

Habilitation à Diriger des Recherches : Modélisation des écosystèmes de connaissances pour une écologie de l'information et de la communication

Samuel Szoniecky

2023-09-14

Table of contents

Mémoire de synthèse en vue de l'Habilitation à diriger des recherches

Université Vincennes - Saint-Denis Paris 8

Samuel Szoniecky

Modélisation des écosystèmes de connaissances pour une écologie de l'information et de la communication

Modeling knowledge ecosystems for an ecological management of information and communication

Sous le parrainage du professeur XXX

Présenté le XXX devant un jury composé de :

XXX, examinateur

XXX, examinateur,

XXX, examinateur,

XXX, examinateur,

XXX, examinateur,

XXX, garant & rapporteur,

Summary

In summary, this book has no content whatsoever.

Introduction générale

« La vision magique du monde
est celle d'un univers de puissances
qui, virtuelles,
ne tendent qu'à devenir actuelles. »

(Thom (1975) p. 362)

« Cette existence d'une polarité perceptive joue un rôle prépondérant dans la ségrégation des unités perceptives [...]. Le sujet perçoit de manière à s'orienter par rapport au monde [...]. Percevoir c'est [...] organiser, maintenir ou inventer une organisation. Il ne suffit pas de dire que la perception consiste à saisir des touts organisés ; en fait, elle est l'acte qui organise des touts ; elle introduit l'organisation en reliant analogiquement les formes contenues dans le sujet aux signaux reçus : percevoir est retenir la plus grande quantité de signaux possible dans les formes les plus ancrées dans le sujet ; ce n'est pas seulement saisir des formes ou enregistrer des données multiples juxtaposées ou successives Percevoir consiste à éléver l'information du système formé par le sujet et le champ dans lequel il s'oriente. [...] C'est en s'orientant dans (une) situation que le sujet peut ramener à l'unité les aspects de l'hétérogénéité qualitative et intensive, opérer la synthèse du divers homogène. (Simondon Gilbert (1989), L'individuation psychique et collective, Paris, Aubier., p. 89-90) » cité par (Quéré, 2020, paragr. 66)

Part I

**Cartographie d'un milieu de
connaissances**

« tous les systèmes de modélisation se valent,
tous sont acceptables,
mais uniquement dans la mesure
où leurs principes d'intelligibilité
renoncent à toute prétention universaliste »
(Guattari 1989, 10)

La démarche auto-réflexive de ce travail commence par une description de l'état actuel de mes recherches. Elle a pour ambition un dévoilement sincère des limites de ma pensée sous la forme d'une cartographie qui trace les frontières numériques (Saleh, Szombieky, and Ghenima 2023) de ce qui est important pour moi aujourd'hui, c'est-à-dire la valeur des rapports que j'entretiens entre ma subjectivité (Guattari 1989) et mon objectivité (Badiou 2018) et qui forme mon milieu de connaissances (Berque 2009a).

Cartographier c'est à la fois explorer, formaliser et guider. Tout commence par la découverte d'un lieu, par l'expérience d'un espace dont on va ensuite décrire les impressions qu'il produit suivant des règles convenues pour que d'autres puissent les retrouver après soi-même.

Les espaces que nous décrivons sont des milieux, c'est à dire un tissu relationnel au sein duquel les connaissances existent, et sans lequel elles n'existeraient pas (Berque 2009a, 146). Ce sont des espaces vivants qui évoluent sans cesse et nous transforment à chaque interaction que nous entretenons avec eux. Suivant le principe d'enaction (Maturana and Varela 1994), le couplage entre l'espace et l'auteur de sa description transforme à la fois l'espace et l'auteur. Cartographier c'est se connaître soi-même en modélisant le milieu de ses expériences et en réfléchissant aux transformations que cela produit.

Peut-on modéliser toutes ses expériences pour les rendre accessibles ? Sont-elles toutes modélisables ? Accessibles ? Même si les cartes utilisent des systèmes de coordonnées calculables et un vocabulaire formel abouti qui les rendent manipulables par des machines et donc potentiellement modélisables, il n'en reste pas moins que la cartographie procède d'une multitude de choix qui sélectionneront dans l'espace à modéliser ce qui sera effectivement dans la carte. Pour transmettre une expérience, il faut la réduire à un ensemble de signes et par la même la soumettre aux contraintes de ce qui les caractérisent : variabilité, associativité, équivocité (G. Deleuze 1981). La transformation de l'expérience en signe est fondamentalement incomplète. L'exhaustivité n'est qu'un leurre. Toutefois, la carte possède une dimension diagrammatique qui donne au signe une potentialité d'action à la manière d'une partition musicale (Stransky and Szombieky 2014). La carte guide vers une expérience sans décrire toutes les connaissances auxquelles l'expérimentation donnera accès. Avec une carte, je sais où trouver une rivière poissonneuse mais pas si je vais effectivement pêcher des poissons et lesquels.

Nous vous invitons à vous laisser guider par les cartographies que nous présentons en espérant que la pêche aux connaissances sera fructueuse pour vous, elle l'est déjà pour moi. Les cartographies que nous vous proposons décrivent mes connaissances en rapport avec mon métier d'enseignant chercheur et les pratiques informationnelles qu'il génère (Thiault and Malingre

2022). Nous suivons un processus d'exploration qui commence par préciser les principes cartographiques Chapter ?? utilisés. Le second chapitre de cette partie présente l'état actuel de mes positions scientifiques dans les sciences de l'information et de la communication et plus globalement dans les sciences humaines Chapter ???. Nous verrons ensuite dans un troisième chapitre quelles sont mes visées scientifiques Chapter ?? par rapport à quatre échelles d'exploration : locale Section ??, globale Section ??, sociale Section ?? et conceptuelle Section ???. Nous développerons ensuite ces perspectives scientifiques en suivant cinq axes de recherches :

- comprendre l'internet des objets Section ??,
- concevoir une générativité générique Section ??,
- formaliser le consensus par une éthique de la discussion Section ??,
- designer les connaissances Section ??,
- développer l'intelligence collective Section ??.

Dans un quatrième chapitre, je présenterais les étapes de mon parcours intellectuel Chapter ?? à travers mes frayages intellectuelles (Citton 2010) qui débutent avec l'histoire de l'art puis s'ancre dans les sciences de l'information et de la communication pour développer un projet de jardinage des connaissances qui mène à une théorisation des écosystèmes de connaissances et finalement à la mise en pratique d'une méthode de modélisation et d'analyse de l'information et de la communication. Le dernier chapitre de cette partie sera consacré à la présentation des résultats Chapter ?? de ce parcours en matière :

- d'utilisation de l'analogie écosystémique Section ??,
- de modélisation d'une ontologie-éthique Section ??,
- de calcul de la complexité info-communicationnelle Section ??,
- de conception diagrammatique et de cible Section ??
- d'interaction avec des cartographies sémantiques Section ??,
- d'outillage pour les humanités numériques Section ??.

En explorant le milieu de connaissances que je vous propose, vous serez peut-être tenté d'y laisser vos propres traces. Si c'est le cas, vous pourrez le faire en utilisant un outil d'expression d'un point de vue comme « CartoAffect » que j'ai développé pour cartographier mon milieu de connaissances. Il suffit de vous rendre à cette page Web et/ou de télécharger l'application mobile. Une fois que vous vous serez identifié avec le compte qui vous a été fourni, vous pourrez participer à la cartographie d'un milieu de connaissances. Bonne exploration.

1 Principes de cartographie des connaissances

« Il faut demander aux algorithmes
de nous montrer
et la route,
et le paysage »
(Cardon 2015, sec. 5)

« Le numérique est donc
à la fois ce qui est
autour de nous,
entre nous,
en nous »
(B. Bachimont 2020)

Les principes de cartographie des connaissances que nous utilisons s'appuient sur des idées inspirées de nombreuses références Chapter ?? à partir desquels nous concevons un design des connaissances et définissons des propriétés graphiques. Il va de soi, que d'autres références définiraient d'autres designs et d'autres propriétés graphiques. Il serait particulièrement intéressant d'analyser en quoi tel ou tel auteur utilise dans ses œuvres des images mentales spécifiques qui en font un plasticien de la pensée. Il nous semble évident que les idées de Descartes et Spinoza n'ont pas la même plasticité, de même celles d'Heidegger et Bachelard ou de Wittgenstein et Deleuze. L'absolu plasticité du numérique permettrait sans doute de montrer ces idées suivant des modes de représentation propres à chacun de ces auteurs afin de les comparer ou des les utiliser pour un design de connaissance spécifique. Nous avons fait ce travail dans les lignes suivantes pour construire notre design des connaissances à partir d'idées dont la plasticité nous semble particulièrement fertile.

Le premier principe sur laquelle nous nous appuyons est celui que les connaissances se produisent suivant un cycle continu d'expériences dans le monde physique et dans le monde de la pensée. Entre les « physicalités » et les « intérieurités » (Philippe. Descola 2005) , les « cycles de sémiose » (µ, Edeline, and Klinkenberg 2015a) canalisent nos perceptions par « anasémiose » pour discerner des informations que nous communiquons par « catasémiose » en donnant forme à l'agir : parler, lire, écrire, gesticuler, ne rien faire...

Le second principe se base sur les travaux de (Hofstadter and Sander 2013) pour qui l'analogie est le « moteur » qui relie le discernement et l'action en gardant le souvenir de cette relation qui devient à force de répétition, une manière d'être en prenant chez Deleuze la forme d'un « pli » qui est notre troisième principe :

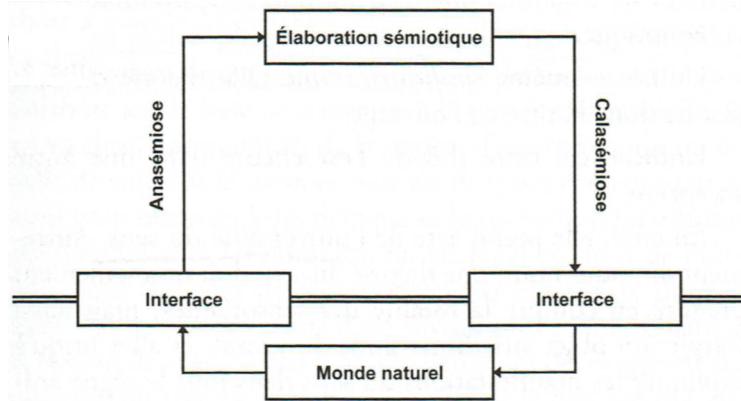


Figure 1. Le cycle de la sémiotique

Figure 1.1: Cycle de la sémiotique (μ, Edeline, and Klinkenberg 2015a, 10)

« L'opération de la perception constitue les plis dans l'âme, les plis dont la monade est tapissée du dedans ; mais ceux-ci ressemblent à une matière, qui doit dès lors s'organiser en replis extérieurs. » (G. Deleuze 1988, 131)

A la manière de deux miroirs qui plient la lumière en se reflétant l'un dans l'autre à l'infini, discernement et action se réfléchissent en pliant les flux d'information. Chaque pli décompose l'information en signes dont les signifiés plongent vers l'intériorité en stimulant l'intuition et dont les signifiants émergent vers des physicalités en stimulant l'expression.

Entre discerner et agir, intuition et expression, c'est dans ce « milieu » qu'Augustin Berque décrit une « pulsation existentielle » mue par la « raison trajective » que nous prenons comme quatrième principe :

« la raison trajective, elle est en effet dans la pulsation existentielle qui, par la technique, déploie notre corps en monde sur la terre, et qui simultanément, par le symbole, reploie le monde en notre chair » (Berque 2009a, 402)

Cette raison pilote la réflexion en modifiant l'inclinaison du pli vers le discernement de signifiés ou vers l'expression de signifiants. Elle procède de processus que nous contrôlons consciemment et d'autres plus imprévisibles et incontrôlables qui se produisent en fonction d'une multitudes de pliages et de leurs capacités à ce faire, ce défaire, ce bloquer suivant un cinquième principe celui du degrés de flexibilité (Clément 2021).

Les cycles de sémioses, les analogies, les plis, les pulsations existentielles, les degrés de flexibilité structurent et produisent nos connaissances tout au long de nos vies en développant trois pouvoirs fondamentaux : discerner, raisonner, agir. Notre hypothèse principale est qu'il est possible de cartographier ces connaissances en représentant les pliages et leurs dynamismes dans trois directions : vers l'intériorité (discerner), en boucles récursives (raisonner)

et vers l'extérieur (agir). À chaque pulsation existentielle, à chaque événement de nos vies, à chaque pli, ces pouvoirs augmentent ou diminuent accentuant ainsi des rapports privilégiés, d'autres, plus secrets, et même certains qui nous restent inconnus. De ce point de vue, nous n'adoptons pas la conjecture au cœur des recherches en IA depuis 1955, formulée à l'occasion de séminaire d'été à Dartmouth par John McCarthy, Marvin Minsky, Claude Shannon et Nathaniel Rochester (Leveau-Vallier 2023, 39) qui prend comme principe que tous les aspects de l'intelligence peuvent être décrits avec une précision telle qu'une machine peut les simuler.

Ainsi, la pulsation varie continuellement, elle est parfois instantanée par exemple quand on rit, elle peut aussi prendre beaucoup de temps quand un souvenir longtemps oublié émerge petit à petit ; elle devient un métier quand à force de pratiquer un geste particulier, celui-ci s'automatise. Ces pulsations se transforment parfois en bêtises ou en inconscience quand le pouvoir d'agir prend le pas sur les pouvoirs de discerner et de choisir en occultant leurs pliages potentiels. Suivant leurs fréquences, les pulsations existentielles forment des ondes dont la vitesse de propagation est fonction de leur longueur (distance séparant deux maxima consécutifs de l'amplitude entre physicalités et intériorité) et du milieu dans lequel elles se déplient. La catégorisation et l'analyse de ces ondes renvoient globalement à une réflexion sur la modélisation de l'esprit qui dépasse le cadre de ce propos mais que nous illustration par le diagramme suivant :

Les chapitres suivants explicitent comment à partir de ces cinq principes et de cette hypothèse principale, nous cartographions dans le Web Section ?? des connaissances qui se développent dans l'espace et le temps Section ?? suivant les pulsations existentielles d'un actant Section ?? entre des espaces matériels Section ?? et conceptuels Section ???. Ces propositions sont le résultat d'un travail de recherche d'une dizaine d'années que nous présentons en détails dans **?@sec-part-modeliserConnaissances**.

1.1 Cartographier dans un environnement Web

Un environnement Web se base avant tout sur une architecture Client / Serveur qui utilise le protocole HTTP pour organiser les échanges de données entre des machines et des utilisateurs via un navigateur (Balpe, Saleh, and Lelu 1996) cf. ci-dessous¹.

Nous ne ferons pas ici une analyse des technologiques de représentation des données (Andry, Kieffer, and Lambotte 2022; Fekete and Boy 2015) préférant nous focaliser sur les outils et les méthodes que nous utilisons dans le cadre de ce travail pour cartographier nos connaissances. Nous ne détaillerons pas non plus tous les éléments qui composent notre environnement² mais

¹

cf. https://samszo.univ-paris8.fr/conf_errance/cours_systeme-information-programmation-internet/slide.html?diapo=5

²

disponibles sur notre forge logicielle : <https://github.com/samszo>

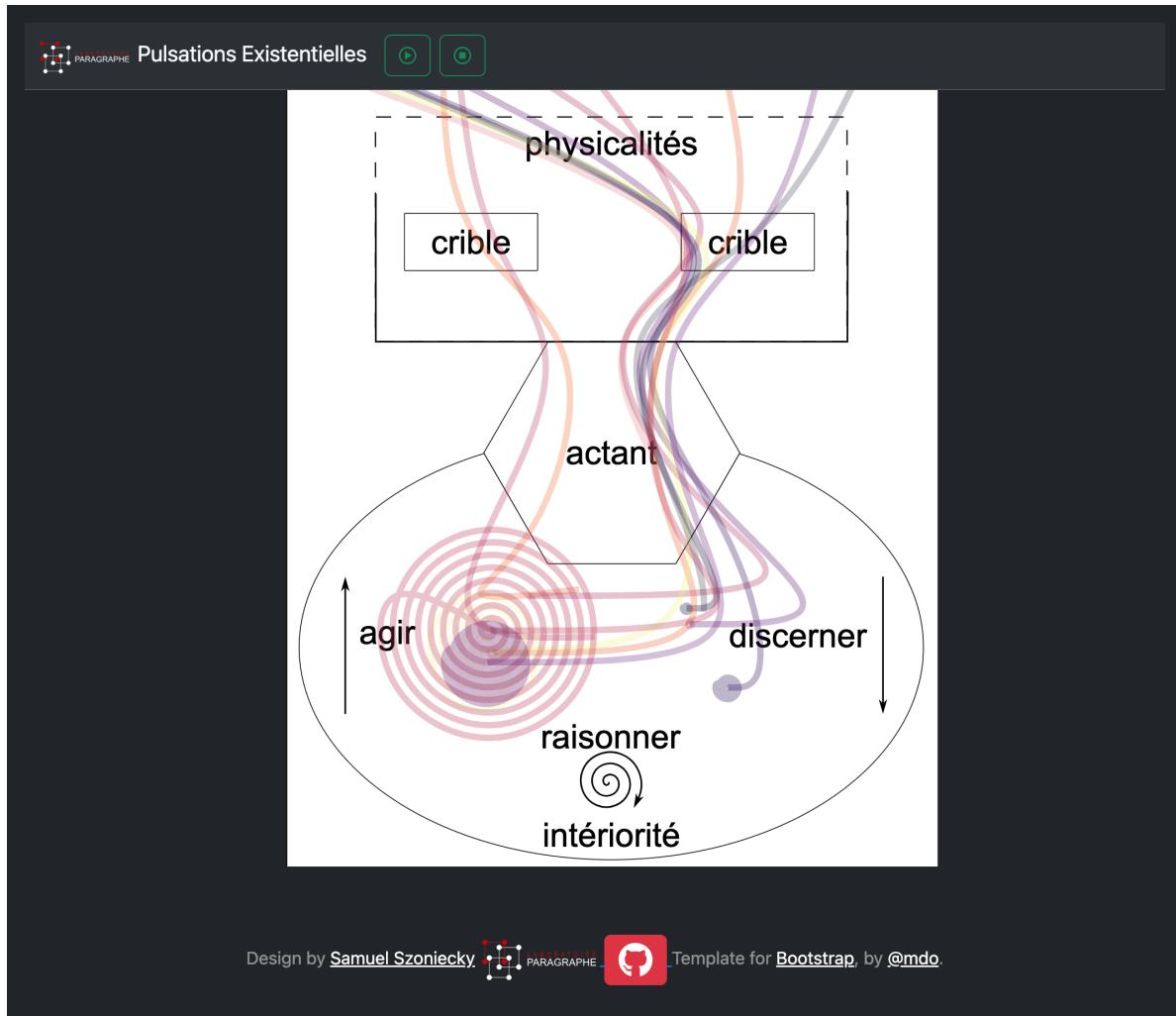


Figure 1.2: Dynamiques pulsations existentielles

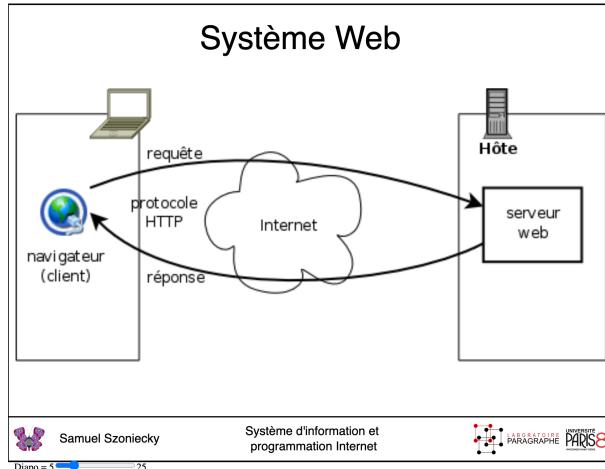


Figure 1.3: Architecture Client / Serveur

uniquement les plus pertinents pour comprendre les principes de cartographie que nous avons mis en place dans ce travail pour gérer les données du coté serveur et naviguer dans leurs représentations du coté client.

1.1.1 Gérer les données sur les serveurs

Les serveurs sont des machines qui fournissent des ressources via une requête spécifique sur une adresse unique dans un environnement Web. Le protocole HTTP définit les conditions d'adressage de ces requêtes et les éventuels paramètres qui lui sont associés. Il existe une multitude de solutions pour gérer les données à partir de ce protocole et des langages informatiques associés comme PHP, Python, Java... Pour nos travaux de recherche, nous avions fait le choix de développer sur nos serveurs, une boîte à outils basée sur PHP et une base de données spécifique (Szonięcky and Bouhaï 2017, 141). Pour des questions de maintenance de l'environnement, de facilités de développement et de diffusion des données de la recherche, nous avons abandonné cette solution pour utiliser depuis quelques années l'environnement Web proposé par le CMS Omeka S³. Cette solution de gestion des archives numériques offrent les fonctionnalités nécessaires pour modéliser une base de données spécifique respectant les principes du Linked Open Data⁴ et les moyens de manipuler ces données avec des vocabulaires, des modèles de ressource, des modules et des thèmes spécifiques. Une fois maîtrisé les

³

<https://omeka.org/s/>

⁴

https://www.w3.org/egov/wiki/Linked_Open_Data

éléments de cet environnement, les données produites par les recherches deviennent accessibles, manipulables et interopérables⁵.

« En utilisant aujourd’hui un tableur ou une base de données ad hoc pour stocker les données, non seulement on se prive de toute la richesse sémantique des LOD et de leur potentiel de traitement, mais encore on risque de ne pas pouvoir réutiliser l’information collectée. La communauté de recherche va ainsi continuer à parcourir mille fois le premier kilomètre, alors qu’une démarche collaborative de collecte de l’information, soutenue par des plateformes de recherche fondées sur les technologies sémantiques, permet de parcourir ensemble des milliers de kilomètres et de disposer, en très peu de temps et en faisant levier sur une curation collective des données, d’un graphe d’information de grande complexité, qualité et richesse. » (Beretta 2023, sec. 15)

Pour chaque projets de recherche et d’enseignement qui nécessitent de manipuler des données, nous avons développé des environnements Omeka S avec le cas échéant des modules et des thèmes spécifiques Section ???. Plus particulièrement, pour ce travail d’HDR, nous avons rassemblé dans un environnement Omeka S les informations concernant notre curriculum vitae et la veille informationnelle que nous menons depuis plus de quinze ans Section ???.

Pour ce faire, nous avons créé :

- 2 vocabulaires spécifiques :
 - Jardin des connaissances : nous utilisons ce vocabulaire pour gérer la modélisation des existences informationnelles dans un écosystème de connaissances⁶
 - Formation Université Paris 8 : ce vocabulaire permet de modéliser l’architecture des enseignements dans l’enseignement supérieur⁷
- 30 modèles de ressource⁸ pour décrire les objets de recherche par exemple :
 - Évènement CV : utilisé pour décrire les événements d’un curriculum vitae

⁵

Lien vers l’API Omeka S des données de ce travail : <https://samszo.univ-paris8.fr/omk/api>

⁶

le vocabulaire au format rdf turtle : <https://raw.githubusercontent.com/samszo/Omeka-S-module-JDC/master/data/vocabularies/jdc.ttl>

⁷

le vocabulaire au format rdf turtle : <https://raw.githubusercontent.com/samszo/Omeka-S-module-JDC/master/data/vocabularies/fup8.ttl>

⁸

la liste exhaustive est ici : https://samszo.univ-paris8.fr/omk/api/resource_templates

- JDC Actant : utilisé pour décrire un actant dans un écosystème de connaissances
- 4 modules spécifiques pour une gestion spécifique des données dans Omeka S:
 - Diigo Import : ce module permet d'importer les signets enregistrés dans une base de données Diigo y compris les copies d'écrans⁹.
 - Zotero Import Plus : ce module¹⁰ basé augmente le module Zotero Import pour importer les notes prises dans Zotero ainsi que les documents associés aux références bibliographiques.
 - JDC : ce module¹¹ fourni les interfaces nécessaires pour modéliser un écosystème de connaissances suivant une ontologie éthique Section ??.
 - CartoAffect : ce module¹² permet de gérer les données pour la modélisation et la présentation des affects en relation avec un écosystème de connaissances.

Les données de cette environnement Omeka S ont été importées dans la base de données avec les modules d'importation Diigo Import, Zotero Import et Bulk Import¹³. Ce dernier module est très pratique pour importer des données à partir de tableurs. Par exemple, nous avons importer les données de notre CV à partir de tableurs¹⁴ ou automatiquement par exemple avec l'outil que nous avons développée pour extraire les informations des dépôts GitHub¹⁵ d'un compte :

Les données de l'écosystème de connaissance que nous avons développé pour ce travail représente une base de données SQL de 75 tables peuplées par plus de 2 000 000 de lignes.

⁹

Dépôt du projet : <https://github.com/samszo/Omeka-S-module-DiigoImport>

¹⁰

Dépôt du projet : <https://github.com/samszo/ZoteroImportPlus>

¹¹

Dépôt du projet : <https://github.com/samszo/Omeka-S-module-JDC>

¹²

Dépôt du projet : <https://github.com/samszo/Omeka-S-module-CartoAffect>

¹³

Dépôt du projet : <https://gitlab.com/Daniel-KM/Omeka-S-module-BulkImport>

¹⁴

Tableur d'importation des données du CV dans Omeka s : <https://docs.google.com/spreadsheets/d/1ap0Q6l8bK9Y8wB21xfCAk9qeJ-Gfa6XAjO-cVk2LxdI/edit?usp=sharing>

¹⁵

Tableur d'importation des données GitHub dans Omeka s : <https://docs.google.com/spreadsheets/d/1W92NdhcMurFO96Et82MOYL43fvFrdOB6nLGF-QeCrkY/edit?usp=sharing>

PARAGRAPHE GitHub Stats [Choisir un export](#) User : API SamSzo [GitHub Stats](#)

	schema:programmingLanguage	schema:ContentSize	schema:size	schema:upvoteCount	
69	HTML JavaScript CSS	147939 11987 807	471	0	
70	?	0	37157	1	
71	?	0	29	0	
72	?	0	4825	4	
73	?	0	5395	0	
74	?	0	689	0	
75		0	132	0	
76		0	196	0	
77	?	0	132	0	
78		0	132	0	
79	?	0	432	0	
80	?	0	260	0	
81	?	0	18285	0	
82	?	0	12322	1	
83	?	0	37560	0	
84	?	0	12874	1	
Affect	?	0	726	0	

Design by [Samuel Szaniecki](#)  PARAGRAPHE  Template for [Bootstrap](#) by [@mdo](#)

Figure 1.4: Outil pour l'extraction des données d'un compte GitHub

Le graphique ci-dessous présente la répartition des objets disponibles dans cet écosystème suivant leur classe¹⁶ :

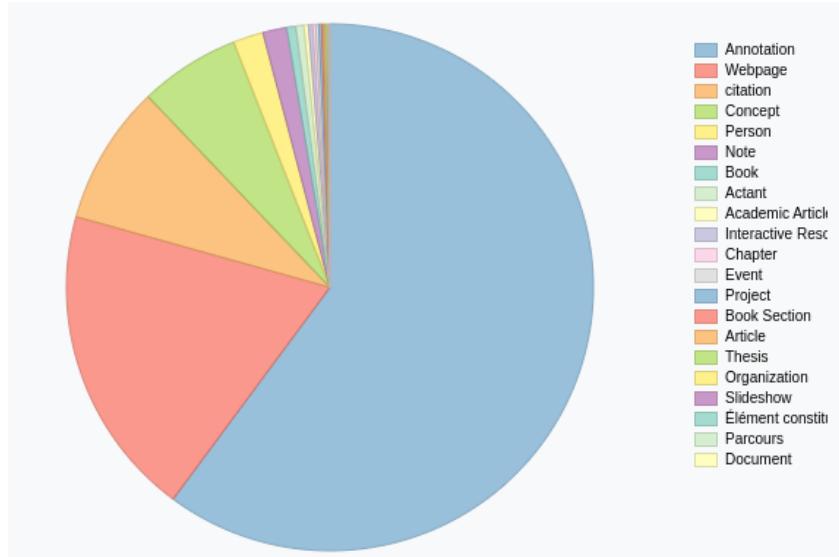


Figure 1.5: Répartition des classes par nombre d'objet dans l'écosystème

Le graphique montre que les deux tiers des objets dans l'écosystème sont des annotations ($61\ 120 = 60\ %$) qui créent un rapport entre un document, un actant et un concept. Nous retrouvons ici les 4 dimensions du modèle que nous utilisons pour modéliser les connaissances (Section ??) : document, actant, concept, rapport. Plus précisément, la dimension physique (documentaire) est composée essentiellement de pages Web ($19\ 491$ items = $19\ %$ ¹⁷), des citations ($8\ 994 = 9\ %$), de médias ($3\ 427 = 3\ %$), des notes ($1\ 488 = 1\ %$) et des livres ($568 = 1\ %$) issues de notre processus de veille. Les autres dimensions de l'écosystème sont les concepts ($6\ 266 = 6\ %$) et les personnes ($500 = 0,5\ %$) associées aux actants ($500 = 0,5\ %$). Le graphique ci-dessous montre cette répartition des objets suivant les dimensions existentielle :

Cette représentation suivant la classe des objets sous estime la complexité de l'écosystème puisqu'elle ne prend pas en compte le détails des valeurs (dimension physique) de chaque propriété (dimension concept) ni l'actant qui exprime les rapports entre propriétés et valeurs encore moins l'évolution de cette complexité au fur et à mesure que l'écosystème se transforme. Pour une meilleure compréhension de l'écosystème, nous considérons chaque donnée comme une existence particulière qui possède sa propre complexité qui s'ajoute à la complexité de

¹⁶

Les données de ce graphique et d'autres statistiques sont disponibles ici : <[omkStats.html](#)>

¹⁷

Liste complète des pages Web : <https://bit.ly/3Qj1NRm>

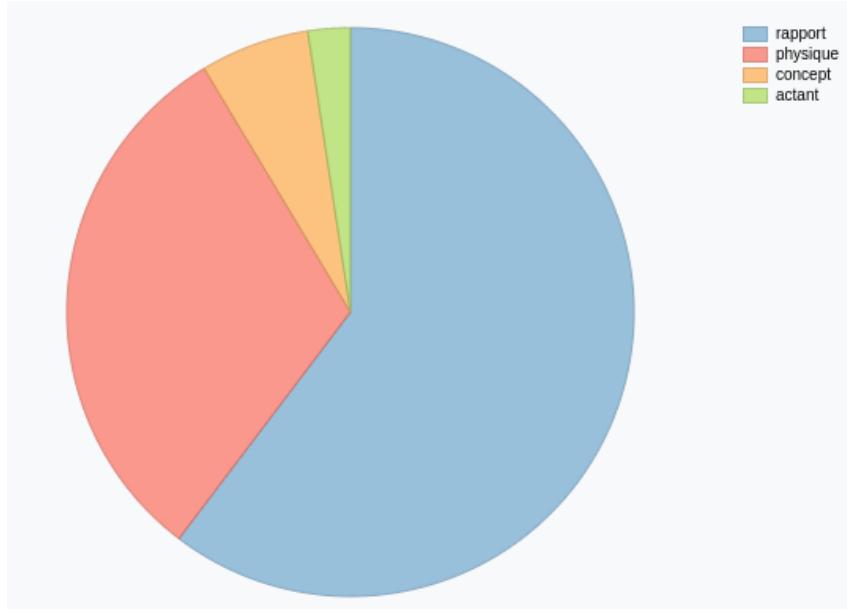


Figure 1.6: Répartition des objets par dimension existentielle

l'ensemble. Cette complexité de l'objet est d'autant plus grande que la valeur d'une propriété est une ressource sous la forme d'une URI vers une page Web ou un lien vers une autre existence de l'écosystème est donc vers une nouvelle complexité qui elle aussi s'ajoute à la complexité globale. A partir des règles génériques pour calculer la complexité existentielle d'un écosystème (Section ??), nous obtenons pour l'écosystème de connaissances de ce travail une complexité de 38.4589 Millions, ce qui est très important en comparaison de la complexité d'une citation d'ouvrage qui varie entre 60 et 3 000 mais ce qui est très peu par rapport à la complexité d'une bibliothèque, de wikipédia ou d'une IA générative comme ChatGPT (Section ??). Ce chiffre prend en compte l'ensemble des existences informationnelles qui peuplent notre base de connaissance pour ce travail d'HDR, il nous est utile pour comparer les connaissances potentielles de ces existences dont on peut représenter la répartition suivant leur niveau de complexité (abscisse) et le nombre d'existence pour chaque complexité (ordonnée) :

Ce graphique montre que la complexité des existences est très diverse puisqu'elle s'établit entre 1 et presque 3 Millions, de même concernant le nombre d'existence ayant la même complexité qui oscille entre 1 et plus de 10 mille. Une analyse des répartitions suivant le type de ressource omeka (media, item, collection, annotation) montre que les ressources les moins complexes sont les médias avec une complexité inférieure à 12 et les plus complexes sont bien évidemment les collections car elles cumulent les complexités des ressources qui la compose.

L'ensemble de ces données sont accessibles via l'API de Omeka S sous un format RDF-JSON utilisé pour l'interopérabilité entre les machines mais aussi via des représentations dédiées à la navigation à l'intérieur de cet écosystème.

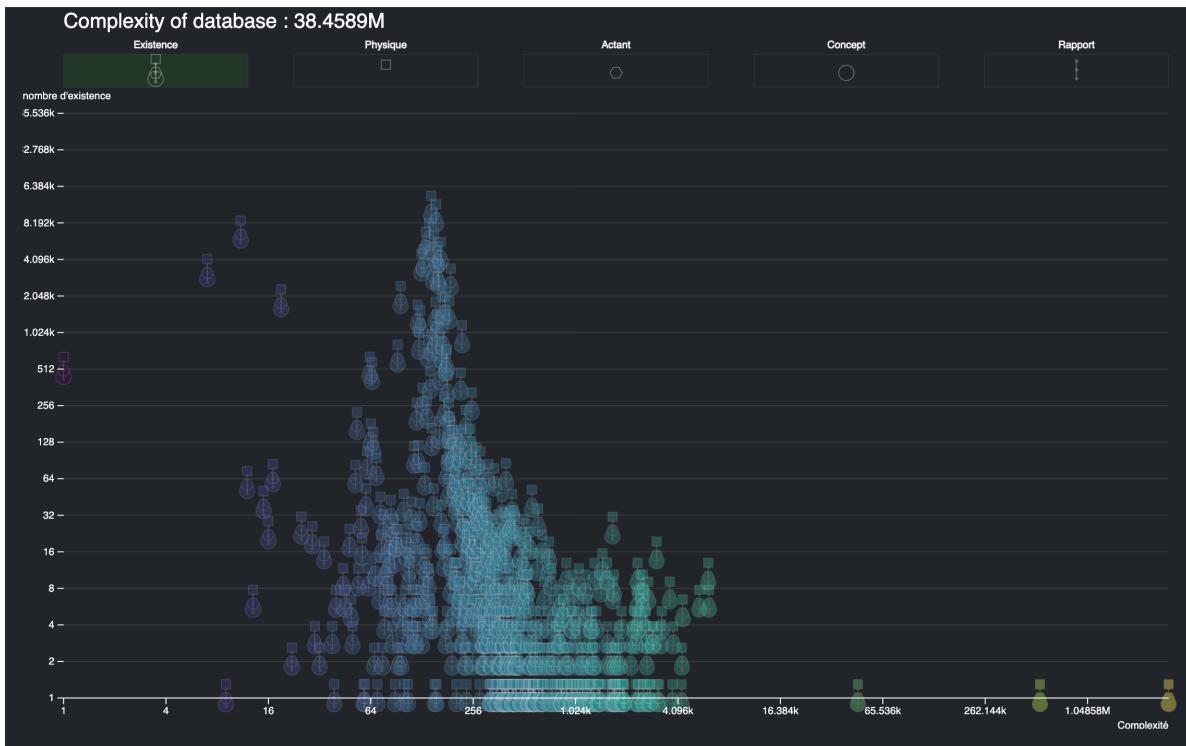


Figure 1.7: Répartition de la complexité des existences dans l'écosystème [voir en ligne](#)

1.1.2 Naviguer dans les représentations

La consultations de notre écosystème de connaissances se fait avec un navigateur Web comme Chrome ou Firefox et passe par des représentations que les utilisateurs explorent suivant les principes hypertextuels. Ces représentations consistent à mettre en relation des données avec un système de coordonnées cartésiennes qui possèdent 2 dimensions (Figure ??) ou 3 dimensions (Figure ??). Ces coordonnées définissent des points qui sont associés pour former des lignes et des plans et ainsi disposer d'un vocabulaire graphique élémentaire (Kandinsky 1991). Toutefois, la réalisation de cartographie en 3 dimensions demande beaucoup de temps et des compétences dont nous ne disposons pas dans le contexte de ce travail (?@sec-pulsaExi3D). Pour les graphiques que nous présenterons, nous avons donc décidé de n'utiliser que le système de coordonnées planaires. Il nous faut donc définir comment utiliser les 2 dimensions (x , y) pour représenter les multiples propriétés de nos données. On peut envisager de nombreuses solutions mais toutes ne seront pas compréhensibles ni facilement manipulables suivant les données et les échelles auxquelles on souhaite les représenter. Nous choisissons donc de multiplier les environnements graphiques en deux dimensions et de les interconnecter les uns avec les autres afin de former un écosystème graphique présentant de manière optimale les multiples propriétés que les données possèdent.

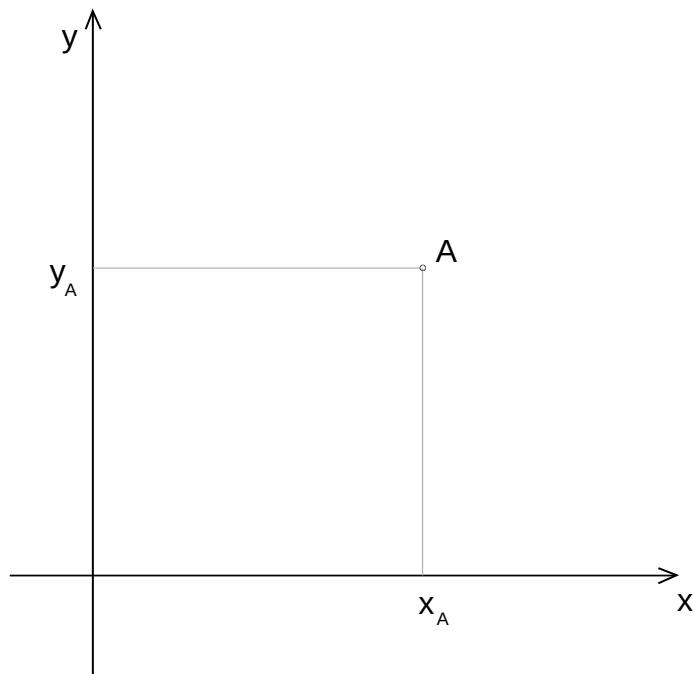


Figure 1.8: Coordonnées cartésiennes : planaires

Dans ce contexte d'écosystème graphique (Ait-Touati, Arènes, and Grégoire 2019; Zreik 2010),

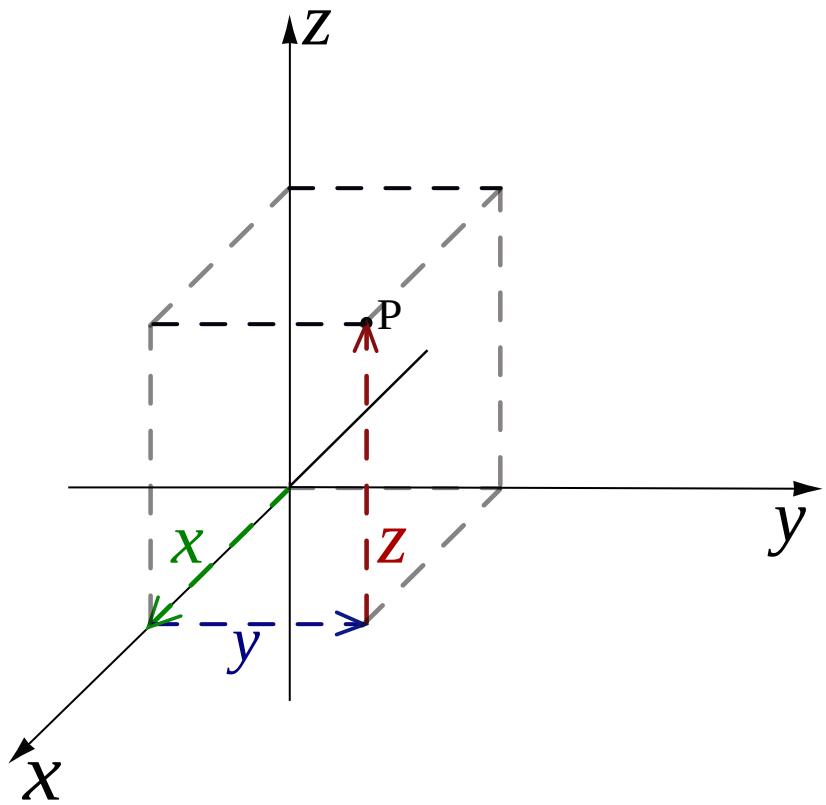


Figure 1.9: Coordonnées cartésiennes : tridimensionnelles

il est très important de disposer des moyens pour créer des graphiques à partir des données mais aussi de manière réciproque gérer les données à partir des graphiques en concevant des interactions riches entre les données, les graphiques et leurs utilisateurs afin d'effectuer les quatre actions fondamentales sur les données : Create Read Update Delete (CRUD). Nous ne sommes pas dans une vision statique de la représentation des données comme pouvait l'être (Bertin 1999) qui prenait comme principe que les graphiques devaient être imprimables. Ce qui compte aujourd'hui c'est la capacité qu'ont les systèmes de visualisation d'être manipulables pour créer les conditions d'une interprétation des données (Drucker, Mignon, and Bortolotti 2020) et l'expression d'une argumentation spécifique (Desfriches Doria and Meunier 2021a).

C'est pourquoi nous avons choisi de travailler dans un environnement Web afin de créer dynamiquement des graphiques à partir d'un flux de données et surtout de rendre ces graphiques interactifs. L'autre choix important que nous avons fait est d'utiliser le langage graphique SVG¹⁸ qui permet de manipuler chaque composant graphique de manière autonome (Fry 2008). Ainsi les points, les lignes et les plans disposent d'une autonomie en terme de propriétés graphiques, événementiels et informationnelles. Grâce à la librairie JavaScript D3.js¹⁹ (Data Driven Document) nous pouvons gérer ces propriétés en pilotant les graphiques à partir des données ou à l'inverse les données à partir des graphiques.

Dans cette environnement Web très ouvert et fertile, les possibilités de dynamisme et d'interaction entre les données, les graphiques et leurs utilisateurs sont potentiellement infinies. Il convient donc de spécifier plus précisément les choix que nous avons fait pour cartographier nos connaissances.

1.2 Représentations spatio-temporelles

Les premières informations à prendre en compte dans la cartographie des connaissances sont le temps et l'espace qui constituent une base fondamental de la recherche en sciences humaines : l'histoire et la géographie. Ce sont les données communes à toutes les analyses en sciences humaines : quand ? Où ?

1.2.1 Cartographier la géographie

Pour réfléchir sur ces informations les humains ont depuis longtemps développé des systèmes de représentations que ce soit pour le temps (Rosenberg and Grafton 2013; Domenget, Miège,

¹⁸

<https://developer.mozilla.org/en-US/docs/Web/SVG>

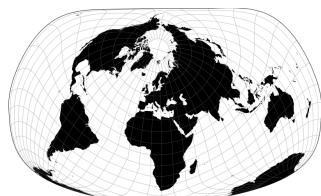
¹⁹

https://observablehq.com/@d3/gallery?utm_source=d3js-org&utm_medium=hero&utm_campaign=try-observable#maps

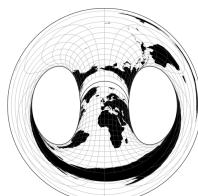
and Pélissier 2017), l'espace (Béguin and Pumain 2017) ou la combinaisons des deux (M. Serres 1997; Giacona et al. 2019; Aït-Touati, Arènes, and Grégoire 2019). Nous ne rentrerons pas ici dans l'analyse de ces représentations cela dépasserais de loin notre propos qui est de présenter nos principes cartographiques. Nous renvoyons le lecteur curieux à la veille que nous faisons depuis plus de dix ans sur cette question²⁰.

Sur notre Terre, les données spatiales sont définis par trois propriétés : une latitude, une longitude et une altitude. Les représentations des données géographiques sont aujourd’hui grandement aidées par les outils qui rendent disponibles pour les concepteurs les fonctionnalités nécessaires à la manipulation des cartes. Le principe de représentation est commun à tous ces outils : $x = \text{longitude}$, $y = \text{latitude}$. Ce qui diffère c'est le type de projection utilisé pour représenter les données suivant un point de vue particulier qui mettra l'accent sur une dimension spatiale. Les exemples ci-dessous montrent comment suivant le type de projection les représentations se transforment :

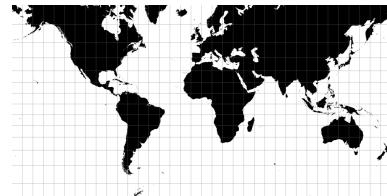
Table 1.1: Exemples de projections géographiques²¹



Jacques Bertin's 1953



Hammer retroazimuthal



Spherical Mercator

Dans notre cas, pour concevoir des cartes géographique en deux dimensions nous utilisons des librairies JavaScript Open Source comme leaflet.js²² ou D3.js qui permettent de manipuler des données géographiques modéliser avec le format GeoJSON²³. Voici par exemple la représentation géographique de mes collaborations dans le monde à partir de mes dépôts dans HAL²⁴

²⁰

<https://www.diigo.com/user/luckysemiosis?query=%23spatiotempo>

²¹

<https://github.com/d3/d3-geo-projection>

²²

<https://leafletjs.com/>

²³

<https://fr.wikipedia.org/wiki/GeoJSON>

²⁴

<https://samszo.github.io/StatsHAL/world.html?q=Samuel%20Szombiecky>

?@sec-datavizActiChercheur :

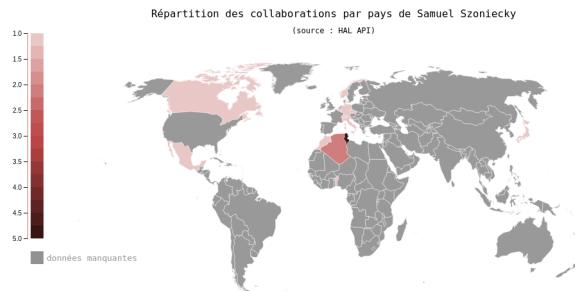


Figure 1.10: Répartition des collaborations dans le monde

Cette carte montre les pays hors de la France où sont publiées mes textes scientifiques et les conférences auxquelles j'ai participées. Parallèlement aux données géographiques, la couleur des pays est proportionnelle au nombre de collaborations. Cette carte montre que mes collaborations se développent essentiellement avec des pays francophones et des pays de l'hémisphère nord.

1.2.2 Cartographier le temps

Pour les informations historiques, nous avons besoin de gérer deux propriétés, une date de début et une date de fin. Notons que la durée n'est pas une propriété nécessaire puisqu'elle se calcule à partir de la différence entre la date de début et la date de fin. Nous posons comme principe qu'une date de fin nulle indique une durée en cours. La frise est sans doute la représentation la plus courante et la plus commode à réaliser puisqu'elle associe une coordonnée graphique avec une échelle de temps, le plus souvent x pour une représentation horizontal et parfois y lorsqu'elle est verticale. Dans notre enfance, nous avons tous réalisé des frises historiques, elles peuplent nos salles de classe et prolifèrent sur le Web²⁵. Nous avons une compréhension évidente de la frise historique, de son fonctionnement et des informations qu'elle diffuse : événements ponctuels, périodes. Voici par exemple la représentation en frise historique de mon activité d'enseignant chercheur :

Cette frise historique²⁶ montre l'évolution de mes activités d'enseignant chercheur suivant plusieurs types d'activité. Comme les outils Web de visualisation des cartes géographiques, cette visualisation fournit des fonctions de zoom, de déplacement et d'hypertextualité pour

25

<https://bit.ly/3r9ROUa>

26

Visualisation conçue à partir du module Omeka S timeline (<https://gitlab.com/Daniel-KM/Omeka-S-module-Timeline>) que nous avons adapté à nos besoins.

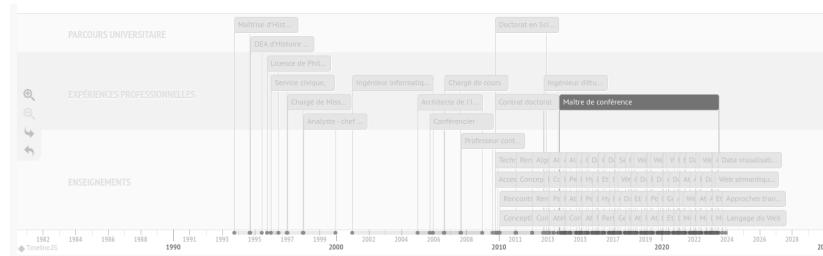


Figure 1.11: Timeline des activités d'enseignant chercheur

faciliter la lecture des données qui si elles sont trop détaillées, ne sont plus visibles. Là encore, la cartographie des connaissances dans le Web est conçu comme un outils de navigation dans les données.

Les connaissances sont toujours en rapport avec l'espace et le temps mais nous posons comme hypothèse qu'entre les connaissances des physicalités et celles des intérieurités, entre l'étendu et la pensée, l'espace et le temps n'ont pas les même modes de perceptions et d'expressions.

« La durée se dit en fonction des parties extensives et se mesure au temps pendant lequel ces parties appartiennent à l'essence. Mais l'essence en elle-même à une réalité ou une existence éternelle ; elle n'a pas de durée, ni de temps qui marque l'achèvement de cette durée. »(G. Deleuze 1968, 291)

Nous suivons sur ce point les principes spinozistes d'une modélisation ontologique corrélée à une éthique en définissant trois dimensions de l'existence corrélées avec trois genres de connaissance²⁷ Section ???. Examinons maintenant comment nous définissons de nouveaux principes cartographiques à partir de ces propositions.

1.3 Espaces matériels : connaissances des chocs

A l'instar de (Bautier 2016) nous pensons nécessaire « de prendre en compte la matérialité de la culture numérique ». Les technologies numériques véhiculent sans doute des idées de dématérialisation à travers des expériences de téléprésence, de virtualisation des échanges et d'autonomisation de la forme logique par rapport à la base matérielle. Mais peut-on encore parler de matière quand le contact avec l'événement se fait à travers des écrans, des réseaux, des milliers de kilomètres, des années, des algorithmes ?

Quoi qu'il en soit de cette « dématérialisation », nos connaissances numériques passent nécessairement par une dimension matérielle car nous sommes nous même constitué de matière :

²⁷

<https://spinoza.fr/les-genres-de-connaissance-extrait-du-cours-de-gilles-deleuze/>

« La sémirose, loin d'être un phénomène sans lien avec le corps, tire son origine de celui-ci. Ce premier aspect de la corporéité du sens peut être qualifié de cognitif : le signe émerge de l'expérience, et ne saurait être étudié qu'à travers les interactions qu'il a avec son contexte » (μ, Édeline, and Klinkenberg 2016, 2)

Les illusions que le numérique procure, tendent pour beaucoup à nous faire croire à la dématérialisation en simulant par exemple des univers immersifs où nous vivons d'autres actualités que celles de notre corps avec des avatars de toutes sortes (Amato and Perényi 2013). Mais en dernière instance nous sommes matière et nous évoluons dans des espaces matériels. Sur ce point nous nous opposons aux spiritualistes qui affirment « qu'il existe une substance spirituelle (l'âme ou l'esprit), indépendante de la matière, et qui serait, en l'homme, principe de vie ou d'action. » (Comte-Sponville 1998, sec. 12).

Les interprétations par Deleuze de L'Étique de Spinoza décrivent ces espaces matériels comme étant la première dimension de l'existence celle des « parties extensives » :

« Ces parties (corpora simplissima) [...] se définissent uniquement par leur déterminisme extérieur, et vont toujours par infinités ; [...] elles constituent la matière modale infiniment variée de l'existence. » (G. Deleuze 2003, 110)

Entre l'infiniment grand et l'infiniment petit (cf. illustration ci-dessous) les parties extensives sont observables et modélisables à toutes les échelles physiques de notre univers. Tout comme le choix d'une projection géographique reflète un point de vue particulier, celui des échelles de représentation contribue lui aussi à l'expression d'une subjectivité spécifique.

Les parties extensives correspondent aux « physicalités » des milieux que nous habitons, elles en sont l'indispensable matérialité. Cette nécessité de la matière est corrélée à des connaissances, elles aussi nécessaires, celles du premier genre de connaissance : les idées inadéquates :

« L'idée inadéquate, c'est l'idée inexpressive et non expliquée : l'impression qui n'est pas encore expression, l'indication qui n'est pas encore explication. » (G. Deleuze 1968, 136)

Pour expliquer à quoi correspondent les connaissances du premier genre, Deleuze décrit dans un de ces cours l'expérience d'une personne au bord de la mer :

« Bien alors c'est quoi la connaissance du premier genre ? Eh bien allez, j'y vais, je me lance, je suis dans le premier genre de connaissance. Je me lance, je barbote, comme on dit. Qu'est-ce que ça veut dire barboter ? Barboter c'est tout simple, ça indique bien, on voit bien que c'est des rapports extrinsèques. Tantôt la vague me gifle, et tantôt elle m'emporte. Ça c'est des effets de choc. C'est des effets de choc, à savoir, je ne connais rien aux rapports qui se composent ou qui se décomposent, je reçois les effets de parties extrinsèques. Les parties qui m'appartiennent à moi, sont secouées, reçoivent un effet de choc des parties qui appartiennent à la vague. » (G. Deleuze 1981)



Figure 1.12: Échelle de l'univers - univers observable <https://htwins.net/scale2/>

Échelle de Planck



Created by Cary Huang - youtube.com/carykh
Ported to PixiJS by Matthew Martori

$10^{-35.0}$

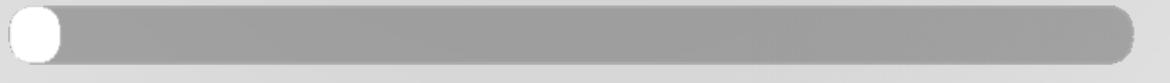


Figure 1.13: Échelle de l'univers - limite de Plank <https://htwins.net/scale2/>

Donnons un autre exemple de ce premier genre de connaissance en vous invitant à faire l'expérience des parties extensives suivantes :

சூரியன் பிரகாசித்து கொண்டு இருக்கின்றது

Figure 1.14: Parties extensives 1

Sauf si vous connaissez le tamoul, le texte ci-dessous est pour vous comme un choc, vous ne connaissez rien des rapports qui se composent ou se décomposent, vous ne voyez que les parties extensives du texte. Pour être plus précis, vous pouvez tout de même discerner des rapports puisque vous savez que l'image est un texte composé de caractères qui composent des mots séparés par des espaces. Par contre, vous n'avez aucune idée des concepts présents dans le texte, vous avez connaissance des signifiants mais pas des signifiés²⁸. D'une certaine manière vous êtes comme un OCR (optical character recognition) capable de reconnaître des caractères et des mots dans une image. Mais à l'inverse d'une machine numérique qui avant la reconnaissance du texte décompose l'image en une multitude de points ayant chacun leurs coordonnées cartésiennes et leurs propriétés de couleur, vous commencez par reconnaître le texte puis vous le décomposez en mots et en caractères. Cette différence entre la machine et l'humain dans le processus de connaissances est au cœur d'une problématique fondamentale de la gestion mécanique du sens :

« il y a un conflit entre l'holisme du sens et le mécanisme de la syntaxe. Le sens d'un texte dépend de son contexte, le sens d'un paragraphe dépend aussi du texte dans lequel il s'intègre, le sens d'un mot du paragraphe qui le contient, etc. : le sens va du global au local, de la compréhension globale vers l'analyse. Or, le formalisme opère de manière inverse : le sens d'une formule logique se construit à partir du sens de ses parties, allant du local au global. » (B. Bachimont et al. 2011, sec. 11)

Ce conflit est d'autant plus flagrant quand le même texte est présenté dans une écriture que vous connaissez (Figure ??). Dans ce cas, vous ne faites plus la décomposition du texte en parties extensives le constituant mais vous accédez directement à sa signification car vous avez appris à lire, c'est-à-dire à discerner les compositions de rapports dans les parties extensives et vous accédez ainsi à un autre genre de connaissance celui des signifiants.

Notre principe de cartographie des espaces matériels consiste à les considérer uniquement en tant que physicalités composées de parties extensives modélisables par leurs propriétés physico-chimique : largeur, hauteur, profondeur, masse, couleur, atome, molécule... Par exemple, dans

²⁸

Précisons toutefois que l'humain est particulière efficace pour la paréidolie (**Dufort2015?**), ce qui lui permet de donner du sens à des formes ambiguës et donc de créer des signifiés quelles que soient les parties extensives qu'il discerne.

Le soleil brille

Figure 1.15: Parties extensives 2

les espaces matériels un livre est considéré du point de vue de sa taille, son nombre de page, son poids, sa matière etc... Dans l'espace matériel, on ne prend pas en compte l'auteur ou la thématique du livre qui respectivement seront cartographiés comme actant (Section ??) et comme élément d'un espace conceptuel (Section ??). Dans les espaces matériels les mots du livre sont des traces de couleur qui génère des connaissances de l'ordre des chocs ; c'est à dire une réaction entre des parties extensives celles de la trace et celles de nos capteurs biologiques ou artificiels. Notons que ces chocs en entraînent d'autres qui eux mêmes se propagent dans un phénomène d'accroissement de l'entropie constitutif de l'univers chaotique du premier genre de connaissances, celui des idées inadéquates qui se répandent sans fin par composition et décomposition :

« qu'est-ce que vous racontez là, mais alors cette nature, c'est un pur chaos ! Pourquoi c'est un pur chaos ? Parce que vous remarquerez que, chaque fois qu'un corps agit sur un autre, il y a toujours composition et décomposition à la fois. Ce n'est pas à ce niveau-là que je pourrais dire, il y a du bon et du mauvais. Pourquoi ? Parce qu'il y a forcément composition et décomposition, les deux l'un dans l'autre. » (G. Deleuze 1981)

Ces compositions et décompositions des corps les plus simples que sont les parties extensives sont modélisables suivant une hiérarchie de parties et de sous-parties. Par exemple le livre est décomposable en parties plus petites : page → paragraphe → phrase → mot → caractère. Ce même livre est aussi composable avec d'autres parties plus vastes : étagère → salle → bibliothèque. La modélisation des espaces matériels est une structure hiérarchiques qui potentiellement se compose jusqu'aux limites de l'univers observable (Figure ??) et se décompose jusqu'à l'infiniment petit de l'échelle de Plank (Figure ??) en passant par l'échelle de l'être humain (Figure ??). Nous verrons plus loin combien le choix de l'échelle cartographique est primordial (**?@sec-illusionExhaustivite**)

Pour cartographier les espaces matériels en tant qu'ensemble des parties extensives définissables par leurs propriétés physico-chimique et leurs compositions vers l'infiniment grand et décompositions vers l'infiniment petit, nous optons pour un modèle de diagramme hiérarchique appelé « treemap » et proposé par (Shneiderman 1998) qui se compose de rectangles imbriqués représentant un élément et ses sous parties et dont la taille des rectangles est proportionnelle à la valeur numérique d'une propriété, par exemple le nombre d'éléments que contient la sous partie .

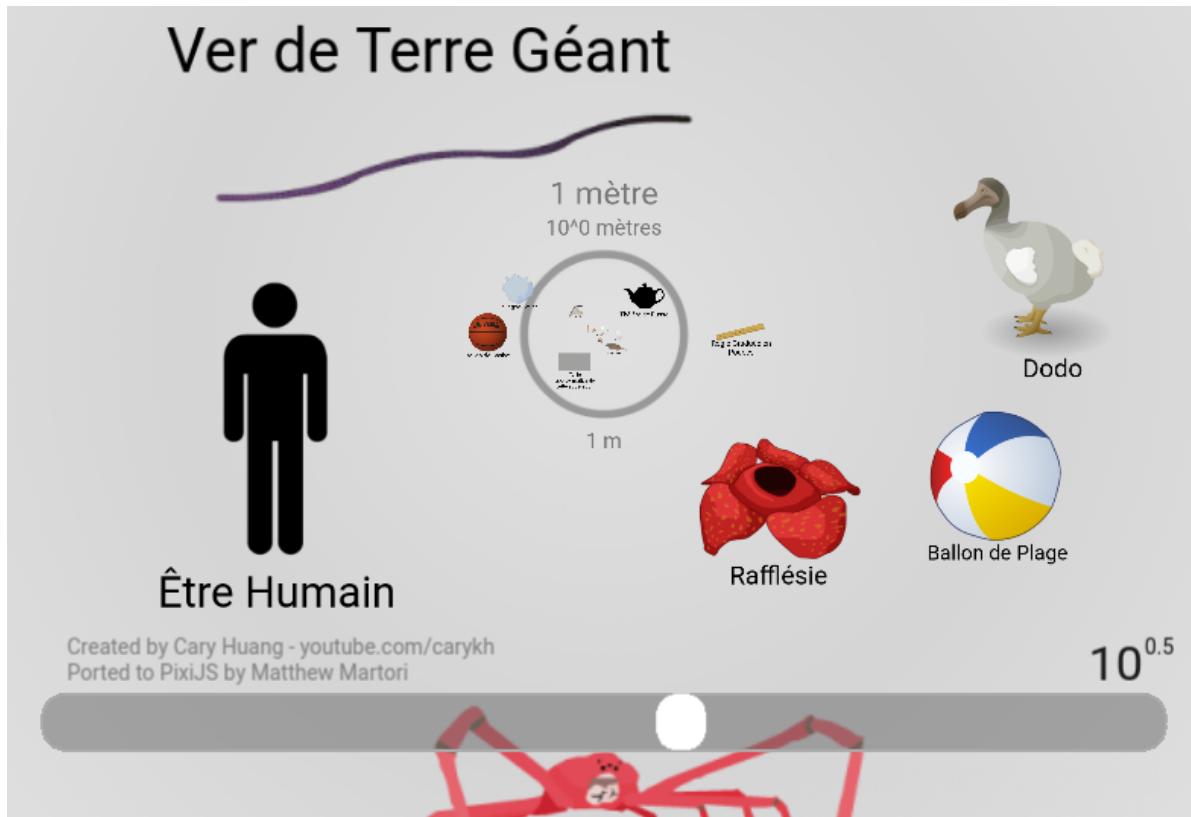


Figure 1.16: Échelle de l'univers - être humain

En utilisant l'objet TreeMap²⁹ de la librairie D3.js, nous avons implémenté ce modèle de diagramme dans un module JavaScript (Section ??) pour le rendre dynamique, interactif et ainsi représenter les espaces matériels que nous cartographions soit à partir de données existantes soit en les créant au fur et à mesure de l'exploration. Notons que pour faciliter la visualisation des dimensions physiques complexes, nous avons implémenté une navigation directe vers une partie ou une sous-partie et une navigation hiérarchique par zoom dans une partie et dé-zoom vers le parent. Par exemple, nous avons cartographié notre CV en utilisant ce modèle de diagramme avec comme paramètre de taille des rectangles la durée d'un événement :

Curiculum Vitae de Samuel Szombieky

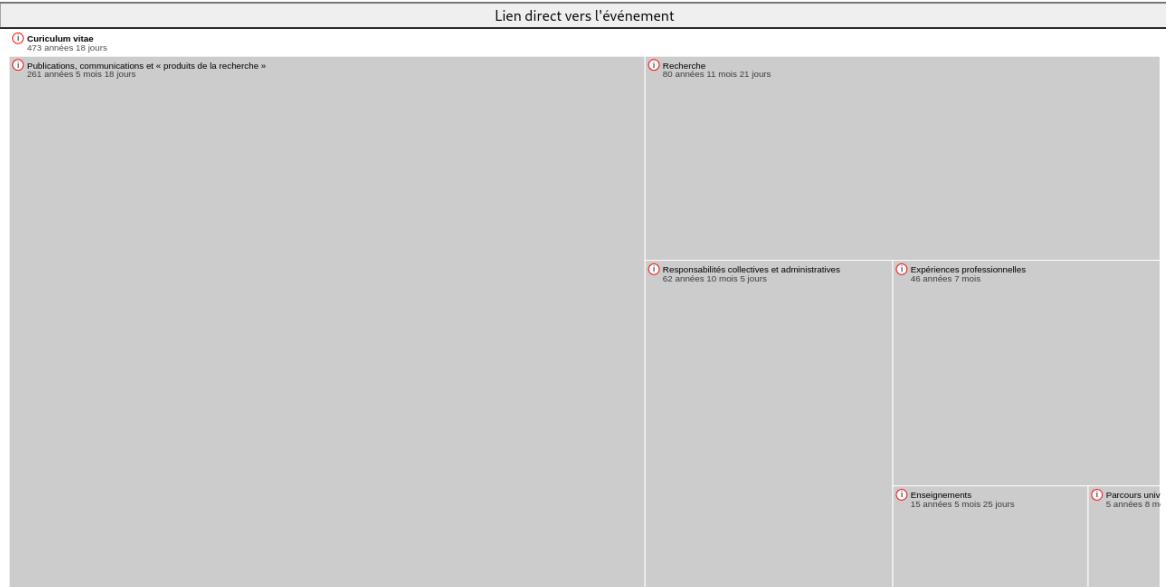


Figure 1.17: Treemap d'un Curiculum Vitae

Dans le cas de ce diagramme, la durée d'un événement cumule l'ensemble des durées des événements qui le compose ce qui explique une durée de plusieurs centaines d'année pour le CV. D'autre part, cette durée exprime une période de travail et ne prend pas en compte les activités parallèles à l'inverse de la frise historique (Figure ??).

Dans les espaces matériels les connaissances sont des chocs qui ne dure qu'un instant, celui du contact entre les parties extensives. On peut les dater plus ou moins précisément, ils peuvent se répéter encore et encore mais ils n'ont pas de durée. Ce qui dure c'est l'onde du choc qui se propage dans les physicalités et dans les intérieurités des acteurs qui participent à l'événement ce qui génère d'autres chocs dans les espaces matériels et des connaissances d'un autre genre dans les espaces conceptuels.

²⁹

<https://d3js.org/d3-hierarchy/treemap>

1.4 Espaces conceptuels : connaissances des essences

A l'opposé des espaces matériels-physiques et de la connaissance des chocs, nos connaissances se composent aussi dans nos intérieurités :

« Par le terme vague d’ “intériorité”, il faut entendre une gamme de propriétés reconnues par tous les humains et recouvrant en partie ce que nous appelons d’ordinaire l’esprit, l’âme ou la conscience - intentionnalité, subjectivité, réflexivité, affect, aptitude à signifier ou à rêver. » (Philippe. Descola 2005, 168)

Comment cartographier ces espaces de connaissances ? Comment mesurer ces espaces ? Combien pèse une âme ?

La question du poids des âmes se pose depuis longtemps comme en témoigne l’iconographie de la pesée des âmes. La psychostasie, nom donné à cette activité de la pesée des âmes, touche historiquement les domaines de la théologie, de la philosophie et de l’éthique mais intéresse aussi les sciences de l’information et de la communication qui cherchent notamment des réponses sur la mesure, l’analyse et la critique de ces espaces informationnels immatériels qui ne sont pas mesurables de la même manière qu’une planche de bois ou qu’une récolte de fruit car ils ne sont pas soumis aux règles physiques de la matérialité tant qu’ils ne sont pas exprimées. Dès le passage de ces espaces intérieurs vers une forme d’expression quelle qu’elle soit (écrit, parole, clic sur un bouton...), ils se transforment en physicalités dont on pourra mesurer les paramètres physiques (hauteur, largeur, vitesse...). Ne peut-on mesurer ces intérieurités qu’une fois exprimées par nos paroles, nos écrits, nos dessins, nos danses, nos activités corporelles... ?

Il faut sans doute passer par une forme d’expression pour que les impressions dans nos intérieurités soient communicables, même s’il existe des connaissances intérieures qui restent secrètes, non par choix de ne pas les exprimer mais par impossibilité de le faire soit parce qu’elles sont inconscientes, soit parce qu’elles relèvent d’une expérience incommunicable. Sans parler des connaissances mystiques qui n’existent que par le fait de les avoirs expérimenté ou non, pensons simplement aux connaissances qui émergent de nos intérieurités à la lecture dans simple mot : aimer. Chacun d’entre nous expérimente la lecture de ce mot suivant ses propres histoires, ses états actuels et ses désirs ; ce que nous en communiquerons révélera ou non une partie de ces expériences que le mot aura fait résonner en nous. Nos intérieurités sont le siège de nos subjectivités et des processus de signification que nous avons abordé en introduction de ce chapitre. Elles sont l'espace des élaborations sémiotiques qui transforment notre pouvoir de discernement en pouvoir d'agir. Mesurer les espaces conceptuels et avant tout un travail de réflexivité individuelle et de concentration sur cette dimension particulière de l'existence que nous ne pouvons explorer que dans la solitude de notre propre conscience. L'enjeu qui nous anime ici est de fournir aux explorateurs de ces espaces conceptuels des outils pour cartographier leurs explorations de manière à les rendre interopérable avec celles menées par d'autres.



Figure 1.18: La pesée des âmes dans le retable polyptyque du Jugement Dernier de Rogier van der Weyden aux Hospices de Beaune, 1443-1452,
<https://commons.wikimedia.org/w/index.php?curid=6028656>

1.4.1 Approches topologique de la cartographie des concepts

Dans le domaine des sciences cognitives, les espaces intérieurs ont été pensé par (Gärdenfors 2001) comme des « espaces conceptuels » en complémentarité des approches symboliques qui modélisent les systèmes cognitifs avec des machines de Turing et des approches connexionnistes qui modélisent avec des réseaux de neurones artificiels. Cet auteur propose de modéliser les espaces conceptuels à partir d'une représentation topologique des similarités qualitatives. La modélisation des espaces conceptuels par des topologies est sans doute une perspective intéressante pour représenter ces espaces car elle permet de concevoir des espaces métriques à partir de la notion simple de voisinage. On peut considérer les concepts comme des points qui définissent un espace dans leurs rapports de voisinage avec d'autre concepts et calculer des distances entre ces points. Toute la difficulté est de définir les valeurs qui seront utilisées pour calculer la distance entre ces points. Gärdenfors propose d'utiliser des valeurs qualitatives pour calculer les distances par exemple les concepts de couleurs seront représentés dans une topologie dont les distances sont calculées suivant les qualités de nuance, d'intensité et de luminosité. L'avantage de cette proposition est de rendre pratiquement objectif la distance entre les concepts car celles-ci résultent d'une mesure physique. Mais de notre point vue, ces qualités font parties de la dimension matérielle que nous avons présentée plus haut (Section ??), elles ne peuvent donc pas être utilisées pour modéliser les concepts qui dans notre modèle relève d'une autre dimension existentielle, celle des essences (Section ??). Il est fondamental de préserver la multiplicité des points de vue à l'intérieur de ces espaces conceptuels et de ne pas les réduire à une mesure physique qui est la même pour tous, à tout moment, en tout lieu. Nous pensons que les espaces conceptuels sont propres à chaque individu; leur cartographie ne peut donc pas relever d'une mesure « universelle » liée à une métrologie physique. La distance qui sépare « aimer » de « haïr » n'est pas la même pour vous ou moi, pour hier, aujourd'hui et demain, ou suivant le lieu de mes rapports avec ces concepts. Nous verrons plus loin comment les dimensions existentielles des actants et des rapports nous permettent de cartographier ces fluctuations temporelles et spatiales (Section ??) retenons juste pour le moment que les principes de cartographie des espaces conceptuels ne peuvent se baser sur une mesure matérielle car nous ne recherchons pas une mesure objective mais tout au contraire, l'expression d'une subjectivité.

Contrairement à Gärdenfors, le langage IEML propose un « filet topologique » (Lévy 2011, 257) dont les espaces métriques sont purement conceptuels puisque les rapports de voisinage sont définies suivant six concepts (être, signe, chose, actuel, virtuel, vide) associés à trois positions conceptuelles (substance, attribut, mode) sur six couches. Il en résulte un grille topologique très vaste : $(6*6*6)^6$ soit $1,015599567 \times 10^1$ positions possibles. Une infime partie de ces positions (3418^{30}) ont été interprétées, classifiées et référencées par Pierre Lévy et ses

³⁰

Le chiffre correspond au nombre d'adresse définie dans le dictionnaire IEML à la date du 25/01/2023. Ce travail toujours entrain de se faire est consultable ici : <https://github.com/plevyieml/ieml-language/>

équipes pour donner du sens à cette topologie et fournir un vocabulaire de base utilisable avec un éditeur dédié à ce langage³¹. Cette solution de cartographie des espaces conceptuels est élégante et très ambitieuse mais elle se confronte à plusieurs difficultés majeures. La première est qu'il n'est pas très facile de comprendre la complexité de ce langage et son utilité par rapport à des outils comme le moteur de recherche Google dont l'usage simplissime demande un effort minimal. IEML s'adresse à un public de spécialistes ayant des besoins très spécifiques et demande un investissement conséquent :

« IEML [...] force à faire un travail d'analyse et de définition des concepts utilisés et fait apparaître de possibles paralogismes dans un raisonnement. » (Vitali Rosati 2021).

La deuxième difficulté porte sur l'usage de ce langage qui ne correspond pas aux habitudes du public de chercheurs auquel il est destiné. Ceux-ci travaillent généralement des textes dans lesquels la définition et la critique des concepts est une part importante mais le référencement de ces concepts par des thésaurus, des vocabulaires normalisés ou des langages formels est considéré comme un travail à la charge des documentalistes, des bibliothécaires ou des « ingénieurs sémantiques », nouveau métier que Pierre Lévy contribue à faire émerger. Le passage par un tiers en charge de traduire un texte écrit en langage naturel dans un langage sémantique comme IEML occasionne une nouvelle difficulté liée à l'économie du processus éditorial qui est déjà soumis à de forte pressions temporelles, financières et humaines. Une autre difficulté que nous avons expérimentée dans notre usage d'IEML depuis une dizaine d'années, est le manque de pérennité des outils mis à disposition pour gérer ce langage³². Cette difficulté inhérente à un travail de recherche « in progress » mais plus généralement aux langages informatiques qui évoluent au fil de temps rend délicat l'investissement important et constant que nécessite l'utilisation d'IEML. Au final, ce magnifique projet mené par Pierre Lévy rejoint sans doute la liste des langues parfaites (Eco 1994) et contribue en tout cas à faire avancer l'utopie d'un dialogue plus fécond entre les humains grâce aux machines.

1.4.2 Modélisations prétopologiques des concepts

Face à ces difficultés, nous proposons de concevoir la cartographie des espaces sémantiques à partir d'outils simples permettant à chacun de construire ses propres représentations conceptuelles et donc de maîtriser le sens de ces représentations. Pour ce faire, nous avons élaboré un outil de conception de cartes sémantiques qui s'appuie sur le principes de la pré-topologie (Belmandt 1993; Thibault 2017; Toumia 2018) pour manipuler des concepts et leurs relations.

³¹

L'éditeur IEML était accessible (<https://dev.intlekt.io/>) mais n'est plus accessible

³²

Pour un historique rapide des différentes implémentations : <https://intlekt.io/histoire/>

Les espaces conceptuels se prêtent particulièrement bien à la modélisation prétopologique car il correspondent à ces deux principes fondamentaux :

« pretopology can be used to represent a system where the relation between an element and a set is not a simple aggregation of the individual relations to the members of the set. In this it is fundamentally different from a graph.

pretopology establishes one single relation between a particular element and a particular group. In this it is different from a multilayer network. » (Laborde 2019, 28)

Nos principes cartographiques utilisent les notions de base de la prétopologie pour guider l'utilisateur dans la construction d'une carte et pas uniquement pour représenter les résultats d'une analyse automatique comme peuvent le faire par exemple les outils de modélisation de graphes comme Gephi (Bastian, Heymann, and Jacomy 2009). L'idée principale de cette démarche est de construire pas à pas des espaces conceptuels relativement simples avec un protocole de formalisation les rendant compréhensibles, interopérables et calculables. Les choix nécessaires à la construction de la carte sont ceux du cartographe et pas ceux d'un algorithme qu'on bricole en jouant avec ses paramètres pour obtenir la représentation désirée. Avec l'outil que nous proposons, le cartographe maîtrise la signification de ces choix ce qui n'est pas toujours le cas quand on applique un algorithme sur une grande quantité de données. L'objectif est d'éviter que la carte serve uniquement d'illustration justifiant un discours par un « preuve » graphique mais soit le discours à part entière.

Le processus de cartographie que nous proposons à partir d'une modélisation prétopologique consiste à définir un espace conceptuel en lui donnant un titre. Cet espace est représenté par une ellipse et par son titre. Dans un deuxième temps, cet espace est peuplé d'un ensemble d'éléments appartenant à l'espace. Par exemple, l'espace conceptuel que nous cartographions porte le titre de « humanités numériques », il se compose des éléments : humains, machines, collaboration, efficace, biais, cognitifs...

Dans un troisième temps, la modélisation prétopologique consiste à créer un ensemble de parties $P(X)$ qui sont des sous-ensembles constitués avec une application d'adhérence qui s'applique aux éléments de l'ensemble.

« On appelle prétopologie sur X , toute application adh de $P(X)$ dans $P(X)$ qui vérifie :

- i - $adh(\emptyset) = \emptyset$
- ii - $A \in P(X), A \subseteq adh(A)$

(X, adh) est appelé espace prétopologique.

adh est encore appelée adhérence. » (Dalud-Vincent 2017, 47)

Dans notre cas, l'application d'adhérence consiste à « conceptualiser » les chaînes de caractères continues pour modéliser des sous-ensemble sous forme de mots : $P(X) = [« humains », « machines », « collaboration », « efficace », « biais », « cognitifs »]$. Ces mots sont eux-aussi

représentés par une ellipse et par un titre ce qui de manière fractale fait que chaque élément de l'ensemble est lui-même un ensemble disposant de propriétés et de méthodes utiles pour sa manipulation cartographique. De même, l'espace conceptuel « humanités numériques » peut-être utiliser comment élément d'un ensemble plus vaste par exemple « sciences humaines ».

1.4.3 Applications prétopologiques des concepts

Pour faciliter les manipulations de concepts nous travaillons à un dispositif de cartographie qui présente un espace dynamique et interactif dans lequel les cartographes pourront utiliser graphiquement les applications prétopologiques pour : créer un espace, le définir, créer des sous ensembles et le mettre en relation avec d'autres espaces suivant des applications prétopologiques spécifiques. Pour faciliter le positionnement des concepts les uns par rapport aux autres en évitant le chevauchement des titres, nous avons fait le choix d'utiliser une grille hexagonale comme le propose (Rodighiero 2021) pour réaliser la carte des affinités d'un laboratoire de recherche ou comme nous l'avons expérimenté pour paramétriser le filtrage des flux d'informations (Szoniecky 2011). Une grille hexagonale permet de représenter les relations d'un élément avec vingt-quatre autres sans aucun chevauchement, ce qui peut paraître faible lorsqu'on pense à l'infini des relations possibles entre les concepts mais qui offre l'avantage de contraindre la cartographie sémantique dans un espace relativement simple et donc facilement compréhensible. De plus, la construction fractale des rapports entre ensembles et éléments rend infini la possibilité d'expression puisque le regroupement des éléments dans un ensemble crée la possibilité de représenter vingt-quatre nouveaux éléments. Pour des raisons d'ergonomie graphique et algorithmique, les espaces conceptuels sont structurés par une grille hexagonale :

« L'hexagone permet de “paver” l'espace en un agencement sans fin, qui, potentiellement, permet de dessiner des réseaux infinis. [...] Choisir une grille hexagonale réduit le bruit numérique et facilite la lecture. » (Rodighiero 2021, 76)

Nous nous sommes inspiré des travaux d'Amit Patel³³ pour mettre en place une grille hexagonale et les fonctionnalités nécessaires pour les applications prétopologiques que nous avons codées dans une librairie JavaScript³⁴ et mises en application dans un module Omeka S (Section ??) afin de gérer les manipulations d'informations dans une base de données.

³³

Pour une explication des grilles hexagonales : <https://www.redblobgames.com/grids/hexagons/>

Pour une proposition d'implémentation algorithmique : <https://www.redblobgames.com/grids/hexagons/implementation.html>

³⁴

Le code est accessible ici : <https://github.com/samszo/HDR/docs/jdcCartoHexa.html>

Une fois connecté à l'application de cartographie, la première action à effectuer est de charger une carte ou d'en créer une. Lorsque l'on crée une carte, l'application propose une espace avec une grille hexagonale de rayon 2 avec positionné au centre de cette grille un concept "vide" qui lui aussi possède une grille hexagonal permettant ainsi e fractaliser la cartographie avec des grilles hexagonales incluant d'autres grille hexagonales etc. Notons que cette grille est créée par défaut avec un rayon de 2 mais qu'elle peut être étendue en déplaçant les bords de l'espace.



Figure 1.19: Cartographie conceptuel dans un espace hexagonal

Les utilisateurs peuvent ensuite utiliser les applications d'adhérence que nous venons de présenter, qui consiste par exemple à « conceptualiser » un espace en lui donnant un titre. Il suffit de cliquer dans l'espace conceptuel pour choisir une position et faire apparaître un formulaire permettant de saisir le titre du concept. Il est important de pouvoir créer tous les concepts pos-

sibles à partir de n'importe qu'elle chaîne de caractères car une des difficultés bien connues dans l'usage des ontologies ou des vocabulaires formalisés est de trouver la « bonne » référence dans un ensemble souvent très vaste dont on ne connaît pas l'ensemble des références et surtout quand on ne trouve pas d'équivalent à ses propres habitudes linguistiques. C'est pourquoi, dans nos principes de cartographie des espaces conceptuels, l'expression des concepts est libre comme c'est le cas dans les folksonomies (Broudoux et al. 2012). Toutefois, lors de la saisie du titre du concept, un processus d'auto-compléTION du champ de saisie renvoie les concepts déjà enregistrée dans la base de données à partir de quelques lettres ; l'utilisateur est informé des concepts existants et peut donc choisir une référence que d'autres utilisateurs ont déjà utilisé ce qui permet de créer une interopérabilité formelle. Notons que cette interopérabilité ne présage pas d'un consensus sur le sens du concept, toutefois elle est grandement utile pour rassembler les différents acteurs qui l'utilise afin qu'il discute de leurs accords et divergences Section ??.

L'application « conceptualiser » enregistre dans une base de données Omeka S la position d'un concept défini par un actant ici et maintenant dans un espace conceptuel de référence en décomposant ces informations dans les propriétés suivantes :

- identifiant de la position
- titre de la position
- identifiant de l'espace de référence
- coordonnées de la position dans l'espace de référence
- identifiant de l'actant
- date du choix de la position
- lieu du choix de la position

Décomposer l'actant, le concept et sa position a pour avantage de partager un concept commun avec d'autres utilisateurs tout en conservant le point de vue de l'actant sur ce concept. Ainsi, « aimer » et « haïr » peuvent être commun à plusieurs personnes mais la distance entre ces deux concepts peut varier suivant les individus, le temps, l'espace... Il y a donc une dé-corrélation entre le concept et ses usages. Le concept est virtuel, c'est une potentialité qui s'actualise dans une « action située » c'est à dire dans des usages ayant leurs propres spécificités ((**actionSitué?**)). De plus, l'enregistrement situé des étapes de construction de la carte rend accessible le processus créatif de son auteur.

Cette définition minimale de l'espace prétopologique peut être enrichie par d'autres applications qui enrichissent l'espace de nouvelles propriété comme celui qui est à l'œuvre avec IEML. Cette activité purement conceptuelle consiste à affiner la cartographie en définissant des rapports entre les concepts à la manière de ce qui se fait lorsqu'on développe une ontologie (Bruno. Bachimont 2007) en utilisant par exemple des propriétés de relation issu du vocabulaire SKOS (Isaac 2011).

Dans notre cas, l'application « intérieur » va nous permettre de définir les espaces conceptuels à l'intérieur d'autres espaces conceptuels en suivant la définition :

« Soit une application $i : P(E) \rightarrow P(E)$ appelée intérieur et définie comme suit :

A, A ⊂ E l'intérieur de A, $i(A) ⊂ E$ est telle que :

– $i(A) = [a(A^c)]^c$ (P1)

– $i(A) ⊂ A$ (P2)

avec A^c le complémentaire de A soit $E - A$. » (Levorato 2008, 40)

Pour utiliser cette application, il suffit de cliquer dans un des hexagones de la grille pour faire apparaître un formulaire permettant de choisir un concept et les relations que ce concept entretient avec le concept dont il est un sous espace. Ces relations sont exprimées avec le vocabulaire des relations de SKOS³⁵ mais sans toutefois limiter les possibilités d'associations aux contraintes logiques. Par exemple, il est illogique que deux concepts possèdent à la fois une relation hiérarchique “plus spécifique” (`skos:narrower`) et une relation “plus générique” (`skos:broader`). Toutefois, il nous semble important de laisser la possibilité aux usagers d'exprimer ces relations illogiques afin de laisser une place aux incohérencies et aux inventions poétiques.

1.5 L'actant : connaissances des dynamiques génératives

Entre la dimension des physicalités organisée en hiérarchies et celle des concepts organisée en topologies, la dimension des actants organise l'espace des connaissances à la fois sous la forme de topos et de chôra (Zamora 2003; Derrida 1997). Topos et chôra sont deux manières complémentaires de définir l'espace soit pour le topos sous la forme d'une identité par exemple “Université Paris 8”, soit pour la chôra en définissant ce que le lieu génère comme activités, ce que sont ses dynamiques génératives, par exemple dans le cas de Paris 8 : recherches et enseignements.

« la chôra [...] relèvent du monde sensible, non du monde intelligible. Inversement, la notion de topos, dans la mesure où elle concorde avec la logique de l'identité du sujet, relève moins de la sensibilité que de l'intelligibilité. » (Berque 2009a, 232).

Il est curieux de voir que dans l'histoire, topos et chôra se sont développés à travers deux manières d'être au monde (Latour 2012a). On pourrait par exemple voir d'un côté une vision occidentale du topos dont les modernistes sont les héritiers et qui plonge ces racines chez Platon, Aristote et que l'on trouve aussi dans la Bible où la première activité de l'homme est de nommer les vivants pour définir leur identité :

³⁵

Pour une présentation des relations sémantiques dans SKOS : <https://www.w3.org/TR/skos-reference/#semantic-relations/>

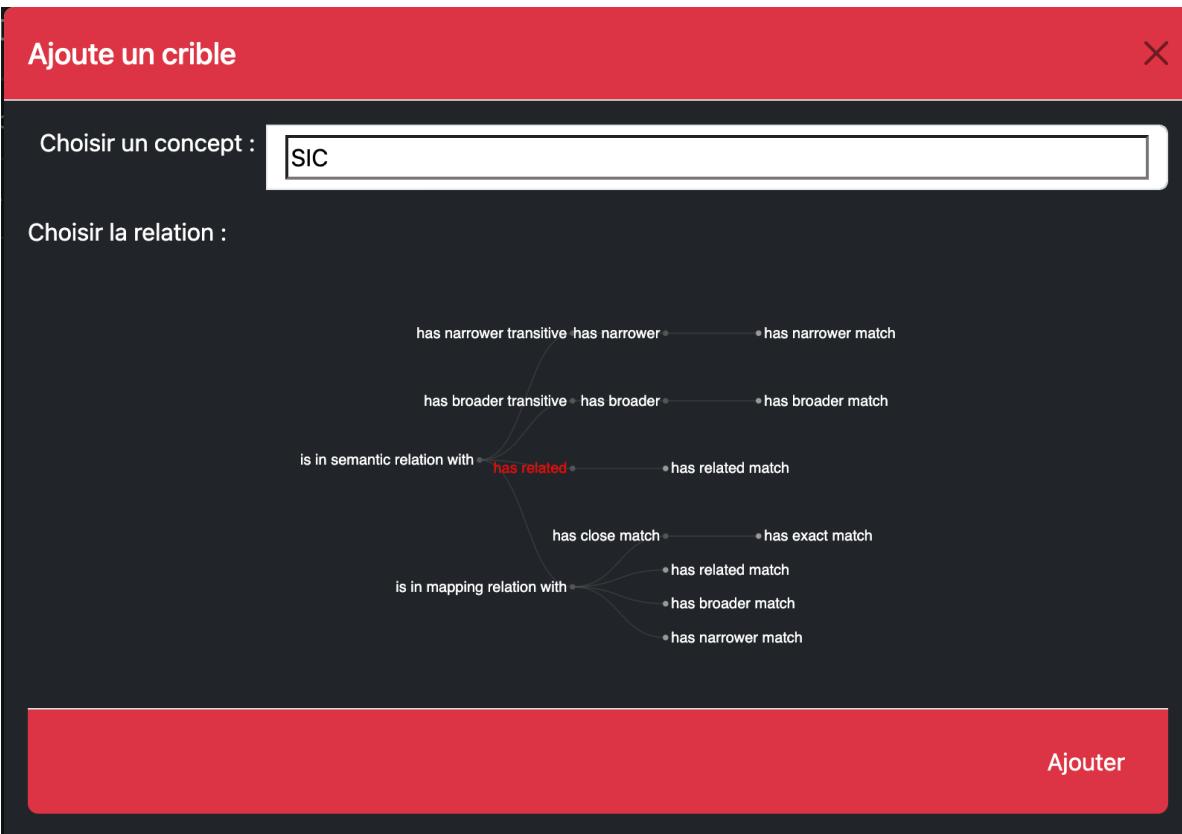


Figure 1.20: Ajout d'un concept et de ses relations

« 2.19 : L'Éternel Dieu forma de la terre tous les animaux des champs et tous les oiseaux du ciel, et il les fit venir vers l'homme, pour voir comment il les appellerait, et afin que tout être vivant portât le nom que lui donnerait l'homme.

2.20 : Et l'homme donna des noms à tout le bétail, aux oiseaux du ciel et à tous les animaux des champs » (Jérusalem 1993, 19).

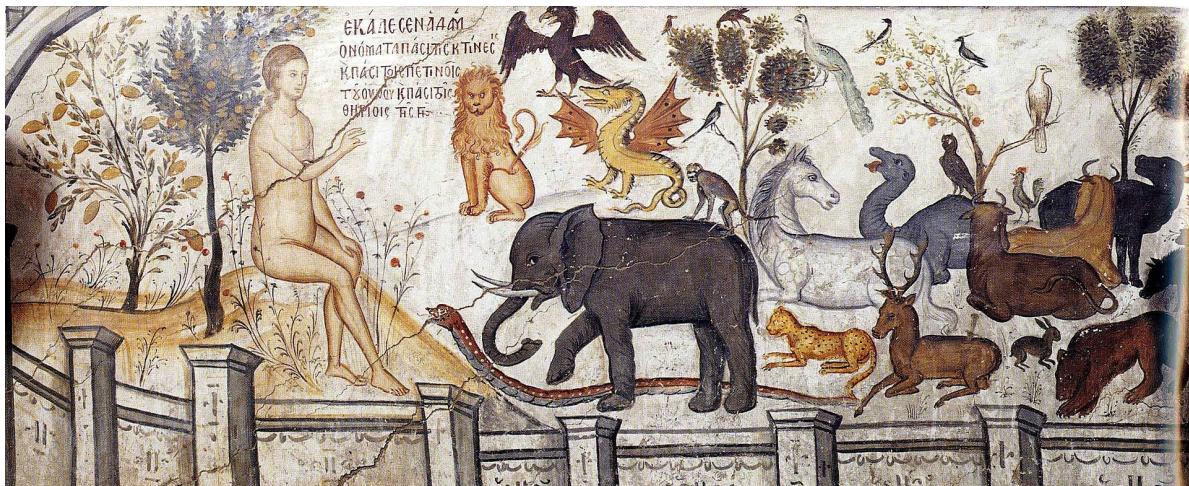


Figure 1.21: Adam nomme les animaux

Le topos fonde « en Occident la logique du tiers exclu » (Berque 2009a, 145) et caractérise les approches Naturalistes dont l'ambition est de décrire le monde de manière universelle (Philippe. Descola 2005). Du coté de l'Orient, la taoïsme apporte quant à lui une approche de l'environnement informationnel différente. Elle ne se base pas sur le nommage de l'identité mais sur l'expérience des flux informationnels, des dynamiques génératives qui caractérisent la chôra, comme en témoigne Lao Tseu six siècles avant notre ère dans les deux premiers paragraphes du Tao Te King où il insiste sur le caractère indicible du Tao :

« 1. Le tao exprimable n'est pas Le Tao. 2. Le nom énonçable n'est pas Le Nom.
» (Saint Girons 2016).

L'opposition entre ces deux approches trouvent aujourd'hui une actualité flagrante entre une IA symbolique qui fonctionne par nommage très précis des identités informationnelles sous la forme d'ontologies et une IA connexionniste qui privilégie l'émergence de modèle par des cycles récurrents d'apprentissages (Masure 2023). Entre identité et dynamiques, les deux approches sont complémentaire notamment quand on cherche à modéliser les actants qui ne sont ni réductibles à leur identité ni à leurs activités. L'actant possède une identité qu'on défini basiquement par un nom et donc par une physicalité mais l'actant est aussi producteur d'activités et donc de dynamismes génératifs en créant des rapports entre des physicalités et des concepts (Section ??). Les actants sont topoi de part les rapports qu'ils entretiennent avec les physicalités et chôra dans les processus internes d'émergence des connaissances intuitives issu

des expériences de l'espace vécu. L'actant est donc le milieu de ces pulsations existentielles que nous décrivions plus haut Figure ??, il est le continuum entre les physicalités et les intérieurités. Modéliser l'actant consiste donc à une double démarche à la fois de définition de son identité par des noms et de caractérisation des ces pouvoirs génératifs spécifiques (discerner, raisonner, agir). En ce sens la représentation de l'actant est à la fois topologique et chorématique (Reymond and Brunet 1996; Brandt 2021).

Pour représenter les actants, nous avons fait le choix de l'hexagone dont la forme est entre le rectangle qui représente les physicalités et le cercle qui représente les concepts. De plus, les six cotés de l'hexagone offrent une possibilité de trois plages de connexion vers le haut où sont représentées les physicalités et trois plages de connexion vers le bas où l'on trouve les concepts. L'hexagone est une forme qui peut être utilisée pour paver l'espace de manière fractale puisque l'intérieur de l'hexagone peut lui aussi être paver avec des hexagones (Figure ??). A l'hexagone nous avons associé un cercle qui défini l'intérieurité d'un actant (Figure ??) ou de plusieurs actants qui se partagent la même intérieurité par exemple dans le cadre d'une institution ou d'un groupe d'individu qui mettent en commun des concepts et des pouvoirs (Figure ??) ou dans le cas d'une modélisation global d'un écosystème qui possède sa propre intérieurité composée de plusieurs actants ayant eux-mêmes leurs propres intérieurités (Figure ??).

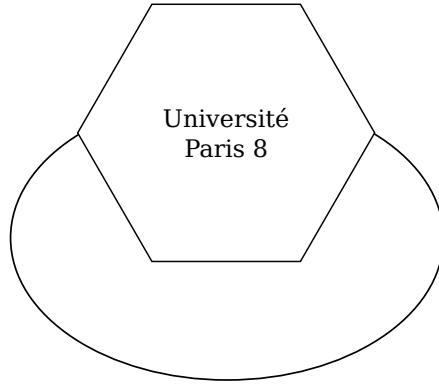


Figure 1.22: Actant individuel

Un des autres grands intérêts d'utiliser des structures hexagonales pour modéliser les actants de l'écosystème de connaissances vient du fait qu'il est possible de créer des formules logiques par positionnement des hexagones les uns par rapport aux autres. Voici par exemple comment utiliser des hexagones pour créer des diagrammes de Venn (Coumet 2020; Venn 1866) afin de modéliser l'intersection entre deux, trois et quatre ensembles d'actants :

Chaque espace défini par les intersections renvoie un ensemble particulier d'actant, par exemple Figure ?? modélise les actants qui appartiennent à l'institution Verte, à l'institution Rouge et ceux qui sont à la fois dans l'une et dans l'autre. Les diagrammes de Venn sont très pratiques pour avoir une vision d'ensemble de toutes les possibilités d'inclusion de plusieurs ensembles. Toutefois, lorsque le nombre d'ensemble est supérieur à cinq, il devient difficile de lire correctement certains cas d'inclusion comme le montre Figure ?? et même très difficile de

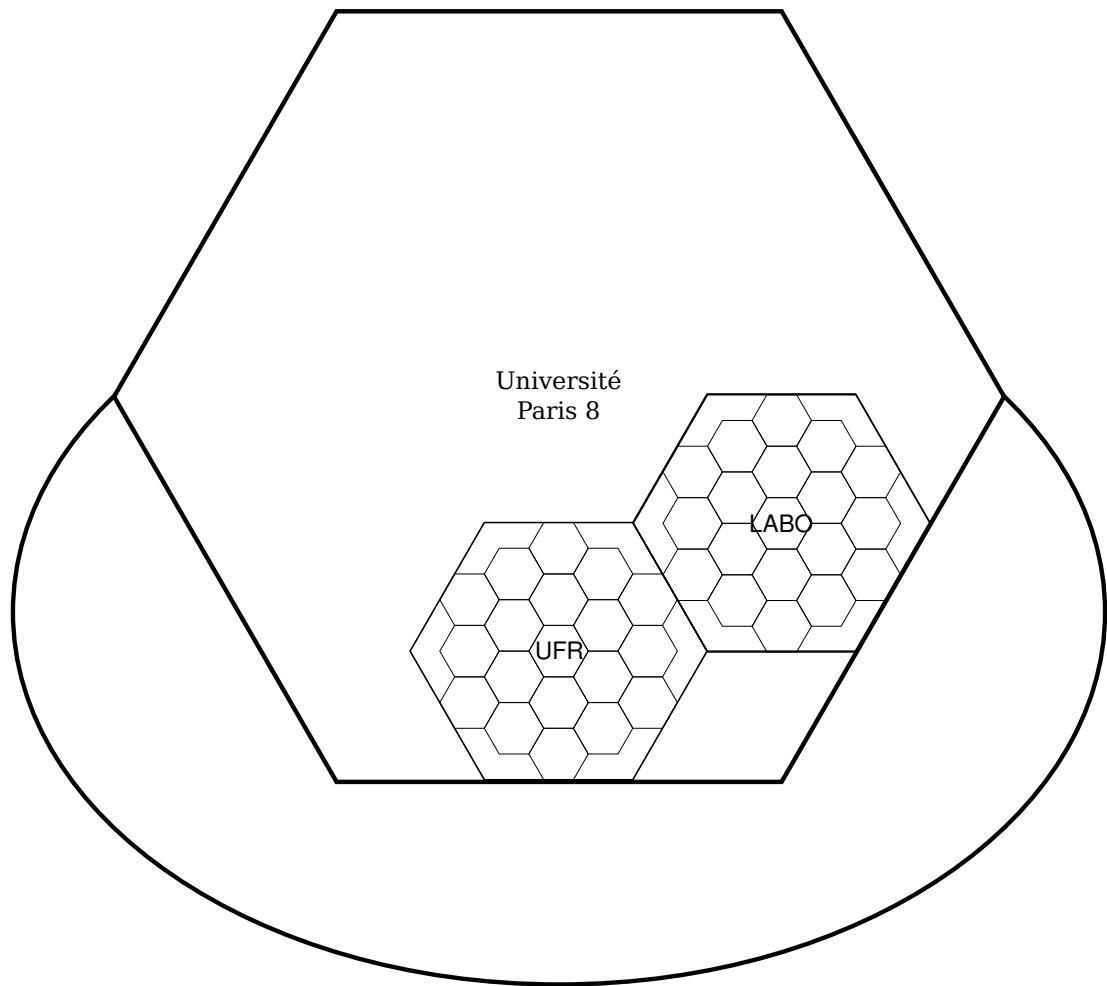


Figure 1.23: Actants d'une communauté

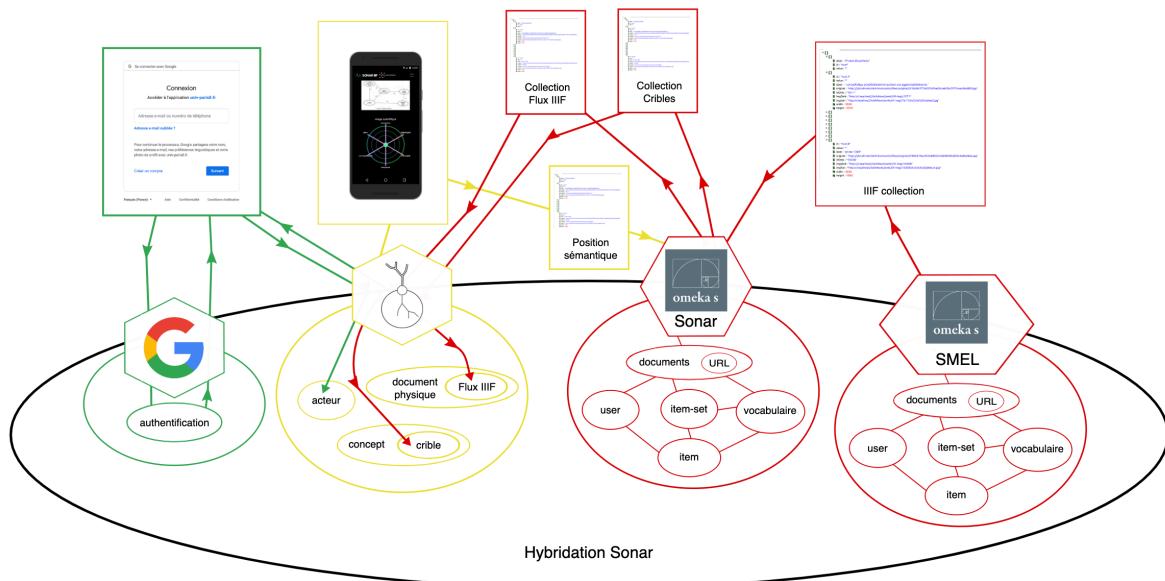


Figure 1.24: Actants dans un écosystème

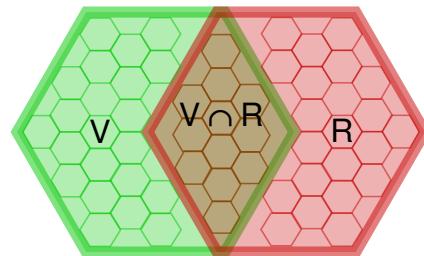


Figure 1.25: Diagramme de Venn à 2 ensembles => 3 espaces hexagonaux

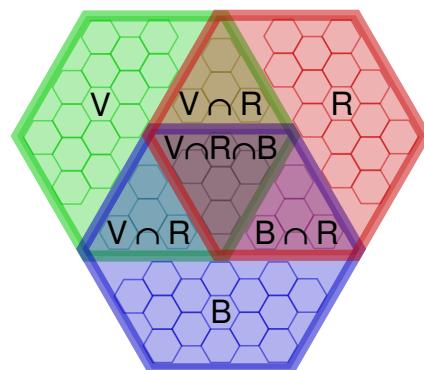


Figure 1.26: Diagramme de Venn à 3 ensembles => 7 espaces hexagonaux

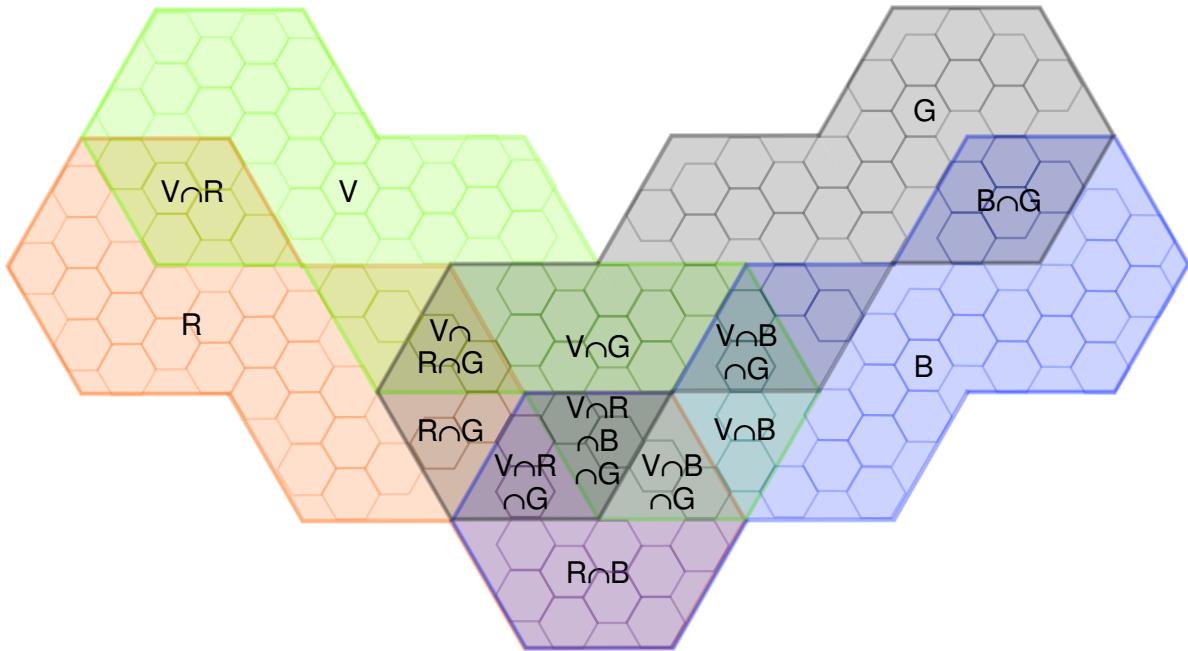


Figure 1.27: Diagramme de Venn à 4 ensembles => 15 espaces hexagonaux

les représenter sous forme hexagonale. Dans le cas où les ensembles d'actant sont nombreux, il convient donc de trouver d'autres solutions graphiques pour représenter l'appartenance d'un actant à une institution, par exemple en représentant une autre dimension existentielle : les rapports.

1.6 Les rapports : connaissances des existences potentielles

Les rapports sont la quatrième dimension de l'existence que nous utilisons pour modéliser les écosystèmes de connaissances. Ils composent la dimension existentielle qui relie les trois autres dimensions, les connaissances et les espaces qui leurs sont associés : dimension physique, connaissances des chocs et espaces matériels hiérarchiques, dimension des concepts, connaissances des essences et espaces topologiques, dimension des actants, connaissances des dynamismes et espaces topographiques chorématiques. Les rapports définissent des espaces temporels, ils transforment une virtualité en actualité. Par exemple, je suis en tant qu'actant actuellement en rapport avec la physicalité de mon clavier et de mon écran qui produisent les concepts que vous lisez. Ces rapports, ne seront plus actuels quand à la fin de cette séance de travail je commencerais une autre activité. Moi en tant qu'actant, mon clavier en tant que physicalités et ce que j'ai écrit en tant que concepts, nous existons toujours comme potentialités de futurs rapports qui s'instancieront à une date et pour une période données. En d'autres