Introduction : Complémentarité des approches symboliques et numériques pour l'étude de dynamiques spatiales

Christine Plumejeaud-Perreau, Mireille Fargette, Julie Gravier, Thérèse Libourel Rouge, Eric Masson, Hélène Mathian, Lucie Nahassia, Laure Nuninger, Xavier Rodier, Lena Sanders, et al.

▶ **To cite this version:**

Christine Plumejeaud-Perreau, Mireille Fargette, Julie Gravier, Thérèse Libourel Rouge, Eric Masson, et al.. Introduction : Complémentarité des approches symboliques et numériques pour l'étude de dynamiques spatiales. Revue Internationale de Géomatique, 2022, 31 (1-2), pp.7-19. 10.3166/rig31.7-19 . halshs-03919029

HAL Id: halshs-03919029
https://shs.hal.science/halshs-03919029

Submitted on 23 Jan 2023

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

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

Introduction

Complémentarité des approches symboliques et numériques pour l'étude de dynamiques spatiales

Plumejeaud-Perreau C.[1], Fargette M.[2], Gravier J. [3], Libourel T. [4], Masson E. [5], Mathian H. [6], Nahassia L. [7], Nuninger L. [8], Rodier X. [9], Sanders L. [10], Saux E. [11]

1. *Migrinter, UMR 7301 CNRS & Université de Poitiers*
   5 av. Theodore Lefebvre,86000 Poitiers, France
   christine.perreau@univ-poitiers.fr

2. *IRD, UMR 228 Espace-Dev,*
   500 rue Jean-François Breton, 38093 Montpellier Cedex 05, France
   mireille.fargette@ird.fr

3. *UMR 8504 Géographie-cités*
   Campus Condorcet, 5 Cours des Humanités, 93322 Aubervilliers – Cedex
   julie.gravier@parisgeo.cnrs.fr

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

5. *Laboratoire TVES ULR 4477, Université Université de Lille,*
   Cité Scientifique, Avenue Paul Langevin, 59650-Villeneuve-d'Ascqc, France
   eric.masson@univ-lille.fr

6. *UMR 5600 Environnement Ville Société*
   ENS, 15 parvis René Descartes – BP 7000 69342 Lyon Cedex 07
   helene.mathian@ens-lyon.fr

7. *UMR 8504 Géographie-cités*
   Campus Condorcet, 5 Cours des Humanités, 93322 Aubervilliers – Cedex
   lucie.nahassia@parisgeo.cnrs.fr

8. *CNRS / University of Bourgogne Franche-Comté – Chrono-Environnement*
   MSHE C.N. Ledoux, 25030 Besançon, France.
   laure.nuninger@univ-fcomte.fr

9. CNRS / Université de Tours – CITERES
   MSH Val de Loire, 37020 Tours, France

Xavier.rodier@univ-tours.fr

10. *UMR 8504 Géographie-cités*
5 cours des Humanités 93322 Aubervilliers cedex
Lena.sander@parisgeo.cnrs.fr

11. *Institut de Recherche de l'Ecole Navale, Ecole navale, BCRM Brest,*
Lanvéoc-Poulmic, CC 600, F-29240 Brest Cedex 9
eric.saux@ecole-navale.fr

RESUME. *Ce texte permet de resituer les contributions respectives des articles de ce numéro spécial au sein d'une problématique scientifique qui a animé les réunions de l'Action Prospective du GdR CNRS MAGIS durant cinq années, de 2015 à 2020. Cette problématique, celle de la possible et nécessaire complémentarité des approches numériques et symboliques pour l'étude de dynamiques spatiales a été discutée en interdisciplinarité. Elle a été analysée d'un point de vue particulièrement réflexif, où les différents participants ont fait part de retours d'expériences et/ou de propositions originales formulées à partir des différentes pratiques vécues en interdisciplinarité.*

ABSTRACT. *This text allows us to resituate the respective contributions of the chapters of this special issue within a scientific problematic that has animated the meetings of the Action Prospective of the GdR CNRS MAGIS during four years, from 2015 to 2020. This problematic, that of the possible and desirable complementarity of numerical and symbolic approaches for the study of spatial dynamics, has been discussed in interdisciplinarity, of course. Furthermore, this question was analysed from a particularly reflexive point of view, where the different participants have shared their feedbacks and/or formulated original proposals from the different practices experienced in interdisciplinarity.*

Mots-clés : ontologie, numérique, statistique, interdisciplinarité, dynamiques spatiales, système, mobilité, temps, dialogue, connaissances, empreinte, trace, transformation

KEYWORDS: *ontology, digital, statistics, interdisciplinarity, spatial dynamics, system, mobility, time, dialogue, knowledge, imprint, trace, transformation*

DOI:10.3166/RCMA.25.1-n © 2016 Lavoisier AR_DOI

1. 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, …), cet ouvrage 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 transdisciplinaire.

2. 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 le long des articles repose sur l'usage d'ontologies. Des ontologies, oui, mais desquelles parle-t-on ? L'ouvrage 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), et qui nous conduit à dire que parmi la multiplicité des définitions du terme "Ontologie" dont on dispose, nous pouvons dire que 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 cet ouvrage 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 êtres opérées sur ces entités."

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 existant 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) ;

[1] https://www.cnrtl.fr/definition/ontologie

– 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 l'ouvrage 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).

3. 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, et 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…), 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 voient une explosion des sujets de thèse et de master incluant l'usage du *deep learning* (apprentissage par réseaux de neurones artificiels), une méthode issue du domaine de l'Intelligence Artificielle qui prend sa revanche sur les approches symboliques qui ont pu prévaloir dans ce domaine (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 recombiner 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 l'article 2. Les données sont obtenues, comme dit Bruno Latour : « Décidément, on ne devrait jamais parler de "données", mais toujours d'"obtenues". » (Latour, 1993). 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 en amont, en particulier dans le cadre d'études de dynamiques spatiales avec des enjeux interdisciplinaires. Les articles 2 et 5 de cet ouvrage discutent abondamment de cette question.

Par ailleurs, concernant les approches numériques pour traiter les données, il y a les habitudes (et les mauvaises habitudes) : en effet, la pratique d'approches statistiques et de l'analyse de données nécessitent 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/infirmer 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.

[2] https://www.lebigdata.fr/data-lake-definition

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.

4. 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 l'ouvrage 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 qu'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 tels que 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 spatio-temporels réellement interopérables, au risque sinon d'une certaine maltraitance de la dimension sémantique des données. Dans cet ouvrage nous avons voulu donc 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 - (Masolo, 2003)) dans un système spatio-temporel 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 renseignés et déconnectés 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

---

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

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 spatio-temporelles (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 ouvrage n'en peut rien dire encore. Il se concentre plutôt sur la démonstration de la complémentarité des approches numériques et symboliques.

5. 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'étude de vestiges et relevés LIDAR en fait la démonstration. Cependant, le défaut de la seule approche symbolique 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 l'ouvrage.

Les approches numériques par l'analyse de données numériques qu'elles développent, si elles développent des modèles parfois implicites, comme expliqué dans l'article 1, si également leur choix est souvent guidé par l'habitus de la discipline (statistique, écologie, géographie, …), 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 enrichir le corpus de connaissances scientifiques, et l'ouvrage, à travers tous les articles essaie de le démontrer.

6. 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, théoriques et opérationnelles. 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 publiées, 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 articles a tout le moins été intensément débattu entre les membres de l'Action Prospective du GdR 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. On ne peut que les remercier de leurs inestimables contributions et leur patience pour la relecture de articles 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 spatiale scolaire. 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, étudiés à différentes échelles, l'article 2 illustre comment un cadre conceptuel partagé concernant la description des différents artefacts que le temps à 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, ils 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 par 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, nous parvenons à inférer 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, nous optons pour cette seconde stratégie d'exploitation du modèle de connaissances.

Dans cet ouvrage 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 articles.

En effet dans un contexte de demande sociétale forte autour de la modélisation des enjeux que pose les 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 généricité. 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 Cloud[4]) permettent la collecte et 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 leur 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 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-échelle et multi-source 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

[4] https://eoscobservatory.eosc-portal.eu/

ontologiquement pour intégrer une 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 6 concepts ayant une portée transdisciplinaire et ambitionnant 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 livre 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 connaissance et traces numérisées, exprimé selon un vocabulaire diversifié, qui est exemplifié dans chacun des articles. De plus, certains articles privilégieront un sens, ou bien l'autre dans leur focus et leurs illustrations. 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*, nº 20, p. 35-41, doi: 10.5311/JOSIS.2020.20.666

Collins et al., S. L. (2011) An integrated conceptual framework for long-term social–ecological research, *Frontiers in Ecology and the Environment*, vol. 9, nº 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, nº 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., « 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, nº 5-6, p. 907-928, doi: 10.1006/ijhc.1995.1081.

Guarino, N (1998) « Formal ontology in information systems ». In Guarino, N., editor, *Proceedings of FOIS'98*, vol. 46, pp. 3–15, Trento, Italy. IOS Press.

Harbelot, B., Arenas, H. & Cruz, C., (2014), A Semantic Model to Query Spatial–Temporal Data, In : V. Popovich, C. Claramunt, M. Schrenk et K. Korolenko (éd.), *Information*

Fusion and Geographic Information Systems (IF AND GIS 2013), Springer Berlin Heidelberg, Berlin, Heidelberg, p. 75-89. doi: 10.1007/978-3-642-31833-7_5

Hess C., E. Ostrom (2006), *Understanding knowledge as a commons,* MIT Press.

Holmes, H. (2001) « The great term robbery [computer jargon] », *Computer*, vol. 34, nº 5, p. 94-96, doi: 10.1109/2.920619.

Isaac, A. H. J. C. A. (2011). Entre thésaurus et ontologies : une affaire d'intéropérabilité 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. & 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. & Fargette, M., (2022), Science paysagère au service de l'observatoire scientifique Sociétés-Milieux 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., S. Bowers, M. Schildhauer, S. Krivov, D. Pennington, et F. Villa, (2007) An ontology for describing and synthesizing ecological observation data, *Ecological Informatics*, vol. 2, nº 3, p. 279-296, doi: 10.1016/j.ecoinf.2007.05.004.

Martinez-Cruz, C., Blanco, I. J. & 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 and 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, « Pour un observatoire des données géographiques du Web: Expérimentation à partir des infrastructures de données géographiques françaises », *Rev. Int. Geomat.*, vol. 29, nº 1, p. 9-30, janv. 2019, doi: 10.3166/rig.2019.00074.

Ollion É. et J. Boelaert, (2015) Au delà des big data. Les sciences sociales et la multiplication des données numériques, *Sociologie*, vol. 6, nº 3, p. 295-310.

Prieur, M. (1999) La Convention d'Aarhus, instrument universel de la démocratie environnementale, *rjenv*, vol. 24, nº 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, nº 1, p. 1-12, doi: 10.1037/a0018326.

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

Tran, B.-H., C. Plumejeaud-Perreau, A. Bouju, et V. V. Bretagnolle, (2015) A Semantic Mediator for Handling Heterogeneity of Spatio-Temporal Environment Data, in *Metadata and Semantics Research*, Manchester, United Kingdom, doi: 10.1007/978-3-319-24129-6.

Ushold M. et M. Gruninger, (1996) Ontology : Principles, methods and applications. Knowledge Engineering Review, vol. 11 n°2, p. 93–126. doi:10.1017/S0269888900007797

Wasserstein, R. L. & 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