

# Du foisonnement des disciplines à la recomposition d'une réalité partagée : élaboration d'une structure de modélisation dédiée à l'intégration de connaissances disciplinaires.

Jean Le Fur<sup>1</sup>

## Résumé

Comme beaucoup d'autres, le domaine de la biologie des populations de rongeurs sauvages est le siège d'une diversification importante de disciplines de plus en plus spécialisées. La réalité observée *in situ* est cependant le reflet intégré de l'ensemble des mécanismes et processus appréhendés séparément par chaque spécialité disciplinaire. On cherche à construire un cadre formel de représentation (de type système multi-agents) permettant d'articuler ces différents savoirs pour rendre compte de ces réalités complexes. La contrainte principale est la construction d'un cadre apte à rendre compte de toutes les facettes et spécificités liées à chaque discipline. Pour construire le cahier des charges du modèle, deux ontologies ont été élaborées aux deux extrêmes du processus de recherche. La première porte sur le questionnement et la perception des différents domaines de recherche. Elle est fondée sur une analyse lexicale d'entretiens réalisés avec des chercheurs. La deuxième traite des savoirs produits, elle a été élaborée à partir d'informations unitaires proposées par les chercheurs du groupe étudié. Les deux classifications ont été comparées puis synthétisées pour produire le cahier des charges du modèle. La plate-forme a été alimentée de façon incrémentale avec des études concrètes de cas contrastées permettant de valider (i) l'articulation entre les problématiques et (ii) la robustesse de l'approche. La correspondance ontologies/plate-forme a été réalisée avec des différences notables telles que l'absence de la notion de protocole dans l'ontologie et l'absence de formalisation des concepts dans la plate-forme. L'approche proposée apparaît potentiellement robuste pour permettre l'enrichissement mutuel des approches propre à la pluridisciplinarité.

## Introduction

Avec le développement de la connaissance, on constate dans des domaines variés une tendance à la multiplication de disciplines dédiées à des descriptions de plus en plus spécifiques d'une réalité observée (*e.g.*, Mathieu, 1991, Rouxel 2002). Dans le domaine de la biologie des petits rongeurs et de leurs parasites par exemple, le seul niveau moléculaire donne lieu à des études en épidémiologie moléculaire, génétique de la conservation, génétique des populations, génomique des populations, immunogénétique, cytogénétique, cytotaxonomie, phylogénie moléculaire, phylogéographie,... Cette démultiplication se constate à tous les niveaux du vivant (cellule, organisme, population, communauté, écosystème), à tous les domaines problématiques, de la biogéographie à la taxonomie en passant par la morphométrie ou l'épidémiologie ainsi qu'à une vaste gamme d'échelles spatiales et temporelles. Il faut enfin noter

---

<sup>1</sup> CBGP, UMR INRAe-IRD-CIRAD-Montpellier Sup Agro, Montferrier-sur-Lez, France, [lefur@ird.fr](mailto:lefur@ird.fr)

l'idiosyncrasie (Pueyo, 2007) de la recherche scientifique pour laquelle, par principe, un scientifique développera systématiquement une recherche originale par rapport à l'existant.

La spécialisation disciplinaire est reconnue comme une nécessité pour décrypter les facettes de plus en plus spécifiques des systèmes étudiés (Newell, 2001) et établir des relations entre des variables très spécifiques dans les différents domaines. La compartimentation de plus en plus aigüe des approches devient cependant problématique car elle conduit à une déconnexion vis à vis des problématiques concrètes (Burger et Kamber, 2003) et se limite de plus en plus à des groupes de plus en plus circonscrits et seuls à entendre un langage très spécialisé (Hursh et al., 1990).

Quelle qu'en soit leur justification, les approches disciplinaires et les connaissances qui en résultent naissent cependant toutes d'une même réalité observée: c'est le même rongeur pris dans un piège sur un terrain d'Afrique ou d'Europe qui servira de support au développement des études abordées par chaque discipline. Autrement dit, la présence du rongeur au lieu et au moment de sa capture est en fait le résultat intégré des multiples processus abordés séparément par chaque discipline, depuis son adaptation phylogénique jusqu'à sa réponse physiologique, biologique, éthologique aux conditions qu'il, ou sa population, a rencontré durant son existence (e.g., Storch et al., 2007). Chaque élément de connaissance produit par les disciplines constitue donc à la fois un élément légitime (Anderson, 1972 dans McMurty, 2009) et une composante de l'ensemble des processus ayant conduit au fait constaté (le rongeur piégé).

D'autre part, en termes systémiques, les phénomènes observés *in situ* sont des patterns émergents issus de processus complexes d'organisation et d'auto-organisation (Warfield et Warfield, 1999, Kauffman, 1993, Newell, 2001, Grimm et Railsback, 2005). Pour pouvoir traiter ces phénomènes, par exemple dans un processus de gestion, il est nécessaire d'appréhender l'aspect multidimensionnel des éléments et de leurs interactions (Meek, 2001) et, conséquemment, de mettre en œuvre une collaboration interdisciplinaire (Wheeler, 2007).

Dans ce contexte, la pluridisciplinarité ne doit pas être seulement présentée comme la combinaison de plusieurs disciplines pour mieux appréhender un problème spécifique, mais aussi comme une tentative d'intégration visant à faire émerger un point de vue "transcendant" à partir des différentes perspectives (McMurtry, 2009). Eu égard à la complexité des processus étudiés, l'interdisciplinarité apparaît alors comme une condition sine qua non pour mieux identifier et donner du sens aux systèmes complexes étudiés (Newell, 2001).

Les modalités de cette intégration ne sont cependant pas claires et posent de multiples questions concernant par exemple:

La prévalence de certains processus; toutes les connaissances disciplinaires acquises ont-elles le même poids pour une observation concrète ?

L'unicité du concret, comparée au nombre et à la diversité des processus et phénomènes (Klein, 2001). Les disciplines doivent-elles ainsi converger toutes vers une même réalité et, si non, quelles réalités sont en présence; peut-on factoriser un sous-ensemble de réalités ?

Le point de vue de l'observation; la diversité des protocoles et des angles de vue retenus par les disciplines est-elle compatible avec la recombinaison d'une réalité partagée ?

On se propose, dans le domaine "restreint" de la biologie des petits rongeurs et de leurs parasites, de chercher les modalités permettant d'élaborer une représentation intégrée de l'apport disciplinaire à la compréhension des phénomènes observés *in situ*. L'approche retenue est fondée sur la construction d'une structure de modélisation (ou plate-forme) dédiée permettant de formaliser conjointement la diversité des approches, questions et connaissances

élaborées séparément par chaque discipline et problématique. Pour assurer l'articulation disciplinaire, l'élaboration de cette structure repose sur l'identification et la formalisation des différentes facettes de la recherche scientifique partagées par les disciplines. Compte tenu de l'aspect patrimonial et hypothétique de l'objectif ainsi que la complexité du problème, cette démarche est mise en œuvre en tant qu'heuristique visant à progresser vers une meilleure appréhension de cet ensemble de questions et une meilleure identification des problèmes concrets qu'elle peut poser.

## Matériel et Méthodes

La contrainte quasi unique imposée au système formel à construire est la robustesse. Pour être valide, il est en effet nécessaire que la plate-forme soit apte à intégrer toute problématique du domaine, quelle que soit sa dimension spatiale, temporelle ou fonctionnelle et quelle que soit son actualité, autrement dit qu'elle corresponde à des problématiques définies il y a dix ans, de nos jours ou dans dix ans.

Les contours de l'étude sont définis par les travaux réalisés par une équipe pluridisciplinaire localisée de chercheurs travaillant sur le domaine<sup>2</sup>. Le sujet d'étude comprend tant la connaissance disponible sur le domaine que le groupe de chercheurs à l'origine du corpus de connaissances.

Pour aborder le problème de l'intégration, l'étude s'est alors basée sur une caractérisation spécifique des recherches réalisées et de leurs résultats puis l'implémentation pas à pas des différentes approches disciplinaires:

Une étude préliminaire du groupe de chercheurs a confirmé la diversité du domaine qui se décline en deux composantes:

1. une diversité d'individualités au sein de laquelle deux scientifiques de la même discipline aborderont les problèmes de façon distincte. Dans ce domaine les recherches développées dans le groupe étudié s'étendent (bouclent) de la seconde au million d'années, du micron à la planète, du gène à l'espèce.
2. une diversité des connaissances construites par ces chercheurs et traduite aussi bien en termes de questionnement, de nature de terrain, d'approche, de traitement et de connaissances produites.

La construction d'une structure robuste aux approches disciplinaires est fondée sur la modélisation de ces deux ensembles diversifiés selon la démarche suivante :

1. En amont de la recherche, l'élaboration d'une classification générique des questions et domaines de recherches abordées par chaque chercheur.
2. En aval de la recherche, l'élaboration d'une classification générique des connaissances produites par les chercheurs.

L'approche retenue dans les deux cas vise à l'élaboration d'une ontologie de type 'facet-ontology' (Madin et al., 2008) dans laquelle la partition est réalisée sur des ensembles disjoints, ou orthogonaux, de concepts tels que discipline, objet biologique, outil,...

---

<sup>2</sup> Groupe rongeur du Centre de Biologie pour la Gestion des Populations, UMR 022 INRA-IRD-Cirad-SupAgro.

## A. Élicitation des questions et domaines de recherche

L'identification des perceptions et domaines de connaissance abordés par les chercheurs est réalisée sur la base d'une analyse lexicale (Le Fur et Hervé, 2009) et selon un protocole formalisé comprenant plusieurs étapes (préparation, interaction, élicitation, réification, articulation).

Préparation de l'entrevue: Le chercheur est sollicité en lui fournissant une matrice de questionnements lui indiquant les contours de l'entretien à venir. Chaque chercheur fournit ensuite un ensemble de documents décrivant son domaine de recherche. Outre les publications clés, les rapports d'activité périodiques des chercheurs ont été mis à disposition et ont constitué des documents de référence pour conduire une première analyse lexicale dont l'objectif principal est de familiariser l'interviewer avec le domaine de questionnement de l'expert et identifier les termes critiques.

La méthodologie de classification est fondée sur les associations multiples entre les mots des textes rédigés et fournis par chaque différent chercheur. Si dans un ensemble de documents, on trouve souvent les mêmes mots à l'intérieur d'un segment de texte, il est probable qu'ils décrivent un même contexte. On ne sait pas exactement ce qu'en dit le locuteur mais on sait de quoi il parle.

Méthode : La classification en thèmes ou en contextes est faite à partir une matrice de co-occurrence entre les mots. Il s'agit d'une méthode statistique qui «imite» en quelque sorte le travail humain de classement. La méthode consiste à calculer les associations au sein d'une "fenêtre" de 10 ou 15 mots. Lorsque deux mots clés sont proches, il est probable qu'ils font partie d'une même unité de sens. Le logiciel utilisé pour réaliser l'analyse est WordMapper (<http://www.grimmersoft.com>).

Représentation (Figure 1) : Les résultats se présentent sous la forme d'un graphique en étoile représentant les différents contextes d'utilisation d'un mot. Le mot étudié est au centre, le chiffre situé sur le trait représente la fréquence de cooccurrence entre les 2 mots. Un mot situé à droite et un mot situé à gauche ne partagent pas le même contexte avec le mot central.

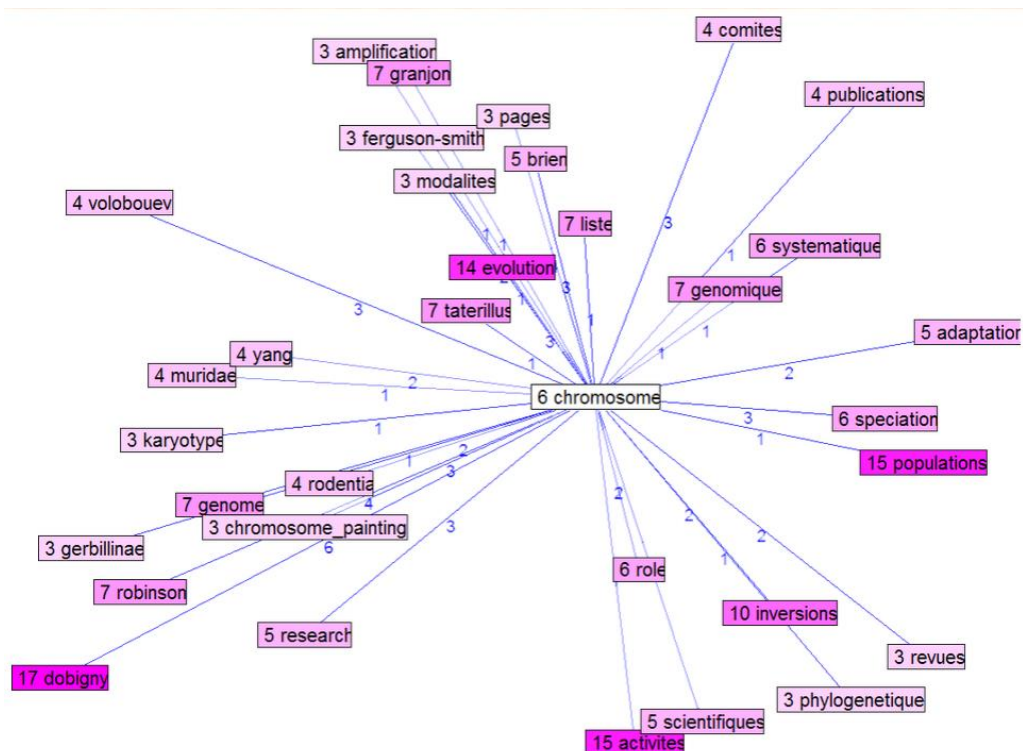


Figure 1 Un exemple de variété et de connexion entre concepts : one scientist's publications provided for the preparation of an interview. Les poids indiquent le nombre de liens dans la série de textes fournie par le chercheur.

L'analyse est fondée sur les associations multiples entre les mots et non sur la syntaxe des phrases. La classification élaborée classe statistiquement les combinaisons de mots en fonction de leur densité et de leur centralité (Courtial, 1990) : la centralité indique l'importance des liens avec les autres combinaisons de termes, la densité indique l'importance des liens qui existent entre les termes des combinaisons.

L'analyse lexicale permet de filtrer la liste des mots composés qui sont en général les plus signifiants, sauvegarder les associations qui paraissent intéressantes ou critiques. Ils constituent le contexte permettant à l'interviewer de réaliser l'étape suivante.

**Interaction/concertation:** L'interview semi-directif est alors réalisé (2 à 6 heures). Le chercheur présente de façon libre son domaine puis les questions sur cette présentation transforment l'entrevue en un échange questions-réponses qui paraît propice à une description plus en profondeur du champ, de ses caractéristiques et des questionnements associés. L'entretien se poursuit par une revue commune des questions notées lors de l'analyse lexicale. Suit alors un inventaire semi-dirigé avec l'expert des descripteurs clés. Cet inventaire se produit après la discussion, l'expert a donc en tête l'ensemble de son domaine et de son questionnement. L'inventaire doit être rapide, pour que seuls les termes saillants ressortent. L'expert confronte enfin son inventaire avec la liste obtenue lors de la pré-analyse lexicale pour obtenir une synthèse consensuelle.

**Elicitation:** L'ensemble de l'entretien fait l'objet d'un enregistrement audio, cet enregistrement est ensuite retranscrit intégralement selon un format normalisé identique pour tous les chercheurs (le texte obtenu fait une dizaine de pages en moyenne) et réalisé par la même personne.

**Réification:** Ce document seul fait ensuite l'objet d'une nouvelle analyse lexicale qui n'identifie que les mots-clés (simples et composés). Cette phase est associée à diverses étapes de nor-

malisation (pluriels-singuliers, homonymes, majuscules, etc.). La liste des mots-clés est ensuite reclassifiée mot par mot dans une ontologie à un niveau où les meta-mots clés ou concepts sont de type « is-a-kind-of ». Les éléments de cette ontologie sont établis de façon incrémentale selon les besoins des nouveaux mots-clés rencontrés (ex: *Senegal* -> *location*, *phylogeography* -> *discipline*, *speciation* -> *phenomenon*, ...).

Articulation (consolidation des types): A chaque nouvelle expertise la classification précédente sert de référence. Lorsque des catégories sont modifiées ou que de nouvelles catégories apparaissent les analyses précédentes sont mises à jour pour assurer la compatibilité arrière ou descendante, c'est-à-dire la cohérence avec les informations précédemment saisies. Les mots-clés propres à chaque chercheur sont de cette façon catégorisés en meta-clés partagés (e.g., discipline, mesure, outil, méthode, concept, ...) définies de façon incrémentale et conduisant à une première typologie.

## B. Élicitation des connaissances produites

La deuxième typologie concerne la nature des savoirs obtenus au terme des travaux de recherche. Elle a été construite à partir d'un système d'information orienté mots-clés et construit ad hoc. L'application logicielle constitue une tentative d'intégration et d'articulation de connaissances et savoirs de toutes nature et formats sur un domaine scientifique donné. L'application élabore à partir d'informations unitaires formatées un ensemble de mots-clés sur lequel est construite une ontologie simple comprenant uniquement des relations du type "est-un-type-de" du domaine. Les trois ensembles informations/mots-clés/ontologie conduisent à un réseau sémantique qui peut être utilisé pour naviguer dans le domaine de connaissance (Figure 2)<sup>3</sup>.

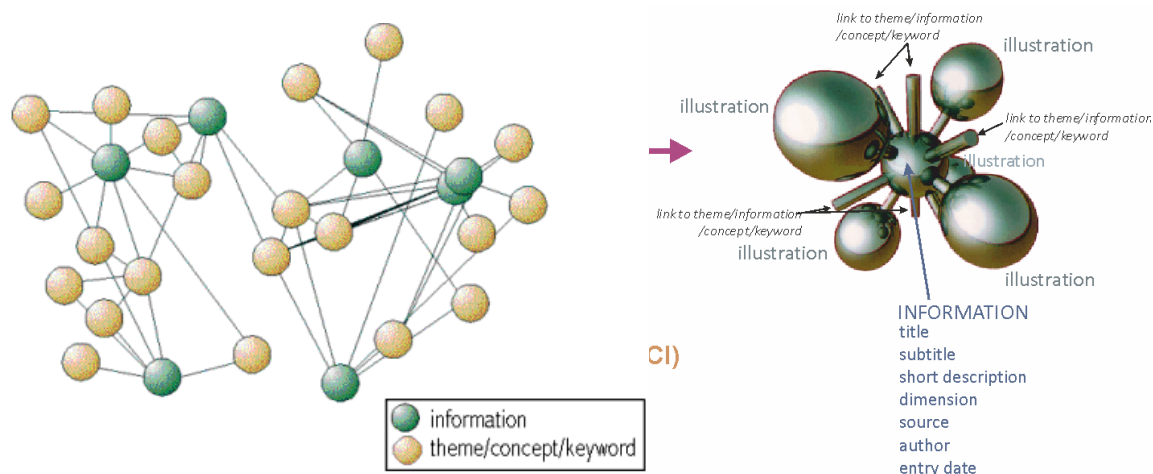


Figure 2 Structure et navigation hypertexte dans le Centre d'Informations (CI): l'utilisateur peut se déplacer d'une information à une autre en prenant n'importe quel chemin. L'ontologie construite au fur et à mesure de l'introduction de nouvelles informations est traduite en diverses tables des matières (thème, support, format, ...) qui facilitent cette navigation (le CI n'est pas un moteur de recherche mais un système de navigation dans un domaine de connaissances).

<sup>3</sup> L'application logicielle proprement dite se présente sous la forme d'un site internet dynamique. Elle utilise un système de données MS Access/SQL, la technologie des servlets java, un transformateur XML (XSLT), un conteneur de servlets (Apache-Tomcat), un serveur virtuel CentOS 5. Elle génère un ensemble de listes (mots-clés, thèmes, sections, supports, types de savoir, sources, auteurs, ...) permettant l'accès au réseau selon différentes modalités. Le système logiciel a déjà été déployé sur quatre domaines de connaissances scientifiques (pêche en Guinée, écosystèmes marins, biologie des rongeurs, étude d'une réserve en Afrique); Les implémentations déjà réalisées sont accessibles sur le portail: <http://centreinfo.science>.

Chaque chercheur du groupe cible a été sollicité pour proposer une/des informations(s) concernant ses résultats. Chaque information fait l'objet d'un formulaire simple contenant une part importante dédiée aux mots-clés. L'identification des mots-clés est entièrement libre. Chaque information est calibrée en plusieurs champs dans une base de données unique et intégrée au système d'information sur internet selon un protocole formalisé.

De la même façon que pour la caractérisation des questions et domaines de recherche, chaque mot-clé est ensuite catégorisé de façon incrémentale dans une nouvelle classification ou typologie/ontologie dont les clés sont définies à la demande (*e.g.*, localisation, groupe taxonomique, échelle, outil, méthode, etc.). La règle ontologique utilisée est aussi du type « is-a-kind-of ». La gestion de base de données assure l'intégrité de la classification et sa compatibilité/articulation avec l'ensemble des informations existantes.

Une distinction importante vis-à-vis de l'approche précédente réside dans l'élaboration de l'ontologie résultante. Celle-ci a été réalisée dans le cadre d'une démarche collégiale pour laquelle un ensemble de cinq chercheurs du groupe a identifié les catégories et classé collectivement la moitié de l'ensemble des mots-clés (160 mots-clés).

### **C. Synthèse et transcription formelle des ontologies du domaine**

On dispose alors de deux typologies, une issue des points de vue des chercheurs et une issue des résultats obtenus par les recherches. Dans un troisième temps, ces classifications sont confrontées : la fusion des deux classifications, obtenues à partir de deux modes distincts d'interaction avec les chercheurs et portant sur les deux extrémités du processus scientifique, est résumée en une seule typologie synthétique de la recherche disciplinaire sur le domaine.

Cette typologie résultante devient alors une base de référence. Elle permet de constituer un cahier des charges de l'architecture informatique permettant de rendre compte et articuler au sein d'un seul schéma potentiellement la plupart des aspects de la dynamique des populations de rongeurs étudiées. Les différentes librairies de programmes assemblées au sein de cette unique plate-forme ont vocation à être le reflet de cette typologie. Les accords et les écarts entre la typologie constituée et la plate-forme proposée constituent des éléments de critique permettant de reconsidérer la problématique.

Le développement de la plate-forme est réalisé par l'implémentation incrémentale d'études de cas spécifiques à chaque chercheur et le plus contrastées possibles (Le Fur, 2013). Pour éprouver la robustesse de l'approche cinq études de cas le plus contrastées possibles, liées notamment à des échelles spatio-temporelles différentes et allant de la cage de laboratoire à l'aire géographique ont été implémentées (Tableau 2).

Tableau 1 étude de cas thématiques modélisées successivement au sein de la plate-forme

Étude de cas	Echelle spatiale	Echelle temporelle
Expérimentation d'hybridation en cage de rongeurs africains	50cm	1 an
Expérimentation d'hybridation en enclos de rongeurs africains	50m	3 ans
Evolution d'une population de campagnols dans un paysage dynamique	500m	10 ans
Expérience de capture-marquage-recapture dans une réserve africaine	500m	20 ans
Transport du rat noir par des véhicules commerciaux au Sénégal	500km	100 ans

Les études de cas successives s'articulent avec les précédentes, en bénéficient, et contribuent au développement global de la plate-forme en l'alimentant avec de nouvelles facettes issues de différentes recherches.

## Résultats

Le panorama des questionnements présentés par chaque chercheur laisse apparaître des différences non prises en compte dans la conception initiale de l'étude, particulièrement les postures des chercheurs dont certains sont spécialisés et abordent un nombre restreint de catégories (ex: une seule zone d'étude ou une seule thématique spécialisée) tandis que d'autres s'avèrent plus généralistes et décrivent des nombres beaucoup plus conséquents de thèmes pour une catégorie donnée. Les différences se soulignent encore lorsque l'on prend par exemple en compte la longueur de la carrière des personnes interrogées.



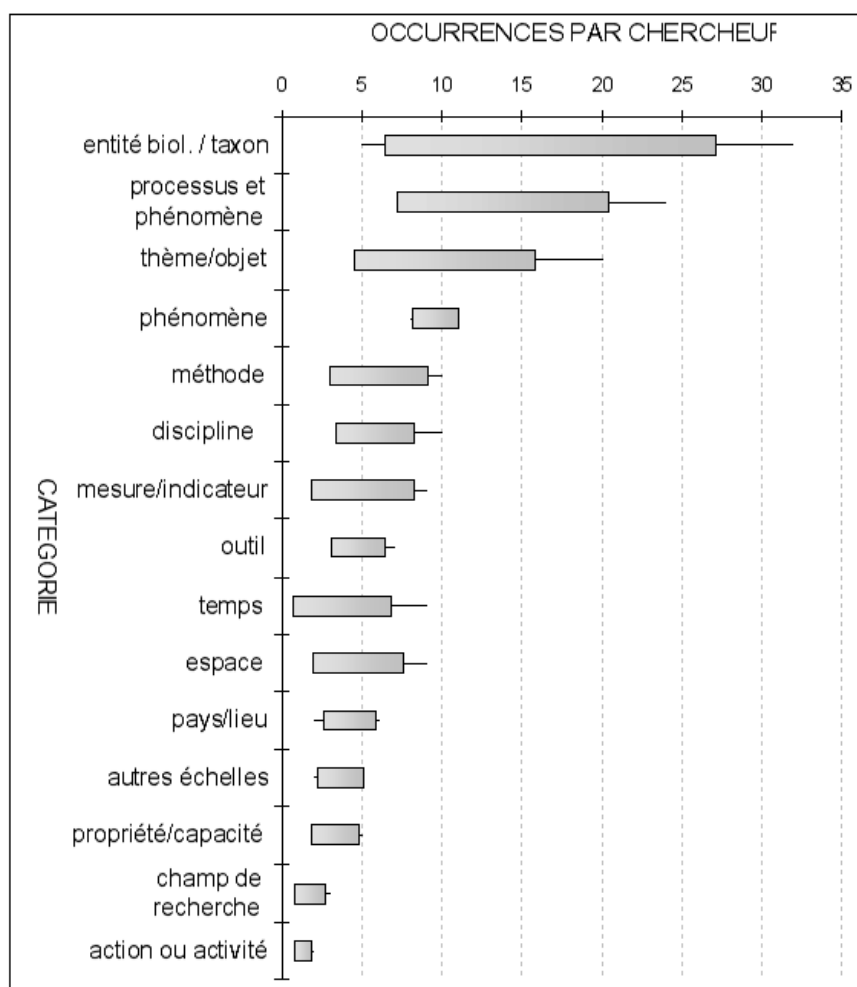


Figure 3 classification des domaines et des questions de recherche. La représentation (max/min  $\pm$  écarts-type) est obtenue à partir de l'analyse lexicale des entretiens avec cinq chercheurs disciplinaires. Les catégories obtenues sont classées par ordre décroissant de la moyenne des occurrences trouvées par chercheur interrogé.

Globalement, l'analyse (Figure 3) montre, dans ce groupe, la prépondérance des thèmes liés à la description des entités biologiques (i.e., diversité biologique). Les processus et phénomènes biologiques, les concepts puis les aspects liés aux terrains d'étude apparaissent ensuite comme des thèmes majeurs. La classification se poursuit par un ensemble de catégories qui cernent les aspects descriptifs (temps, espace, outil, ...) et se termine par des clés peu représentées qui traitent d'aspects particuliers comme des comportements ou des terrains spécifiques.

La classification obtenue en aval de la recherche à partir des connaissances produites est réalisée sur la base de 26 informations élaborées dans le centre d'informations (CI) et aboutissant à 330 mots-clés distincts. La catégorisation obtenue (Figure 4) est proche de la précédente. Les catégories peu représentées (référence bibliographique, projet, ...) apparaissent ici aussi accessoires. Par contre la typologie est dominée par la catégorie 'taxon' qui se détache nettement et beaucoup plus fortement que pour la typologie liée aux questionnements. On note de même que les 26 informations conduisent à l'identification de 36 disciplines distinctes, cette catégorie se trouvant en deuxième position des descripteurs d'information. Les catégories intermédiaires semblent décroître uniformément sans que se distingue d'autre groupe de catégories.

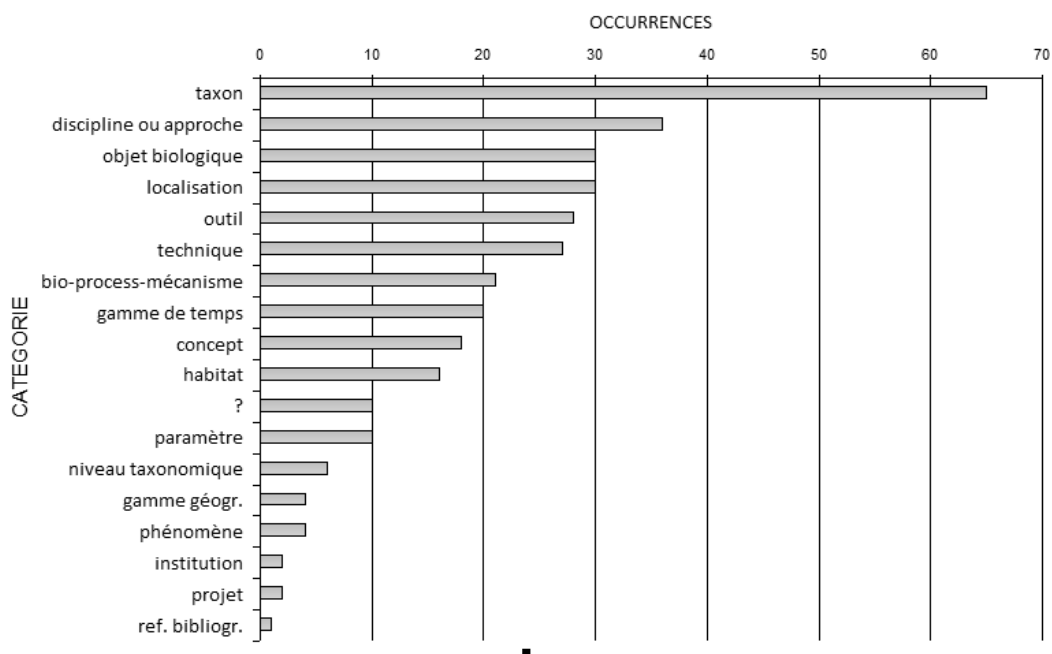


Figure 4. Classification obtenue collégalement à partir des connaissances disciplinaires insérées dans le centre d'informations (CI) et classées par ordre décroissant d'occurrence pour chaque clé.

Le rapprochement des ontologies est présenté sur la Figure 5. Les modèles biologiques (taxons) et la biodiversité constituent indéniablement la catégorie dominante dans ce domaine de recherche. Dans le reste du classement, les principales catégories se retrouvent identifiées dans les deux classifications mais positionnées à des rangs différents: prévalence des disciplines dans les résultats, des processus et phénomènes dans les questionnements. En bas de la figure, les catégories accessoires ne sont pas partagées.

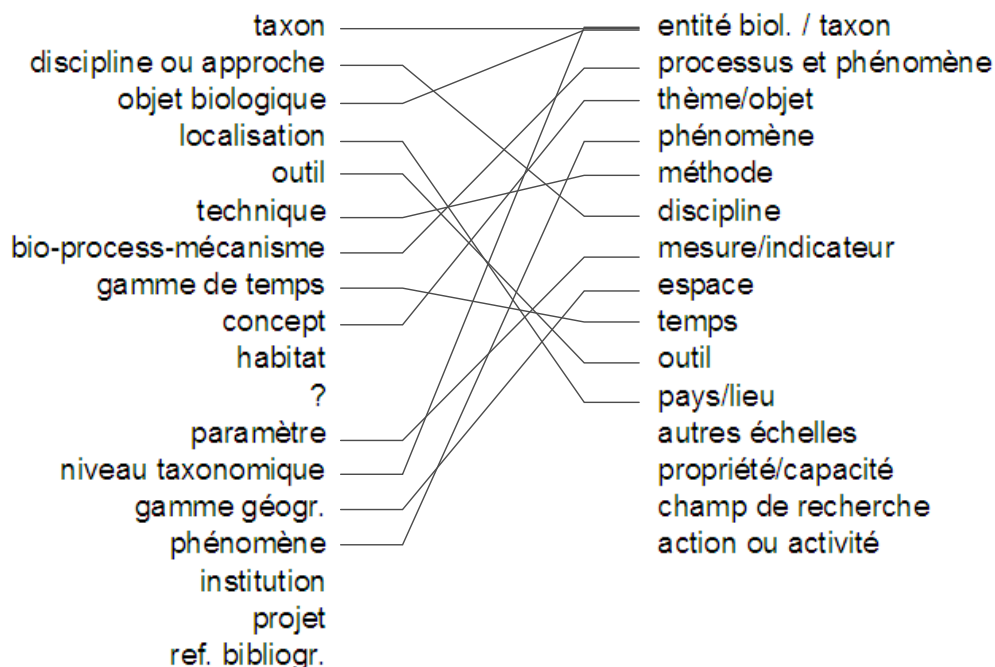


Figure 5 comparaison des classifications obtenues à partir des connaissances disciplinaires (gauche) et des domaines et questions de recherche exprimées par les chercheurs (droite). Chaque typologie est classée par rang décroissant du nombre d'occurrences citées pour chaque type. Les catégories non liées (le plus souvent en bas de tableau) dénotent des catégories sans correspondance dans l'autre classification.

Les deux classifications révèlent de nombreux points communs, elles ont pu être effectivement fusionnées. La typologie résultante (Tableau 2) comporte un ensemble de 11 classes

décrivant (i) les questions de recherche et les points de vue, (ii) la caractérisation des sujets de recherche, (iii) les modalités d'acquisition et de restitution des connaissances.

Cet ensemble sert de point de départ et de cahier des charges pour le design de l'architecture de modélisation. L'insertion de chaque cas d'étude dans la plate-forme de modélisation a été réalisée en collaboration avec les chercheurs thématiques concernés. Chaque étude de cas a permis d'aborder la représentation des différentes catégories mises en avant dans la typologie de référence, en mettant l'accent sur l'un ou l'autre aspect à des degrés divers selon que l'accent était mis sur l'expérimentation en animalerie, le croisement de génomes, l'invasion à grande échelle ou le comportement fouisseur des rongeurs étudiés (Figure 6).

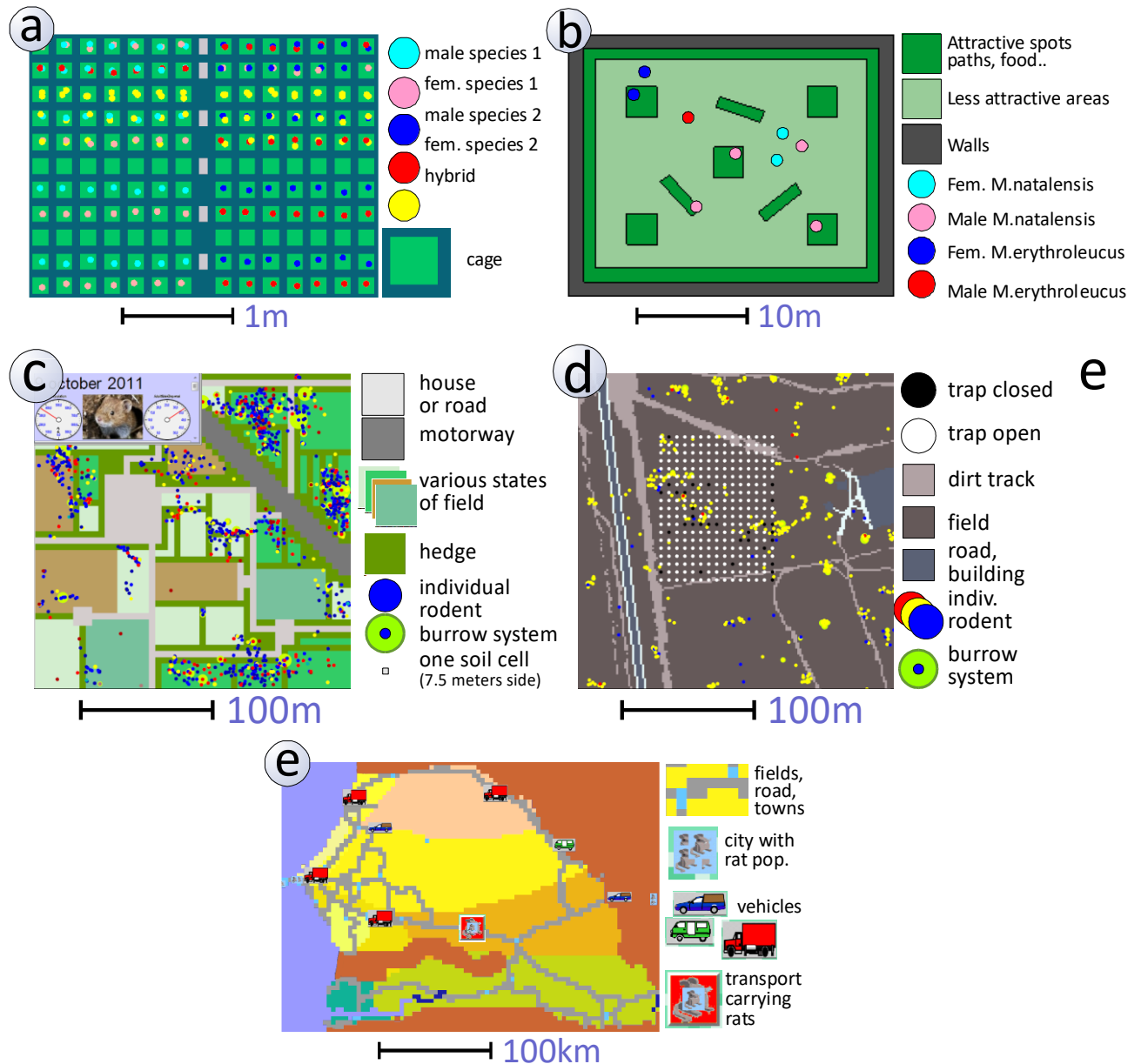
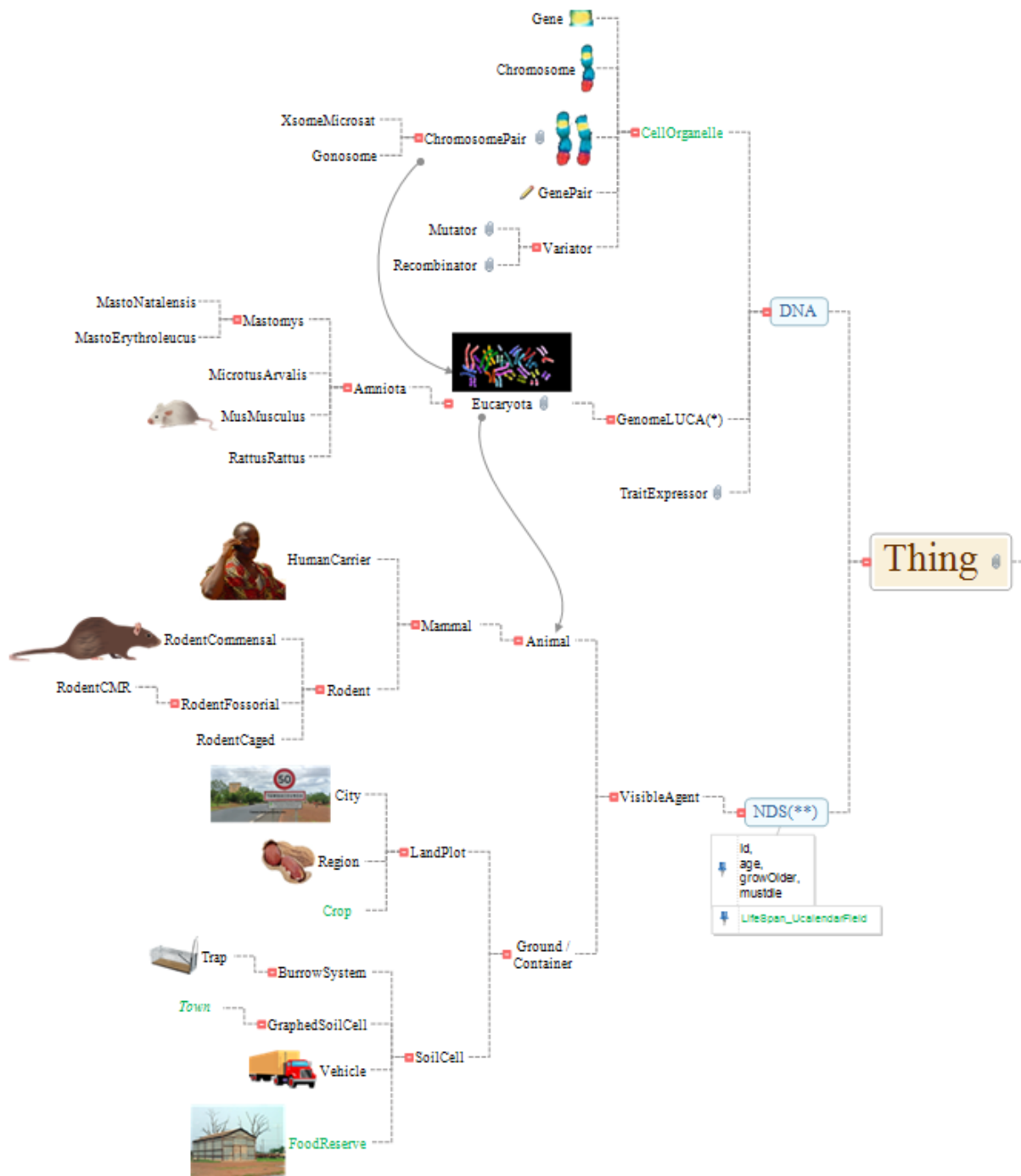


Figure 6 Cinq études de cas intégrées dans le modèle proposé. Expérience d'hybridation d'espèces jumelles (a) en cage et (b) en enclos ; (c) population de campagnols évoluant dans un paysage agricole fragmenté dynamique, (d) expérience de capture-marquage-recapture dans une réserve africaine, (e) colonisation du rat noir au Sénégal lors du siècle écoulé.

Le modèle issu de la recherche est présenté sur la Figure 7. L'arborescence est construite pour être extensible sans remise en cause majeure de l'architecture; il est ainsi envisageable dans ce schéma d'introduire des problématiques sur e.g., les virus, les prédateurs de rongeurs.



(\*) LUCA: Last Universal Common Ancestor, (\*\*) NDS: Nearly decomposable system (Simon, 1962) : terme utilisé comme proxy pour les systèmes complexes organisés au sein d'une hiérarchie, En vert : nœud non explicitement formalisé.

Figure 7 Arborescence objet du modèle telle que définie par les études de cas représentées. Chaque entité correspond à une classe d'objet au sens informatique du terme. Chacune peut être ou non spécifiée selon l'étude de cas sélectionnée. Chacune dispose de fonctionnalités propres (fusion de chromosomes dans le cas d'un génome, creusement de terrier pour un rongeur fouisseur...).

Le module *thing* qui est présenté constitue l'objet de l'application ; il est connecté à trois autres composants d'importance identique qui font aussi l'objet d'une recherche méthodologique sur la pluridisciplinarité

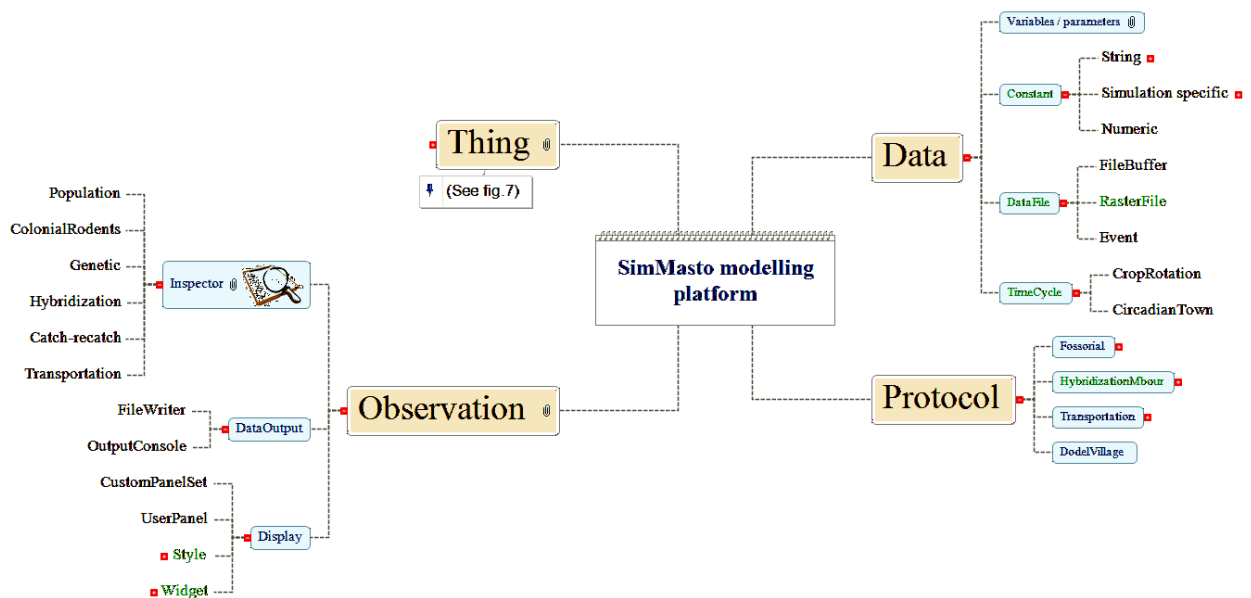


Figure 8 : organigramme des classes associées à la gestion des simulations formalisées dans le package thing, (Figure 7) telles que définies par les besoins des études de cas représentées.

1. Le module *data* qui gère les constantes et les données issues des projets de recherche. Particulièrement, un système minimaliste de chronogramme a été mis au point pour traiter indistinctement d'événements disciplinaires (géographie, biologie, climatologie, histoire, ...).
2. Le module *protocol* qui permet de caractériser et autonomiser chaque étude de cas agrégée dans le modèle (choix des échelles, entités, indicateurs, données d'entrée ; mise en œuvre...)
3. Le module *observation* qui permet de restituer de façon multimodale (affichages, stockage de données, connexion à d'autres logiciels) le fonctionnement du système. Ce module reprend les principes des systèmes épiphytes ou conseillers (Giroud et al., 1995, Pachet, 1997, Richard et Tchounikine, 2004) qui modélise les actions de prise de données (enquêtes, piégeages, comptages...) effectivement réalisées dans les projets. Il est représenté par un ensemble simulé d'observateurs spécialisés chargés de la collecte de données sur le monde modélisé et de l'élaboration des indicateurs de calcul.

Tableau 2 Synthèse des caractéristiques transdisciplinaires issues d'une typologie synthétique (b) et conduisant à la spécification formelle d'un modèle (c). La typologie synthétique est une combinaison de classifications construites à partir de (i) la connaissance produite par les disciplines et articulées au sein d'un système d'information, (ii) les questionnements et domaines de recherche présentés lors d'entretiens avec les scientifiques. Les catégories sont classées par leur rang d'importance décroissant combiné pour l'ensemble des deux typologies d'origine.

a) Exemples de mots-clés	b) Typologie synthétique	c) Exemples de classes, variables ou procédures informatiques
arenavirus, Cricetides, rodents, ...	<b>Élément biologique</b>	<u>gene</u> , <u>chromosome</u> , <u>genomeEucaryote</u> , <u>genomeRattus</u> , <u>genomeHomo</u> , ...
epidemiology, taxonomy, ecophysiology, ...	<b>Discipline</b>	<u>getFixationIndex</u> , <u>getCitiesList</u> , <u>getHybridRate</u> , <u>ploughing</u> , <u>incrPbEpistasy</u> , <u>makeBurrowSystem</u> , <u>makeGametes</u> , <u>updateUserParameters</u> , ...
estivation, regulation, species adaptation, ...	<b>Processus bio-écologique</b>	<u>updatePhysiologicalStatus</u> , <u>growOlder</u> , <u>mate</u> , <u>interact</u> , ....
classification, survey, training, ...	<b>Méthode</b>	<u>addTrapSystem</u> , <u>nextGaussian</u> , <u>openTrap</u> , <u>changeCage</u> , ...
biotope, local scale, equilibrium, ...	<b>Concept</b>	
Habitat, village, area, continent, forest fragment, island, km2, ..	<b>Descripteur d'espace</b>	<u>container</u> , <u>rasterScape</u> , <u>soil cell</u> , <u>crop</u> ,... <u>identifyLandPlots</u> , <u>cellWidth</u> , <u>superDijkstra (graphs &amp; networks)</u> ...
Africa - South America, Sudano-sahelian eco-region, un-localized, ...	<b>Localisation</b>	<u>rasterUrl</u> , <u>makeRasterFromChrono</u> , <u>getMap</u> , <u>geographyName</u> , ...
karyotype,reference list, biological indicator, ...	<b>Outil</b>	<u>trap</u> , <u>tagRodent</u> , ...
1903-2007, 2003, a-temporal, geological time, ...	<b>Descripteur temporel</b>	<u>chronoReader</u> , <u>manageOneEventLine</u> , ...
abundance, parasite load, evolution rate, allele frequency, specific richness, karyotype structure, ...	<b>Mesure ou indicateur</b>	<u>maxFemaleDispersal</u> , <u>computeIndicators</u> , <u>getTraitValue</u> , ...
aridity, agriculture, dry season, ...	<b>Phénomène</b>	<u>GenerateCropTransitionMatrix</u> , <u>agriculturalPractice</u> , ...

## Discussion

Les composantes liées aux différents domaines modélisés ont pu grâce à ce canevas être articulées d'une étude de cas à l'autre. En effet, chaque étude de cas a clairement à la fois enrichi et bénéficié des mécanismes et des composants impliqués dans le simulateur. Ceci a conduit par exemple à l'addition progressive de processus cellulaires, comportements sociaux, structures d'espace, divers types de mouvements ou le partage d'indicateurs dans la plate-forme globale. Les dispositifs d'observation (indicateurs de dynamique) ont pu dans de nombreuses situations être réutilisés d'une étude de cas à l'autre. Des éléments de protocoles ont pu être aussi partagés. Par contre, les échelles spatiales et temporelles distinctes ont dû être, à ce stade, considérées de façon séparée.

L'approche a permis d'individualiser des interfaces déterminantes pour la formalisation interdisciplinaire des thématiques. Par exemple, une des clés pour la définition d'une interface entre les disciplines est la prise en compte de la perception des objets concrets appréhendés (Wheeler, 2007). En Science, la perception du monde dépasse les aspects sensoriels (vue, odorat, ...) et inclut la panoplie d'outils et de collecteurs utilisés pour acquérir des données sur les systèmes appréhendés. La catégorie "outil" (e.g., marqueur génétique, piège, base bibliographique, ...) rassemble ces modalités dans les ontologies. Dans la plate-forme, elle est formellement individualisée ; soit globalement par le système d'observateurs spécialisés simulés (système épiphyte) qui récupère différents types de listes et d'indicateurs, soit spécifiquement dans un paquetage "objet" pour formaliser des objets fonctionnels concrets tels que des pièges.

L'approche formelle ne couvre cependant pas l'ensemble du processus. Tout d'abord, la catégorie rassemblant les concepts (Tableau 2) n'est pas représentée dans le modèle. Cette catégorie rassemble des thèmes (*e.g.*, *biotope*, *candidate gene*, *equilibrium*, *habitat*, *public health*, *quality*, *resistance*, *risk*, ...) qui se situent dans le domaine de l'interprétation, c'est à dire en amont ou en aval du processus de recherche formalisé.

D'autre part, le développement de la plate-forme a conduit à générer une nouvelle interface clé pour articuler les études de cas. Celle-ci rassemble, dans un paquetage informatique spécifique intitulé 'protocole', la définition des lieux et des populations mises en jeu, la façon de les aborder ou l'échelle à considérer. Bien qu'il s'agisse d'une structure déterminante de la plate-forme interdisciplinaire, aucune des approches typologiques présentées n'a révélé, d'une façon ou d'une autre, la notion de protocole. Ce point souligne une faiblesse de la démarche ontologique. Cette dernière en effet, bien qu'orientée (*i*) questionnement, (*ii*) domaines et (*iii*) savoirs produits, n'a pu traiter cet aspect important du travail scientifique.

L'identification des catégories pour l'élaboration des ontologies peut enfin apparaître globalement comme un biais important. Dans le premier cas (enquête auprès des chercheurs), elle a été déterminée par l'auteur, dans le deuxième (centre d'informations), collégialement lors d'une réunion ad hoc lors de laquelle l'appartenance de certains termes a été (longuement) discutée. Malgré cette subjectivité incontournable, dans l'ensemble, le consensus était acquis sur l'identification de la plupart des catégories. Un travail similaire sur des connaissances liées à une population de rongeurs africains a de même été réalisé par un groupe d'étudiants sénégalais (Diakhate et al., 2015). Dans ce domaine qui est proche de celui ici étudié, la typologie résultante a conduit à des résultats différents dans le détail mais avec des constantes (temps, espace, outil, technique, discipline, taxon, ...) que l'on retrouve dans les deux études. Dans le cadre du présent travail, la fusion des typologies permet de gommer ces différences de détail et de conserver les grandes lignes du savoir scientifique et de son acquisition.

## Conclusion

L'approche proposée doit encore être considérée à un stade de mise en place compte tenu de plusieurs problèmes non résolus, dont notamment les questions d'échelles temporelles, les composantes non prises en compte tels que les protocoles dans les typologies ou les concepts dans la plate-forme formelle. Elle permet cependant de commencer à discerner quels sont les composantes partagées, partageables ou inconciliables et, peut-être, les modalités d'une intégration mutuellement bénéfique de points de vue disciplinaires distincts.

Newell (2001) caractérise l'usage disciplinaire en relation avec la nature des systèmes complexes : interdisciplinarité (*viz.* synthèse) lorsque les facettes du système sont en cohérence, multidisciplinarité (*viz.* intégration) si non (et disciplinarité lorsque que le système ne comporte pas de facettes). L'approche retenue ici visait dans un premier temps à combiner les approches disciplinaires au sein d'un seul formalisme (pluridisciplinarité). Les résultats obtenus à partir de l'accrétion des études de cas montrent que chaque approche nourrit les autres pour aboutir à un système intégré. L'enjeu est cependant de parvenir à ce que l'articulation puisse révéler de nouveaux processus ou dynamiques issus de l'articulation disciplinaire (McMurtry, 2009) via par exemple la mise en cohérence des échelles spatio-temporelles variées. Les résultats présentés ne permettent pas encore de constater cette synergie.



## Remerciements

Je remercie C. Brouat, N. Charbonnel, Y. Chaval, J. F. Cosson, A. Dalecky, G. Dobigny, J. M. Duplantier, P. Gauthier, L. Granjon, J. Michaux, A. Ndiaye, M. Pagès, A. Poulet, J.P. Quéré, N. Sarr, B. Sicard and C. Tollenaere, membres de l'équipe scientifique qui ont aimablement collaboré à l'alimentation du 'Centre d'Informations (CI)', participé aux entretiens étendus et réalisé les études de cas thématiques; S. Piry, L. Thibaut, M. Ito, M. Fofana, C. Ba, Q. Baduel, J.E. Longueville, A. Realini, A. Comte, P.A. Mboup, M. Ndiaye, S. Causse, A. Pagès pour leur contribution au développement du 'centre d'informations' (CI) et de la plate-forme de simulation SimMasto (<http://simmasto.org>).

## Références

- Burger, P. et R. Kamber (2003) Cognitive integration in transdisciplinary science: knowledge as a key notion. *Issues in integrative studies*, 21:43-73.
- Courtial, J.P. (1990) Introduction à la scientométrie. De la bibliométrie à la veille technologique, *Anthropos-Economica*, Paris, 126p.
- DeAngelis, D. & W.M. Mooij (2005) Individual based modeling of ecological and evolutionary processes. *Annu. Rev. Ecol. Evol. Syst.*, 36:147-68
- Ferber, J. (1995) Les systèmes multi-agents : vers une intelligence collective. InterEditions, 1995, 522p.
- Giroux S., Pachet F., Paquette G. et Girard J. (1995) Des Systèmes Conseillers Epiphytes, *Revue d'Intelligence Artificielle*, 9(2):165-190,1995.
- Grimm, V. et S.F. Railsback (2005) *Individual-based Modeling and Ecology*. Princeton Univ. Press, 448p.
- Hursh, B., Haas, P. et M. Moore (1990) An Interdisciplinary Model to Implement General Education. *Issues in Integrative studies*, 8 :133-150.
- Kauffman S.A. (1993) *The origins of Order: Self-Organisation and Selection in Evolution*. Oxford University Press.
- Kelly, K. (1994) *Out of Control: The New Biology of Machines, Social Systems and the Economic World*. Addison Wesley 1994.
- Klein, J.T. (2001) Interdisciplinarity and the prospect of complexity: the tests of theory. *Issues in integrative studies*, 19:43-57.
- Le Fur, J. and P. Simon (2009) A new hypothesis concerning the nature of small pelagic fish clusters. An individual-based modelling study of *Sardinella aurita* dynamics off West Africa. *Ecological Modelling*, 220(9-10), 17 May 2009, Pages 1291-1304
- Le Fur, J. (2013) Extending life concepts to complex systems. *Interdisciplinary Description of Complex Systems*, 11(1), 37-50.
- Le Fur J. (2013) A formal computer framework for linking multidisciplinary multiscale knowledge. A case study on rodent population dynamics and management – Proc. Eur. Conf. Complex Syst. Lisbonne, sept 2013, 5p.
- Le Fur, J. et D. Hervé (2009) Diversité des modélisations à l'interface entre natures et sociétés. Diversité des objectifs, des disciplines, des objets et des outils. In: Hervé, D. et F.Laloë (sci. eds) *Modélisation de l'environnement : entre natures et sociétés*. Quae ed., coll. Indisciplines, mai 2009, chap.I, 35-50
- Madin, J.S., Bowers, S., Schildhauer, M.P. et M.B. Jones (2008) Advancing ecological research with ontologies. *Trends in Ecology and Evolution*, 23(3) :159-168.
- Mathieu, J. (1991) Les médiations du passé: à la recherche d'un carrefour. In Mathieu, J. (ed.) *Les dynamismes de la recherche au Québec*. Les Presses de l'Université Laval: 45-62
- McMurtry, A. (2009) Knowers and phenomena : two different approaches to interdisciplinarity and interprofessionalism. *Issues in integrative studies*, 27 :1-16.



- Meek, J. (2001) The practice of interdisciplinarity : complex conditions and the potential of interdisciplinary theory. *Issues in Integrative studies*, 19 :123-136
- Newell, W.H. (2001) A theory of interdisciplinary studies. *Issues in integrative studies*, 19 :1-25.
- Nisbet, R.M., Muller, E.B., Lika, K and S.A.L. Kooijman (2000) From molecules to ecosystems through dynamic energy budget models. *J. Animal Ecol.*, 69: 913-926.
- Pueyo, S. (2007) The maximum entropy formalism and the idiosyncratic theory of biodiversity. *Ecology Letters* 10: 1017–1028
- Rouxel, A. (2002) Identité disciplinaire : cohérence ou éclatement ? *Tréma n°19, Littérature enseignée et reconfigurations*, Montpellier: 29-36.
- Simon, H.A. (1962) The Architecture of Complexity. *Proceedings of the American Philosophical Society*, Vol. 106, No. 6. (Dec. 12, 1962), pp.467-482
- Storch, D., Marquet, P.A. and J.H. Brown (2007) Scaling diversity – what is the problem ? In: Storch, D., Marquet, P.A. and J.H. Brown (eds.) *Scaling biodiversity*. Cambridge Univ. Press. 2007:1-10.
- Strand E, Huse G, Giske J. (2002) Artificial evolution of life history and behavior. *The American Naturalist* 159:624-644.
- Warfield, J.N. et R.Warfield (1999) A role for formalisms in integrative studies. *Issues in integrative studies*, 17 :21-54.
- Wheeler, T.J. (2007) Analysis, Modeling, Emergence and Integration in Complex Systems: a Modeling and Integration Framework and System Biology. *Complexity*, 13(1), 60:75.
- Pachet, F. (1997) Représentation de connaissances et langages à objets. Habilitation à diriger des recherches, Université Paris 6, 1997. Collation, 64 p.
- Richard B., Tchounikine P. (2004) Enhancing the Adaptivity of an Existing Website with an Epiphyte Recommender System, In: *The New Review of Hypermedia and Multimedia*, Vol. 10 n°1, p. 31- 52.