

Infrastructures de recherche numériques et Organisation des Connaissances : Apports du Linked Open Data pour les communautés épistémiques de l'ESR

Digital research infrastructures and Knowledge Organization: Linked Open Data contributions for ESR epistemic communities

Edwige Pierot¹

¹ IMSIC, Aix-Marseille Université, epierot[at]free.fr

RÉSUMÉ. Dans cet article, nous nous intéressons aux activités d'organisation des connaissances des spécialistes de l'Enseignement Supérieur et de la Recherche dans la mesure où elles contribuent à la circulation des connaissances et à l'évolution des systèmes d'organisation des connaissances (SOC). Les domaines scientifiques auxquels ces spécialistes sont institutionnellement rattachés se recomposent dans des communautés épistémiques qui conduisent des activités similaires tant en matière de recherche d'information, de production de connaissances, que d'accès à l'information. Pour effectuer ces activités, ces communautés, notamment en sciences humaines et sociales, disposent d'outils, de services et de dispositifs qui participent du mouvement des « Humanités numériques ». Les deux grandes agences bibliographiques françaises (ABES et BnF) prennent part à ce mouvement en appliquant aux données le modèle entité-relation qui s'étend progressivement à l'indexation matière (répertoires d'autorités matières, thésaurus, etc.). Les connaissances produites par les communautés épistémiques sont donc étroitement liées aux infrastructures numériques de recherche basées sur la normalisation et la mutualisation qui composent un socle technologique distribué et commun. Tandis que la communauté scientifique exprime le caractère pluridisciplinaire, multidisciplinaire, voire transdisciplinaire des objets qu'elle étudie, nous nous intéressons aux conditions dans lesquelles le web de données peut répondre aux besoins de structuration et d'organisation des connaissances de ces communautés épistémiques. Dans quelle mesure les SOC conservent-ils une place prépondérante au sein des infrastructures des « Humanités numériques » et de leurs dispositifs sociotechniques ? Comment valoriser les propriétés heuristiques des SOC lors du processus d'indexation effectué par les spécialistes de ces communautés épistémiques ? Afin de répondre à ces questions nous étudions les orientations techniques, les composants et le fonctionnement de trois Infrastructures de Recherche (HAL, ISIDORE, Persée). Nous constatons que la production et la circulation des connaissances sont fondées sur la puissance technologique des dispositifs et soulignons la présence de nombreux SOC et une valorisation fragmentée de l'indexation matière. Aussi, nous nous intéressons aux conditions de possibilités d'une fertilisation croisée entre les SOC et proposons de nous appuyer sur le Linked Open Data (LOD) pour construire un modèle d'interopérabilité entre les SOC des communautés épistémiques afin de composer un continuum documentaire.

ABSTRACT. In this article, we focus on the knowledge organization activities of specialists in Higher Education and Research insofar as they contribute to the circulation of knowledge and the evolution of Knowledge Organization Systems (KOS). The scientific fields to which these specialists are institutionally attached are recomposed in epistemic communities that carry out similar activities in terms of information research, knowledge production and access to information. To make these activities, communities, particularly in the human and social sciences, have tools, services and devices that are part of the "Digital Humanities" movement. The two major French bibliographic agencies (ABES and BnF) are taking part in this movement by applying the entity-relation model to the data, which is gradually extending to subject indexing (subject headings, thesauri, etc.). The knowledge produced by epistemic communities is therefore closely linked to digital research infrastructures based on standardization and mutualization, which form a distributed and common technological foundation. While the scientific community expresses the multidisciplinary, multidisciplinary, even transdisciplinary nature of the objects it studies, we are interested in the conditions under which the Web of data can meet the needs of structuring and organizing the knowledge of these epistemic communities. To what extent do KOS maintain a prominent place within the "Digital Humanities" and their socio-technical devices? How to value the heuristic properties of KOS during the indexing process carried out by specialists in these epistemic communities? In order to answer these questions, we study technical orientations, components and functioning of three Digital Research Infrastructures (HAL, ISIDORE, Persée). We note that the production and circulation of knowledge is based on the technological power of the devices and highlight the presence of many KOS and a fragmented valuation of subject

indexing. So, we are interested in the conditions of possibilities for cross-fertilization of KOS and propose to use Linked Open Data (LOD) to build an interoperability model between the KOS of epistemic communities in order to compose a documentary continuum.

MOTS-CLÉS. Communautés épistémiques, Humanités numériques, Systèmes d'organisation des connaissances, Linked Open Data, Interopérabilité, Continuum documentaire.

KEYWORDS. Epistemic communities, Digital Humanities, Knowledge Organization Systems, Linked Open Data, Interoperability, Documentary continuum.

Introduction

Les spécialistes scientifiques de l'Enseignement Supérieur et de la Recherche (ESR) participent activement à la production et la diffusion de données descriptives (mots-clés, éléments de catégorisation) qui viennent compléter la production des connaissances scientifiques proprement dites. Ces opérations se présentent comme des activités d'organisation des connaissances auxquelles les chercheurs se livrent sur les demandes des éditeurs scientifiques à l'occasion de la publication d'articles, d'ouvrages, chapitres d'ouvrages, etc. Bien que moins normalisées que les opérations d'indexations que réalisent les professionnels de l'information et de la documentation lors du catalogage et du dépouillement bibliographique, elles sont similaires même si situées bien en amont des pratiques professionnelles de l'indexation (Jacobi 1993, Holzem 1999, Amar 2000, Maniez 2002, Celotti & Musacchio 2004, Menon 2007). Complémentaires, ces pratiques tendent à une double amélioration : celle de l'identification, l'organisation, la diffusion et la circulation de connaissances et celle des systèmes historiques d'organisation des connaissances (SOC) [Hodge 2000] vers de nouveaux SOC que les bouleversements numériques en matière de traitement de l'information documentaire ont contribué à faire émerger [Ertzscheid et Gazellot 2006, Le Deuff 2006, Papy 2009, Zacklad 2010].

Les organisations scientifiques de la Recherche et de l'Enseignement Supérieur se réorganisent virtuellement, au-delà des domaines scientifiques institutionnels (CNU, CoNRS), en communautés épistémiques au sein desquelles de grandes similarités existent dans les pratiques de production, de diffusion, d'accès et de recherche d'informations [Cohendet et al. 2003, Lièvre et Laroche 2014]. Participant de la constitution des SOC via le paradigme documentaire [Casemajor-Loustau 2011, Hudon 2013, Favier 2017], et produisant des instruments dédiés à l'usage [Papy et Leblond 2007], leurs pratiques d'élaboration de l'indexation matière ont trouvé une place centrale dans l'environnement documentaire (« literary warrant »). Ces descripteurs apparaissent sous la forme de « signes passeurs » [Bonaccorsi 2016] dans les dispositifs sociotechniques [Guyot 2009] que développent les Infrastructures de Recherche numériques [André et Charnay 2007, Humanum 2014, Favier 2017].

Les outils, services et dispositifs sociotechniques du mouvement des « Humanités numériques » [Dacos 2010, Presner et al. 2009, Vinck 2016] tendent à prendre place dans le web de données [Roumieux 2011, Kalfon 2016, Aymonin 2017] que les deux grandes agences documentaires que sont la BnF et l'ABES associent au revirement des modes de représentation des données catalographiques [BNF 2011, 2012, 2017] en se tournant vers le modèle entité-relation et les ontologies, pour signaler, traiter, diffuser et exposer, ce qui deviendra à terme, d'immenses volumes de données initialement bibliographiques et catalographiques. Aujourd'hui engagé pour les données d'autorité [Persee 2016, Clavel 2018] ce revirement inclut, à terme, l'indexation matière [Riva et al. 2017, Ménard 2017] traditionnellement corrélée à des SOC (thésaurus ou répertoires d'autorité matière) fondés sur une structure hiérarchique, une terminologie et une distribution sémantique [Maniez 2002, Menon 2007, Mustafa El Hadi et Hudon 2010, Zacklad 2010, Hudon, 2013].

Dès lors que la communauté scientifique exprime le caractère pluridisciplinaire, multidisciplinaire voire transdisciplinaire des objets qu'elle étudie, notamment dans les sciences humaines et sociales, la problématique de l'organisation des connaissances se pose à nouveaux frais : comment répondre, à l'appui du web de données aux besoins de structuration et d'organisation des connaissances pour

l'activité de ces communautés épistémiques ? Dans quelle mesure les SOC, dotés de propriétés hiérarchiques, de codification et d'énonciation conservent-ils une place prépondérante au sein des « Humanités numériques » et de leurs dispositifs sociotechniques ? Comment valoriser les propriétés heuristiques des SOC lors du processus d'indexation effectué par les spécialistes de la communauté épistémique ?

Nous explicitons dans un premier temps la notion de communauté épistémique et la positionnons dans l'environnement des Infrastructures de Recherche puis nous attardons sur les orientations techniques et le fonctionnement des dispositifs sociotechniques afin de mettre en évidence quelques indicateurs de l'usage de l'indexation matière et les conditions de mise en œuvre d'un continuum documentaire. Nous décrivons enfin une forme possible d'expression de ce continuum documentaire.

1. Les communautés épistémiques de l'ESR dans l'écosystème des infrastructures numériques

1.1. Une communauté épistémique, productrice et consommatrice des connaissances

L'ensemble des chercheurs et enseignants-chercheurs français est identifié comme appartenant à la communauté de l'ESR, celui-ci est réparti selon un système des sciences [Le Moigne 2007] qui peut être divisé selon des sections [Conseil National des Universités¹, Centre National de la Recherche Scientifique²]. Les membres de l'ESR relèvent tous de disciplines scientifiques reflétant un système de connaissances exprimé dans les classifications documentaires à l'instar de la Classification Décimale Dewey (CDD) : économie, épistémologie, sciences informatiques, arts, littérature, etc. [Béthery, 2013]. Les spécialistes de cette communauté sont dotés de compétences complémentaires [Gensollen 2004, Lièvre et Laroche, 2014] et liés par un projet épistémique. Ils forment une communauté épistémique qui s'identifie par la production de connaissances de façon délibérée et délibérative [Goldenberg, 2011] :

« Un groupe de représentants partageant un objectif cognitif commun de création de connaissances et une structure commune permettant une compréhension partagée. La finalité de la communauté épistémique se trouve placée au-dessus des membres de la communauté. [...] La validation de l'activité cognitive d'un représentant de la communauté est faite suivant les critères fixés par l'autorité procédurale. L'objet de l'évaluation concerne la contribution individuelle à l'effort vers le but collectif à atteindre. Dans une communauté épistémique, les représentants sont unis par leur responsabilité à mettre en valeur un ensemble particulier de connaissances » [Cohendet et al. 2003, p. 105-106]

La « structure commune » inhérente à la communauté est exprimée dans les formes normatives, techniques et sémiotiques de la médiation dont la figure procédurale [Jeanneret 2009] convoque les dispositifs et les formes de l'interaction. [Peraya 1999]. Les pratiques [Paul et Perriault 2004], rendues visibles dans les processus de peer-reviewing, de publication et diffusion scientifiques, révèlent les dynamiques qui matérialisent la légitimation, la circulation et la visibilité des connaissances produites dans cette communauté épistémique.

Le besoin intensif en connaissance est assuré depuis l'intérieur de la communauté en premier lieu par la production ou la participation à des monographies, articles, collaboration à des ouvrages de référence, etc. Cette étape de production de connaissances est complétée en mobilisant des « arts de faire », « une manière de penser investie dans une manière d'agir, un art de se combiner indissociable d'un art d'utiliser » (De Certeau 1990, p. XL) pour fournir des éléments de paratexte : index ou métadonnées d'indexation matière (mots-clés) destinés à l'identification et au repérage des connaissances produites. Cet apport descriptif s'effectue à la demande des éditeurs et des institutions

¹ <http://www.cpcnu.fr/listes-des-sections-cnu>

² <http://www.cnrs.fr/comitenational/sections/intitsec.php>

documentaires dans le cas des thèses de doctorat, et témoigne de l'importance de ces métadonnées « sujet » notamment pour l'identification, l'accès et la circulation des connaissances produites.

Le besoin intensif en connaissance de cette communauté épistémique est également satisfait par l'intervention des professionnels de l'information-documentation qui établissent les notices bibliographiques consultables à partir des OPAC³, des bases de données ou des bibliothèques numériques. Alors, la communauté de pratique [Wenger 2005] des professionnels de l'information-documentation et les spécialistes de la communauté épistémique forment :

« une communauté intensive en connaissances [, est] une communauté dans laquelle une part non négligeable de ses membres produit et reproduit la connaissance, dont les contours délimitent un espace public (ou semi-public) de circulation des savoirs et où l'usage des nouvelles technologies d'information et de communication a radicalement réduit les coûts de codification et de distribution de la connaissance. » [David et Foray, 2002, p.19].

Les pratiques de cette communauté intensive en connaissances sont mises en œuvre dans un environnement technique renforcé par la stabilité des technologies d'interopérabilité [Miller 2000, Quint 2003, Gandon 2012, Papy 2015, Direction Interministérielle du Numérique et du Système d'Information et de Communication de l'Etat 2015].

Les Infrastructures de Recherche sont positionnées dans l'écosystème technique, intellectuel et économique de la communauté épistémique et tendent à lui apporter des moyens destinés à la structuration, au traitement et à la ré-exploitation de données, etc. [Humanum 2015] auxquels s'ajoutent la production d'outils (archives institutionnelles ouvertes, édition en ligne, ...) et de méthodes (linguistique de corpus, bibliométrie) qui se situent au plus près de ses activités.

1.2. Des infrastructures inscrites dans un écosystème de données

La notion d'« Infrastructures de Recherche » résulte de l'apport de l'informatique dans les sciences humaines et sociales [Svensson 2010, Gefen 2015, Granjon et Magis 2016] ; cet apport a conduit au déploiement de méthodes, d'objets et d'outils d'analyse, de traitement, de visualisation pour des données numériques. Le phénomène s'est développé à partir des « *humanities computing* » (calcul statistique, représentation des données géographiques, etc.) pour évoluer vers les « *blogging humanities* » avec l'introduction de fonctionnalités sociales dans les dispositifs numériques (blogs de chercheurs, carnets de recherche) ; cette deuxième étape ouvre la voie aux « *multimodal humanities* » [Mac Pherson 2008], entendues comme espace de création pour de nouveaux outils (recours aux hypermédias pour l'enseignement, dispositifs sociotechniques dédiés au travail coopératif, plateformes de crowdsourcing, etc.). Le périmètre international (Canada, Etats-Unis, France, Suisse) des « Humanités numériques » [Presner et al. 2009, Dacos 2010] se décline à l'échelle européenne avec DARIAH⁴ (Digital Research Infrastructure for the Arts and the Humanities) et son réseau d'infrastructures nationales fondé sur le partage de bonnes pratiques. Comme le souligne L. Favier (2017), « *les définitions des infrastructures mentionnent toutes l'étroite imbrication des niveaux technologiques, organisationnels et de celui des contenus* » (Favier 2017, p.5).

L'engagement politique et institutionnel français en faveur du mouvement des « Humanités numériques » s'est traduit par la mise en œuvre de très grandes infrastructures de recherche (TGIR) (le TGE Adonis⁵ devenu Humanum) visant à encourager et promouvoir la standardisation des modes de production et de diffusion des données numériques en SHS⁶ [CNRS 2011, p.9] et d'Infrastructures de

³ Online Public Access Catalog

⁴ <https://www.dariah.eu/>

⁵ Très Grand Equipement ADONIS

⁶ https://www.huma-num.fr/sites/default/files/presentation_du_tge_adonis.pdf

Recherche (IR) dédiées à des questions précises : édition en ligne (CLEO⁷), archives institutionnelles ouvertes (HAL⁸), etc.

« La notion d'infrastructure est utilisée pour caractériser l'évolution des sciences dans un environnement numérique en réseau dont l'une des conséquences est la possibilité et la nécessité de partager les données de la science ainsi que celle d'en assurer la pérennité. »
[Favier 2017, p.4]

Le recours aux outils et services proposés par les Infrastructures de Recherche s'échelonnent sur une palette allant de l'incitation à déposer les résultats de la recherche dans les archives institutionnelles ouvertes (HAL) à la contrainte de production de Data Management Plan (DMP) pour les projets de recherche financés sur les programmes européens (H2020). Les infrastructures des « Humanités numériques » proposent des outils dédiés à la production scientifique (par exemple des outils techniques de balisage en vue de composer des corpus pour l'analyse de données), le recours aux institutions compétentes en matière de conservation numérique, et la diffusion, par interopérabilité entre des plateformes (HAL, Persée, Open Edition, Canal U), des ressources produites par les spécialistes scientifiques de la communauté scientifique.

« La recherche sur les infrastructures conçues pour les humanités s'est alors portée, non sur la recherche d'un langage commun, mais sur la modélisation des pratiques savantes »
[Favier, op. cit., p. 12]

Les outils et services de collecte, de traitement et d'exposition des données sur le web développés depuis la TGIR Humanum⁹ [Humanum 2014, 2015] sont complétés par l'archivage des ressources par le CINES¹⁰. C'est à partir de pratiques inscrites dans les méthodes historiques de production de connaissances que se sont construites les expertises de description et de normalisation inhérente aux dispositifs documentaires secondaires [Fayet-Scribe 2000, Hudon 2013]. Les pratiques des professionnels de l'information-documentation ont d'abord consisté dans la normalisation des descriptions conçue pour échanger des informations bibliographiques (ISBD, MARC, protocole Z39.50), le protocole OAI-PMH¹¹ a ensuite permis leur exposition directement sur le web. L'actuel mouvement de l'Open Data tend à rendre librement disponibles des jeux de données et des référentiels (afin notamment de restituer largement les résultats de l'activité de production de connaissances et permettre la réutilisation de ressources existantes à des fins de création de nouvelles connaissances) tandis que le Linked Data investit un autre axe destiné à exprimer les données en milieu semi-ouvert en les instanciant à des modèles de description (schémas de métadonnées et ontologies) [Gandon et al., 2012, Bermès et al. 2013] afin de valoriser et favoriser la visibilité des données produites auxquelles sont conférées, en vertu de l'expertise des producteurs, des garanties de qualité. Cette étape la plus récente de l'interopérabilité technique [Miller 2000, Quint 2003, Papy 2009, 2015, Gandon et al. 2012] est investie dans les travaux actuels d'alignement entre les référentiels d'autorités¹² [Clavel 2017 Ménard 2017] pour disséminer les données grâce au concours du calcul algorithmique appliqué à la sémantique formelle. D'abord fondées sur le modèle de la notice bibliographique, les descriptions des ressources sont bouleversées par l'arrivée du modèle entité-relation¹³ [BNF 2012, 2017, Claud et Duquenne 2017] qui s'appuie sur des modèles d'information (FRBR¹⁴, CIDOC-CRM¹⁵, LOM¹⁶),

⁷ <http://cleo.openedition.org/>

⁸ <https://hal.archives-ouvertes.fr/>

⁹ Très Grande Infrastructure de Recherche, <https://www.huma-num.fr/>

¹⁰ Centre Informatique National de l'Enseignement Supérieur

¹¹ Open Archive Initiative – Protocol Metadata Harvesting : http://www.bnf.fr/documents/intro_oaipmh.pdf

¹² Persée, programme national de numérisation et de diffusion de collections de documents scientifiques :

<http://www.persee.fr/documents/10099/0/Pers%C3%A9e+s%E2%80%99engage+dans+%E2%80%99%C3%A8re+du+web+s%C3%A9mantique.pdf/40431ef5-35f9-4589-8954-366080857e02>

¹³ http://www.bnf.fr/fr/professionnels/modelisation_ontologies/a.modele_FRBR.html

¹⁴ Fonctionnalités requises des notices bibliographiques

¹⁵ Modèle conceptuel de référence pour l'information muséographique

interdépendants des schémas de métadonnées (DC, MODS¹⁷, etc.). Ce modèle entité-relation est destiné à être complété par la scénarisation graphique et interactive des données dans des interfaces web.

Afin d'identifier les modes d'intervention des Infrastructures de Recherche françaises pour la production et la circulation des connaissances, nous nous intéressons à trois d'entre elles : ISIDORE¹⁸, HAL et PERSEE¹⁹.

2. Outils et services des infrastructures, quelle opérativité pour la production et la circulation des connaissances

2.1. Orientations technique et institutionnelle des Infrastructures de Recherche

L'offre technologique des Infrastructures de Recherche pour les sciences humaines et sociales est fondée sur des principes de normalisation et de mutualisation comme l'atteste la notion de « bonnes pratiques » directement liée aux recommandations émises par le consortium W3²⁰ et à la volonté de diffuser et enrichir, à l'échelle européenne à partir de DARIAH, un socle technologique distribué et commun.

C'est à partir de ce dessein que les Infrastructures de Recherche françaises ont pour mission d'accompagner la communauté épistémique dans l'usage des technologies d'infrastructure et des outils de traitement des données et résultats de la recherche scientifique. Concrètement la « grille de services » vient apporter des supports technologiques et des méthodes « *de transformation, d'analyse, d'inférence de visualisation et d'interaction [...] [pour] une double démarche de mise à disposition des données de recherche (ouverture des données, des métadonnées) et d'interopérabilité des métadonnées (normalisations, API, interface d'accès aux données)* » [Humanum, 2015].

Le Guide des bonnes pratiques [Humanum 2015] stipule que les projets émanant de la communauté épistémique doivent décliner des logiques de décision, d'organisation, de numérisation, de structuration, d'exploitation, de diffusion et de pérennisation des données et des ressources. Ces ressources ne sont plus identifiées comme des documents au sens bibliothéconomique historique²¹ et doivent être décrites à l'appui de jeux de métadonnées (techniques, descriptives, génériques, etc.). Le langage à balise XML²² sert de support technologique pour la description des schémas de métadonnées. Destiné à contribuer à l'infrastructure ISIDORE²³, plateforme de traitement et d'enrichissement des données [Humanum 2014], le schéma Dublin Core²⁴ (DC) est plébiscité parce qu'il a fait l'objet d'une normalisation (ISO 15836-2003) et que sa structure, centrée sur quinze éléments, est suffisamment générique pour accueillir la description de multiples ressources numériques. Les jeux de métadonnées du schéma DC (« simple » ou « qualifié ») sont exprimés dans le langage XML, ce qui les rend exploitables par le protocole OAI-PMH, dans des flux RSS²⁵ ou dans des modèles de description de métadonnées en RDF.

Le champ « sujet » du DC nous intéresse pour son lien direct aux questions d'organisation des connaissances puisqu'il permet d'instancier des descriptions thématiques exprimées par des mots-clefs,

¹⁶ Learning Object Metadata

¹⁷ http://www.bnf.fr/fr/professionnels/f_mods/s.mods_presentation.html

¹⁸ <https://www.rechercheisidore.fr/>

¹⁹ <http://www.persee.fr/>

²⁰ World Wide Web Consortium : <https://www.w3.org/Consortium/mission>

²¹ Toute métadonnée devient une ressource du point de vue du web de données et de l'informatique documentaire et plus encore du web de données.

²² Hypertext Markup Language

²³ <https://www.huma-num.fr/sites/default/files/guide-isidore.pdf>

²⁴ <http://dublincore.org/documents/dces/>

²⁵ Really Simple Syndication

des expressions ou des codes de classification. Le principe de bonnes pratiques associées à cet élément renvoie à l'usage d'un vocabulaire contrôlé tel que le répertoire RAMEAU²⁶. Le traitement et l'enrichissement des ressources scientifiques peuvent également être effectués à partir de schémas de métadonnées « locaux » exprimés avec XML. Proche des modes historiques de description documentaire, le schéma DC, également utilisé dans l'archive institutionnelle ouverte HAL, constitue un support privilégié pour la description, la circulation des ressources et de connaissances scientifiques collectées puis exposées dans les dispositifs des Infrastructures de Recherche.

Les fondations technologiques de ces infrastructures nous permettent d'identifier les processus sur lesquels reposent la caractérisation des ressources, les traitements qui leur sont appliqués et les restitutions auxquelles procèdent les Infrastructures de Recherche dans des interfaces web. La participation directe des chercheurs à la structuration des données en vue de leur exposition et de leur partage entre les plateformes nous permet également d'observer la place des référentiels d'indexation matière, leur mode de mise à la disposition pour la communauté épistémique ainsi que leurs effets sur les usages de dispositifs d'accès aux ressources.

2.2. Composants et fonctionnement des Infrastructures de Recherche

ISIDORE, plateforme de traitement et d'enrichissement des données directement issue de la « grille de services » d'Humanum, a également pour dessein de proposer un accès unifié aux données numériques de la recherche en SHS, cette plateforme contribue à promouvoir la communication directe entre chercheurs. Les données exploitées dans ISIDORE sont caractérisées à partir de référentiels. « *Les référentiels sont des vocabulaires, thésaurus ou listes d'autorités qui sont utilisés pour indexer les données dans le but de constituer des espaces de navigation dans les données, de faciliter les recherches et d'enrichir les données* » [Humanum 2014, p. 19]. Le guide des bonnes pratiques précise que les référentiels ne sont pas moissonnés mais déposés par leur fournisseur aux conditions d'une conformité éditoriale et technique. Compte tenu de l'évolution des référentiels, le choix de ne pas les moissonner nécessite de définir une fréquence de mise à jour afin d'écartier tout risque d'obsolescence dans un environnement par essence producteur de nouvelles connaissances. Les référentiels déposés dans ISIDORE sont le répertoire d'autorité matière RAMEAU, le thésaurus Pactols²⁷, les référentiels de HAL-SHS²⁸, la taxonomie de Calenda.org²⁹, GeoNames³⁰, GEMET³¹ et GeoEthno³², ils sont décrits avec le vocabulaire SKOS³³ qui a pour particularité de disposer de propriétés d'alignement entre référentiels. En outre, toutes les données apportées dans ISIDORE sont converties en RDF [Humanum 2014, p. 17].

Pour toute collecte de données scientifiques et dépôts de référentiels, la caractérisation automatique des ressources s'effectue par le biais d'un algorithme fondé sur une analyse morphologique des termes (comparaison du texte intégral et d'une sélection de métadonnées descriptives avec les termes contenus dans les référentiels déposés). Si une équivalence entre un élément de métadonnée de la ressource et une entrée de référentiel est identifiée, alors la ressource sera reliée à (aux) l'entrée(s) de référentiel(s). Ce mécanisme de recherche et d'appariement est destiné à promouvoir une présentation par facette (à partir des éléments du DC) sur la plateforme ISIDORE et à permettre d'affiner les résultats d'une recherche [Humanum 2014].

La contribution de la communauté épistémique à l'indexation matière est instanciée à l'appui de l'élément `dc:subject` du schéma DC. Celui-ci est répétable de sorte qu'il est possible de multiplier les

²⁶ <http://rameau.bnf.fr/>

²⁷ Peuples et cultures, Antroponymes, Chronologie relative, Toponymes, Œuvres, Lieux, Sujets

²⁸ Laboratoires, domaines scientifiques, auteurs

²⁹ Centre pour l'édition électronique ouverte

³⁰ Référentiel géographique issu du projet du même nom

³¹ Thésaurus multilingue sur le thème de l'environnement

³² Thésaurus géographique conçu pour l'indexation géographique de documents dans le domaine de l'ethnologie

³³ Simple Knowledge Organization System : <https://www.w3.org/TR/skos-reference/>

entrées matière pour caractériser une ressource ; ce sont donc autant de mots-clefs auteur et de descripteurs contrôlés qui vont constituer les accès matière des ressources décrites. Toutes les données entrantes étant converties en RDF, les URIs des différents référentiels matière peuvent techniquement être alignés en vertu des propriétés spécifiques du SKOS.

Lors de la consultation de l'interface publique de la plateforme ISIDORE l'utilisateur peut choisir entre le mode « recherche » (à partir du langage naturel) et la « navigation ». Les intitulés du mode navigationnel sont appariés aux éléments du schéma DC de même que la recherche en langage naturel. Ces deux modes vont explorer les éléments du schéma de métadonnées (sujet, titre, auteur, contributeur, description, relation ou couverture) de manière transparente pour l'usager.

Les traitements effectués sur les ressources produites consistent à compléter l'indexation matière apportée par les spécialistes de la communauté épistémique, qui s'apparente à une folksonomie [Trant 2008], selon le procédé précédemment décrit. Les SOC associés au dc:subject sont exploités de manière indépendante comme le montrent ci-dessous la notice ISIDORE et l'extrait de la notice RAMEAU correspondant à l'un des mots-clefs auteur : la vedette RAMEAU « Littérature française » n'est pas associée à cette ressource.

Capture d'écran : « Formes kaléidoscopiques : l'hybridité chez Chloé Delaume »
Source : ISIDORE (consulté le 27/02/2018)

Extrait de la notice RAMEAU pour la littérature française
Source : Bibliothèque Nationale de France (consulté le 27/02/2018)

HAL est « une infrastructure nationale mutualisée hébergeant des archives institutionnelles » [Bauin 2014, p. 3]. Créé en 2001 pour promouvoir la diffusion de la production scientifique des communautés de recherche, elle s'inscrit dans la bibliothèque scientifique numérique (BSN³⁴). HAL est une archive ouverte et repose sur une plateforme qui a pour spécificité d'être multidisciplinaire et de fédérer de multiples portails (HAL-SHS, INRIA³⁵, INSERM, IRD³⁶, etc.)

“Each specific repository has its own interface with its own design and its own search criteria; it also has a deposit interface adapted to its own needs where additional metadata related to the scientific field or the institution can be entered. For example, the MESH thesaurus is linked to the deposit of the publications of INSERM researchers” [André et Charnay 2007, p. 5].

Les sciences humaines et sociales disposent ainsi d'un portail et d'une interface dédiés. L'archivage à long terme du portail HAL-SHS³⁷, comme des autres portails, est assuré par le CINES et l'interopérabilité technique de l'ensemble de HAL est mise en œuvre à partir du protocole OAI-PMH³⁸. L'orientation bibliographique du schéma DC sert à nouveau de support à la description des ressources notamment effectuée par les auteurs. La démarche des spécialistes de la communauté consiste à promouvoir la visibilité, la circulation des connaissances, et leur accessibilité directe lorsque le texte intégral est associé à la notice.

L'observation de l'interface unique de dépôt (que celui-ci soit réalisé par un auteur, un professionnel de l'information-documentation ou un éditeur) met en évidence une correspondance entre les champs proposés et le schéma DC. Du point de vue de l'identification du thème du contenu de la ressource, les champs « mots-clés » et « classification » renvoient à l'élément `dc:subject`. Bien que les bonnes pratiques relatives à l'usage de ce champ incitent, comme nous l'avons vu, à recourir à des SOC pour effectuer les indexations analytique et synthétique, l'instrumentation offerte par la plateforme HAL ne permet au déposant d'accéder, ni à un thésaurus, ni aux autorités matières (ou aux « grappes terminologiques » de RAMEAU [Mingam 2005], ni aux classes, divisions et sections d'une classification décimale.

La consultation de l'interface publique de la plateforme HAL met en évidence la correspondance entre les éléments du schéma de description des ressources et les champs de recherche (auteur, titre, langue, collection, projet, type de document, identifiant pérenne, etc.) ; lors des requêtes au serveur ce sont ces appariements de contenus qui sont analysés. HAL propose des fonctionnalités avancées de recherche d'information avec le recours aux API³⁹ pour composer et spécifier des requêtes (sur des portails, collection, champ spécifique, etc.) ; ces requêtes par API permettent d'identifier la présence de classifications référencées (d'institutions –AERES-, d'éditeurs –Elsevier, ACM-, de producteur de base de données –PASCAL-). La recherche d'information fondée sur le recours aux API met en évidence le préjudice causé par l'absence de référentiels associés au `dc:subject` pour guider le membre de la communauté épistémique. En effet, celui-ci ne dispose pas nécessairement de l'expertise métier des deux autres acteurs des dépôts dans l'archive institutionnelle HAL (professionnels de l'information-documentation ou éditeurs).

PERSEE est une bibliothèque numérique [Fargier 2014] initiée par Ministère de l'Education Nationale, de l'Enseignement Supérieur et de la Recherche en 2003, elle s'inscrit dans le mouvement de diffusion numérique des publications scientifiques mis en œuvre par les éditeurs commerciaux [Tosello-Bancal 2007] et participe d'une réponse à l'Initiative de Budapest de 2002. Le portail Persée

³⁴ <http://www.bibliothequescientifiquenumerique.fr/>

³⁵ Institut National de Recherche en Informatique et en Automatique

³⁶ Institut de Recherche et Développement

³⁷ <https://halshs.archives-ouvertes.fr/>

³⁸ <https://api.archives-ouvertes.fr/docs/oai>

³⁹ Application Interface Programming

se donne « *pour ambition de donner à l'exploitation et à l'archivage pérenne [...] une place centrale afin de garantir l'utilisabilité et l'accessibilité à long terme du patrimoine scientifique. [...] Persee.fr est appelé à devenir une plateforme de recherche dédiée à la valorisation du patrimoine documentaire* » [Fargier 2014, p.2] et de valoriser la recherche francophone à l'international. Depuis 2014 Persée bénéficie d'un label d'Infrastructure de Recherche, CollEx-PERSEE, en raison de son rôle d'opérateur pour l'organisation de programmes de numérisation et d'administration de la plateforme de diffusion et d'accès aux collections⁴⁰.

La patrimonialisation des collections de revues des sciences humaines et sociales est effectuée selon un double processus de numérisation qui permet d'utiliser le portail pour la navigation et pour la recherche [Tosello-Bancal 2007]. Si la conservation de l'identité visuelle des revues est assurée par la numérisation en mode image, la numérisation des ressources en mode texte et leur structuration au moyen des langages d'interopérabilité du Web promus par le W3C favorise une importante granularité dans la collection et rend possible la recherche en texte intégral.

Les exigences d'interopérabilité, liées aux activités partenariales de Persée, ont conduit les concepteurs à s'appuyer sur des schémas de métadonnées (DC, MODS⁴¹ et marcXML⁴²) pour élaborer les descriptions des ressources. Bien qu'elle soit qualifiée de bibliothèque numérique par ses concepteurs, la collection de Persée n'est reliée à aucun SOC, de sorte que les ressources ne bénéficient pas d'accès matière. Les fonctionnalités de recherche offertes à l'utilisateur reposent dès lors sur leur positionnement dans les collections (revue, fascicule, article) et sur l'exploration du texte intégral.

Entre 2005 (ouverture du portail) et 2015, la progression du nombre de revues numérisées s'accompagne de l'émergence de deux tendances : « *la nécessité d'une performance accrue des outils et des services et un déplacement de l'intérêt pour les documents vers une exploitation des données* » [Fargier 2014, p.4]. Le recours à la technologie RDF, l'annonce de l'ouverture du triplestore data.persee.fr⁴³ puis la fourniture de pages « autorités »⁴⁴ pour les auteurs illustrent une transition : au portail, initialement valorisé comme une bibliothèque numérique, les administrateurs de Persée souhaitent superposer une plateforme offrant un large corpus de recherche inscrit dans le web de données et ainsi fournir des moyens complémentaires de production de connaissances à la communauté épistémique des spécialistes scientifiques. La visée d'exploitation et d'exposition des données est engagée par les collaborations relatives à l'alignement de référentiels : autorités auteurs de l'ABES, fiche auteur de data.bnf.fr⁴⁵, référentiel international de biodiversité, référentiel des monuments classés du Caire. Délaissant les questions d'organisation des connaissances, Persée exploite efficacement la recherche en texte intégral et puise dans des référentiels externes pour accroître sa visibilité sur le web. En s'inscrivant au plus près du mouvement de coopération entrepris entre les institutions documentaires historiques (BnF, ABES) pour élaborer un fichier national des autorités (FNE), Persée bénéficie d'une visibilité accrue censée drainer de nouveaux flux d'utilisateurs et ouvrir la voie à de nouveaux usages des ressources en ligne.

La consultation de l'interface web de Persée, permet d'identifier deux modes d'accès aux ressources : le mode navigationnel prend pour point de départ une répartition disciplinaire des revues et peut être complété par la formulation d'une requête dans les divers niveaux de granularité de la structure intellectuelle du portail. Lors des requêtes, l'absence d'indexation matière est compensée par la recherche dans le texte intégral à chacun des niveaux de granularité. Le puissant appareillage

⁴⁰ <http://www.enseignementsup-recherche.gouv.fr/cid99665/collections-d-excellence-pour-la-recherche-%C2%96-persee-coll-ex-persee.html>

⁴¹ http://www.bnf.fr/fr/professionnels/f_mods/s.mods_presentation.html

⁴² http://www.bnf.fr/fr/professionnels/formats_catalogage/a.f_marc_xml.html

⁴³ <http://data.persee.fr/ressources/le-triplestore-de-persee/>

⁴⁴

<http://www.persee.fr/documents/10099/0/Pers%C3%A9e+s%E2%80%99engage+dans+l%E2%80%99%C3%A8re+du+web+s%C3%A9mantique.pdf/40431ef5-35f9-4589-8954-366080857e02>

⁴⁵ <http://data.bnf.fr/>

technologique de cette infrastructure ne saurait cependant compenser le besoin d'identification du caractère multidimensionnel des ressources attaché à la complexité des objets de recherche des spécialistes scientifiques de la communauté épistémique.

Malgré des exploitations incomplètes et des pratiques laissées à la discrétion des spécialistes de la communauté épistémique, l'actualité des questions d'organisation des connaissances reste manifeste dans les Infrastructures de Recherche observées : absence compensée par un traitement morphologique (Persée), caractérisation effectuée à l'aide du langage naturel, co-présence de langages documentaires et libre (HAL, ISIDORE). En outre, le rôle de médiation [Peraya 1999] des outils d'indexation matière est confirmé par leur scénarisation sous la forme de « signes passeurs » [Bonaccorsi 2016] dans les notices. Une approche sémiotique nous permet d'identifier la présence d'attentes d'usage vis-à-vis de ces ressources dédiées à l'organisation des connaissances. En écho au mouvement d'alignement de référentiels (Inist-BnF, ABES-Persee, ABES-BnF) et au projet de constitution d'un fichier national des autorités (FNE), la production et la circulation des connaissances sont massivement appréhendées à partir de la puissance technologique des dispositifs. Les « Humanités numériques » ayant vocation à faire dialoguer les disciplines des sciences humaines et sociales, notamment par l'étude d'objets de recherche [Davallon, 2004], nous nous attachons à comprendre les formes, les conditions et les effets de cette médiation dans les Infrastructures de Recherches dédiées à la communauté épistémique.

3. De la flexibilité des points de vue au continuum documentaire

3.1. Quelques indicateurs de l'usage dans les Infrastructures de Recherche

Le concept de dispositif sert de support à l'étude de l'objet technique et permet d'observer et de décrire les activités (partagées entre l'action et le contrôle) et les formes ergonomiques, anthropologiques et cognitives sur lesquelles repose l'interactivité avec l'utilisateur [Peeters et Charlier 1999, Peraya 1999, Caelen 2003]. La médiation qui s'opère dans l'environnement procède d'inscriptions normées, techniques, sémiotiques et symboliques qui qualifient les usages possibles et réels de dispositifs sociotechniques [Jeanneret 2009 ; Appel et al. 2010, Papy, 2015].

Les dispositifs sociotechniques adossés aux Infrastructures de Recherche résultent de directives politiques empruntées des vues mondiales attribuées à l'économie de la connaissance [Foray 2009]. Celle-ci est entendue comme le point de pivot entre des potentialités technologiques et organisationnelles, et des contenus à forte valeur cognitive permettant de produire continuellement, et dans des cycles courts, de nouvelles connaissances [Powell et Snellman 2004]. Les attentes de production de connaissances, en prise directe avec les activités et les usages de la communauté épistémique, nous invitent à interroger le processus social de la médiation sociotechnique [Jeanneret 2009, Peraya 1999] en nous fondant sur les besoins d'opérativité de la communauté épistémique. Dès lors, nous inscrivons notre analyse des dispositifs sociotechniques dans la théorie de l'activité [Leontiev 1975, Clot 1999, Desfriches-Doria 2015] et prenons appui sur la théorie instrumentale et dialectique [Rabardel 1995] pour discerner les registres de l'usage dans les Infrastructures de Recherche.

Ne disposant pas de résultats d'enquête ethnographique et de logs de connexion d'un panel d'utilisateurs, nous observons l'accessibilité manipulative des interfaces selon une approche sémiotique [Chevalier 2013, Seammer 2015, Bonaccorsi 2016]. Les interfaces des dispositifs sociotechniques HAL, ISIDORE et Persée exploitent la technique de l'hypertexte pour restituer leur modèle structurel. Afin de procéder à une recherche d'information, chacune des trois infrastructures explicite son modèle conceptuel en recourant à des « facettes ». Cette présence, sous une forme fléchée, du système d'organisation des connaissances rangianthien est scénarisée à l'appui de « signes passeurs » [Bonaccorsi 2016]. Dans chacun des dispositifs les « facettes » interviennent également pour raffiner les résultats d'une recherche. Dans Persée par exemple, les « facettes » vont permettre de spécifier l'instanciation des ressources selon des typologies éditoriales, linguistiques, actoriales, temporelles ou

illustratives. Dans HAL l'instanciation des données dans le schéma DC permet de naviguer dans les typologies de ressources, les domaines scientifiques, les organismes de rattachement, des auteurs, laboratoires, langues, mots-clefs, types de dépôt ou années. La mise en forme de ce raffinement exprime une « rhétorique du texte numérique » [Seammer 2015] destinée à scénariser et organiser les résultats du traitement morphologique des lemmes ou de l'analyse des instances des éléments du schéma DC. En outre, cette « rhétorique du texte numérique » présente une forte homothétie avec les caractéristiques manipulatoires déployées dans d'autres dispositifs numériques, qu'ils soient à vocation ludique, administrative, ou économique, et ce malgré l'existence de différences structurelles propres à chaque site web [Chevalier 2013].

Les fonctionnalités et actions dédiées à l'organisation des connaissances dans les Infrastructures de Recherche observées sont centrées, dans ISIDORE, sur la syndication de contenus depuis la requête vers une plateforme extérieure et sur la visualisation des occurrences de termes descripteurs des différents référentiels thématiques. L'activation du « signe passeur » apposé sur le descripteur permet d'accéder à la documentation du concept (définition, forme préférentielle multilingue, concept lié, notes). Sur la plateforme ISIDORE il est également possible d'effectuer des rebonds à partir des mots-clefs auteur, du référentiel de disciplines ou la taxonomie de Calenda.

Dans HAL les fonctionnalités dédiées à l'organisation des connaissances sont principalement orientées vers le raffinement de la requête en vertu de liens hypertextuels apposés sur les instances du `dc:subject` (domaines, mots-clefs) et la constitution d'une bibliothèque personnelle. Cette dernière fonctionnalité, possible pour un usager inscrit, permet de créer des classes pour organiser des notices bibliographiques. L'utilisateur effectue cette opération sans que l'instrumentation lui propose l'assistance d'un référentiel d'indexation.

Dans Persée, l'absence d'indexation empêche toute possibilité de proposer des fonctionnalités associées aux outils d'organisation des connaissances. Afin d'apporter néanmoins des éléments de réponse à ce besoin, Persée prend appui sur l'élément `dc:source` pour exposer, sous forme de « signes passeurs », les intitulés lexicaux des fascicules thématiques et amener à la visualisation des intitulés des contenus scientifiques des numéros thématiques. Bien que ce substitut apporte une opérativité à la recherche d'information, il ne permet pas d'accéder à une description du thème du contenu et aucunement à la description du caractère multidimensionnel des articles contenus dans le fascicule. De manière à nouveau substitutive, la lemmatisation des contenus textuels permet d'effectuer des recherches par « concept » ou « mot-clef » dans le texte intégral à partir de la barre de recherche (par l'analyse des chaînes de caractères de la requête). Ni ISIDORE ni Persée ne proposent d'espace personnel à leurs usagers, cette absence d'instrumentalisation dédiée aux « arts de faire » inhérents à la communauté épistémique (sélection et collecte de ressources) renvoie les actions d'organisation des connaissances vers d'autres dispositifs.

Cette succincte lecture sémiotique, centrée sur les questions d'organisation des connaissances dans les interfaces de trois dispositifs sociotechniques dédiés aux « Humanités numériques » confirme la valorisation fragmentée de l'indexation matière malgré la présence de nombreux SOC : « facettes », référentiels de domaines, taxonomies, folksonomies, classifications professionnelles. La coexistence constatée de modes distincts d'indexation nous conduit à souligner qu'« *en favorisant la flexibilité des points de vue, le risque s'accroît de mettre de l'avant, là encore, une perspective et une vision locale et de faire obstacle à l'interopérabilité sémantique des systèmes d'information* » [Hudon et Mustafa El Hadi 2010, p. 26]. Nous nous intéressons dès lors aux conditions de possibilité d'une fertilisation croisée de ces points de vue et du continuum documentaire qui en résulterait.

3.2. Les conditions de mise en œuvre d'un continuum documentaire

Dans les communautés épistémiques de l'ESR, toutes les caractérisations associées aux ressources produites s'inscrivent résolument dans le système des sciences. Cet enracinement commun aux membres des communautés épistémiques autorise à parler d'une co-production contigüe d'indexation.

Les référentiels d'indexation identifiés dans les dispositifs sociotechniques des « Humanités numériques » (HAL, ISIDORE) restent cependant encore disjoints les uns des autres (classifications, répertoire de vedettes matières, thésaurus, folksonomie). En conséquence, les membres de la communauté épistémique ne peuvent pas encore exploiter toute la puissance heuristique des SOC [Hodge 2000, Zacklad 2010, Hudon 2013] : celle-ci s'exprime à l'appui d'une structure hiérarchique, d'un lexique et de relations sémantiques qui permettent le repérage, l'orientation, la navigation, l'indexation, la recherche et le raffinement de la recherche d'information [Arsenault 2006, Holzem 2006, Feyler 2007, Ihadjadène et Favier 2008]. L'indexation analytique que véhiculent les folksonomies [Gazellot et Ertzscheid 2006, Le Deuff 2006, Trant 2008] « [désigne] un processus d'organisation collaborative à l'aide de mots-clés librement choisis. Ce mode d'indexation dite sociale, par opposition à l'indexation normée (métadonnées normées, listes d'autorités, thésaurus, ...) trouve ses origines dans les pratiques classiques (organisation et classification des connaissances) mais aussi dans des pratiques métier d'ordre techniciste (conservation et accès aux documents, métadonnées, Dublin Core) » [Hudon et Mustafa El Hadi 2010, p. 30-31]. Les recherches portant sur les folksonomies se sont notamment attachées à identifier leurs composants structurels : Golder et Huberman (2006) ont souligné la présence d'étiquettes qui qualifient des catégories existantes, et dessinent, en creux, la présence d'une hiérarchie. Les recherches sur les folksonomies ont également souligné plusieurs caractéristiques : l'étiquetage comme processus reposant sur le choix du vocabulaire, la folksonomie comme résultat d'un vocabulaire collectif, l'inscription de l'indexation sociale dans un contexte sociotechnique [Trant 2008, citant Marlow et al 2006], confirmant l'appartenance des folksonomies à la catégorie des systèmes d'organisation des connaissances [Francis et Quesnel 2007, Zacklad 2010, Hudon et Mustafa El Hadi 2010, Durieux 2010].

La fonction d'accès reconnue aux SOC, leur présence dans les dispositifs sociotechniques et leur rôle décisif pour la conduite des actions d'organisation des connaissances questionnent leur accessibilité. Cette dernière recouvre un aspect juridique essentiel : en effet, le passage d'une co-production contigüe d'indexation à un continuum documentaire ne peut être effectif que par l'usage de SOC librement accessibles dans un environnement ouvert reposant sur des données libres de droits, structurées et disponibles dans un format ouvert. Cet environnement, préambule nécessaire à un continuum documentaire entre des SOC, commande d'adopter les principes du Linked Open Data (LOD) [Bermès et al. 2013] qui fixe cinq conditions indispensables à la réutilisation démocratisée et pérenne de ces SOC, et à leurs alignements. L'existence puis l'accessibilité juridique, par exemple de la folksonomie produite par les spécialistes de la communauté épistémique, requièrent leur participation ainsi que leur adhésion pour dépasser les enclosures liées à la propriété intellectuelle [Dujol 2017] : en acceptant d'inscrire l'activité d'indexation sous licence libre⁴⁶, la communauté épistémique participerait de la production de biens communs informationnels qu'elle restituerait, en vertu du LOD, à l'ensemble de sa communauté et au-delà. L'inscription de SOC dans le linked Open Data est déjà mise en œuvre pour des référentiels tels que le répertoire d'autorités matières RAMEAU ou le thésaurus de l'Union Européenne EUROVOC.

L'existence d'un continuum documentaire, en plus de la communauté épistémique et du recours à des SOC librement disponibles et réutilisables, requiert de mobiliser des dispositifs sociotechniques [Appel *et al.* 2010, Guyot 2009, Ihadjadène et Chaudiron 2008] ainsi que les technologies d'interopérabilité [Quint 2003, Amar et Menon 2011, Papy 2015]. Ces dernières peuvent être envisagées sous le contrôle de l'usage et ouvrir la voie à une nouvelle hybridité permettant d'élaborer un dispositif sociotechnique instrumental [Rabardel 1995]. En effet, ces technologies (les URI, les flux RSS, le formalisme RDF), distribuées dans les plateformes existantes, sont adaptées à l'utilisation dynamique et modulable des outils d'organisation des connaissances et peuvent permettre d'en accroître l'utilisabilité (ISO 9241-1998). Enfin, l'existence d'un continuum documentaire, s'il commande la présence de la communauté épistémique, appelle aussi la prise en compte de ses « arts de faire » [De Certeau 1990], ces actions et méthodes déployées pour effectuer les actions d'organisation

⁴⁶ Licence ouverte de l'Etat, licences Creative Commons, Licence Open Data Commons Attribution
© 2019 ISTE OpenScience – Published by ISTE Ltd. London, UK – openscience.fr

des connaissances : choix de mots-clefs pour un article, sélection des entrées de l'index d'une monographie. Ces « arts de faire » peuvent être soutenus et accompagnés à l'appui de propriétés instrumentales inscrites dès la conception des dispositifs sociotechniques [Caelen 2003, Agostinelli 2009, Mustafa El Hadi et Favier 2015, Verlaet 2015]. Alors la relation entre l'outil et l'activité des spécialistes de la communauté serait fondée sur une logique d'utilisation où l'outil accompagne le projet de l'utilisateur pour en devenir l'instrument [Rabardel 1995] : cette relation entre activité et fonctionnalités du dispositif sociotechnique, entre les propriétés de l'objet et les propriétés de l'action est aux fondements de la conception de dispositifs sociotechniques « instrumentaux ».

Cette convergence entre communautés, technologies démocratisées, dispositif sociotechnique à teneur instrumentale et SOC librement disponibles augure un appareillage numérique à l'épistémè que M. Foucault définissait en 1969 comme « *l'ensemble des relations qu'on peut découvrir, pour une époque donnée entre les sciences quand on les analyse au niveau des régularités discursives* » [Foucault, 2008 : 250]. Ces conditions de mise en œuvre d'un continuum documentaire par les membres des communautés épistémiques de l'ESR permettent d'envisager l'actualisation d'une interopérabilité de type épistémique caractérisée par une instrumentation de la distribution de la continuité documentaire. Cette instrumentation prend appui sur des fonctionnalités dédiées au développement, à l'exploration et à l'alignement de SOC dont les périmètres complémentaires peuvent répondre aux différents besoins d'indexation d'objets de recherche devenus multidimensionnels. Cet appareillage numérique de l'épistémè définie par M. Foucault peut prendre forme dans un modèle hybride d'indexation dont nous proposons d'explicitier une possible mise en forme.

4. Le modèle de jonction sémantique

Afin de concrétiser le continuum documentaire, les différents référentiels servant de support à l'indexation doivent être reliés et répondre à l'exigence d'interopérabilité sémantique, cette

« capacité qu'ont certains agents, services, systèmes et applications d'échanger des données, de l'information et des connaissances en préservant l'intégrité et la signification pleine de celles-ci. L'interopérabilité sémantique concerne le langage et le vocabulaire utilisés dans un processus de communication (entre systèmes informatiques ou entre humains et systèmes informatiques) ; elle facilite pour l'utilisateur le repérage et le partage d'information, peu importe le vocabulaire et le langage utilisés. L'interopérabilité sémantique passe par le contrôle plus ou moins strict des significations, contrôle associé aux systèmes d'organisation des connaissances (SOC) de types verbaux ou symboliques » [Hudon, 2012, 217].

Le vocabulaire SKOS [W3C 2009] est un support approprié pour effectuer cette nécessaire interconnexion entre les langages documentaires : ses primitives ont notamment pour fonction de décrire la structuration de la sémantique produite par les professionnels de l'information ou par les spécialistes de la communauté épistémique. Les propriétés d'alignement vont ensuite permettre d'aller au-delà de l'interopérabilité sémantique en accédant à une interopérabilité de type épistémique distribuant le continuum documentaire.

4.1. Composants théoriques du modèle de jonction sémantique

Le modèle de jonction sémantique relève des paradigmes usager [Favier 2002, Ibekwe-SanJuan 2012, Vidal 2012] et acteur [Flichy 2008, Kennel 2013]. Du point de vue théorique, il s'agit d'un modèle hybride d'indexation qui s'inspire de la théorie instrumentale et dialectique de Rabardel (1995) : en mettant à la disposition de l'utilisateur des fonctionnalités qui explicitent les systèmes d'organisation des connaissances, le modèle de jonction sémantique cherche à produire de la signification à des fins de repérage, d'orientation, de navigation, d'indexation et de recherche d'informations. L'instrumentation dégage une valeur heuristique en guidant et supportant l'aboutissement du dessein :

identification et affichage d'un descripteur potentiellement identique et de son langage documentaire d'appartenance, accès au réseau sémantique de ce descripteur, etc. Quant à l'instrumentalisation (partie de l'activité tournée vers le sujet), elle tend à générer des interactions menant à une documentarisation épistémiquement riche propulsant le dispositif au rang d'instrument cognitif : elle permet par exemple d'acter la liaison d'un descripteur (depuis l'espace de noms duquel il provient) à un descripteur appartenant à un autre langage documentaire, améliorant à la fois le réseau épistémique et le profil sémantique du descripteur. Le modèle de jonction sémantique illustre ici la puissance d'une architecture cognitive articulant des SOC dans la mesure où elle supporte des fonctions d'orientation et de contextualisation pour accompagner l'interprétation signifiante de l'espace intellectuel d'instanciation d'une ressource.

En outre, les outils d'organisation des connaissances, lorsqu'écrits dans les standards du web sémantique deviennent eux-mêmes un composé de ressources à la fois enchevêtrées et individuelles, stabilisées par leurs URI. Ils sont exploitables par la classification à facettes qui partage avec les autres systèmes d'organisation des connaissances le principe de la dépendance conceptuelle : « *Evidence is presented that the facet concept has claim to be considered as method of subdivision that is cognitively available to human beings, regardless of language, culture or academic discipline. The possibility that faceting is a universal method of subdivision enhances the claim that facet analysis is an unusually useful design principle for information retrieval in any field. This possibility needs further investigation in an age when information access across boundaries is both necessary and possible* » [Beghtol, 2006, cité par Mustafa El Hadi 2013, p. 728]. Dans le contexte des « Humanités numériques », alors que les spécialistes scientifiques des sciences humaines et sociales expriment le caractère pluridisciplinaire, multidisciplinaire voire transdisciplinaire des objets qu'ils étudient, la reconnaissance du caractère multidimensionnel d'une ressource, son explicitation par l'indexation matière, pourrait être restituée à l'appui de facettes :

« l'utilisation de facettes pour la représentation des sujets dans un domaine permet de rapprocher les techniques d'analyse documentaire des besoins et des stratégies de recherche des usagers / chercheurs » [Hudon et Mustafa El Hadi 2017, p. 17]

Compte tenu de l'inscription de nos travaux dans les standards du web sémantique (LOD), les entités, désignées par leurs URI, peuvent être organisées en facettes, catégories correspondant aux différents aspects des sujets décrits ; pour ce faire, il conviendra de faire évoluer les primitives du SKOS [Pierot 2017]. L'inscription des sujets, dans des facettes offre dès lors la possibilité de valoriser cette caractérisation contextualisée, de compléter les accès en soulignant le « profil épistémique » des ressources.

Au plan méthodologique, le modèle de jonction sémantique inscrit son action dans un web socio-sémantique [Cahier et al. 2006] et vise à accompagner la caractérisation des ressources numériques selon trois axes : le réemploi de descripteurs existants, la création de descripteurs, le multilinguisme.

Le premier axe est dédié aux référentiels existants, tel le répertoire d'autorité matière RAMEAU, librement disponible selon les standards du web sémantique. Le travail de révision systématique a mis en exergue l'importance de la « grappe terminologique » [Mingam 2005] composée de la vedette et de ses termes exclus. La description de RAMEAU en SKOS / RDF a permis de conserver sa structure, notamment en maintenant les formes rejetées via des labels alternatifs. Si la constitution du FNE et la mise en œuvre du modèle FRBR⁴⁷ questionnent quant au maintien de la sémantique du vocabulaire RAMEAU (laquelle participe manifestement de ses propriétés heuristiques) il n'en demeure pas moins que l'inscription des descripteurs dans des facettes permettra de compléter le réseau sémantique des sujets.

⁴⁷ Fonctionnalités requises des notices bibliographiques

Le deuxième axe concerne l'enrichissement sémantique pour lequel nous identifions deux étapes : la création d'une entité, c'est-à-dire la création et la description des propriétés d'un `skos:concept` dans l'espace de noms de la folksonomie et l'enrichissement sémantique par liaison au(x) descripteur(s) d'un ou plusieurs autres SOC (propriétés d'alignement du vocabulaire SKOS). Cette fertilisation croisée permet de conserver l'intégrité et l'autonomie de chacun des référentiels.

Enfin, l'axe tourné vers le multilinguisme, indissociable des environnements web et des activités de la communauté épistémique, peut être pris en charge dans le modèle de jonction sémantique puisque des thésaurus, tel Eurovoc ou le Thesaurus for the Social Sciences sont déjà disponibles selon les principes du LOD. Ici encore, les propriétés d'alignement du SKOS permettent d'effectuer cette fertilisation croisée de différents langages documentaires [Gandon 2012, p. 127, W3 2009].

Du point de vue de l'action et des « arts de faire », il ne s'agit pas de créer systématiquement de nouveaux descripteurs mais d'identifier leur existence et, dans le cas opportun, de les employer. À cette fin, le modèle de jonction sémantique devra disposer de l'auto-complétion et faire apparaître les descripteurs existants dans les SOC. Parallèlement, un second point méthodologique nous semble crucial, il concerne la visualisation de la sémantique existante dans les SOC et la scénarisation des réseaux sémantiques que nouent les termes d'indexation. Les utilisateurs du modèle de jonction sémantique pourront ainsi bénéficier de toute la puissance heuristique de SOC « étendus ». Ces directions théoriques, méthodologiques et pratiques esquissent les contours d'un dispositif sociotechnique à teneur instrumentale pouvant être construit avec les technologies d'interopérabilité disponibles.

4.2. Instrumentation technique du modèle de jonction sémantique

Les principes de séparation entre forme et contenu, la modularité et l'extensibilité des langages à balises fondés sur XML autorisent la construction de nombreuses combinaisons technologiques [Quint 2003, Papy 2009 et 2015], cette souplesse et la garantie de pérennité et d'accessibilité aux données fondent notre choix de ce langage générique (et sa galaxie) pour construire le modèle de jonction sémantique afin que soient préservées ses qualités d'indépendance, d'interopérabilité et d'évolutivité. Le modèle de jonction sémantique repose sur une architecture distribuée dont les ressorts sont largement éprouvés : modèle de structure guidant le processus de création des documents, protocoles de requêtes, structuration des métadonnées, entrepôts de données, langages web, technologies web.

Comme indiqué, la structure de concepts SKOS, nous intéresse pour ses propriétés de description et d'alignement puisque le modèle de jonction sémantique veut interconnecter différents langages documentaires. Les propriétés du SKOS serviront à déclarer les attributs et valeurs à teneur conceptuelle des SOC. Le modèle de jonction sémantique pourrait être exploité à partir d'une API⁴⁸ [Nanard et Nanard, 2002] reliée à plusieurs noyaux fonctionnels auxquels l'interface abstraite pourra expédier les requêtes selon les protocoles et méthodes adéquats. La technologie RSS⁴⁹ pourra également être mobilisée pour appeler les référentiels selon plusieurs niveaux de granularité (une classe, une section, une division de classification par exemple).

Le DOM⁵⁰ interviendra pour qu'XSL⁵¹ effectue les traitements nécessaires à la restitution des composants (facettes, scénarisation multi-média) dans l'interface de l'instrument *via* les langages web. Ce pourtour technologique et la dynamique véhiculée par les technologies et langages du web, mobilisés pour l'usage, apportent sans conteste un ensemble propice à la mise en œuvre de genèses instrumentales [Rabardel 1995].

⁴⁸ Application Programming Interface

⁴⁹ Really Simple Syndication

⁵⁰ Document Object Model

⁵¹ eXtensible Stylesheet Language

Conclusion

La communauté de l'Enseignement Supérieur et de la Recherche est définie à partir de la notion de communauté épistémique en raison des actions de production et de participation à la circulation des connaissances (fourniture d'index et de mots-clefs auteur). L'informatisation des activités et la maturité des technologies informatiques ont introduit de nouvelles méthodes éditoriales, d'exposition et de recherches pour la communauté épistémique des spécialistes scientifiques. Leur mise en œuvre est accompagnée par les Infrastructures de Recherche des « Humanités numériques » qui ont élaboré des outils et des services destinés à renforcer la visibilité et l'exploitation des résultats de la recherche. Elles entendent valoriser les données et exposer les ressources à partir de leur description dans des schémas de métadonnées. Pour ce faire, elles élaborent des dispositifs sociotechniques et associent la description des ressources numériques (schéma Dublin Core) à de puissantes fondations technologiques. Nous nous sommes intéressés à trois infrastructures de recherche (ISIDORE, HAL et Persée) et avons pu y souligner la mise à disposition à *minima* des outils d'organisation des connaissances. Dans ces infrastructures, les descriptions thématiques des ressources révèlent une co-production contiguë d'indexation matière : les référentiels historiques des institutions documentaires ne bénéficient pas de l'interopérabilité sémantique et restent isolés des folksonomies. En conséquence, la communauté épistémique ne profite pas de toutes les potentialités inhérentes aux systèmes d'organisation des connaissances (SOC) malgré la présence des technologies d'interopérabilité.

Alors que la communauté épistémique exprime le caractère pluridisciplinaire, multidisciplinaire voire transdisciplinaire des objets de recherche qu'elle étudie, notamment dans les sciences humaines et sociales, nous avons observé les conditions du passage de l'actuelle co-production contiguë d'indexation matière à un continuum documentaire en mettant en relief les ressorts humains, intellectuels, juridiques, technologiques, et instrumentaux qu'il requiert. Nous avons identifiés et décrits les contours d'un instrument cognitif « instrumental » permettant qu'une interopérabilité épistémique distribue le continuum documentaire.

Bibliographie

- Amar M. (2000). *Les fondements théoriques de l'indexation : une approche linguistique*. Paris : ADBS éditions.
- Amar, M., Menon, B. (dirs.) (2011). « Web sémantique, web de données, quelle nouvelle donne ? ». In, *Documentaliste-Sciences de l'Information*, vol. 48, n° 4, p. 20-38.
- Andre, F., Charnay, D. (2006). « Support of open archives at national level. The HAL experience ». In, *Institutional archives for research: experiences and projects in Open Access, Nov 2006, Italy*, p. 15-21. Disponible à : https://archivesic.ccsd.cnrs.fr/sic_00187260/document
- Appel, V., et al. (2010). *Les dispositifs d'information et de communication : concepts, usages et objets*. Bruxelles : De Boeck éditions.
- Arsenault, C. (2006). « L'utilisation des langages documentaires pour la recherche d'information ». In, *Documentation et bibliothèques*, vol. 52, n° 2, p. 139-148.
- Aymonin, D. (2017). Autorités, identifiants, entités. L'expansion des référentiels. In, *Arabesques*, n° 85.
- Bauin, S. (2014). *L'open Access à moyen terme : une feuille de route pour HAL (Hyper Articles en ligne) « Au printemps 2013, signature unanime de la convention de partenariat en faveur des archives ouvertes et de la plateforme mutualisée HAL. Et après ? CNRS. Disponible en ligne à : <http://www.cnrs.fr/dist/zoutils/documents/CNRS%20DIST%20Rapport%20Bauin%20sur%20CCSD%20et%20HAL%20septembre%202014.pdf> (consulté le 10/03/2018)*
- Bermès, E. et al. (2013). *Le web sémantique en bibliothèque*. Paris : Cercle de la librairie éditions.
- Béthery, A. (2013). *Guide de la classification décimale Dewey*. Paris : Cercle de la Librairie éditions.
- Bibliothèque nationale de France (2012). *Fonctionnalités requises des données d'autorité matière (FRSAD) : un modèle conceptuel*. Disponible à : http://www.bnf.fr/documents/frsad_rapport_final.pdf

- Bonaccorsi, J. (2016). « Approches sémiotiques du web », dans Barats, C. (dir.) *Manuel d'analyse du web*. Paris : Armand Colin éditions, p. 133-150.
- Caelen, J. (2003). « Systèmes interactifs : essai de caractérisation », dans Miège, B. (dir.) *Communication personnes-systèmes informationnels*. Paris : Hermès Lavoisier, p. 41-58.
- Cahier, J.-P. (2010). « Document et modèle pour l'action, une méthode pour le web socio-sémantique. Application à un web 2.0 en développement durable ». In *Document numérique*, vol. 13, n° 2, p. 75-96.
- Capelli et al. (2014). *Comment contribuer à ISIDORE avec ses données numériques ?* Disponible à : <https://www.humanum.fr/sites/default/files/guide-isidore.pdf>
- Casemajor Loustau, N. (2011). « La contribution triviale des amateurs sur le web : quelle efficacité documentaire ? ». In, *Études de communication*, vol. 36, n° 1, p.39-52.
- Celotti N., Musacchio M. (2004). « Un regard diachronique en didactique des langues de spécialité ». In, *Revue de didactologie des langues-cultures et de lexiculturologie*, n° 135, vol. 3, p. 263-270.
- Charlier, P. Peeters, H. (1999). « Le dispositif ». In, *Hermès, La Revue*, vol. 3, n° 25, p. 15-23.
- Chevalier A. (2013). *La conception des documents pour le web*. Villeurbanne : Presses de l'ENSSIB éditions.
- Claud, J., Duquenne, I. (2017). *Les nouveaux modes de coopération entre l'Agence bibliographique de l'enseignement supérieur et les établissements dans le cadre du système de gestion de bibliothèque mutualisé. Rapport à madame la ministre de l'enseignement supérieur, de la recherche et de l'innovation*. Disponible à : http://cache.media.enseignementsup-recherche.gouv.fr/file/2018/96/7/Rapport_Abes-reseau_SGBM_887967.pdf
- Clavel, T. (2017). « 8e journée nationale RAMEAU ». In, *Bulletin des bibliothèques de France (BBF)*, n° 13. Disponible à : http://bbf.enssib.fr/tour-d-horizon/8e-journee-nationale-rameau_68036
- Clot, Y. (dir. (1999). *Avec Vygotski*. Paris : La Dispute éditions.
- CNRS (2011). *Une infrastructure numérique. Le TGE Adonis*. Disponible à : https://www.humanum.fr/sites/default/files/presentation_du_tge_adonis.pdf
- Cohendet, P. et al. (2003). « Innovation organisationnelle, communautés de pratique et communautés épistémiques : le cas de Linux ». In, *Revue française de gestion*, n° 146, vol. 5, p. 99-121.
- Dacos, M. (2011). *Manifeste des Digital Humanities*. Disponible à : <http://tcp.hypotheses.org/318>
- Davallon, J. (2004). « Objet concret, objet scientifique, objet de recherche ». In, *Hermès, La Revue*, vol. 1, n°38, p. 30-37.
- David, P.-A., Foray, D. (2002). « Une introduction à l'économie et à la société du savoir », in *Revue internationale des sciences sociales. ERES*, vol. 1, n° 171, p. 13-28.
- De Certeau, M. (1990). *L'invention du quotidien I. Arts de faire*. Paris : Gallimard éditions.
- Desfriches-Doria, O. (2015). « Théories de l'activité en sciences de l'information et de la communication et conception de dispositifs orientés utilisateurs ». In *Questions de communication*, n° 28, p. 43-64.
- Direction interministérielle du Numérique et du Système d'Information et de Communication de l'Etat (2015). *Référentiel général d'interopérabilité. Standardiser, s'aligner et se focaliser pour échanger efficacement*. Version 2. Disponible à : http://references.modernisation.gouv.fr/sites/default/files/Referentiel_General_Interoperabilite_V2.pdf
- Dujol, L. (dir.). (2017). *Communs du savoir et bibliothèques*. Paris : Cercle de la Librairie éditions.
- Durieux V. (2010). « Collaborative tagging et folksonomies : l'organisation du web par les internautes ». In, *Les Cahiers du numérique*, vol. 6, n° 1, p. 69-80.
- Fargier, N. (2014). « Persée, une bibliothèque numérique par et pour les chercheurs ». *La francesistica italiana à l'ère du numérique, Sep 2014, Gênes, Italie*. Disponible à : <https://hal.archives-ouvertes.fr/hal-01187830/document>
- Favier, L., Martin-Juchat, F. (2002). « Les sciences de l'information face à de nouveaux paradigmes : prise en compte de la dimension sociale de la recherche d'information et remise en cause de la figure de l'utilisateur », dans *Recherches récentes en SIC : convergences et dynamiques. Actes du colloque MICS-LERASS, 21-22 mai 2002, Toulouse*, p. 255-266.
- Favier, L. (2017). « Les humanités numériques et l'évolution des infrastructures de recherche. Quels enjeux pour l'organisation des connaissances ? », dans Chevry-Pébayle, E. (dir.) *Systèmes d'organisation des connaissances et humanités numériques : actes du 10e colloque ISKO France 2015, 6 et 7 novembre 2015, Collège doctoral européen, Strasbourg*. Londres : Iste éditions, p. 3-15.

- Favier, L., Mustafa El Hadi, W. (2015). « La continuité d'une problématique sur l'organisation des connaissances en lettres et sciences humaines ». In *Les Cahiers du Numérique*, vol. 11, n° 1, p. 83-102.
- Fayet-Scribe, S. (2000). *Histoire de la documentation en France : culture, science et technologie de l'information 1895-1937*. Paris : CNRS éditions.
- Feyler, F. (2007). « De la compatibilité à l'interopérabilité en matière de repérage d'information pertinente » In *Documentaliste-Sciences de l'Information*, vol. 44, n° 1, p. 84-92.
- Flichy, P. (2008). « Technique, usage et représentations ». In, *Réseaux*, n° 148-149, p. 147-174.
- Foray, D. (2009). *L'économie de la connaissance*. Paris : La Découverte éditions.
- Foucault, M. (1977). « Le jeu de Michel Foucault », in *Ornicar*, n°10, p. 62-105.
- Foucault, M. (2008). *L'archéologie du savoir*. Paris : Gallimard éditions, nouvelle édition.
- Francis, É., Quesnel, O. (2007). « Indexation collaborative et folksonomies », In *Documentaliste-Science de l'Information*, vol. 44, n° 1, p. 58-63.
- Gandon et al. (2012). *Le web sémantique. Comment lier les données et les schémas sur le web*. Paris : Dunod éditions.
- Gazellot, G., Ertzscheid, O. (2006). « Etude exploratoire des pratiques d'indexation sociale comme une renégociation des espaces documentaires. Vers un nouveau big bang documentaire ? », dans
- Ghislaine Chartron, G., Broudoux, E. (dirs.) *Document numérique et société, 1ère édition, Actes de la conférence organisée dans le cadre de la Semaine du document numérique à Fribourg (Suisse) les 20 et 21 septembre 2006*. Paris : ADBS Éditions, 2006, p. 291-306.
- Gefen, A. (2015). « Les enjeux épistémologiques des humanités numériques ». In, *La nouvelle revue des sciences sociales*, n° 4, p. 61-74.
- Gensollen, M. (2004). « Biens informationnels et communautés médiatées ». In *Revue d'économie politique*, n° 113, p. 9-40.
- Goldenberg, A. (2011). *La négociation des contributions dans les wikis publics : légitimation et politisation de la cognition collective*. Thèse de l'Université du Québec à Montréal, Montréal (Québec, Canada), Doctorat en communication.
- Golder, S. A., Huberman, B. A. (2006). « Usage patterns of collaborative tagging systems ». In, *Journal of Information Science*, vol. 32, n° 2, p. 198-208.
- Granjon, F., Magis, C. (2016). « Critique des humanités numériques. Pour une approche matérialiste de « l'immatériel ». In, *Variations. Revue internationale de théorie critique*, n° 19. Disponible en ligne : <http://journals.openedition.org/variations/748>
- Guyot, B. (2009). Introduction. *Études de communication*, vol. 33, n° 2, p. 9-18.
- Hodge, G. (2000). *Systems of knowledge organization for digital libraries*. Disponible à : <https://www.clir.org/pubs/reports/pub91/pub91.pdf>
- Holzem M. (1999). Terminologie et documentation, Paris : ADBS éditions.
- Holzem, M. (2006). « L'organisation des connaissances à l'ère du village planétaire : un point de vue global sur le monde ». In, *Matériaux pour l'histoire de notre temps*, vol. 2, n° 82, p. 82-86.
- Hudon, M. (2013). *Analyse et représentation documentaires. Introduction à l'indexation, à la classification et à la condensation de documents*. Québec : Presses de l'Université du Québec.
- Hudon, M., Mustafa El Hadi, W. (2010). « De l'organisation hiérarchique centralisée à l'organisation sociale distribuée ». In, *Les Cahiers du Numérique*, vol. 6, n° 3, p. 9-38.
- Hudon, M. (2012). « ISO 25964 pour le développement, la gestion et l'interopérabilité des langages documentaires ». In, *Documentation et bibliothèques*, vol. 58, n° 3, p. 130-140.
- Humanum (2015). *Le guide des bonnes pratiques numériques*. Disponible à : https://www.huma-num.fr/sites/default/files/guide_des_bonnes_pratiques.pdf
- Ibekwe-SanJuan, F. (2012). *La science de l'information : origines, théories, paradigmes*. Paris : Hermès-Lavoisier éditions.
- Ihadjadène, M., Favier, L. (2008). « Langages documentaires : vers une « crise de l'autorité ? », dans Couzinet, V., Chaudiron, S. (coords.) *Organisation des connaissances à l'ère numérique*. In, *Sciences de la société*, n° 75, p. 11-21.

- Ihadjadene, M., Chaudiron, S. (2008). « L'étude des dispositifs d'accès à l'information électronique : approches croisées », dans Papy, F. (dir.) *Problématiques émergentes dans les sciences de l'information*. Paris : Hermès Lavoisier éditions, p.183-207.
- Jacobi, D. (1993). « Les terminologies et leur devenir dans les textes de vulgarisation scientifique ». In, *Didaskalias*, n° 1, p. 68-83.
- Jeanneret, Y. (2006). « Désigner, entre sémiotique et linguistique », dans Timimi, I., Kovacs, S. (coords.) *Indice index indexation. Actes du colloque international organisé les 3 et 4 novembre 2005 à l'Université Lille-3 / par les laboratoires CERSATES et GERIICO*. Paris : ADBS éditions, p. 17-36.
- Jeanneret, Y. (2009). « La relation entre médiation et usage dans les recherches en information-communication en France ». In, *RECIIC, Electronic Journal Communication Information and Innovation in Health*, vol. 3, n° 3. Disponible à : https://www.reciis.icict.fiocruz.br/index.php/reciis/article/view/753/pdf_140
- Kalfon, J. (éd.) (2016). Créer du lien, faire sens, un nouveau souffle sur les données. In, *Arabesques*, n° 83. Disponible à : <http://www.abes.fr/Publications-Evenements/Arabesques/Arabesques-n-83>
- KENNEL, S. (2011). « Espace de formation en ligne et culture de l'information : un nouveau paradigme pour l'usage et la médiation ? », dans Kiyindou, A., Bautista, R. A. (dirs.) *Nouveaux espaces de partage des savoirs : dynamiques des réseaux et politiques publiques*. Paris : L'Harmattan éditions, p.93-108.
- Le Deuff, O. (2006). « Folksonomies », In, *Bulletin des bibliothèques de France (BBF)*, 2006, n° 4, p. 66-70. Disponible à : <http://bbf.enssib.fr/consulter/bbf-2006-04-0066-002>
- Le Moigne, J.-L. (2007). *Les épistémologies constructivistes*. Paris : PUF.
- Leontiev, A. (1984). *Activité conscience personnalité*. (traduction française) Moscou : Editions du Progrès.
- Lièvre, P. et Laroche, N. (2014). « Retour sur la notion de communauté épistémique ». In *7è colloque GeSCO LEST CNRS, Université d'Aix-Marseille 4-5-6 juin 2014*. Disponible à : https://www.researchgate.net/publication/281279076_RETOUR_SUR_LA_NOTION_DE_COMMUNAUTE_EPISTEMIQUE
- Mac Pherson, T. (2008). *Dynamic Vernaculars: Emergent Multimodal Forms of Scholarship*, Umeå. Disponible à <http://stream.humlab.umu.se/index.php?streamName=dynamicVernaculars>
- Maniez, J. (2002). *Actualité des langages documentaires : les fondements théoriques de la recherche d'information*. Paris : ADBS éditions.
- Ménard, F. (2017). *Rapport du groupe de travail national sur la syntaxe RAMEAU*. Préconisations et pistes d'évolution. Disponible à : http://rameau.bnf.fr/chantier_syntaxe/pdf/rapport_final_syntaxe_rameau.pdf
- Menon, B. (2007). « Les langages documentaires. Un panorama, quelques remarques critiques et un essai de bilan ». In *Documentaliste-Sciences de l'Information*, vol. 44, n° 1, p. 18-28.
- Miller, P. (2000). « Interoperability : What Is It and Why Should I Want It? », In *Ariadne*, n° 24. Disponible à : <https://goo.gl/OM3d32>
- Mingam, M. (2005). « Rameau : Bilan et perspectives ». In, *Bulletin des bibliothèques de France*, vol. 50, n° 5, p.38-47.
- Mustafa El Hadi W. (2013). « Faceted classification as model for universal classifications. Knowledge Organization from Libraries to the Web ». In, *Journal of Information Management, SRELS (Sarada Ranganathan Endowment for Library Science), Bangalore*, vol. 50, n° 6, p. 720-734.
- Nanard M., Nanard, J. (2002). « Interfaces graphiques – Architecture ». In, *Techniques de l'ingénieur*. Disponible sur abonnement.
- Papy, F., Leblond, C. (2007). « L'interface de recherche d'information du Visual...Catalog : un outil innovant à double détente ». In *Documentaliste-Sciences de l'Information*, vol. 44, n° 4, p.288-298.
- Papy, F. (2009). *Technodocumentation : des machines informationnelles aux bibliothèques numériques*. Paris : Hermès Lavoisier éditions.
- Papy, F. (2015). *Bibliothèques numériques : interopérabilité et usages*. Londres : Iste éditions.
- Paul, V., et Perriault, J. (2004). « Introduction. Pratiques d'information et de communication, l'empreinte du numérique ». In, *Hermès, La Revue*, vol. 2, n° 39, p. 9-16.
- Peraya, D. (1999). « Médiation et médiatisation : le campus virtuel ». In, *Hermès, La Revue*, vol. 3, n° 25, p.153-167.

- Pierot, E. (2017). « Prise en charge de la classification à facettes par SKOS. Quelle représentation pour les sujets complexes ? ». In, *Les Cahiers du Numérique*, vol. 13, n° 1, p. 89-114.
- Presner, T. et al. (2009). *The Digital Humanities Manifesto 2.0*, UCLA. Disponible à : <http://manifesto.humanities.ucla.edu/2009/05/29/the-digital-humanities-manifesto-20/>
- Quint, V. (2003). « Hypermédia et technologies web ». In, *Actes de la conférence H2PTM'03*. Disponible à : http://ticri.univ-lorraine.fr/ticri-h2ptm.fr/index.php/H2PTM_%282003%29_Quint
- Rabardel, P. (1995). *Les hommes et les technologies : approche cognitive des instruments contemporains*. Paris : Armand Colin éditions.
- Riva, P. et al. (2017). Transition mappings. *User tasks, entities, attributes, and Relationships in FRBR, FRAD, FRSAD mapped to their equivalents in the IFLA Library Reference Model*. IFLA. Disponible à : https://www.ifla.org/files/assets/cataloguing/frbr-lrm/ifla-lrm-august-2017_rev201712.pdf
- Roumieux, O. (éd.) (2011). Web sémantique, web de données... Quelle nouvelle donne ? In, *Documentaliste-Sciences de l'Information*, Vol. 48, n°4.
- Saemmer, A. (2015). *Rhétorique du texte numérique : figures de la lecture, anticipation de pratiques*. Villeurbanne : Presses de l'Enssib éditions.
- Svensson, P. (2010). « The Landscape of Digital Humanities ». In, *Digital Humanities Quarterly*, Vol.4, n°1. Disponible à : <http://digitalhumanities.org/dhq/vol/4/1/000080/000080.html>
- Tosello-Bancal, J.-E. (2007). PERSEE : Programme de valorisation des revues francophones en sciences humaines et sociales mis en œuvre par le MENESR., dans Papy, F. (dir.) *Usages et pratiques dans les bibliothèques numériques*. Paris : Hermès Lavoisier éditions, p. 171-182
- Trant, J. (2008). "Studying social tagging and folksonomy : A Review and Framework". In, *Journal of Digital Information*, vol. 10, n° 1, p. 2-44.
- Verlaet, L. (2015). « La deuxième révolution des systèmes d'information : vers le constructivisme numérique », *Hermès, la Revue*, vol. 71, n° 1, p. 249-254.
- Vidal, G. (dir.) (2012). *La sociologie des usages : continuités et transformations*. Paris : Hermès Lavoisier éditions.
- Vinck, D. (2016). *Humanités Numériques. La culture face aux nouvelles technologies*. Paris : Le Cavalier Bleu éditions.
- W3C (2009). *Simple Knowledge Organization System*, World Wide Web Consortium. Disponible à <https://www.w3.org/TR/skos-reference/>
- Zacklad, M. (2010). « Evaluation des systèmes d'organisation des connaissances ». In, *Les Cahiers du Numérique*, vol. 6, n°3, p. 133-166.