. scalaire
Adaptation	Struct. adaptative	Fonct. adaptative	Connex. adaptative	Phase adaptative	Échel. adaptative	Meta adaptation

Pour comprendre ce passage de la 4D aux six dimensions de l'espace de phases nous prenons l'exemple du croisement de la structure des deux objets A et B dont la méta structure (structure de structure) correspond à l'une des six dimensions de notre proposition OOO. Soit une structure A, administrative, « département » qui est un assemblage d'objets « cantons », eux même assemblages d'éléments/objets

« communes ». Un département est une structure, de structures cantonales, de structures communales... Il est possible de dimensionner cet objet administratif en fonction de la quantification (mesure) du nombre d'éléments qui s'agrègent dans la méta structure administrative d'un département. Soit une autre structure B, arborée, « boisement » qui est un assemblage d'objets « arborés » (forêts, bois, bosquets, etc.), eux même assemblages d'éléments/objets « arbres ». Il est également possible de dimensionner cet objet boisement en fonction de la quantification (mesure) du nombre d'éléments qui s'agrègent dans leurs structures (ici arborée). Un boisement est une structure, de structures arborées, de structures arbres.

Si nous croisons la structure de l'objet administratif « département » avec l'objet arboré « boisement ». Nous pouvons extraire une nouvelle structure d'objets « département.boisement » qui segmente par intersection OO les deux structures. Nous obtenons ainsi une nouvelle structure, de structures « cantons.boisement », de structures « comunes.boisement ». Nous appelons « métastructure » (tableau 7) cette structure de structures et nous postulons que celle-ci est une dimension de la structure du SES département.boisement. Par transposition aux 5 autres concepts, nous postulons également que les méta fonction, connexion, phase, échelle et adaptation sont elles aussi des dimensions systémiques qui sont donc applicables, à partir des connaissances disponibles, au SES « département.boisement » (pour l'exemple simple utilisé ici). Nous postulons également qu'un modèle relationnel basé sur le dimensionnement des objets d'un SES à partir de ces six dimensions « méta » relève d'un domaine théorique de connaissances s'appuyant sur une OOO 6.0. Enfin, nous postulons que ces six dimensions ne sont pas disciplinaires et qu'elles peuvent donc servir comme outils de dialogue pour la modélisation plurivoire transdisciplinaire.

Les cinq premières dimensions ont une proximité sémantique indéniable avec les concepts couramment utilisés dans la modélisation spatiotemporelle des SES : structure pour entité, fonction pour fonction, connexion pour relation, phase pour temporalité, échelle pour réduction (de la complexité étudiée). Il y a cependant un décalage sémantique volontaire :

- 1) pour limiter les confusions avec le modèle entité-attribut-relation-temps qui est couramment utilisé en géomatique et en modélisation de base de données ;
- 2) parce que notre proposition OOO 6.0 accepte des attributs (propriétés) sémantiques, sur les six dimensions des objets ;
- 3) parce que les objets 6.0 de l'espace de phases incluent des attributs (propriétés) de positionnement spatiotemporel pour leur visualisation-représentation-mise à jour dans l'espace physique 4D.

En revanche, lors des échanges scientifiques autour de cette proposition OOO 6.0 dans les séminaires de l'AP Ontologies et Dynamiques spatiales du GDR MAGIS entre 2017 et 2020, la question de la sixième dimension « adaptation » a fait débat justifiant un effort de clarification.

4.3. Discussion sur l'adaptation

Pourquoi une sixième dimension plutôt que d'en rester à cinq ? Pourquoi ne pas ajouter plus de dimensions ? Pour répondre à ces deux questions il est d'abord nécessaire de différencier la dimension adaptation de la dimension fonction dans notre proposition.

La dimension fonction modélise l'assemblage des processus, opérations, actions, rôles... d'un système tels qu'ils sont connus et décrits par l'utilisateur. Dans une approche purement déterministe, une ou plusieurs entrées connues d'une fonction produisent une ou plusieurs sorties déterminées, anticipées, prévisibles... Une fonction est donc strictement délimitée dans ses interactions avec le système. Dans un système purement physique une fonction obéit donc à des lois déterminées et connues par le modélisateur. Or, dans le monde réel, ce déterminisme « fonctionnel » ne permet pas de prédire, de modéliser précisément, le comportement d'un SES car les capacités d'autodétermination, d'apprentissage, d'évolution des éléments du système sont précisément des capacités d'adaptation relatives à leur niveau de cognition. Certains éléments d'un SES sont très adaptatifs et autodéterminés quand d'autres peuvent au contraire répondre à des lois de comportement parfaitement prédictives. Certains éléments des SES sont sensibles aux forçages, d'autres moins... Il y a donc, de notre point de vue, nécessité d'une autre dimension, une sixième, qui modélise un intervalle/une variété de comportements possibles, non linéaires, pour une fonction, mais aussi pour les quatre autres dimensions de notre proposition. Les dimensions structure, fonction, connexion, phase et échelle sont toutes adaptatives (figure 3 et tableau 7), relativement aux interactions entre et à l'intérieur des objets d'un SES qui, par rétroactions peuvent être indéterminées voire indéterminables.

Ainsi dans la figure 3, nous considérons les relations entre les six dimensions comme des interactions de type 6.0 et nous considérons que l'adaptation exerce également une rétroaction (positive, négative ou neutre) sur l'ensemble des dimensions y compris elle-même (voir tableau 7). L'adaptation d'une structure (un assemblage parmi tous les assemblages) d'un SES aura donc, par rétroaction un effet de redimensionnement sur une ou plusieurs des six dimensions du modèle. Les six dimensions étant interactives, chaque objet est lui-même un assemblage structurel-fonctionnel-relationnel-phasé-scalaire-adaptatif en interaction avec les autres objets qui l'environnent (externalités comme internalités). La complexité croît très rapidement dans ce modèle d'abstraction à six dimensions. En contrepartie, elle propose une « portabilité transdisciplinaire » pour la modélisation des connaissances utiles dans l'analyse d'un SES. Elle ne répond cependant pas à la question : est-il nécessaire d'utiliser un modèle complexe pour résoudre l'analyse d'un système complexe ?

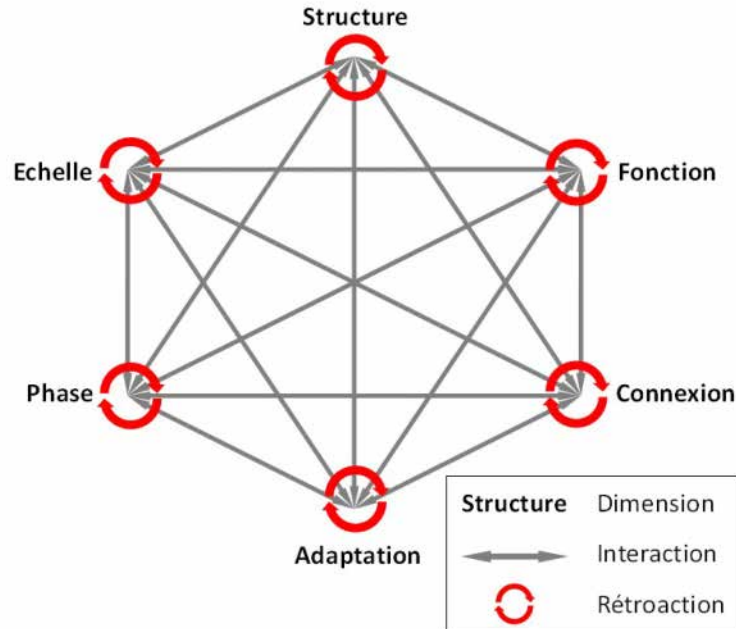


Figure 3. Schéma relationnel en 6D

Sans une dimension adaptation, un SES serait nécessairement déterministe. L'ensemble des objets répondraient aux lois de comportement définies par l'utilisateur sans laisser de place à des phénomènes tels que l'émergence, ou l'autodétermination. Si certains systèmes physiques peuvent être considérés comme strictement déterministes, les SES sont quant à eux extrêmement auto-adaptatifs donc incertains. Ces derniers procèdent à leurs propres phénomènes d'émergence, d'autorégulation, d'évolution, d'intentionnalité et d'intelligence. Pour le dire plus simplement ils s'adaptent selon des modalités systémiques plus ou moins complexes auxquelles ils sont en interrelations (internalités et externalités). Avec cette sixième dimension notre proposition OOO 6.0 ressemble à un modèle systémique chimérique. Il n'est pas possible de représenter physiquement, graphiquement ou numériquement ses six dimensions. Raisonner avec un tel modèle de connaissances n'est pas aisé. Il est donc nécessaire d'accompagner cette proposition ontologique par une analyse des relations qui existent entre ces six dimensions afin de comprendre comment l'utiliser pour agréger les connaissances utiles à la modélisation d'un SES.

4.4. Une topologie à six dimensions

Dans une ontologie basée sur un modèle de réalité 4D, les objets sont des volumes temporels que l'on peut représenter par une animation temporelle associée à une géométrie 3D. Par exemple un volume d'eau envahi, puis libère une zone inondable en fonction de la temporalité d'une crue. Si le système étudié n'est pas seulement l'écoulement de crue mais le risque inondation, alors la question des vulnérabilités qui y est associée implique une plus grande complexité dans la modélisation du système étudié. Cette complexité est fortement dépendante des décisions individuelles prises par les populations concernées qui n'obéissent pas à des lois comportementales aussi déterministes que celles de l'hydraulique fluviale. En fonction de critères multiples tels que la mémoire du risque, l'anticipation (préparation à la crue), la mitigation du risque, la mobilité journalière (travail) ou saisonnière (vacances)... les habitants des zones inondables sont non seulement des « agents autodéterminés » qui développent des stratégies de réponses complexes mais aussi des éléments du système risque. Une modélisation réaliste des différentes actions réelles est très complexe à modéliser au niveau de l'individu ou d'une communauté locale. Notre proposition ne s'appuyant pas sur une modélisation 3D temporelle, elle permet d'intégrer plus de complexité que l'interaction spatiotemporelle d'entités physiques ou abstraites décrites par leurs propriétés dans un modèle de connaissances sous SIG.

Nous postulons que les objets d'un SES utilisant notre modèle de connaissances peuvent être extraits par segmentation OO à partir d'une base de connaissance construite sur le modèle d'OOO 6.0 proposé. Nous postulons également qu'une topologie OO à 6 dimensions peut aussi être implémentée pour réaliser une analyse des connaissances d'un SES qui intègre plus de complexité que la relation spatiotemporelle 3D (surface-temps) ou 4D (volume-temps). Mener à bien une telle analyse impose de réinterpréter en dimension 6.0 les modèles de relations topologiques DE-9IM (Clementini et Di Felice, 1996 ; Proteau, 2011 ; Strobl, 2008), RCC-8 (Randell *et al.*, 1992), temporels (Allen, 1983) et booléens (Egenhofer, 1989 ; Egenhofer *et al.*, 1989). Les treize relations de bases de James Allen (Allen, 1983) pour la modélisation des relations entre deux segments temporels (figure 4) en est un autre très bon exemple. En effet, ce modèle relationnel par J. Allen pour la dimension temporelle, donc en 1D, peut directement s'appliquer à chacune des six dimensions de notre modèle de connaissances OO. Là où Allen figure des durées par un segment de droite de dimension variable, nous pouvons figurer une étendue entre la limite interne et externe des objets sur chacune de leurs six dimensions. Chaque étendue correspond donc à un segment sur chacune des six dimensions du volume 6.0 de la même façon que les segments de longueur, largeur et profondeur sont les trois dimensions d'un objet physiques dans l'espace 3D. En d'autres termes, dans notre proposition OOO 6.0, un objet est un hypervolume de dimension 6 qui est l'assemblage de ses six étendues dimensionnelles.

Relation	Interprétation	Réciproque	Visualisation de la relation de temporalité
$X < Y$	X est avant Y	Y est après X ($Y > X$)	
$X m Y$	X rejoint Y	Y rejoint X ($Y mi X$)	
$X o Y$	X recouvre partiellement Y	Y recouvre partiellement X ($Y oi X$)	
$X s Y$	X débute avec Y	Y débute avec X ($Y si X$)	
$X d Y$	X pendant Y	Y enveloppe X ($Y di X$)	
$X f Y$	X fini avec Y	Y fini avec X ($Y fi X$)	
$X = Y$	X égale Y	Y égale X ($Y = X$)	

Figure 4. Modèle de relations entre deux segments temporels X et Y
(modifié de Allen, 1983)

Comme pour le modèle topologique 1D de Allen, il est possible de décliner des modèles de dimensions supérieures 2D, 2D temporels et 3D temporels... Une topologie 6.0, orientée objet n'est donc pas une chimère mais bien un modèle relationnel, méréotopologique (Smith, 1996), qui peut s'appuyer sur des bases théoriques disponibles et qui permet d'intégrer plus de complexité tout en dépassant le carcan de la modélisation 3D temporelle qui s'accommode très mal des interactions complexes et anthropocentrées présentes dans les SES.

5. Perspectives ontologiques

Cette proposition OOO 6.0 est singulière à plus d'un titre par la position épistémique qu'elle propose d'occuper. Il s'agit en effet de changer de paradigme pour la modélisation et l'analyse des SES en proposant une alternative aux ontologies contemporaines qui sont principalement basées sur des entités temporelles, physiques (3D) ou intentionnelles (formes), associées à leurs propriétés, relations et attributs. Cette alternative, une OOO 6.0, ambitionne de répondre aux prédicats suivants :

- l'interopérabilité conceptuelle entre plusieurs domaines de connaissances issues des disciplines de l'environnement et de la société sur la base des six concepts non disciplinaires proposés ;
- transcender la dichotomie ontologique entre entité (l'être) et propriétés (l'étant) qui sont ici considérées comme des assemblages, des construits, des représentations,

etc. à partir d'éléments de connaissances réorganisés dans une ontologie orientée objet s'appuyant sur les six concepts proposés ;

– la modélisation orientée objet dans un espace de phases de dimension 6 où la construction/déconstruction ontologique des SES s'appuie sur une logique booléenne (figure 5) appliquée à la théorie des ensembles qui sont ici des assemblages d'éléments de connaissance ;

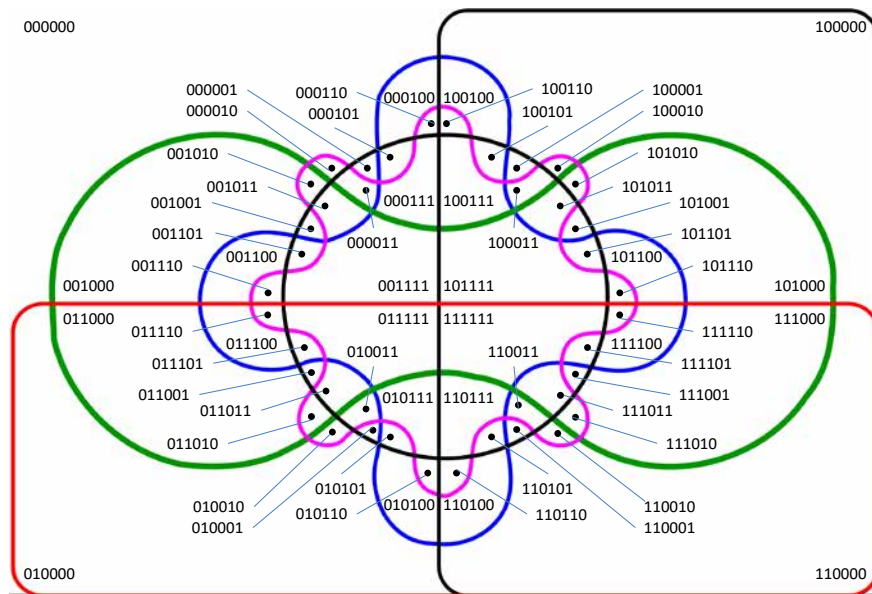


Figure 5. Diagramme de Venn à 6 classes (modifié de Edwards, 2004 ; figure sous licence CC BY-SA 3.0). Un objet de dimension 6.0 peut présenter 2^6 éléments logiques (table de vérité) pour instancier un assemblage de connaissances

– de compatibilité avec la logique des ensembles flous par le jeu d'interaction et d'assemblage des éléments connaissances par intégration de leurs dimensions de phase, d'échelle et d'adaptabilité ;

– l'analyse systémique associant des parties, des objets éléments de connaissance, pour construire un tout, l'objet système étudié, en réinterrogeant la dynamique spatiotemporelle des SES avec une matrice de six concepts de premier ordre (structure fonction, connexion, phase, échelle, adaptation), de trente concepts de second ordre (tableau 7) et de six « méta concepts » (tableau 7), ces derniers étant considérés comme les six dimensions de notre proposition d'OOO ;

– d'alternative théorique à l'organisation des connaissances basée sur un modèle disciplinaire descendant et arborescent, tel que celui proposé par Melvil Dewey (Dewey, 1876) qui s'appuie exclusivement sur une segmentation par discipline-

sujets/généralités-singularité, ou celui proposé par le modèle RDF, bien plus récent car lié au web sémantique et à l'indexation des connaissances sur la base d'un multigraphe-orienté-annoté, s'appuyant sur le triplet sujet-prédictat-objet,

– d'extraction des lacunes de connaissances d'un SES étudié à partir de l'empreinte inverse du volume de connaissances explicitées, ces dernières correspondant aux conditions limites de l'ontologie déployée.

L'OOO 6.0 propose donc des concepts généraux, applicables dans toutes les disciplines, permettant de modéliser les entités et leurs propriétés comme des objets de second ordre. Notre proposition ontologique est encore à l'état conceptuel tout en se positionnant dans le registre des ontologies fondationnelles, de haut niveau, selon les critères définis par Guarino (1997) ou Arp *et al.* (2015). Plusieurs défis apparaissent encore comme autant d'étapes à franchir pour envisager son opérationnalité :

- passer d'une ontologie conceptuelle à une ontologie formelle ;
- compléter son axiomatisation en s'appuyant sur les travaux déjà réalisés dans le cadre d'autres ontologies formelles de haut niveau ;
- assembler une base de données susceptible de tester/valider son opérationnalité pour un modèle de SES suffisamment complexe et documenté pour alimenter une OOO 6.0.

Remerciements

L'auteur tient à remercier les nombreux collègues qui ont contribué par leur écoute, leurs critiques constructives et leurs encouragements à la réalisation de ce travail théorique : Pr P. Meire (University of Antwerp), Dr D. Amatya (USDA-FS, Santee Experimental Forest), Pr L. Lavkulich (University of British Columbia), Pr C. De Jong (université de Strasbourg), Pr A. Vadineanu (universitatea din Bucuresti), Pr M. Keiler (universität Bern), Pr Y. Guermond (université de Rouen), Pr O. Blanpain et Pr A. Leprêtre (université de Lille). Cet article doit également beaucoup aux membres de l'Action Prospective « Ontologies et dynamiques spatiales » du GDR 2340 MAGIS du CNRS dont les nombreux échanges scientifiques ont été déterminants. Enfin, nous remercions sincèrement les relecteurs pour leurs suggestions constructives qui ont permis la finalisation de cet article.

Bibliographie

- Allen C.R., Holling C.S. (Eds) (2008). *Discontinuities in Ecosystems and Other Complex Systems*, Columbia University Press.
- Allen J.F. (1983). Maintaining knowledge about temporal intervals. *Communication to the ACM*, vol. 26, n° 11, p. 832-843.
- Alter N. (2003). Movement and dyschrony in organizations. *L'année sociologique*, vol. 53, n° 2, p. 489-514.

- Arenas H., Trojahn C., Comparot C., Aussenac-Gilles C. (2018). Un modèle pour l'intégration spatiale et temporelle de données géolocalisées. *Revue internationale de géomatique*, vol. 28, n° 2/2018, p. 243-266.
- Armstrong D.M. (1997). *A World of States of Affairs*, Cambridge University Press.
- Arp A., Smith B., Spear A.-D. (2015). *Building Ontologies with Basic Formal Ontology*. The MIT Press
- Arvor D., Durieux L., André S., Laporte M-A. (2013). Advances in geographic object-based image analysis with ontologies: A review of main contributions and limitations from a remote sensing perspective. *ISPRS Journal of Photogrammetry Remote Sensing*, vol. 82, p. 125-137.
- Baartman J.E.M., Melsen L.A., Morre D., van der Ploeg M.J. (2020). *On the Complexity Of Model Complexity: Viewpoints Across the Geosciences*, Catena, 186.
- Bak P. (1996). *How Nature Works. The Science of Self-Organized Criticality*, Springer Verlag.
- Barnaud G. (2000). Identifier et caractériser les zones humides une diversité de points de vue. *Fonctions et valeurs des zones humides*, Fustec E., Lefeuvre J.C. *et al.*, Éditions Dunod Paris, coll. « Environnement », p. 39-59.
- Bausch U. (1979). Extraktion von Objekten aus Luftbildern durch objektspezifische Verfahren mit stufenweiser Verbesserung der örtlichen Genauigkeit. *Angewandte Szenenanalyse, DAGM-Symposium, Karlsruhe, 10-12 Okt.*, p. 50-62.
- Blaschke T. (2010). Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 65, n° 1, p. 2-16.
- Blaschke T., Hay G.-J., Kelly M., Lang S., Hofmann P., Addink E., Quieroz Feitosa R., van der Meer F., van der Werff H., van Coillie F., Tiede D. (2014). Geographic object-based image analysis – Towards a new paradigm. *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 87, January, p. 180-191.
- Bouillé F. (1975). Structuration et saisie des données cartographiques. *Journée d'étude : Acquisition et structuration de l'information graphique*, Comité français de Cartographie (CFC), Paris.
- Bouillé F. (1977). Un modèle universel de banque de données simultanément portable, répartie, Thèse d'État ès sciences (spécialité : mathématiques, mention : informatique) Paris, Université Pierre et Marie Curie-Paris VI.
- Bouillé F. (1978). Structuring cartographic data and spatial processes with the hypergraph-based data structure. *First International Symposium on Topological Data Structures for GIS*. Cambridge, G. Dutton (Ed.), Mass.: Laboratory for Computer Graphics and Spatial Analysis, Harvard University.
- Brentano F. (1976). *Philosophische Untersuchungen zu Raum, Zeit und Kontinuum*, hrsg. von S. Körner and R. M. Chisholm, Hamburg: Meiner (cited according to the English translation by B. Smith, *Philosophical Investigations on Space, Time and the Continuum*, London: Croom Helm, 1988).

- Burrough P.A., Frank A.U. (eds.) (1996). *Geographic Objects with Indeterminate Boundaries*, London and Bristol, PA: Taylor and Francis.
- Carnap R. (1922). *L'espace. Une contribution à la théorie de la science*. Trad. Wagner P., Editions Gallimard, coll. "NRF".
- Casati R., Varzi A.C. (1999). *Parts and Places. The Structures of Spatial Representation*. The MIT Press.
- Clementini E., Di Felice P. (1996). A model for representing topological relationships between complex geometric features in spatial databases. *Information Sciences*, 90, p. 121-136.
- Clementini E., Lejdel B., Mazzaguo S., Laurini R. (2021). Homological relations: A methodology for the certification of irregular tessellations. *Transactions in GIS*, 25, p. 491-515.
- Daniels S.E., Walker G.B. (1996). Collaborative learning: Improving public deliberation in ecosystem-based management. *Environmental Impact Assessment Review*, vol. 16, n° 2, p. 71-102.
- De Sède-Marceau M.-H., Moine A.T.S. (2011). Développement d'observatoires territoriaux, entre complexité et pragmatisme. *L'Espace Géographique*, vol. 2, p. 117-126.
- Dewey M. (1876). *A Classification And Subject Index For Cataloguing And Arranging: The Books And Pamphlets Of A Library*. Amherst, Mass.
- DiGiuseppe N., Pouchard L., Noy N. (2014). SWEET ontology coverage for earth system sciences. *Earth Science Informatics*, December, vol. 7, n° 4, p. 249-264.
- Edwards A.W.F. (2004). *Cogwheels of the Mind - The Story of Venn Diagrams*. Johns Hopkins University Press.
- Egenhofer M.J. (1989). A formal definition of binary topological relationships. *Lecture Notes in Computer Science*, vol. 367, p. 457-472.
- Egenhofer M.J., Frank A.U., Jackson J.P. (1989). A topological data model for spatial databases. Design and implementation of large spatial databases. SSD 1989. *Lecture Notes in Computer Science*, Buchmann A.P., Günther O., Smith T.R., Wang Y.F. (Eds) vol. 409, Springer, p. 271-286.
- Ehresmann A.C., Vanbreemsch J.-P. (2007). Memory evolutive systems: Hierarchy, emergence, cognition. *Studies in Multidisciplinarity*, Elseviers, vol. 4.
- Esch T., Thiel M., Bock M., Roth A., Dech S. (2008). Improvement of image segmentation accuracy based on multiscale optimization procedure. *Geoscience and Remote Sensing Letters, IEEE*, vol. 5, p. 463-467.
- Folke C. (2006). Resilience: The emergence of a perspective for social-ecological systems analyses. *Global Environmental Change*, vol. 16, n° 3, p. 253-267.
- Folke C., Colding J., Berkes F. (2003). Synthesis: Building resilience and adaptive capacity in social-ecological systems. *Navigating Social-Ecological Systems: Building Resilience for Complexity and Change*, Berkes F., Colding J., Folke C. (Eds), Cambridge University Press, p. 352-387.

- Fu K.S., Mui J.K. (1980). A survey on image segmentation. *Pattern Recognition*, vol. 13, p. 3-16.
- Fustec E., Lefeuvre J.C. (ed.), (2000). *Fonctions et valeurs des zones humides*, Dunod, coll. « Environnement ».
- Gangemi A., Guarino N., Masolo C., Oltramari A., Schneider L. (2002). Sweetening ontologies with DOLCE. Proceedings of the 13th *International Conference on Knowledge Engineering and Knowledge Management. Ontologies and the Semantic Web. EKAW'02*, Springer-Verlag, London, UK, p. 166-181.
- Garcia L. F., Abel M., Perrin M., Alvarenga R. S. (2020). The GeoCore ontology: A core ontology for general use in Geology. *Computers & Geosciences*, 135 (2020), 104387.
- Gervais S.J., Bernard P., Klein O., Allen J. (2013). Toward a unified theory of objectification and dehumanization. *Objectification and (De)Humanization, Nebraska Symposium on Motivation*, Gervais S.J. (Ed.), vol. 60, p. 1-23.
- Girel M. (2006). Relations internes et relations spatiales : James, Bradley et Green. *Archives de la Philosophie*, 69, p. 396-414.
- Grenon P. (2008). A primer on knowledge management and ontological engineering. *Applied Ontology. An introduction*, Munn K., Smith B. (Dir.), Ontos Verlag, p. 57-81.
- Gruber T.R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, vol. 5, n° 2, p. 199-220.
- Gruber T.R. (1995). Toward principles for the design of ontologies used for knowledge sharing. *International Journal of Human-Computer Studies*, 43, p. 907-928.
- Gruenfeld D.H., Inesi M.E., Magee J.C., Galinsky A.D. (2008). Power and the objectification of social targets. *Journal of Personality and Social Psychology*, vol. 95, n° 1, p. 111-127.
- Guarino N. (1997). Some organizing principles for a unified top-level ontology. *AAAI Technical Report*, SS-97-06.
- Guizzardi, G., 2005. Ontological foundations for structural conceptual models. *CTIT Ph.D. Thesis Series*, vol. 05-74, Universal Press, Enschede, The Netherlands.
- Gunderson L.H., Holling C.S. (eds) (2002). *Panarchy. Understanding Transformation in Human and Natural Systems*. Island Press.
- Halioua B. (2017). *Le procès des médecins de Nuremberg*, Erès, coll. « Espace éthique ».
- Harman G. (2018). *Object Oriented Ontology. A New Theory of Everything*. Pelican Book.
- Heller M. (1984). *Temporal Parts of Four-Dimensional Objects*. *Philosophical Studies*, 46, p. 323-334.
- Heller M. (1990). *The Ontology of Physical Objects: Four-Dimensional Hunks of Matter*, coll. "Cambridge Studies in Philosophy", Cambridge University Press.
- Hennig B. (2008). What is formal ontology? *Applied Ontology. An introduction*, Munn K., Smith B. (Dir.), Ontos Verlag, p. 39-56.

- Herre H. (2010). General formal ontology (GFO): A foundational ontology for conceptual modelling. *Theory and Applications of Ontology: Computer Applications*, Springer, p. 297-345.
- Hervé D., Guermond Y., Frieberg C., Masson E., Laloë F. (2016). Modélisation des systèmes complexes et interdisciplinarité. *Interdisciplinarités entre Nature et Sociétés*, Actes du colloque de Cerisy, Hubert B. et Mathieu N. (dir.), coll. « EcoPolis », Peter Lang, p. 167-178.
- Hess G. (2017). Réconcilier l'éthique environnementale et l'écologie politique : une analyse méta-éthique. *La pensée Écologique*, n° 1.
- Holling C.S., Carpenter S.C., Brock W.A., Gunderson L.H. (2002). Chapter 15 – Discoveries for sustainable futures. *Panarchy-Understanding Transformations in Human and Natural Systems*, Gunderson L.H., Holling C.S. (Ed.), p. 395-417.
- Husserl E. (1900). Logische Untersuchungen. *Prolégomènes à la logique pure*, trad. Elie H. et al. (1969), PUF.
- Johansson I. (1992). Intentionality and tendency: How to make Aristotle up to date. *Language, Truth and Ontology*, K. Mulligan (Ed.), Kluwer, p. 180-192.
- Johansson I. (2004). *Ontological Investigations. An Inquiry Into Categories of Nature, Man and Society*. Ontos Verlag, 2nd ed.
- Jorgensen S.E., Fath B.D., Bastianoni S., Marques J.C., Muller F., Nielsen S.N., Patten B.C., Tiezzi E., Ulanowicz R.E. (2007). *A New Ecology. System Perspective*, Elsevier.
- Kuipers B.J. (1978). *Modeling Spatial Knowledge. Cognitive Science*, 2, p. 129-153.
- Kuipers B.J. (1994). An ontological hierarchy for spatial knowledge. *AAAI Technical Report FS-94-03*, p. 93-97.
- Laurini R. (1990). Ingénierie des connaissances spatiales : le cas de la géomatique. *Espace Géographique*, tome 19-20, 1, p. 24-40. doi: <https://doi.org/10.3406/spgeo.1990.2941>.
- Le Moigne J.-L. (1999). *La modélisation des systèmes complexes*, Dunod.
- Le Moigne J.-L., Morin E. (dir.) (2007). *Intelligence de la complexité. Epistémologie et pragmatique*, L'Aube, essai.
- Liu J., Dietz T., Carpenter S.R., Alberti M., Folke C., Moran E., Pell A.N., Deaman P., Kratz T., Lubchenko J., Ostrom E., Ouyang Z., Provencher W., Redman C.L. Schneider S.H., Taylor W.W. (2007). Complexity of coupled human and natural systems. *Science*, n° 317-5844, p. 1513-1516.
- Loireau M., Fargette M., Desconnets J.-C., Khiari H. (2017). Observatoire scientifique en appui aux gestionnaires de territoire, entre abstraction OSAGE et réalité ROSELT/OSS. *Revue internationale de géomatique*, vol. 27, n° 3, p. 303-333.
- Malczewski J., Jankowski P. (2020). Emerging trends and research frontiers in spatial multicriteria analysis. *International Journal of Geographical Information Science*.
- Marx K. (1992 [1844]). Economic and philosophical manuscripts. *Early Writings*, Trad. Livingstone R., Benton G., Peguin Books p. 279-400.

- Mc Mullen J.M., Meacham P.A. (1996). A comparison of wetland boundaries delineated in the field to those boundaries on existing state and federal wetlands maps in central new york state. *Wetlands: Environmental Gradients, Boundaries, and Buffers*, Mulamootil G., Warner B.G., McBean E.A. (Eds), CRC, Lewis Publishers, Proceedings of an international symposium, April 22-23, 1994, p. 193-205.
- Meunier C.L., Malzahn A.M., Boersma M. (2014). A new approach to homeostatic regulation: Towards a unified view of physiological and ecological concepts. *PLoS ONE*, vol. 9, n° 9.
- Morin E. (2008 [1986]). La méthode. Tome 3. La connaissance de la connaissance. *La méthode*, E. Morin, vol 1, coll. « Opus », Seuil, p. 1169-1466.
- Munn K., Smith B. (dZir.) (2008). *Applied Ontology. An introduction*. Ontos Verlag.
- Niles I., Pease A. (2001). Towards a standard upper ontology. *Proceedings of the International Conference on Formal Ontology in Information Systems*, vol. 2001, ACM, p. 2-9.
- Nussbaum M. (1995). Objectification. *Philosophy & Public Affairs*, vol. 24, n° 4, p. 249-291.
- Partelow S. (2018). A review of the social-ecological system (SES) framework: Applications, methods, modifications and challenges. *Ecology and Society*, vol. 23, n° 4.
- Partridge C., Stefanova M. (2003). Building a foundation for ontologies of organisation. *The Ontology and Modelling of Real Estate Transactions: European Jurisdictions* ("International Land Management" Series), p. 141-149.
- Peterson G.D. (2008). Self-organization and discontinuities in ecosystems. *Discontinuities in Ecosystems and Other Complex Systems*, Allen C.R., Holling C.S. (Ed.), Columbia University Press, p. 20-30.
- Pirot F., Saint-Gérard T. (2005). La Géodatabase sous ArcGIS, des fondements conceptuels à l'implémentation logicielle, *Géomatique Expert*, n° 41/42, février-mars, p. 62-66.
- Prewitt J.M.S. (1971). Object enhancement and extraction, *Picture Processing and Psychopictorics: Pictorial Pattern Perception*, Lipkin B.S., Rosenfeld A. (eds), 1st edition, Academic Press Inc, p. 75-149.
- Proteau J.C. (2011). Matrices de Clementini et Prédicats spatiaux de l'OGC. Ministère de l'Écologie, du Développement durable, des Transports et du Logement, SG/SPSSI/PSII.
- Radig B. (1978). Description of moving objects based on parametrized region extraction. *Proceedings of the 4th International Conference on Pattern Recognition*, Kyoto Japan; 7-10 Nov., p. 719-721.
- Randell D.A., Cui Z., Cohn A.G. (1992). A spatial logic based on regions and connection. *KR* 92, p. 165-176.
- Raskin R. G. and Pan M. J. (2005). Knowledge representation in the semantic web for Earth and environmental terminology (SWEET). *Computers & Geosciences*, vol 31, n° 9, Nov., p. 1119-1125.
- Saint-Gérard T. (2002). SIG : Structures conceptuelles pour l'analyse spatiale. HDR Université de Rouen.

- Servigne S. (2010). Conception, architecture et urbanisation des systèmes d'information. *Encyclopædia Universalis*, p. 1-15.
- Smith B. (1994). *Austrian Philosophy: The Legacy of Franz Brentano*. Open Court.
- Smith B. (1996). Mereotopology: A theory of parts and boundaries. *Data and Knowledge Engineering*, 20, p. 287-303.
- Smith B. (1999). *Ontology: Philosophical and Computational*, non publié, <http://wings.buffalo.edu/philosophy/faculty/smith/articles/ontologies.htm>
- Smith B., Mark D.M. (2003). Do mountains exist? Towards an ontology of landforms. *Environment and Planning B: Planning and design*, vol. 30, p. 411-427.
- Smith B., Varzi A.C. (2000). Fiat and bona fide boundaries. *Philosophy and Phenomenological Research*, vol. 60, n° 2, p. 401-420.
- Smith B., Woodruff Smith D. (dir.) (1995). *The Cambridge Companion to Husserl*. Cambridge University Press.
- Sowa J-F. (1995). Top-level ontological categories. *International Journal of Human-Computer Studies*, vol. 43, n° 5-6, Nov., p. 669-685.
- Steneck R.S. (1986). The ecology of coralline algal crusts: convergent patterns and adaptative strategies. *Annual Review of Ecology and Systematics*, vol. 17, p. 273-303.
- Strobl C. (2008). Dimensionally extended nine-intersection model (DE-9IM). *Encyclopedia of GIS*, Shekhar S., Xiong H. (Eds), Springer, Boston, MA.
- Tissot C., Cuq F. (2004). Apport des SIG pour la modélisation spatio-temporelle d'activités humaines. *Revue Internationale de Géomatique*, vol. 14, n° 1, p. 83-96.
- Tran B.H., Plumejeaud-Perreau C., Bouju A., Bretagnolle V. (2017). Système d'information spatiotemporel pour l'intégration et l'exploitation de données environnementales. *Revue internationale de géomatique*, vol. 27, n° 3, p. 423-443.
- Vogt L., Grobe P., Quast B., Bartolomaeus T. (2011). Top-level categories of constitutively organized material entities. Suggestions for a formal top-level ontology. *PLoS ONE* vol. 6, n° 4, e18794.
- Von Bertalanffy L. (1993 [1968]). *General System Theory*. Trad. Chabrol J.-B. (1973), Dunod.
- Walker B., Holling C.S., Carpenter S.R., Kinzig A. (2004). Resilience, adaptability and transformability in social-ecological systems. *Ecology and Society*, vol. 9, n 2.
- Ye J., Stevenson G., Dobson S. (2011). A top-level ontology for smart environments. *Pervasive and Mobile Computing*, 7, p. 359-378.