



CLASSIQUES
GARNIER

NOYER (Jean-Max), « Prolégomènes à des infrastructures intellectives pour l'édition numérique scientifique », *Études digitales*, n° 4, 2017 – 2, *Immersion*, p. 159-206

DOI : [10.15122/isbn.978-2-406-09288-9.p.0159](https://doi.org/10.15122/isbn.978-2-406-09288-9.p.0159)

La diffusion ou la divulgation de ce document et de son contenu via Internet ou tout autre moyen de communication ne sont pas autorisées hormis dans un cadre privé.

© 2019. Classiques Garnier, Paris.
Reproduction et traduction, même partielles, interdites.
Tous droits réservés pour tous les pays.

NOYER (Jean-Max), « Prolégomènes à des infrastructures intellectives pour l'édition numérique scientifique »

RÉSUMÉ – Le développement des mémoires numériques dans le domaine des sciences, couplé avec la crise des humanités et la difficulté de faire apparaître les relations internes des savoirs, imposent de nouvelles exigences aux infrastructures de l'édition. Sont en jeu les écritures du web, les technologies intellectives qui rendent productifs ces nouveaux “milieux” d'intelligence. L'article met en évidence les principaux points d'application des forces pour surmonter *l'encyclopédisme en éclats*.

MOTS-CLÉS – Humanités digitales, milieu, écologie des savoirs, infrastructures, Web sémantique, hybridation, hypertexte, archive

NOYER (Jean-Max), « Prolegomena to intellectual infrastructures for scientific digital publishing »

ABSTRACT – The development of digital memories in the sciences, coupled with the crisis in the humanities and the difficulty of revealing the internal relations of knowledge, impose new requirements on publishing infrastructures. At issue is web writing, the intellectual technologies that make these new intelligence “environments” productive. This article highlights the main points of application of these forces to overcome *splintered encyclopedism*.

KEYWORDS – Digital humanities, environment, knowledge ecology, infrastructures, Semantic Web, hybridization, hypertext, archive

PROLÉGOMÈNES À DES INFRASTRUCTURES INTELLECTIVES POUR L'ÉDITION NUMÉRIQUE SCIENTIFIQUE

INTRODUCTION

L'édition numérique scientifique est consubstantielle au développement du réseau Internet. La question des infrastructures accompagne le bouillonnement des Humanités Digitales pour reprendre l'expression de Pierre Mounier¹ et la transformation profonde de la variation des couplages cerveaux-écritures-technologies intellectives. Elle va de pair avec le développement et l'utilisation dans la plupart des domaines et disciplines de traitements algorithmiques liés avec la croissance des empiries numériques. Cette irrésistible montée de l'algorithmie est à présent notre milieu².

Françoise Thibault (DGRI et Bertrand Jouve (CNRS, InSHS) ont dans leur rapport de 2012 dressé un premier tableau des infrastructures « molaires » de recherche en sciences humaines et sociales. (Rapport du groupe Infrastructures de l'Alliance Athéna)³. Dans le même temps nous avons commencé à explorer certaines des nouvelles transformations affectant le champ des savoirs et de ses conditions de « production-circulation-exploitation ». En effet la plasticité des écritures associée à la plasticité des mémoires hypertextuelles-hypermédiés en réseaux, a

1 (<http://press.openedition.org/241>)

2 L'irrésistible montée de l'algorithmie, Méthodes et concepts en SHS, *les Cahiers du numérique* 2014/4 (Vol. 10) <http://cache.univ-tln.fr:2057/revue-les-cahiers-du-numerique-2014-4-page-63.htm> et Pierre Lévy, <http://pierrelevyblog.com/2013/02/17/le-medium-algorithmique/>

3 Les infrastructures de recherche en sciences humaines et sociales, Rapport du groupe Infrastructures de l'Alliance Athena, Françoise Thibault (DGRI) Bertrand Jouve (CNRS, inSHS) Septembre 2012.

engagé assez rapidement l'invention associée de nouvelles « technologies intellectives ». La différenciation des intelligences collectives a accompagné le processus et la fabrication des nouveaux agencements collectifs d'énonciation des savoirs, a pris des formes de plus en plus hétérogènes.

Dans quelle mesure cela était-il en train (et continue) de transformer les « milieux⁴ » d'intelligence, la capacité à remonter vers la fabrication des problèmes, dans quelle mesure, les pratiques analogiques, abductives, transductives, ont-elles été (et continuent-elles) affectées, les modes de construction de la preuve transformés ? Et ce dans le cadre de pratiques d'écriture-lecture en constante évolution ? Enfin la production de vastes empiries numériques (production-capture de traces sémantiques, comportementales, temporelles complexes, spatiales-géolocalisées) que nous laissons à l'occasion de nos pratiques dans le plissement numérique du monde, a une influence profonde dans le domaine des sciences en général et des SHS en particulier. Mise en évidence (explicitation) certes partielle, mais importante, du vaste réseau de relations internes dont nous sommes l'expression et l'exprimé, nouvelles pratiques cartographiques des processualités dont nous sommes tissés, d'entrelacs de plus en plus étroits entre Mathématique et Sciences humaines, entre Raison Statistique et Prédiction. Ainsi de nouvelles relations semblent constituer la trame d'une alliance que l'on avait perdue depuis Kant entre les sciences (mathématique-biologie-cerveau-physique) la Philosophie et les SHS.

Dans ce contexte et parce que le monde des savoirs paraît constitué de lignes de forces et de champs de plus en plus divergents, la question d'un nouvel encyclopédisme se pose et ce d'autant plus que les grandes crises en cours de nos écologies radicalisent nos tourments face à l'étrange étrangeté de la « nécessité de la Contingence » et face à l'exigence d'une complexité sans précédent des hétérogènes. Nous avons appelé cela « *l'Encyclopédisme en éclats*⁵ » comme nouvelle Frontière, comme agencement où la question de la productivité des frontières et

4 Au sens de G. Simondon.

5 *L'encyclopédisme en éclats*. L'édition scientifique numérique face aux nouvelles mémoires et intelligences en procès Jean-Max Noyer et brigitte juanals <https://hal.archives-ouvertes.fr/halshs-00512157> in Papy F. et Guyot B. *L'édition scientifique : analyses et perspectives*, Lavoisier Hermès Science, INIST-Wiley (UK), p. 203-221, 2008, coll. STI – Sciences et Techniques de Encyclopédisme en éclats : réflexions sur la sortie des « essences » l'Information et <http://www.cosmopolis-rev.org/2014-3-4-f>

leur différenciation sont centrales. Nouvelles manières d'habiter les dissensus et leurs rapports, les controverses, le cosmopolitisme dans les savoirs (y compris scientifiques) et la « Traduction Générale » sont là comme enjeux cognitifs mais aussi stratégiques.

C'est la raison pour laquelle en prenant pour colonne vertébrale incertaine la question du développement de l'édition numérique scientifique (open ou non), nous proposons de mettre l'accent sur les points de faiblesse qu'une certaine répétition des modes éditoriaux impose (encore une fois que les modes soient open ou non). La persistance voire le maintien d'une contre-révolution de l'ordre des essences dans le champ documentaire général contre les puissances de la Renaissance engendrée par les écritures hypertextuelles, leurs baroques labyrinthes leur pyrotechnie conceptuelle, est à l'œuvre.

Nous proposons aussi de mettre en évidence des points de bifurcation, des écritures et combinatoires qui font face à la communauté des œuvres comme « incomplétude en procès de production », aux hétérogénèses qui la parcourent. Pour viser une édition processuelle faisant place aux dissensus et controverses, et qui mettrait la variation et la navigation dans des espaces à n-dimensions, sinon au cœur du processus éditorial, en tout cas de ne pas la laisser sur les marges, pour incarner encore ce que Michel Foucault déclarait dans un entretien avec Didier Éribon, « vous savez à quoi je rêve ? Ce serait de créer une maison d'édition de recherche. Je suis éperdument en quête de ces possibilités de faire apparaître le travail dans son mouvement, dans sa forme problématique. Un lieu où la recherche pourrait se présenter dans son caractère hypothétique et provisoire⁶ ». De la « Derrida base⁷ » au « deconstructed journal⁸ » en passant par le « livre bombe » de Foucault encore⁹ et le

6 M. Foucault, D. Eribon « Pour en finir avec le mensonge », Le Nouvel Observateur, 1985.

7 G. Bennington, J. Derrida, *Jacques Derrida*, Édition du Seuil, Collection Les contemporains, 1991, Paris. Jacques Derrida, *Ulysse gramophone, Deux mots pour Joyce*, Édition Galilée, 1987, Paris.

8 Smith J. W T. (1999). The deconstructed journal – a new model for academic publishing. *Learned Publishing*, 12, 79-91. Voir sur ces questions Gabriel Gallezot, Jean-Max Noyer De la numérisation des revues à l'expérimentation d'une édition de recherche processuelle, technologies de l'information et intelligences collectives, Hermès-Lavoisier, p. 263-280, 2010.

9 En 1978, devant des étudiants californiens, Michel Foucault rêvait à haute voix de « livres bombes » : ils ne tueraient personne, mais « disparaîtraient peu de temps après qu'on les aurait lus ou utilisés ».

« Liquid Journal » et les modes d'édition du même nom¹⁰, sortir des modes classificatoires essentialistes, fermés. En suivant, dans le creusement des modes d'écritures, dans l'exploitation des corpus des savoirs scientifiques en particulier les modes algorithmiques et l'utilisation d'une mathématique ciblant le qualitatif¹¹, nous donnerons à voir à la fois la persistance des modes représentatifs des métriques qui vont avec, mais aussi leur productivité (connectivité, traductibilité, ouvertures), leur puissance à générer des devenirs au cœur de l'activité de recherche, au cœur de la pensée¹². La question du dedans-dehors des corpus est plus que jamais au centre ou plutôt, la possibilité d'aller à cheval sur les zones frontières de la recherche, de travailler dans les zones d'indétermination où la pensée puise et ne s'épuise pas.

Cette article tentera d'indiquer que nombre de voies de développement sont dans cette tension et nous montrerons que des technologies intellectives hybrides pourraient faire face à ces problèmes dans des milieux d'intelligence compliqués, pleins d'hétérogènes.

COMMENT ÊTRE « TRANSVERSAL ET TRAVERSÉ » ?

La manière dont nous parcourons les dynamiques de ces configurations techno-socio-cognitives sera aimantée par la question de savoir comment accroître la « transdisciplinarité-transversalité » en sciences et définir plus

10 Liquid Journals : Knowledge Dissemination in the Web Era, Position Paper by Marcos Baez, Fabio Casati, Dipartimento di Ingegneria e Scienza dell'Informazione, University of Trento.

11 Voir Bruno Latour et alii et aussi Jean-Pierre Courtial : « Structure fractale des connaissances scientifiques ». Essai de résumé des travaux de Rafael Bailon Moreno (Communication personnelle) et The pulsing structure of science : Ortega y Gasset, Saint Matthew, fractality and transfractality in Scientometrics, 2007 DOI : 10.1007/s11192-007-1600-8; Rafael Bailón-Moreno University of Granada, Rosario Ruiz-Baños University of Granada, Encarnación Jurado Alameda University of Granada, Jean-Pierre Courtial University of Nantes Enfin voir, R. Bailon-Moreno, Ingeniera del conocimiento y vigilancia tecnologica aplicada a la investigacion en el campo de los tensioactivos. Desarrollo de un modelo ciencimetrico unificado, 2003, Grenade : Thèse Université de Grenade.

12 On tentera d'utiliser de manière indirecte l'hypothèse développée dans le contexte d'une réflexion générale sur la Valeur par Brian Massumi Intitulé « 99 Theses on the Revaluation of Value. A Postcapitalist Manifesto ».

précisément les buts et contraintes pour les infrastructures de l'édition scientifique numérique aujourd'hui, pour un « *encyclopédisme fragmenté* », des points de vue, des réseaux-rhizomes, des morphogenèses, dit autrement et plus exactement un a-encyclopédisme ouvert et ne cessant de créer les conditions de son propre dépassement, de son propre effondrement. Un encyclopédisme comme Machine Anaphorique.

La mise en crise des trois écologies, (environnement-ressources, esprit et socius)¹³, la question de la possibilité de faire émerger de nouveaux modes d'existence, dans le cadre de changements affectant les processus de différenciation internes au système des relations (Bateson 1986 ; 1977) entre les êtres, les nations, les cultures, rendent plus nécessaire que jamais que nous nous interroguions sur ce qui affecte les conditions d'une nouvelle Renaissance à partir de la transformation du système des savoirs, que ces savoirs soient scientifiques, religieux. Dans ce cadre, la question de l'Esprit est capitale et la recherche d'une ou plusieurs « noopolitiques¹⁴ » susceptibles d'aider à la mise en place de nouvelles conditions matérielles et idéelles en vue de la dite Renaissance permettant de faire face aux crises et défis évoqués plus haut, est traversée par les transformations des systèmes d'écritures, des dispositifs de mémoires. Il ne s'agit nullement ici de poser la question de la Renaissance à partir d'un point de vue techniciste, mais bien au contraire de l'aborder à partir de la variation du « couplage cortex-silex¹⁵ » des effets de l'adoption des nouveaux « couplages cerveaux, technologies intellectuelles¹⁶ » sur les collectifs de pensée et les complexes relationnels.

Le développement de la strate anthropologique Internet est à présent une partie importante de notre milieu associé et l'intérêt que nous devons porter à la créativité, au sens de A. N. Whitehead¹⁷ tout autant et même peut-être plus qu'à l'innovation même, ne cesse de grandir. Si nous voulons éviter de voir se développer les symptômes morbides, il est

13 Felix Guattari, *Les Trois Écologies*, Ed. Galilée, 2008.

14 Nous prenons ici « Noopolitique » dans le sens que lui donne Bernard Stiegler à savoir, « Politique de l'Esprit », associé à « Psychopouvoir » comme lieu central aujourd'hui de l'anthropologie politique. Nous le prenons aussi dans le sens de John Arquilla et David Ronfeldt, « Noopolitik » (opposé à Realpolitik) appliqué cette fois à la redéfinition des rapports de puissance dans le domaine de la géopolitique et de la stratégie, à partir des intelligences collectives et des subjectivités en conflit.

15 Bernard Stiegler, 1994.

16 Jean-Max Noyer, 1995.

17 A. N. Whitehead, *Procès et Réalité*, éd. Gallimard, Paris 1995.

urgent de penser la Créativité pour elle-même au cœur de l'ensemble ouvert des agencements des savoirs.

VERS UNE ÉCOLOGIE GÉNÉRALE DES SAVOIRS ?

La question des humanités digitales risque d'être bien mal engagé, si l'on n'accorde de l'importance qu'au second terme. Il y a certes l'irruption des mémoires hypertextuelles-hypermédias en réseaux, des écritures et des algorithmes qui vont avec, la montée du quatrième paradigme¹⁸ mais il y a aussi accompagnant cette irruption ce que l'on a appelé la crise des humanités, ce que l'expression « les humanités » recouvre. À partir du milieu des années soixante, sont ici et là formulées interrogations et inquiétudes sur ce que l'on appelle tantôt les humanités, tantôt les sciences humaines et sociales, ces expressions ne se valant pas cependant.

Énoncée de manière encore générale, la transformation des savoirs, de leurs modes de production, différenciation-fragmentation mais encore de circulation, les rapports entre eux et la fabrication des problèmes, tout cela pose à nouveau la question de comment « encyclopédier » à nouveau les savoirs spécialisés, éparpillés, de comment faire émerger un nouveau rationalisme et donc une nouvelle culture, surmontant aussi ce que Bruno Latour en particulier nomme le « grand partage¹⁹ ».

18 Jim Gray, *The Fourth Paradigm : Data-Intensive Scientific Discovery*, Microsoft. 2009. Et l'irrésistible montée de l'algorithmique, Méthodes et concepts en SHS, *Les Cahiers du numérique* 2014/4 (Vol. 10) <http://cache.univ-tln.fr:2057/revue-les-cahiers-du-numerique-2014-4-page-63.htm>

19 Bruno Latour, *Le grand partage*, Comment redistribuer le Grand Partage ? 1983 In *Revue de Synthèse*, n° 110, Avril/Juin, p. 203-236, 1983 [Reprint in *La Revue du Mauss* N° 1, p. 27-65, 1988.

LES AGENCEMENTS POUR DE NOUVELLES HUMANITÉS

Les divers problèmes liés à l'évolution des organisations des connaissances, les diverses incarnations du web et ses devenirs, les transformations des systèmes socio-cognitifs, sont pris dans le creusement intensif des collectifs de pensée. Et la mise en place, certes incertaine, des conditions de développement et de propagation, stabilisée dans le temps, de ce qui pourrait constituer les contours d'une nouvelle culture générale, ouverte et processuelle, constitue un enjeu majeur pour ceux qui souhaitent définir les formes d'un nouveau rationalisme ouvert, expression et exprimé à la fois, des dissonances culturelles, philosophiques et religieuses, des dissonances issues de « l'inégal développement » du rationalisme attaché à la sphère technoscientifique de l'Occident et du vaste système de relations internes en quoi consistent les agrégats d'intelligences collectives. Agrégats et intelligences qui constituent la trame d'un vaste et très différencié tissu sans couture voilant les faces du monde.

Une des questions possibles pourrait être celle-ci : quel système du savoir est en train de se développer aujourd'hui, à partir de la sphère scientifique et technique et de ses rationalités ? Pour reprendre l'analyse de Bertrand Saint Sernin, « l'ensemble des savoirs va-t-il continuer de se présenter sous la forme d'un agrégat de spécialités dont les relations internes demeureront imprécises ? Ou bien leur unité organique est-elle en train de devenir visible ? Dans la première hypothèse, nous aurons au mieux une classification des sciences ; dans la seconde, nous disposerons d'un système du savoir. Chacun des deux partis a ses avocats : dans le passé, le premier a eu pour défenseurs Mach et Duhem ; le second, Cournot et Whitehead ». (Saint Sernin, 2007)²⁰.

20 Bertrand Saint-Sernin, *Le rationalisme qui vient*, Collection Tel (n° 346), Gallimard, 2007.

LES PLISSEMENTS DU SAVOIR

Les réflexions sur l'importance de ces infrastructures sont relativement nombreuses. Toutefois elles se sont concentrées sur les problèmes du stockage, de l'accès, de l'interopérabilité et des conditions économiques générales et sur les questions – très souvent biaisées – liées à la légitimité, valeur scientifique des productions. Dès 2004 le rapport de la « Commission on Cyberinfrastructure for the Humanities and Social Sciences Berkeley » sous la direction de Michael K. Buckland avait mis l'accent sur un certain nombre de contraintes et en particulier sur le rôle central des metada : « *Combining diverse digital resources depends on an interoperable infrastructure of metadata. This is as much infrastructure as hardware or policies. But national standards for format and content of these tools don't exist, even for gazetteers ! An investment in the development of scholarly "best practices" is needed²¹* ».

Notre réflexion semble reprendre certaines de ces perspectives mais elle se déploie dans le cadre du plissement numérique des savoirs, des intelligences et en particulier des intelligences dites scientifiques ; comme nouvelle écologie de l'intellect, de l'esprit.

Elle vient à la suite de nombreux efforts publics et privés, nationaux, européens et mondiaux pour définir et construire « des infrastructures » pour la recherche en général et pour la recherche en sciences humaines et sociales en particulier.

Les manières de « fabriquer » ces infrastructures dependent certes des modèles d'économies politiques des savoirs qui leur servent de milieu. Entre modèles centralisateurs, planificateurs et modèles décentralisés, a-centrés, entre modèles hybrides plus ou moins complexes. Ces infrastructures se différencient aussi selon la manière dont elles conçoivent les contraintes d'interopérabilité, les types de normes, sous la double contrainte générale de respecter à la fois l'hétérogénéité des champs de recherche, leur singularité et de permettre leurs articulations et leurs

21 Commission on Cyberinfrastructure for the Humanities and Social Sciences. Berkeley, 2004. Metadata as Infrastructure ; Interoperability ; and the Larger Context. Voir aussi <http://www.people.virginia.edu/~jmu2m/Cyberinfrastructure.RLG.html> et Cyberinfrastructure for the humanities and social sciences : advancing the humanities research agenda, <https://dl.acm.org/citation.cfm?id=1255217&dl=ACM&coll=DL>

traductions relatives. Elles se différencient encore selon qu'elles sont molaires ou moléculaires et selon leur rapport aux devenirs minoritaires des sciences et des champs de recherches qu'elles laissent vivre ou favorisent.

On sait de manière générale et pour suivre ici Bruno Latour, Geoffrey C. Bowker²², L. Starr, l'importance d'une infrastructure informationnelle. Starr et Ruhleder ont montré aussi qu'une telle infrastructure présente cinq caractéristiques majeures :

- elle s'inscrit elle-même dans un système enchevêtré et distribué d'autres structures ;
- elle est transparente (« elle n'a pas à être réinventée à chaque fois ») ;
- elle est étendue et transversale ;
- elle a une dimension essentiellement collective ;
- elle repose sur un ensemble de conventions et pratiques concrètes dont elle est à la fois l'expression et l'exprimé.

Que les intelligences soient toujours déjà machinées, que les savoirs émergent fussent-ils les plus singuliers au milieu d'agencements collectifs plus ou moins importants, que les technologies intellectives, les dispositifs, les espaces et les moyens de l'échange et de circulation des traces... que tout cela ait été présent depuis très longtemps ne change rien au fait que depuis le début du xx^e siècle les conditions de production, circulation, l'incessant travail de commentaires et de ré-écritures, la construction des preuves et le travail sur les données (les obtenues), les variations théoriques affectant paradigmes et fabrication-invention des problèmes se sont transformés en profondeur. Il en va de même pour les conditions des modes de lecture, d'écriture, des conditions pour l'attention²³.

22 Geoffrey C. Bowker University of Illinois at Urbana-Champaign : « L'Histoire des Infrastructures Informationnelles » in revue Solaris, 1995, accessible à <http://gabriel.gallezot.free.fr/Solaris/d04/4bowker.html>W<

23 Yves Citton, Comment nous lisons Lecture rapprochée, hyperlecture, lecture machinique, <http://www.yvescitton.net/wp-content/uploads/2017/09/HAYLES-LireEtPenserEnMilieuxNumeriques-Chap3-CommentNousLisons-2016.pdf> et Maryanne Wolf Proust and the Squid : The Story and Science of the Reading Brain ? 2008.

LE TRAVAIL
« PAR » ET « DANS » LES INFRASTRUCTURES

Indexer, Classer, Cataloguer, Relier, Cartographier, Naviguer... sont toujours au cœur des pratiques cognitives. Ces fonctions sont mêmes amplifiées, elles se sont différenciées et reposent, aux échelles où elles sont à l'ouvrage, sur des écritures, des algorithmes plus ou moins complexes. Dès la naissance du processus dit des archives ouvertes visant à transformer le stockage et l'accès à la littérature scientifique sous la forme d'ArXiv²⁴. Paul Guinsparg son créateur insiste sur les effets de ce nouveau type d'infrastructures. Des effets « d'une participation spontanée maximale, nous pouvons donc nous attendre non seulement à une interopérabilité de plus en plus automatisée entre les bases de données et à une disponibilité croissante de ressources en ligne pour la collecte algorithmique – articles, jeux de données, notes de cours, multimédias et logiciels (des balises, des liens, des commentaires, des corrections, des contributions à des ontologies et des liens à forte intensité d'expertise, tous activement organisés, deviendront de plus en plus importants, agissant pour coller des bases de données et des textes dans une structure de connaissances plus puissante²⁵. »

Et il continue, le but est bien « la création d'une structure de connaissance semi-supervisée et auto-entretenu, pilotée par des concepts synthétiques, débarrassée de toute redondance et de toute ambiguïté, recherchée, authentifiée et mise en avant pour la nouveauté. Notre navigation dans la littérature sera alors bien plus complète, guidée par des algorithmes permettant d'accéder à nos propres comportements et à ceux des utilisateurs ; et notre lecture (sera) guidée par des liens vers des ressources explicatives et complémentaires liées à des mots, des équations, des figures et des données²⁶. »

24 www.arXiv.org

25 Paul Guinsparg, *It was twenty years ago today, Physics and Information Science*, Cornell University, in arXiv (Submitted on 14 Aug 2011 (v1), last revised 13 Sep 2011 (this version, v2))

26 *Idem.*

Comme depuis toujours « la communauté des œuvres, et celles des sciences sont comme incomplètes en procès de production²⁷ et sous les conditions du Numérique elles se présentent comme un vaste domaine ouvert de relations internes, domaine où la notion même d'objet documentaire numérique considéré du point de vue des essences ne peut plus opérer et où trajectoires et morphogènes des relations, transformations et déformations des relations sont à la source de toutes les sémiotiques, entre métastabilité et fluidité, entre émergence et effacement, entre attraction et évanescence.

Des nœuds et des liens, des langages de balisage de plus en plus puissants et souples jusqu'aux ontologies (« all the way down » ?) fabriquées de relations spécifiques et de leur cortège de traductions et de frontières, d'interstices et de fentes pour des pratiques cartographes ouvrant à la fois la création de nouvelles connexions et ouvertures temporelles infiniment variables, faisant exploser les libertés analogiques au cœur des « milieux des sciences ».

D'où, au passage, la nécessité d'une conception documentaire renouvelée où les documents sont pensés et traités comme « complexes relationnels », flux et événements. Les dits complexes relationnels se définissant par leur puissance de connectivité, de traduction, par leurs pouvoirs attracteurs.

Chaque document est donc une variété de relations associée à une capacité combinatoire. Il se définit par le système de relations plus ou moins stable dont il est à la fois l'expression et l'exprimé. De manière différente, on peut dire qu'un document est une écologie avec ses frontières + « n-points de vue » qui expriment l'histoire de son dehors, c'est-à-dire des plis et des déplis qui ne cessent de le transformer. Il est donc le fruit d'un couplage entre son écologie interne et son écologie externe, entre ce qui converge vers lui, ou se laisse capturer et ce qui sort de lui. C'est la raison pour laquelle l'expression « complexe relationnel » qui a aussi pour corrélat une véritable énergétique intellectuelle, nous paraît rendre compte de la transformation.

27 Pour suivre Whitehead A.N. (1929). *Procès et réalité : Essai sur la cosmologie*. Gallimard, 1995.

LES ÉCRITURES DU WEB

Dès le début des mémoires hypertextuelles de l'édition scientifique, est inscrit son devenir sémantique. Et la formulation par Tim Berners Lee²⁸ de la question encyclopédique, des nouvelles écritures pour prendre en charge modes de lectures et de navigation dans les savoirs, vient relativement rapidement après l'explosion de l'édition électronique et lorsqu'il expose le projet du Web sémantique il le résume sous la forme de son célèbre "Cake".

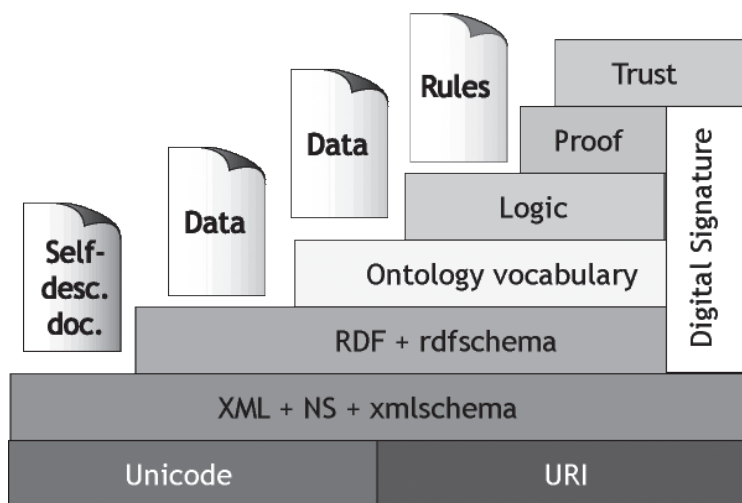


FIG. 1 – Le « Cake » de TIM Berners Lee.

Dans ce schéma de Tim Berners Lee il est décrit les besoins du Web sémantique, on trouve une organisation en couches de différents langages : dans le premier niveau, la norme UNICODE de codage des caractères qui s'impose pour toutes les langues du monde et le mode d'adressage des pages (Unified Resource Identifier) du Web déjà en œuvre pour le Web normal forment le socle du Web sémantique.

28 Tim Berners-Lee, Fischetti, Mark, *Weaving the Web*, HarperSanFrancisco, 1999 et *The Semantic Web A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities* by Tim Berners-Lee, James Hendler and Ora Lassila, Scientific American : Feature Article : The Semantic Web : May 2001.

La couche suivante, « XML + NS + xml schema », fournit les langages nécessaires à l'écriture des documents sur le Web, à savoir, le langage XML, les espaces de noms (NS pour name space) pour permettre de former des groupes de balises indépendants et uniques (en préfixant les noms des balises) et les Schéma XML qui permettent de décrire la structure d'un document XML et d'en typer son contenu. À ce niveau, on a les bases nécessaires (en termes de langages normés) à la mise en place et l'échange de documents textuels sur le Web. La couche « RDF + rdf schema » correspond à un changement de paradigme annoncé : elle permet de passer d'un document à des données (de self descriptive document à data). Cette couche permet de décrire des concepts et des relations entre eux (avec des triplets, {concept, relation, concept}) et donc de décrire des informations complexes sur le monde. Ces informations peuvent être associées à un document grâce à ce même langage et être des métadonnées – appelées aussi annotations formelles – de ce même document. En d'autres termes, elle assume ce que l'utilisation des Schémas XML de la couche précédente ne fait pas totalement, à savoir le remplacement du texte par sa représentation conceptuelle.

La couche « Ontology vocabulary » est, comme son nom ne le précise pas complètement, la fourniture des langages qui permettent de représenter des ontologies et de raisonner dessus (en particulier, faire des inférences sur les héritages de propriétés de concepts grâce à la relation de subsumption et sur les types de concepts).

On choisit ainsi son langage en fonction des besoins d'expressivité et de décidabilité. Les ontologies exprimées en OWL fournissent des métadonnées pour la couche précédente. Les 3 couches suivantes, « logic », « proof », « trust » ne sont pas encore normalisées et permettraient respectivement, a) d'appliquer explicitement des règles d'inférences que l'on se sera donné sur les concepts de l'ontologie, b) de se donner les mécanismes nécessaires à faire la preuve des algorithmes et c) de se donner les mécanismes nécessaires à calculer et exprimer la croyance en une assertion au sein d'une application. Ces couches reprennent les intentions de certains travaux de l'Intelligence artificielle. Enfin, le rectangle de droite représente les méthodes de signatures électroniques et (ce n'est pas sur cette version du schéma) la disponibilité de méthodes de cryptage. Ces deux addenda complètent le « cake » pour fournir les mécanismes d'échange de documents et données de tous types²⁹.

29 Synthèse tirée de Roger T. Pédaque, « Le Document à la lumière du numérique ». Bulletin des bibliothèques de France (BBF), 2007, n° 4, p. 122-123. Et Roger T. Pédaque,

Le Web sémantique propose des normes de codage logique des informations. Son but est de constituer une sorte de système d'exploitation des données du Web qui serait principalement au service des *moteurs de recherche* et des « *agents intelligents* ». En bref, comme l'indique le Cake, le Web sémantique repose sur un certain nombre d'outils, de langages de balisage. Le langage XML (eXtended Mark-up Language), dérivé du langage SGML de Charles Goldfarb, autorise la description de la structure des données (RDF, Ressource Description Framework) et permet le catalogage des données du Web. Le langage OWL (Ontology Web Language) est utilisé pour décrire les « ontologies », c'est-à-dire la structure conceptuelle des divers domaines de connaissances. Ces outils – descripteurs et marqueurs – ont pour but de favoriser l'automatisation des traitements dans la recherche des données et l'exécution des opérations confiées aux agents intelligents ou robots logiciels.

Le Web sémantique se spécialise donc dans la définition consensuelle de normes favorisant l'interopérabilité en ligne. Son efficacité repose principalement sur une vision réductrice et fermée des pratiques cognitives, des situations d'échange de traduction, des processus réels de travail et des différenciations dans les phénomènes essentiels de recherche, de navigation ou de lecture-écriture. Et les ontologies sont, de ce point de vue, dans l'ombre portée des « systèmes experts³⁰ » avec leur écologie inférentielle et propositionnelle.

Notre but cependant, n'est pas ici, d'explicitier en profondeur les débats et problèmes portés par l'approche ontologique, inférentielle et formelle. Ils ont été relativement bien identifiés à défaut d'avoir été surmontés³¹.

Il s'agit de rappeler quelques points concernant directement le spectre des écritures, des programmes et des technologies intellectives ainsi

Document : forme, signe et médium, les re-formulations du numérique. 2003 https://archivesic.ccsd.cnrs.fr/sic_00000511/document https://en.wikipedia.org/wiki/Semantic_Web_Stack

30 Système expert : https://fr.wikipedia.org/wiki/Syst%C3%A8me_expert et Jean-Max Noyer : « L'expertise stratégique face aux développements de l'intelligence artificielle », *Études internationales* Volume 19, numéro 4, 1988 <https://www.erudit.org/fr/revues/ei/1988-v19-n4-ei3035/702416ar.pdf>

31 Sur ces points voir parmi beaucoup de travaux : *Ontologies pour le web sémantique*, Jean Charlet- Bruno Bachimont, Raphaël Troncy, http://www.researchgate.net/publication/228905296_Ontologies_pour_le_web_semantique ; *What Is an Ontology ?* Nicola Guarino, Daniel Oberle, Steffen Staab, https://iaoa.org/isc2012/docs/Guarino2009_What_is_an_Ontology.pdf

que leurs rapports différentiels, leur hybridation possible (les ontologies n'étant qu'une partie des options) dans le cadre d'infrastructures pour la transversalité.

DES ONTOLOGIES « ALL THE WAY DOWN » ?

Concernant de la fabrication des ontologies (modélisation parmi d'autres de l'agencement productif des savoirs) en vue d'automatiser un certain nombre de procédures et afin de rendre plus aisé des fonctions élémentaires ou complexes des pratiques de la recherche, il convient toutefois de rappeler un certain nombre de points.

La fabrication de ces ontologies se fait en s'appuyant sur les **modes d'exposition** des recherches sur des conditions documentaires de ces dernières. Elles viennent dans un après coup de l'activité même de recherche.

Reflexives elles participent cependant à l'amélioration des pratiques de recherche internes au domaine dont elles sont issues et relativement, en s'insérant dans des ensembles documentaires plus vastes et hétérogènes aux divers modes de cheminement, aux nomadismes des concepts et des méthodes, aux processus de traduction, à l'exploration-chevauchement des zones frontières.

Cette fabrication des ontologies est complexe et repose sur un travail collectif qui oscille entre consensus et dissensus, entre modélisation du système de relations internes (inferences, analogies, abductions, construction et validation de la preuve, régimes de vérité) et « passages en force » contre, tout contre les dimensions spéculatives des sciences, des savoirs, et ce que l'on pourrait appeler leur monde imaginal³².

Mais les ontologies dans leur prétention à vouloir dire le processus de recherche et de pensée sont dans une fermeture logico-sémantique forte. (C. Hewitt³³)

32 Henry Corbin *L'Imagination créatrice dans le soufisme d'Ibn 'Arabî*, Édition Entreal, Paris, 2006.

33 Carl Hewitt, « Formaliser le raisonnement de bon sens pour une coordination évolutive de l'information incohérente et robuste à l'aide du Direct Logic™ Reasoning et du modèle Actor. » ; *Formalizing common sense reasoning for scalable inconsistency-robust*

Il faut donc, d'emblée s'inquiéter des risques de maladies auto-immunes engendrées par ces approches : res renforcés, frontières réductrices...

En effet, comment les ontologies définissent-elles leur rapport aux « dehors » ? Ce que l'on peut appeler le « petit dehors » qui est défini par le couplage entre champ ou discipline et son propre environnement, dehors qui est aussi celui des hétérogènes internes associées aux hétérogènes à cheval sur les frontières et ce que l'on pourrait appeler « le grand dehors », c'est-à-dire ce qui fait une fente dans ce que l'on pourrait appeler un dôme chaotique de buckminster fuller des savoirs et qui laisse entrer du chaos, à partir de disciplines et corpus plus larges, hétérogènes, instables. Dit autrement, y-a-t il une politique étrangère possible pour les ontologies ?

Charlet, Bachimont et bien d'autres ont, dès le début suggéré³⁴ un certain nombre de solutions pour faire face à ces questions y compris la co-existence productive entre ontologies hétérogènes. Ces solutions, selon eux, passent par deux types de propositions :

- des méthodologies séparant explicitement les termes et les concepts d'un domaine et c'est ce qui est expérimenté en médecine dans des serveurs de terminologie ou dans d'autres domaines dans ce que les chercheurs appellent des thesaurus sémantiques ;
- des méthodologies s'écartant de la formalisation des ontologies et recherchant des proximités conceptuelles dans les termes d'un domaine permettant d'en appréhender intuitivement la complexité.

“Cette dernière approche interroge directement les chercheurs du Web sémantique. Soit continuer à développer des ressources formelles, des ontologies, pour le Web sémantique, soit passer par des ressources moins formelles, ces deux possibilités n'étant d'ailleurs pas exclusives”³⁵.

information coordination using Direct Logic™ Reasoning and the Actor Model, <https://hal.archives-ouvertes.fr/hal-01148501v12>, 2017.

34 « Ontologies pour le web sémantique », Jean Charlet Bruno Bachimont, Raphaël Troncy, https://www.researchgate.net/publication/228905296_Ontologies_pour_le_web_semantique

35 Idem.

LE VERTIGE IEML COMME
“LINGUA CHARACTERISTICA UNIVERSALIS”

Ces problèmes sont même à la source d’un projet tel celui de Pierre Lévy, avec IEML (Information Economy MetaLanguage). « Je propose (écrit ce dernier) la construction d’une sixième couche – basée sur IEML – au-dessus du Web sémantique. IEML propose un système de coordonnées sémantique indépendant des langues naturelles, capable d’adresser une infinité de sujets différents et aptes à servir de base à des calculs de relations entre concepts. IEML a été conçu pour traduire les unes dans les autres les ontologies les plus diverses et pour interconnecter disciplines et points de vue divergents au sein du même système d’adressage.

Le langage IEML utilise XML et traduit des ontologies. Il n’est donc pas le concurrent d’un Web sémantique sur lequel il repose, au moins sur le plan technique. IEML a pour ambition de résoudre les problèmes de communication entre ontologies et de compatibilité entre architectures de l’information locales que le Web sémantique a permis de poser mais ne peut régler au niveau où il se situe. En somme, le langage IEML, avec le protocole de l’intelligence collective (CIP) qui organise son adressage numérique, veut constituer une nouvelle couche logicielle du cyberspace, ouvrant la voie à une informatique cognitive renouvelée (calculs sémantiques et pragmatiques) ainsi qu’à de nouveaux usages de l’Internet orientés vers le développement de l’intelligence collective, le pilotage distribué de l’économie de l’information et la gouvernance auto-organisatrice d’un développement humain multifactoriel et interdépendant³⁶. »

Les développements partiels et relativement fragmentés n’en sont pas moins suffisamment avancés pour converger vers des dispositifs locaux, des infrastructures localement opératives pour changer certaines des conditions d’exercice de la recherche et de la réflexion. Les données liées, les grammaires des metadata, la fabrication d’ontologies aux frontières plus ou moins productives, et de thesaurus complexes conçus à partir de données (au sens le plus extensif : données numériques, textes, documents

36 Pierre Lévy « IEML : finalités et structure », Working paper, Archives SIC, 2006. En ligne Voir aussi <https://pierrelevyblog.com/tag/ieml/>

hybrides, hypermédias) structurées, semi-structurées et non structurées, tout cela offre des possibilités de recherche navigation, d'association, mais encore analogiques étendues.

Une des incarnations les plus récentes qui fait converger un certain nombre des propositions précédentes hors IEMML, et qui tente de surmonter certaines des difficultés évoquées, est celle du projet "Open research knowledge graph" (ORKG)³⁷.

Ce projet prend acte que « l'échange d'informations devient de plus en plus sémantique et structuré dans un certain nombre de domaines : des graphes pour la connaissance encyclopédique, tel il y a environ une décennie, DBpedia, Yago et Freebase, ont créé les premiers graphes de connaissance pour la représentation de la connaissance encyclopédique, qui ont considérablement évolué ces dernières années et fournissent des milliards de faits sur la connaissance encyclopédique dans divers domaines. Après que DBpedia et Yago aient montré la valeur des graphes de connaissances encyclopédiques, Wikidata fournit désormais un graphe de connaissances organisé par la communauté de Wikipedia, utilisé pour élaborer des connaissances structurées dans Wikipedia de manière indépendante du langage et de qualité-certifié ».

« Le Open Research Knowledge Graph (ORKG) fournit des fonctions d'interconnexion, d'intégration, de visualisation, d'exploration et de recherche. Il permet aux scientifiques d'avoir une vue d'ensemble beaucoup plus rapide des nouveaux développements dans un domaine spécifique et d'identifier les problèmes de recherche pertinents. Il représente l'évolution du discours scientifique dans les différentes disciplines

37 « Du flux d'information fondé sur les documents au flux d'information fondé sur les connaissances en science et technologie Le TIB – Leibniz Information Centre for Science and Technology a publié un article de référence sur exposé de position sur l'*Open Research Knowledge Graph* », « *From document-based to knowledge-based information flows in science and technology The TIB – Leibniz Information Centre for Science and Technology has published a position paper on the Open Research Knowledge Graph* », <http://www.tib.eu/en/service/news/details/tib-publishes-position-paper-on-open-research-knowledge-graph> Bref ce type de « graphe ouvert » (ORKG) doit être constitué et organisé à partir de quatre sources complémentaires : « Les métadonnées, les données, les taxonomies, les ontologies et les modèles d'information existants, car les procédures automatisées à elles seules ne permettent pas d'obtenir la couverture et la précision nécessaires ; la "curation" entièrement manuelle prend trop de temps ; les bibliothécaires n'ont pas l'expertise nécessaire dans le domaine ; et les scientifiques manquent des compétences nécessaires en matière de représentation des connaissances. En combinant les quatre stratégies de manière significative, ils peuvent converger de manière performante. »

et permet aux scientifiques de rendre leurs travaux plus accessibles aux collègues et aux utilisateurs potentiels de l'industrie par le biais d'une description sémantique. (...) C'est donc un écosystème socio-technique pour la communication scientifique, le service ORKG est conçu comme un environnement de développement ouvert. À la base, il s'agit d'une infrastructure de gestion de données évolutive avec un modèle de données flexible basé sur des graphes, accessible via des API simples. »

AUTRES APPROCHES ET HYBRIDATION

Toutefois dans la recherche “d'infrastructures pour la transversalité” nous devons étendre notre effort vers d'autres approches – bottom-up – pour des onto-éthologies variables (“mettre la variation au cœur des ontologies” ?) dominées par des méthodes statistiques et ce pour que le mouvement inventif, spéculatif puisse se nourrir dans les agencements de rhizomes socio-cognitifs et de leurs interstices par où les mouvements de la recherche et de la pensée se produisent.

Bref, le lieu des chaosmoses de la pensée, des associations trouées et multifractales, de ce par quoi il y a de la créativité, de ce par quoi il y a au sein des pâtes plus ou moins homogènes de la recherche et de ses ralentissements (instances logiques, combinatoires avec leurs conditions et leurs systèmes de transmission, leurs mondes algorithmiques) des hétérogenèses plus ou moins contrôlées.

C'est la raison pour laquelle, nous pensons qu'il faut plus que jamais comprendre ce que doit être l'agencement des technologies intellectives que nous devons viser, dans le cadre des infrastructures des mémoires numériques. Et, selon les niveaux d'échelles où les pratiques se déploient, nous devons tenir compte des rapports entre métastabilité(s) et instabilité(s), entre normes pour produire de l'homogène et autres normes outils pour entrer encore une fois, ou aller à cheval sur les zones frontières. Entre ce qu'il convient de synchroniser et la diachronisation qu'il convient de maintenir. *Les boucles écritures-lectures sont puissantes aussi parce qu'elles ne cessent de produire des événements dans la pensée ou dans la recherche.*

Une autre conception tente de prendre en compte les pratiques communicationnelles « associés à la conduite d'interactions éphémères entre utilisateurs distants tout en offrant des représentations, souvent de nature graphique, des réseaux sociaux ainsi constitués ». Cette posture s'oppose donc à l'approche logiciste du Web sémantique formel. Elle défend une conception pragmatique des processus informationnels et communicationnels, tout en envisageant la linguistique et la sémiotique de manière plus ouverte. « Selon cette vision, le web est appréhendé avant tout comme un instrument de gestion documentaire facilitant des transactions coopératives interpersonnelles, éventuellement très asynchrones et distribuées entre des acteurs individuels et collectifs engagés dans des échanges, des débats, des controverses relevant de domaines très variés. Selon cette vision du Web, les modalités et les outils de gestion des documents doivent être pour partie conçus par les acteurs engagés dans une coopération active. Parmi ces outils le Web socio-sémantique préconise des cartes de thèmes ou réseaux de description que l'on peut considérer comme relevant des ontologies sémiotiques³⁸. »

Dans la perspective du déploiement des nouveaux agencements associés à l'exploitation créative de la logique hypertextuelle, nous pensons qu'il y a donc un grand intérêt à ne pas laisser le champ libre au seul formalisme évoqué par le « cake » de T. Berners-Lee.

En effet, son efficacité est subordonnée à une « fermeture sémiotique », c'est-à-dire à une réduction et à une standardisation des comportements et des pratiques.

La posture inverse, que nous défendons, prend radicalement en compte la sociologie en acte des pratiques et des usages, ainsi que les phénomènes de co-construction des connaissances, afin que « les langages syntaxiquement formels » puissent précisément être efficaces. Le problème est de concevoir des méthodes qui puissent représenter de telles structures sémiotiques et socio-cognitives, de manière à ce qu'un *formalisme faible* rende possible, à travers de nouvelles écritures, une approche pragmatique forte.

Nous verrons plus loin combien la position de Michel Callon et Bruno Latour marque de ce point un tournant. Et nous constaterons combien le programme informatique exposé est « pour » le programme sociologique

38 Zacklad M., « Introduction aux ontologies sémiotiques dans le Web Socio Sémantique ». Archives SIC, 2005. En ligne : [http://archivesic.ccsd.cnrs.fr/sic_00001479]

voire philosophique tel que ce dernier est exprimé de manière radicale dans « Irréductions³⁹ ».

Et il est en effet souhaitable de sortir, ou contre balancer l'approche logiciste du web sémantique, ce type de formalisation reposant sur des schèmes linguistiques logiques trop réductionnistes quant à la prise en compte des usages des communautés. C'est la raison pour laquelle il est important de discuter, de manière critique, l'élaboration de ces nouveaux alphabets, de leurs contraintes combinatoires et de leurs grammaires. Il convient également de réfléchir à des nouvelles manières « non-documentaires » de produire des « onto-éthologies » ouvertes et dynamiques, et ce afin de se rappeler que les écritures s'évaluent et s'imposent à partir de ce qu'elles ouvrent de créativité et d'inventions, de ce qu'elles portent de nouveaux modes combinatoires comme autant d'herméneutiques possibles.

« CAMINANTE NO HAY CAMINO
SE HACE CAMINO AL ANDAR⁴⁰ »

À la recherche donc de chaînes infométriques : quand les cartes des dynamiques de la science et leurs réseaux ne servent pas seulement à représenter mais à ouvrir les mouvements de la lecture et des cheminement, des analogies, des simulations : « mapping in the reading », « mapping in the writing » en quelque sorte. La création de ce type d'outils est essentielle afin de déployer de nouvelles configurations de lecture « transversales », lectures nomades et percolatives, de nouvelles pragmatiques internes aux modes de lectures⁴¹.

Les textes, les pratiques d'écritures et de lectures dont ils sont l'expression et l'exprimé, sont toujours des machines labyrinthiques, à n dimensions, qui ne cessent de créer les conditions de leur propre

39 Bruno Latour : *Irréductions in Les Microbes : guerre et paix, suivi de Irréductions*, 1984 Ed. La Découverte.

40 « Toi qui marches, il n'y a pas de chemin, le chemin se fait en marchant » Antonio Machado.

41 Stephen Ramsay, *Reading Machines : Toward an Algorithmic Criticism*, University of Illinois, Urbana, Chicago and Springfield, 2011.

démantèlement, c'est-à-dire de ré-écriture, re-lecture, de travail interprétatif, qui ne cessent d'ouvrir vers un nombre toujours plus grand de trouées, percées, chemins virtuels dont seul pourtant certains s'actualiseront. Ils ne sont jamais blocs denses et pleins, ils sont comme le cube de Serpienski ou l'éponge de Menger, territoires à la superficie potentiellement infinie, ouverts et connectables vers le hors-champ de chacun de nos mondes, des textes qui constituent notre milieu associé, notre niche éco-cognitive. Ils sont des architectures « différentielles » hypercomplexes créant les conditions matérielles et idéelles d'une tension permanente au milieu des coupures, des limites, des zones frontalières, des trous et des vides. Pleines et entières positivities de ces machines à vides, à fractures, brisures, par qui le mouvement de la pensée s'engendre, contre, tout contre les combinatoires et leurs contraintes, des signes, des traces.

Pleine et entière positivité des processus de chaotisation d'où émergent (auto-organisations souveraines), sous les conditions de production de ces machines textuelles, les ordres locaux, les formes métastables de la pensée. De la lecture-écriture donc comme art(s) complexe(s) des cartographies réelles et imaginaires pour un territoire étrange qui ne lui préexiste pas, sinon comme milieu virtuel associé des textualités non encore connectées, entre l'éclatante et noire positivité de l'écriture, des inscriptions répétées et l'obscur, tantôt glaciale et volcanique positivité des vides, des espaces deux fois troués qui leur sont couplés.

Tel est donc le travail de la lecture : à partir d'une linéarité ou d'une platitude initiale, cet acte de déchiffrer, de froisser, de tordre, de recoudre le texte pour ouvrir un milieu vivant où puisse se déployer le sens ? C'est en le parcourant, en le cartographiant que nous l'actualisons. Mais pendant que nous le replions sur lui-même, produisant ainsi son rapport à soi, sa vie autonome, son aura sémantique, nous rapportons aussi le texte à d'autres textes, à d'autres discours, à des images, à des affects, à toute l'immense réserve fluctuante de désirs et de signes qui nous constitue. Ici, ce n'est plus l'unité du texte qui est en jeu, mais la construction de soi, construction toujours à refaire, inachevée (Cette fois-ci, le texte n'est plus froissé, replié en boule sur lui-même, mais découpé, pulvérisé, distribué, évalué selon les critères d'une subjectivité accouchant d'elle-même⁴²).

42 Pierre Lévy, *Qu'est-ce que le virtuel ?* La lecture, ou l'actualisation du texte, (la virtualisation du texte) 1998, Ed. La Découverte.

C'est la raison pour laquelle, face à ces processus, on risquerait de manquer l'enchevêtrement des logiques, types de causalités, et couplages, des jeux combinatoires et associatifs, des rapports différentiels entre pleins et vides si l'on ne prenait la mesure des niveaux d'échelle des pratiques d'écriture-lecture indéfiniment ouverts, de la granularité complexe et mouvante des ensembles de documents et d'acteurs-auteurs associés, des cartographies et aides à l'orientation en devenir, de la métastabilité et processualité des hyperdocuments, ainsi que des communautés d'œuvres.

Dans ce contexte, on ne peut penser l'auteur, que comme « multiplixité » expression et exprimé de ses milieux associés des communautés interprétatives dans laquelle il est inclus ou qui le prennent pour cible. Les diverses formes de “text mining” appliquées aux échanges de type Twitter internes aux communautés scientifiques de même que les analyses dites de Clickstream⁴³ pour suivre les mouvements plus ou moins erratiques (sérendipité comprise) des cheminements pourraient être pour partie dévoilées. Cela posant au passage un problème majeur concernant la constitution de ces corpus. Ces derniers ouvrant vers ce que l'on peut appeler l'intimité des cérébralités en acte à l'occasion des travaux de recherches. Dans quelle mesure cela est-il souhaitable ?

Les nouvelles écritures rendent de plus en plus visible le fait qu'il est (seul ou à plusieurs) agencement collectif d'énonciation, milieu au milieu d'autres agencements collectifs, à leur traversée et les traversant. InSCRIPTION partielle donc de « cela », c'est-à-dire de cet agencement collectif, de cette mémoire collective en acte, distribuée selon des diagrammes spécifiques, des réseaux hétérogènes, hybrides, mais aussi de cet attracteur complexe des traces, trajectoires conceptuelles ou autres, entrelacées, qu'est le texte (dans ces deux sens), attracteur négociant vis-à-vis d'autres, sa puissance de capture et de traduction des sémiotiques en mouvement.

La pensée se déploie alors dans et à partir de cette zone incertaine formée à la jointure des mouvements de subduction et de convection engendrés par les couplages entre la productivité des contraintes combinatoires neuronales et la productivité des contraintes combinatoires liées aux modes d'écritures, aux modes linguistiques, sémiotiques... et aux modes sociaux de transmission de ces contraintes.

43 Johan Bollen, Herbert Van de Sompel, Aric Hagberg, Luis Bertencourt, Ryan Chute, Marko A. Rodriguez, Lyudmila Balakireva, Clickstream Data Yields High-Resolution Maps of Science, www.plosone.org, 2009.

Mouvements de subduction par lesquels vont se déployer en une alchimie noire les combinaisons chaotiques des pensées les plus frêles, des trajectoires les plus incertaines, des rapports de vitesse et de lenteur les plus subtils. Mouvements de convection par lesquels, contre et tout contre la matérialité des traces, de leur répétition, de leur combinatoire, vont se déployer coupures, failles, fractures, anfractuosités, lignes de fuites et entres elles des liens, des connexions, des ponts... Contre la pâte étouffante de l'homogène, contre l'éther anesthésiant du chaos, toute détermination est négation.

LES SCIENCES DE L'INFORMATION... ?

Il est à noter que les sciences de l'information, communication, prenant appui entre autres sur les sciences non-linéaires et la théorie du chaos, peuvent apporter, à côté des sciences de la cognition, de l'éducation, leur contribution à la définition des problèmes portés par cette complexité. Comment décrire et se représenter de telles pratiques cognitives distribuées ? Comment décrire et penser le statut des nouvelles médiations ?

Les sciences de l'information, communication – et plus encore ce qui se regroupe sous l'expression « information science » – se sont en effet, (nous l'avons aperçu) appliquées depuis une dizaine d'années au moins, à exploiter, dans leur partie fortement mathématisée les modèles et algorithmes de la géométrie en particulier et de la théorie du chaos⁴⁴ ». Les travaux produits, par exemple, par la scientométrie et l'infométrie, ont ainsi commencé à tirer parti des possibilités d'améliorer les analyses quantitatives, qualitatives concernant les relations dynamiques qui unissent des collectifs d'objets et/ou d'actants au sein des diverses communautés de recherche, et leurs modes de représentations graphiques. (Analyse des communautés, des fronts de recherche etc.) Des évolutions proches sont perceptibles aussi du côté de l'ingénierie documentaire, renforcées par l'apparition des gigantesques mémoires

44 Rafael Bailón-Moreno, Encarnación Jurado Alameda, Rosario Ruiz-Baños, Jean-Pierre Courtial *The pulsing structure of science* : Ortega y Gasset, Saint Matthew, fractality and transfractality *Scientometrics* 2007.

numériques portées par Internet. L'importance croissante en effet, des méthodes d'analyse statistiques et la nécessité de décrire et de discriminer les phénomènes d'émergence de formes stables, métastables, de formes fluides et instables au cœur de corpus hétérogènes ont renforcé l'usage de ces modèles. (Graphes conceptuels, mises en évidence des phénomènes de convergence, divergence participant de la création des réseaux d'acteurs réseaux complexes et dynamiques, mais aussi d'attracteurs). L'espace des mémoires numériques portées par Internet a déjà suscité des approches montrant que l'on pouvait considérer ce dernier comme un puissant système auto-organisateur.

Les hypertextes numériques en réseau sont en effets caractérisés par de multiples niveaux d'organisation systémiques, reliés par de nombreuses boucles récursives, de rétroaction. Par l'intermédiaire des multiples interfaces, logiciels, mémoires, mécanismes de répétition et de redondance dont ils sont porteurs, ils sont en effet marqués par une utilisation de structures hautement récursives. Ils se caractérisent aussi par un nombre important de règles et de pratiques locales, qui, par leur application répétée, favorisent l'émergence de structures plus vastes, plus complexes. Et les boucles de rétroaction agissent à l'intérieur et entre les niveaux de codes, de sémiotiques, de mémoires, de textes et ce, parce qu'ils sont habités par une grande variété de commandes et de modes d'écriture, largement distribuées sur un niveau d'échelle important, voir à tous les niveaux d'échelles. Ces dispositifs donnent aux méthodes, pratiques ascendantes, une place prépondérante.

Dans le même ordre d'idées, les pratiques, les comportements, les usages se développent en interaction directe avec l'environnement plutôt que par l'intermédiaire de modèles abstraits. Enfin les dispositifs de type hypertextuel incluent les utilisateurs dans leurs diverses boucles et les processus d'altération/création, d'association, peuvent se produire en des lieux multiples et ouverts. Pour aller à l'essentiel : à l'intérieur-extérieur des textes, des mémoires, à l'intérieur/extérieur des programmes, des logiciels, des interfaces enfin à l'intérieur-extérieur des pratiques collectives singulières et spécifiques ou de ce que l'on pourrait appeler "les cheminements synaptiques" liés aux multiples pratiques d'écriture/lecture, pratiques singulières et/ou collectives. Pour finir entre les multiples niveaux de production de sens et de représentation et les niveaux des codes, algorithmes utilisés pour produire ce sens, ces représentations.

Notons encore une fois, que ces caractéristiques montrent que l'on touche à un point critique lorsqu'on se concentre sur les phénomènes collectifs et sur ce fait que des savoirs, des connaissances peuvent être présents à un niveau d'échelle tout en étant absents au niveau inférieur. Dans le cadre de la cognition distribuée cela signifie que l'on ne peut faire l'impasse sur la multiplicité des logiques, des usages impliqués et que le statut des médiations, des technologies intellectuelles, les possibilités de connexion, de filtrage, de représentation qu'elles permettent, l'étendue de leur dissémination, jouent un rôle essentiel. Et ce d'autant que ces phénomènes collectifs mettent en jeu des hybrides comme agencements complexes de couplages d'éléments humains et non-humains. Certains travaux qui sont à l'œuvre dans le domaine des formes organisationnelles sont de ce point de vue, radicalement concernés et impliqués.

Autre secteur important des sciences de l'information-communication, la théorie "mémétique", principalement développée aux États-Unis propose, d'un point de vue néo-darwinien, une approche prometteuse quant à la compréhension des phénomènes de transmission duplication, propagation, répétition altération des unités de sens (au sens large "memes"). Cette posture pose un certain nombre de problèmes, en raison principalement de deux facteurs. Premièrement la question des échelles n'est pas clairement posée et deuxièmement, les concepts utilisés viennent en droite ligne se loger sous l'autorité non questionnée du programme fort de la génétique.

Programme dont le dogme central s'inscrit dans le fil du discours des essences et s'appuie sur l'affirmation que la transmission d'information est unidirectionnelle selon le schéma classique émetteur-récepteur ; point de processus ouverts, reprise classique du débat entre déterminisme biologique et environnementalisme, renoncement au modèle hasard-sélection dans le contexte général d'une émergence statistique du sens, des formes.

Au-delà de ces remarques, les théoriciens de la "memetic" offrent une ouverture vers les approches auto-organisationnelles et participent au développement des nouvelles technologies intellectuelles permettant d'explorer les problèmes complexes de la morphogenèse du sens.

Enfin ils participent au développement d'une activité de simulation particulièrement intéressante sur le plan cognitif où la notion d'attracteur vient approfondir par exemple la notion d'intentionnalité.

Ces modèles et ces algorithmes de simulation qui font appel pour une large part aux ressources des sciences non-linéaires permettent de mieux comprendre les divers types de mécanismes, de couplages impliqués dans les phénomènes interactionnels participant à l'émergence de formes organisationnelles plus ou moins stables et se révèlent de plus, être des outils d'écriture porteurs de nouvelles formes et pratiques cognitives⁴⁵.

Dans ce contexte, les questions et problèmes de la morphogenèse du sens, des processus d'inscription, de répétition, de transmission, de traduction, les modèles non-essentialistes de la communication, la critique des schèmes traditionnels de la théorie de l'information (menée par la cybernétique de seconde génération (*autopoïésis*), et les nouvelles conceptions de la biologie) se déploient « à cheval » sur les frontières disciplinaires.

Et la migration traduction, percolation des concepts et outils théoriques, qui n'est en aucune manière une transgression (c'est-à-dire une installation dans un au-delà radicalement nouveau, ou le dépassement illégitime de norme(s)), s'opère de manière plus forte et créatrice dans des régimes de discours, d'écritures très variés.

Ce qui semble le plus manifeste c'est tout d'abord, la prise de conscience d'avoir à penser autrement la question du collectif, de lui faire une place prépondérante, centrale et ce quelque que soit le niveau d'échelle où l'on considère les phénomènes.

Deuxièmement c'est l'introduction au cœur des multiplicités hétérogènes, comme conditions de leur émergence plus ou moins stables, d'une grande variété de type de « couplages » permettant la prise en compte de la participation permanente de phénomènes collectifs, hybrides au cours desquels de nombreux micro-événements, micro-actants, médiations sans rapport les uns avec les autres, ne cessent de se produire de manière plus ou moins désordonnée, de converger diverger au cours de processus d'actualisation et différenciation complexes.

Troisièmement, c'est la tentative de penser l'emboîtement de ces couplages, des modes ascendants et descendants et les divers modes de processus qui en découlent, avec des régimes de redondances et d'altération création spécifiques. Quatrièmement ce sont les efforts menés

45 Voir sur ces questions les travaux de David Chavalarias, David Chavalarias, Jean-Philippe Cointet *Phylomemetic Patterns in Science Evolution—The Rise and Fall of Scientific Fields*.

pour étendre ces réquisits et axiomes dans un nombre toujours vaste de phénomènes et processus (de recherche) conçus comme incomplétudes en procès de production⁴⁶.

Il s'agit d'appréhender le domaine des dynamiques et dispositifs informationnels communicationnel, comme champ processuel, toujours ouvert, où les notions, de stabilité, métastabilité morphogénétique prennent la place des essences. De ce point de vue tous les modèles interactionnistes hérités et encore dominants fondées sur le grand partage et les ontologies monovalentes sont de plus en plus fortement contestés.

Nombreux sont donc les associationnismes impliqués dans cette vaste hyperpragmatique au sein de laquelle la multiplicité des pratiques et dispositifs communicationnels aboutit à des instanciations singulières de cette hyperpragmatique généralisée selon les actants, médiations et niveaux d'échelles. Ces apports, qui ne sont pas les seuls, sont loin d'être négligeables afin d'appréhender les évolutions en cours du projet encyclopédique numérique. C'est là un ensemble de raisons pour lequel les caractères du document numérique doivent être « creusés », dès lors que l'on considère les ensembles de mémoires numériques associées en réseau de type hypertextuel, et que l'on veut en comprendre les modes d'exploitation en devenir.

LES PRINCIPES DU MODÈLE HYPERTEXTUEL

Pierre Lévy⁴⁷ a, au tout début des années quatre-vingt-dix, caractérisé « afin d'en préserver les chances de multiples interprétations » le modèle hypertextuel, par six principes abstraits :

« *Le principe de métamorphose* : le réseau est sans cesse en construction et en renégociation. . . Son extension, sa composition et son dessein sont un enjeu permanent pour les acteurs concernés, que ceux-ci soient des humains, des mots, des images. . .

⁴⁶ Whitehead A. N., *Procès et Réalité*, Éd. Gallimard, Paris, 1995.

⁴⁷ Pierre Lévy, *Les technologies de l'intelligence. L'Avenir de la pensée à l'ère informatique*, Paris, La Découverte, coll. Sciences et société, 1990.

Le principe d'hétérogénéité : les nœuds et les liens d'un réseau hypertextuel sont hétérogènes... Le processus sociotechnique mettra en jeu des personnes des groupes, des artefacts, des forces naturelles de toutes tailles, avec tous les types d'association que l'on peut imaginer entre ces éléments.

Le principe de multiplicité et d'emboîtement des échelles : l'hypertexte s'organise sur un mode "fractal", c'est-à-dire, que n'importe quel nœud ou n'importe quel lien, à l'analyse, peut lui-même se révéler composé d'un réseau, et ainsi de suite indéfiniment, le long de l'échelle des degrés de précision... »

Le principe d'extériorité : le réseau ne possède pas d'unité organique, ni de mot interne. Sa croissance et sa diminution, sa composition et sa recomposition permanente dépendent d'un extérieur indéterminé...

Le principe de topologie : dans les hypertextes, tout fonctionne à la proximité, au voisinage. Le cours des phénomènes y est affaire de topologies, de chemins...

Le principe de mobilité des centres : le réseau n'a pas de centre, ou plutôt, il possède en permanence plusieurs centres qui sont comme autant de pointes lumineuses perpétuellement mobiles, sautant d'un nœud à l'autre, entraînant autour d'elles une infinie ramification de radicules, de rhizomes ; fines lignes blanches esquissant un instant quelque carte aux détails exquis, puis courant dessiner plus loin d'autres paysages du sens⁴⁸. »

Ces dimensions du document numérique, dans les espaces-temps des mémoires numériques ne cessent donc de se creuser de manière intensive, suivant les niveaux d'échelle considérés. La notion de granularité « sémantique » s'impose donc pour penser ces nouvelles textures. En effet, la processualité des documents numériques se définit entre autres, par les divers types, les diverses formes et tailles des documents eux-mêmes, ainsi que par les caractéristiques des nœuds et des liens, autrement dit par les modes d'association et par cela même qui est associé. Ces modes d'associations sont très ouverts, potentiellement infiniment ouverts, dans l'espace du réseau. Toutefois, les liens entre des nœuds très variables, ouvrent un espace de jeux d'écriture asymétriques et complexes, un espace interprétatif très ouvert. De plus les états successifs (états métastables) des documents, parties de documents, fragments de documents mis en jeu au cours du processus d'écriture deviennent plus visibles.

48 Idem.

Les modes éditoriaux s'en trouvent donc affectés, transformés, étendus. Ils ne consistent plus à mettre en forme stable, finie, les résultats d'un travail, quels que soient les dispositifs de validation, légitimation qui disent la limite et la norme, qui annoncent l'arrêt suspensif du procès de recherche, figent le mode d'exposition et enfin disent : "ceci est achevé et donc peut être mis en jeu, exposé comme texte de référence".

À présent, en étant en mesure, d'exhiber de manière relativement stable, les éthologies « amont » conceptuelles ou autres, qui convergent, enveloppent le travail des textes et auxquelles les formes traditionnelles d'édition avaient renoncé, en étant en mesure de montrer les éthologies « aval » qui se développent à travers l'incessant travail de re-prise, de commentaire, de citation, de nouveaux modes éditoriaux voient le jour.

Ces nouveaux modes éditoriaux expriment de plus en plus précisément les dynamiques de construction des textes, le caractère de toute façon toujours transitoire des formes stables, leur fonction d'attracteur-transformateur, à la durée variable. En rendant donc plus visibles les sociologies qui sont à l'œuvre au cours des processus d'écriture vers et à partir des formes textuelles métastables, ils engagent un mouvement de contestation des dispositifs éditoriaux hérités et donc aussi des modes de fonctionnement des communautés, des liens qui les font être. Ils opèrent au cœur même de l'Économie Politique des Savoirs. Ils élargissent le mouvement déconstructif en ce qu'ils déplacent de manière concrète, le concept d'écriture, le généralisent et l'arrachent d'un certain point de vue à la catégorie de communication « si du moins on l'entend (cette dernière) au sens restreint de communication de sens⁴⁹ ».

ONTO-ÉTHOLOGIES ET ÉMERGENCE

Nous venons d'explicitier pour partie l'état actuel, dans l'univers des mémoires numériques des sciences, de la mise en mouvement des savoirs à la traversée plus ou moins complexe les uns des autres peut passer (peut-être)

49 Œuvre de Jacques Derrida à partir de *De la grammatologie* et de manière directe pour notre propos « Signature, Évènement, Contexte », 1971 disponible à <http://laboratoirefig.fr/wp-content/uploads/2016/04/SIGNATURE.pdf>

par des agencements de métalangages ou par une grammaire géo-graphique permettant de naviguer dans l'espace hétérogène des ontologies ou des « onto-éthologies » qui décrivent les savoirs spécifiques constitutifs du savoir scientifique général et processuel disponible sur le Web. Qui permettent d'établir de nouvelles connexions. Car, comme nous l'avons déjà dit, il est préférable à la place de faire porter tout le poids à la modélisation sous forme d'ontologies, d'accéder à la définition des « onto-éthologies » : elles expriment les structures socio-cognitives⁵⁰ portées par les corpus, les traductions et les processualités à l'œuvre au cœur même des communautés.

La « structuration » (formalisation) des textes et des documents, de même que leur filtrage, doivent être envisagés, dans leurs aspects techniques, sous une double contrainte. Il faut pouvoir traiter des populations de textes numériques, susceptibles d'être en permanence re-composées et trans-formées, d'une part ; il faut fabriquer des outils d'exploration et d'exploitation intellectuelles de ces populations, des outils de représentation de leurs processualités constitutives qui favorisent, nous l'avons déjà dit, des capacités analogiques, associationnistes et combinatoires, selon des niveaux d'organisation multiples, d'autre part.

Du programme scientométrique⁵¹ à la « Machine de Hume⁵² » en passant par des algorithmes visant la logique des associations et fondés

50 De « Mapping the Dynamics of Science and Technology, Sociology of Science in the Real World » Callon, Michel, Rip, Arie, Law, John 1986 à Collaboration and cognitive structures in social science research fields. Towards socio-cognitive analysis in information systems, Peter Mutschke, Anabel Quan Haase *Scientometrics*, Vol. 52, No. 3 (2001) 487–502 et Science models as value-added services for scholarly information systems Peter Mutschke • Philipp Mayr • Philipp Schaer • York Sure, *Scientometrics* (2011) 89:349–364. – Plus récemment, dans la suite de la sociologie des sciences latourienne voir, « Great Expectations méthodes quali-quantitative et analyse des réseaux sociaux » par Tommaso Venturini ; « Hyphe, a Curation-Oriented Approach to Web Crawling for the Social Sciences » par Mathieu Jacomy, Paul Girard, Benjamin Ooghe-Tabanou, Tommaso Venturini ; et « Quantitative and Qualitative STS : The intellectual and practical contributions of scientometrics, par Sally Wyatt, Staša Milojević, Han Woo Park and Loet Leydesdorff.

51 Xavier Polanco, « Aux sources de la scientométrie », *Revue Solaris*, in dossier n° 02, Bibliométrie, Scientométrie, Infométrie, sous la direction de Jean-Max Noyer, 1995.<http://gabriel.gallezot.free.fr/Solaris/d02/2polanco1.html> et Callon, Courtial et Penan (1993). *La Scientométrie*, Paris : PUF ; Courtial, J. P. (1990). *Introduction à la scientométrie. De la bibliométrie à la veille technologique*. Paris : *Anthropos Economica* ; Courtial, J. P. (ed.). (1994). *Science cognitive et sociologie des sciences*. Paris : PUF.

52 Geneviève Teil & Bruno Latour, *The Hume machine, can association networks do more than formal rules? How can we give qualitative analyses in social science mechanical means for dealing with large bodies of heterogeneous data? SEHR, volume 4, issue 2 : Constructions of the Mind 1995.*

sur le principe général de co-occurrences⁵³, c'est là l'effort qui depuis le début des années quatre-vingt, a sous-tendu les approches « bottom-up » pour exprimer les structures socio-cognitives des sciences, en étant éloigné des approches « symboliques » et inférentielles lourdes.

Face à de tels enjeux et problèmes il convient de se mettre en situation de pouvoir décrire les acteurs-réseaux et les actants⁵⁴ constitutifs du champ de recherche et de l'espace de déploiement des réseaux de savoirs. Bref, il faut pouvoir rendre compte des agencements collectifs, hétérogènes de production, d'énonciation, de légitimation des énoncés de la science, de leurs rapports différentiels et conflictuels. Il faut pouvoir les faire émerger à partir des traces exprimées, inscrites par les acteurs participant à l'élaboration de ces réseaux (Latour Bruno. – Callon Michel)

LEXIMAPPE ET CANDIDE COMME REPÈRES

Pour se donner des visibilitées différentes, plus variées, de viser les flux, les processus, les relations, les médiations, les intermédiaires, les traducteurs, les relais, les transformateurs qui définissent les rapports de forces en devenir ; et ce sinon à toutes les échelles, en tout cas selon le plus grand nombre d'échelles possibles et en utilisant les capacités offertes par les nouveaux dispositifs de la mémoire, la matière numérique et les techniques intellectuelles. Il faut apprendre à utiliser les traces et les systèmes de traces, d'indices produits et laissés par les acteurs eux-mêmes. Les modes de coopération-affrontement, bref, de mises en relation entre ces différents acteurs-réseaux (à l'intérieur des organisations ou entre elles), se matérialisent entre autres par la circulation de toute une série de textes, documents, programmes de recherches, déclarations... Ces textes constituent, individuellement ou pris ensemble, des dispositifs qui établissent des branchements et des connexions de toutes sortes avec d'autres textes et d'autres inscriptions littéraires, d'autres

53 M. Callon, J. P. Courtial, F. Laville, Co-word analysis as a tool for describing the network of interactions between basic and technological research : The case of polymer chemistry ; <https://link.springer.com/article/10.1007/BF02019280>

54 « Actant » est un terme sémiotique utilisé pour effacer la différence entre les acteurs humains et les non-humains objets. Il désigne tout (ce qui peut être dit) pour agir dans une histoire.

dispositifs... « *Au texte clos sur lui-même et auquel on opposerait de manière classique son contexte, il faut substituer ce texte, sans intérieur ni extérieur mais ramifié, feuilleté, opérant des branchements et des mises en relation, dispositif définissant des entités hétérogènes et les associant d'une manière particulière.* » (Callon Michel) Bref, « *partir de ce qui circule nous conduit, à ce qui est décrit par ce qui circule. (Décrire aux sens de description et de décrire une orbite)*⁵⁵ ».

L'élaboration de cartographies exprimant, les mondes conceptuels, les divers dispositifs collectifs de production et d'énonciation des sciences est donc centrale. Ces cartes pouvant à leur tour servir de filtres sur des écologies plus larges, ces cartographies fonctionnant alors comme outil d'exploration et d'interprétation d'univers complexes, suscitant la prise en considération de point d'entrées, de points de vue, d'états d'échelles multiples sur le ou les domaines, champs de forces considérés.

Recherche des traces, indices exprimant les conditions d'énonciation, mise en évidence des réseaux d'influence, des voies de circulation – presque de percolation des concepts, notions, relais, objets, systèmes hommes-machines, à travers entre autres l'analyse des mots associés etc. sont ici des axes précieux. La possibilité de traiter, d'examiner de gros corpus documentaires, des données hétérogènes, rend possible le développement « *d'analyses relationnelles qui consistent à imaginer des indicateurs fournissant la morphologie des interactions et qualifiant les éléments en interaction*⁵⁶. »

Nous optons donc, à la suite de Michel Callon et Bruno Latour pour une « modélisation faible », une modélisation émergente, c'est-à-dire que nous mettons en place les conditions ouvertes de suivre au plus près la richesse des processus de production-circulation des savoirs et des connaissances en tenant compte des nouveaux modes de collecte, stockage, traitement des informations, à partir de nouveaux modes d'organisation du travail intellectuel, de Réseaux Hybrides Intelligents.

Selon Latour et Teil, tel qu'ils exposent le devenir en cours dans « la Machine de Hume », « ce qui se passe, c'est que nous nous rapprochons de plus en plus des techniques de la narration – un outil essentiel pour les historiens, les ethnologues, les sociologues de terrain et les naturalistes – et de la description de réseaux d'associations constitués d'un fouillis de bases de données hétérogènes. Dans cette fusion de qualités littéraires qualitatives et de la puissance du traitement quantitatif, nous attendons

55 Michel Callon, *La dynamique des acteurs réseaux sociaux économiques*, CSI, Working paper1990.

56 idem.

un renouvellement des méthodes et des explications en sciences humaines. Il n'existe pas d'explication plus puissante que l'analyse des circonstances contingentes des réseaux d'association, mais pour le moment, le seul moyen d'obtenir cette forme de signification consiste à utiliser un récit limité aux terrains restreints. Jusqu'à présent, le seul remède à cette limitation consistait à recourir aux tableaux statistiques et aux analyses quantitatives. C'était au prix d'une rupture avec le tissu délicat des réseaux et des circonstances, d'où l'interminable débat entre sociologie du terrain et sociologie des structures, entre histoire et sociologie, entre histoire économique et théorie économique, entre arts et sciences, entre histoire de philosophie de la science et de la construction de modèles. La machine Hume ouvre un itinéraire alternatif. Qualitatif et quantitatif à la fois⁵⁷.

Nous proposons rapidement de décrire un tel mode d'approche qui laisse entrevoir comment il est possible de faire émerger des dispositifs hétérogènes impliqués. Suivant les types de traces pris en compte, suivant la taille des corpus documentaires établis et suivant les modes d'indexation et les outils mathématiques utilisés, on peut ainsi faire apparaître les réseaux selon des états d'échelles différents et selon des points de vue variables tels qu'ils sont exprimés par la pratique et son cortège d'inscriptions, de médiations.

Le principe de calculabilité de la Machine de Hume et la méthode des mots associés : « La robustesse des relations structurées ne dépend pas des qualités inhérentes à ces relations mais du réseau d'associations qui en constitue le contexte. Le principe de départ de la construction de la machine Hume est un principe de calculabilité différent de celui des machines de Turing, mais qui occupe la même position stratégique pour notre projet que le sien.

Le raisonnement est le suivant :

- toute forme est un rapport de force ; B) toute relation de force est définie dans un procès ; C) tout essai peut être exprimé sous la forme d'une liste de modifications d'un réseau ; D) tout réseau peut être résolu en une liste d'associations d'actants spécifiques et contingents ; E) cette liste est calculable.

57 *The Hume machine, can association networks do more than formal rules ?*, In Stanford Humanities Review Vol. 4, n°2 (with G. Teil), p. 47-66, 1995 Geneviève Teil & Bruno Latour SEHR, volume 4, issue 2 : *Constructions of the Mind Updated 4 June 1995*.

Il n'existe donc pas de concept formel plus riche en informations que celui d'une simple liste d'actants spécifiques et contingents. On a tendance à croire que nous sommes mieux lotis avec des catégories formelles qu'avec des faits circonstanciels, mais les formes ne sont que le résumé d'un réseau : c'est-à-dire du nombre et de la répartition des association⁵⁸.

L'analyse des mots associés est une technique statistique introduite conjointement par le Centre de Sociologie de l'Innovation (CSI) de l'École des Mines de Paris, et le Centre de Documentation Scientifique et Technique (CDST) du CNRS (Courtial J. P, Michelet B., Turner W.). Contrairement aux approches lexicométriques, l'approche des mots associés vise l'analyse des cooccurrences de termes. Elle permet d'exprimer les convergences d'intérêts entre acteurs, actants (au sens le plus extensif de ces termes), de faire apparaître des agrégats, des agencements, de les hiérarchiser selon la nature et la taille du corpus sur lequel on travaille, corpus qui rappelons-le doit être validé, clôturé. La question des traces et de leur pertinence est plus que jamais essentielle

Brièvement donc, cette méthode est fondée sur la co-occurrence des termes indexés. L'association de deux mots-clés (i, j) – dont l'occurrence (C_i, C_j) est égale ou supérieure à 2 – est mesurée en fonction de leur nombre d'apparitions communes (co-occurrence) dans les documents qu'ils indexent (C_{ij}). L'indice statistique qui mesure la valeur de l'association entre les mots-clés est le coefficient d'équivalence. Si l'on note C_{ij} la co-occurrence entre i et j , et C_i, C_j leur occurrence, l'indice peut s'écrire sous la forme suivante : $E_{ij} = (C_{ij} / C_i) \times (C_{ij} / C_j) = C_{ij}^2 / C_i \times C_j$.

Les associations sont ainsi affectées d'une valeur. Tous les couples de termes obtenus sont triés par valeurs décroissantes. Ensuite, on parcourt

58 "The hume machine, can association networks do more than formal rules?" In Stanford Humanities Review Vol. 4, n°2 (with G. Teil), p. 47-66, 1995. Geneviève Teil & Bruno Latour *SEHR, volume 4, issue 2 : Constructions of the Mind Updated 4 June 1995*. Notre procédure (disent Latour-Teil) est liée aux idées utilisées dans le domaine des réseaux de neurones et par les connexionnistes. Par exemple, voir David E. Rumelhart, James L. McClelland et le PDP Research Group, *Traitement parallèle distribué : Explorations dans la microstructure de la cognition* (Cambridge, MA : MIT Press, 1988). Les réseaux de neurones résolvent avec élégance certaines difficultés de calcul et de reconnaissance de forme en court-circuitant la programmation traditionnelle par étapes et par règles. Cependant, ils présupposent l'existence de formes très structurées, soit au début, soit dans leur sortie. Même s'ils offrent certains avantages pour la programmation de la machine Hume, ils n'ont aucun moyen de résoudre le problème fondamental de son architecture : il n'y a pas de formes structurés pour commencer, même pour récompenser le réseau.

séquentiellement la liste classée des couples pour construire les clusters. Tous les éléments à clusteriser (agréger) forment initialement un seul grand réseau d'associations. Il s'agit d'un réseau valué, c'est-à-dire, d'un système de relations dans lequel les mots sont reliés par des liens plus ou moins forts.

Le découpage du réseau d'associations en clusters (en agrégats) se fait en fonction d'un critère de lisibilité, à savoir la taille : fixant le nombre maximal et minimal de composants et le nombre de liens. Si un couple de termes appartient au même cluster, le lien entre ces termes est considéré comme lien interne au cluster. Si les termes d'un couple appartiennent à deux clusters différents, leur lien est considéré comme lien externe, c'est-à-dire un lien entre clusters. Sont ainsi exprimées la position respective des agencements les uns par rapport aux autres et la stabilité plus ou moins grande de ces mêmes agencements, leur capacité « d'affecter ou d'être affecté ». Les « clusters » sont donc situés dans un espace à deux dimensions, et sont situés dans un plan défini par un coefficient de cohérence interne du thème et par un coefficient de centralité. La « cohérence interne » d'un agencement est ici la moyenne des valeurs des associations « internes ». Cette valeur est notée sur l'axe vertical *y* de la carte. Plus la moyenne est élevée, plus il est considéré comme étant un agencement bien structuré et reconnu. La « centralité » d'un agencement (cluster) est mesurée par la moyenne du nombre total d'associations externes qui existent entre l'agencement donné et les autres. Cette valeur est notée sur l'axe horizontal *x* de la carte. Plus la moyenne est élevée, plus l'agencement en question est considéré comme un point de référence (centralité) pour l'ensemble.

Outre les coefficients de Densité et de Centralité – qui expriment : pour le premier, la solidité, le degré d'homogénéisation, la stabilité, la richesse d'un cluster ; et pour le second, la capacité de connexion, la puissance de capture, la capacité d'affecter ou d'être affecté du même cluster –, le profil d'association de chaque « actant » est le profil de l'entité linguistique de référence et qui permet de définir cette dernière de manière univoque au sein du réseau, peut être interprété de la même manière. « Chaque point dans le réseau peut être défini par la liste de ses associations directes à d'autres points du réseau. (...) » De la sorte « on peut dresser une typologie descriptive des points à partir de leur profil... de “grands” points ont un profil très long et se connectent à de

nombreux autres points dans le réseau. De “petits” points occupent une position très localisée voire marginale dans le réseau⁵⁹. »

Au terme – provisoire – de ce dévoilement des configurations à l’œuvre, aux échelles déterminées à la fois par le type de traces pris en compte (moment de la constitution du corpus et légitimation de sa fermeture) et les seuils d’associations sélectionnés pour calculer le réseau d’association lui-même, nous nous sommes donnés en outre, exploitant les capacités hypertextuelles émergentes (représentation cartographique et semi-dynamique du réseau qualifié des clusters – exploration libre), non seulement la possibilité d’interpréter sous des conditions de visibilité nouvelles, mais encore de relancer le processus d’exploration et d’écriture des dispositifs impliqués. En effet sont, de facto, indiqués, pointés les domaines ou les actants vers lesquels l’exploration peut être poursuivie et amplifiée. Par exemple, les réseaux de chercheurs, les collègues invisibles à l’œuvre, les liens avec des domaines de recherche très fins, les entreprises liées au complexe militaro-industriel... analyses de contenu, description d’acteurs-réseaux hétérogènes invitant à parcourir selon des niveaux d’échelles variables, les diagrammes complexes, à n dimensions définissant notre champ de recherche. D’une autre manière, ce sont d’autres corpus à constituer qui sont suggérés, d’autres types d’indices, de traces à capturer, à inventer. De même dans les contextes variés de veilles techno-scientifiques et stratégiques, l’analyse répétée dans le temps, permet de suivre le déplacement des actants selon les axes de centralité et de densité, leur apparition et/ou leur disparition.

SORTIR DE L'IMPÉRIUM BINAIRE :
QUANTITATIF / QUALITATIF :
LA MACHINE DE HUME FRAYE SA VOIE

Comme la variation de la précision dans la Mesure ouvre vers de nouveaux mondes qualitatifs, de nouveaux paradigmes conceptuels, perceptifs, intensifs, la variation des couplages structurels Pensée, écritures,

59 Geneviève Teil, *Candide, un outil de sociologie assistée par ordinateur pour l’analyse quali-quantitative de gros corpus de textes*, www.theses.fr/1991ENMP0291

algorithmes, dans les va-et vient de leur co-détermination, est le foyer transductif en mouvement. C'est la raison pour laquelle l'ouverture humienne dans sa simplicité et humilité est radicale.

Bruno Latour et Geneviève Teil exprime cette radicalité en écrivant « que le seul point de départ nécessaire nous est offert par l'ordinateur lui-même. Il s'agit d'un réseau d'associations entre des contingences, n'ayant d'autres caractéristiques a priori que celles qui sont différentes les unes des autres, et qu'elles peuvent être traitées. Est-ce que l'ordinateur est aveugle ? Mais nous aussi. Est-ce que l'ordinateur a des règles formelles pour commencer, non ? Nous non plus. L'ordinateur ne traite-t-il pas les abstractions ? Nous non plus. Est-ce que l'ordinateur va cheminant, d'un essai à l'autre, d'une circonstance à l'autre ? Nous aussi, et nous n'en demandons plus. »

Ils enfoncent le cou et mettent en évidence « le paradoxe des discussions sur les possibilités des ordinateurs est qu'ils se voient attribuer des qualités qu'ils n'ont pas – le formalisme, le rêve épistémologique de l'homme – alors qu'ils se voient refuser les capacités herméneutiques qu'ils possèdent déjà. En réalité, l'ordinateur est déjà un système ouvert. Il respecte les aléas et les spécificités bien plus que ne le croient les humains qui tentent de le programmer. Nous renonçons ensuite à pouvoir jamais en sortir ce dont il est déjà capable, à la seule condition d'abandonner nos fantasmes épistémologiques. Cette conception de ce que font les ordinateurs est clairement le résultat de l'application de notre argument concernant l'origine des formes structurées, puisque c'est cet argument qui nous permet de contester des projections anthropomorphes sur des ordinateurs. Mais ce résultat est à son tour ce qui nous permettra d'utiliser immédiatement l'ordinateur pour prouver notre argumentation concernant l'origine réseau de ces formulaires »

La conclusion est donc la suivante : « la machine Hume est une machine associationniste et n'est que cela. Il ne vient avec aucune catégorie logique, aucune forme syntaxique, aucune structure. C'est, osons dire, aveuglément empiriste. Ce que nous espérons tirer de notre stratégie, c'est de trouver, au lieu du chaos auquel on pourrait s'attendre, toutes les propriétés émergentes dignes de notre attention et qui nous permettront de circuler entre les micro-théories⁶⁰.

60 Carl Hewitt, *Formalizing common sense reasoning for scalable inconsistency-robust information coordination using Direct Logic™ Reasoning and the Actor Model*, disponible à <https://hal.archives-ouvertes.fr/hal-01148501v12> 2017.

Tout le travail qui se développe à partir de ce cadre conceptuel très général afin de rendre plus efficace l'analyse des réseaux dans le cadre du web et dans le secteur des Sciences sociales est ici le bienvenu et favorable à l'étude des communautés de chercheurs et des agencements des sciences.

Si nous récapitulons nos exigences et notre ambition les mémoires de l'édition électronique et leurs infrastructures, dès lors qu'elles visent à développer les milieux d'intelligence dont elles porteuses sont confrontées à un certain nombre de contraintes et de logiques de développement. Les dimensions collaboratives et coopératives étant au cœur des relations qui tissent ces milieux depuis le début, deux logiques en effet co-existent plus ou moins harmonieusement : « une logique normative, ou "top-down", promue par les partisans du Web Sémantique, logique qui garantit une certaine interoperabilité et la facilité d'analyse », logique qui s'actualise et se complique avec ce que l'on appelle les domaines des ontologies, (enchâssées dans un cadre documentaire reflet des consensus de chaque champ de recherche) et qui reposent sur des modèles conceptuels, relationnels, linguistiques, logiques, forts⁶¹ ; et de l'autre une logique émergente, qui garantit un champ d'application beaucoup plus vaste (prenant l'information, structurée ou non ; tous les usages, prévus ou non). Toutefois, pour suivre François Bourdoncle, ces logiques ne s'opposent pas de manière frontale et il y a même tout intérêt à mettre en œuvre des approches hybrides pour les raisons déjà évoquées. Les normalisations, les systèmes classificatoires (même hérités), les inférences élémentaires ou complexes et stabilisés de certains champs de la connaissance, peuvent rendre plus aisé les approches émergentes, le travail de structuration de l'information non-structurée ou semi structurée. Les approches pragmatiques ne rechignent pas loin de là à s'appuyer sur les diverses technologies du Web sémantique. On peut imaginer renforcer la finesse mises à jour des rhizomes plus ou moins stables des lieux de savoirs grâce à des pratiques de social-tagging sur les contenus et l'approfondissement des modes de lectures⁶². Sur ce dernier point on sait que les nouvelles écritures numériques transforment les pratiques de lectures. De plus si nous prenons en compte le fait de la co-présence des mémoires « papier » avec les mémoires numériques, on assiste là une différenciation des modes de lecture, à leur enchevêtrement. La tension entre le « lire de près » (supposé être attaché au document papier) et le « lire de loin » (attaché au numérique) est trop imprécise, radicale⁶³. Toutefois dans l'univers des livres « le près » est relatif et il n'est pas séparé de la profondeur de champ due aux

61 François Bourdoncle, *L'intelligence collective d'usage*. Co-fondateur, *Exalead*, Dassault in *Technologies de l'information et intelligences collectives*, Jean-Max Noyer, Brigitte Juanals, éd. Hermes-Lavoisier 2011.

62 Yves Citton, *Comment nous lisons ? Lecture rapprochée, hyperlecture, lecture machinique* <http://www.yvescitton.net/wp-content/uploads/2017/09/HAYLES-LireEtPenserEnMilieuxNumeriques-Chap3-CommentNousLisons-2016.pdf>

63 Stephen Ramsay *Reading Machines : Toward an Algorithmic Criticism* (Topics in the Digital Humanities) 2011.

processus des hypertextes internes au livre lui-même et relayé vers le lointain par les configurations de citabilité et des références. Il y a du lointain dans la lecture de près. Ce que les mémoires numériques rendent possible c'est, d'une part l'extension et la différenciation des associations, des plis et des déplis par quoi se définit pour une large part, pour reprendre Pierre Lévy⁶⁴, la Lecture, mais c'est aussi la possibilité de produire des logiciels, des algorithmes pour mettre à jour les réseaux, la structure fractale voire multifractale de ces réseaux, pour étendre la lecture en amont et aval des documents ; en zoomant en quelque sorte sur les toiles d'araignées à n-dimensions de nos lectures, au gré des plis et déplis qui rendent la lecture plus productive⁶⁵.

Les algorithmes qui rendent cela possible peuvent être dits « réflexifs ». D'un point de vue général un grand nombre de logiciels de datamining, de textmining développés depuis les années soixante (dans le cadre du programme scientométrique) ont en vue l'explicitation des réseaux conceptuels, des réseaux collaboratifs entre chercheurs, entre laboratoires... afin de mieux habiter ces milieux d'intelligence, ces « collectifs de pensée », ces agencements qui produisent des savoirs.

Penser les rapports entre « les près » et les « lointains » de la lecture. Et la numérisation complique les rapports entre les « essences » documentaires, les blocs d'écritures et la granulosité sémantique, les trajectoires associatives, analogiques, les processus de transformation sémiotique et les pragmatiques qui vont avec dans le cadre d'une nouvelle alliance entre les images, les textes et les sons. Il y a dans cette capacité à produire des diagrammes, l'espoir que nous « pourrions apprendre beaucoup sur la forme roman et ses avatars, ainsi que sur les forces à l'œuvre dans l'« écosystème » littéraire et le monde social qui contribuent à leur transformation⁶⁶ » Cela vaut aussi depuis que la sociologie de la science s'est occupée d'accéder aux « collèges invisibles » aux fronts de recherche...

De telles infrastructures incluant les technologies intellectives pour des approches émergentes donneraient consistance à « des intelligences collectives d'usage » dans le domaine des agencements de savoirs, ses socles pour « travailler » dans leurs hétérogénéités.

64 Pierre Lévy, *Qu'est ce que le virtuel ?* La lecture, ou l'actualisation du texte, (la virtualisation du texte) 1998, Éd. La Découverte, disponible à : http://www.manuscritdepot.com/edition/documents-pdf/pierre-levy-le-virtuel_01.pdf et aussi <http://hypermedia.univ-paris8.fr/pierre/virtuel/virt0.htm>

65 Franco Moretti, *Graphes, cartes et arbres. Modèles abstraits pour une autre histoire de la littérature, Les prairies ordinaires*, 2008.

66 Pour reprendre ici Marc Escola dans un commentaire de l'ouvrage de Franco Moretti.

De l'intérêt des modèles et technologies déployées dans le monde industriel à travers les Systèmes d'Informations d'entreprise (dit à l'époque « intranet ») et qu'il y aurait à en appliquer (en l'adaptant) le modèle⁶⁷.

C'est là nous semble-t-il en effet que certaines des approches développées dans le domaine industriel proposent à travers leur puissance la possibilité de produire cartes et metadata en temps réel de manière émergente et ce selon des échelles variables dans le cadre de ce que l'on peut appeler des recherches contextualisées

Le principal enjeu pour cela en suivant la voie proposée par François Bourdoncle⁶⁸ est de concevoir « des dispositifs qui sachent extraire les faits les concepts, les actants et les systèmes plus ou moins différenciés de leurs relations et qui sachent “interpoler” de manière automatique ou semi-automatique des thesaurus de faits, de concepts, d'actants et de relations au-delà des corpus ». Et en raison de la taille des mémoires de l'édition numérique scientifique, la gestion locale ou globale des corpus structurés ou pas doit « supporter des ressources linguistiques et sémantiques de taille elles-mêmes considérables. »

Les technologies utilisées pour parvenir à ces résultats ne peuvent pas être des technologies purement linguistiques, car ces techniques certes très précises, souffrent de deux défauts rédhibitoires. Tout d'abord, elles sont très coûteuses en temps de calcul, quasiment de deux ordres de grandeur trop coûteux, ce qui rend leur utilisation impossible à grande échelle et ce pour un futur encore assez lointain. La seconde raison c'est que les thesaurus linguistiques sont très pauvres vis-à-vis de la richesse terminologique que l'on trouve dans (les corpus des savoirs) « Il est donc indispensable « d'interpoler les thesaurus par des techniques d'apprentissage » (« machine learning » en anglais). Ces techniques d'interpolation partent d'une idée simple : observer le contexte dans lequel certains mots « semblables » (par exemple des noms de couleurs) apparaissent, modéliser ce contexte, et ensuite détecter toutes les occurrences de ce contexte avec des mots inconnus pour compléter la liste⁶⁹.

67 François Bourdoncle, « Internet et les systèmes d'information (SI) de l'entreprise », *Réalités Industrielles*, Novembre 2010.

68 François Bourdoncle L'intelligence collective d'usage Co-fondateur, *Exalead*, Dassault in *Technologies de l'information et intelligences collectives*, Jean-Max Noyer, Brigitte Juanals, éd. Hermes-Lavoisier 2011.

69 Idem.

LIVE TOPICS – EXALEAD

Dans un bref mais séminal article François Bourdoncle et Patrice Bertin⁷⁰ proposent de définir de nouvelles fonctionnalités pour les moteurs de recherche. Il s'agit pour eux de sortir d'une approche essentialiste de la recherche (liste de documents à la suite d'une requête) pour viser ce que l'on pourrait appeler de manière plus précise que « recherche contextualisée », recherche par « clusterisation » ou par micro-agencements. Certes ils n'avaient pas forcément en tête la conception sémio-linguistique de Deleuze-Guattari et leur critique des postulats de la linguistique.

Mais comme le notait Tracy Stanley en 1997⁷¹ dans une excellente analyse de Live Topics, en cherchant à surmonter le problème de la surcharge d'informations et des listes documentaires primitives qui ne permettent pas d'en comprendre le contenu et de trouver des informations réellement pertinentes, les concepteurs de LiveTopics pour Alta Vista (le moteur de référence des années quatre-vingt-dix avant l'avènement de Google) « avaient pour objectif de contourner ce problème en utilisant une méthode connue sous le nom de *catégorisation dynamique des résultats*. Cela signifiait qu'une fois la recherche effectuée, Alta Vista (pouvait) regrouper les résultats dans un certain nombre de sujets ou de thèmes clés différents, organisant ainsi des pages Web ayant un contenu similaire sous le même en-tête. Ce regroupement n'est pas prédéfini, mais est effectué "à la volée" en fonction des résultats de la recherche. Les pages sont automatiquement analysées pour leur contenu et l'analyse statistique est utilisée pour regrouper les sujets en thèmes ; ceci est basé sur le nombre d'occurrences de mots dans chaque document faisant partie des résultats de la recherche. Le grand avantage de cela est que la catégorisation des documents peut changer au fur et à mesure que leur contenu change, au lieu d'être figée. L'utilisateur peut ensuite passer en revue ces sujets et affiner leur recherche en choisissant ceux qui semblent le plus appropriés et adaptés à leurs besoins. Des informations non pertinentes peuvent être exclues de la recherche. »

70 François Bourdoncle, Patrice Bertin, « Le Web, et encore le Web, Recherche d'aiguilles dans une botte de liens », *La recherche* N° 328, 2 2000.

71 Tracey Stanley Search Engines Corner Alta Vista LiveTopics, 19 May 1997 in ARIANDN, <http://www.ariadne.ac.uk/issue9/search-engines#author1>

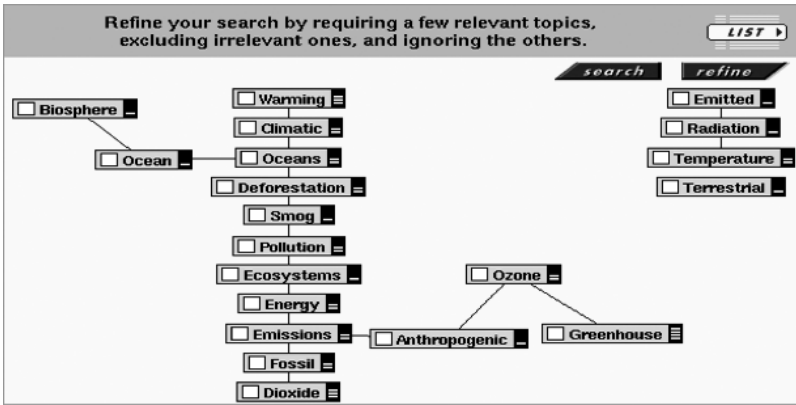


FIG. 2 – Live topics 1996-1997 (Document personnel).

FIG. 3 – Live topics 1996-1997 (Document personnel).
 Source <http://www.ariadne.ac.uk/issue9/search-engines>

Comme on peut le voir ci-dessus, à chaque requête le logiciel génère une carte de mots-clés reliés entre eux (cluster) à chaque mot-clé dominant du cluster est associé un certain nombre d'autres mots (constituant un sous cluster). La carte exprime ainsi de manière certes grossière, le réseau-agencement dans lequel est inclus le concept ayant servi de requêtes et produisant simultanément le corpus associé. Ceci étant fait en temps-réel la conception de LiveTopics permettra de rajouter (Exalead) la fonction « REFINE » elle-même permettant de faire, à partir du premier réseau, varier la requête de multiples manières, générant ainsi un nouveau corpus et sa carte associée.

Cela nous semblait-il, était un embryon d'outil de navigation dans la géographie à n-dimensions des corpus quelle que soit leur taille ou échelle et pour des données structurées, semi-structurées ou non-structurées.

Ce que visait LiveTopics au-delà d'une localisation rapide d'information, mettre en place les conditions pour les pratiques et besoins élémentaires en termes de veille concurrentielle-Intelligence économique-veille scientifique et technique, veille informationnelle et de conception des intranets.

D'une manière générale on peut dire que cette capacité à faire du « *mapping in the making* » au cœur des systèmes d'information d'entreprise (à partir de la reprise des « Intranets » et du dataming de l'hypermarketing)⁷² exploite massivement cette i en temps réel inspirée de la scientométrie des années soixante. Scientométrie qui avait pour but en exploitant les grandes bases de données de produire de nouvelles conditions pour la définition des politiques scientifiques. Les bases créées du SCI et SSCI l'ont été dans cet esprit⁷³. C'est aussi pour développer de nouvelles visibilitées concernant les dynamiques des sciences, les réseaux d'acteurs que s'est développé dès les années soixante donc le programme scientométrique orienté vers la sociologie des sciences. Ce n'est pas un hasard si dans l'article séminal évoqué plus haut, il est fait référence aux travaux d'E. Garfield sur les citations, aux travaux (accrochés au programme scientométrique⁷⁴) des représentants du centre de sociologie de l'innovation

72 Jean-Max Noyer, *Les vertiges de L'hypermarketing : dataming et production sémiotique* : <https://books.openedition.org/pressesmines/1662?lang=fr>

73 Science Citation Index et Social Science Citation Index.

74 Xavier Polanco, « Aux sources de la scientométrie », *Revue Solaris*, in dossier n° 02 Bibliométrie, Scientométrie, Infométrie, dir. Jean-Max Noyer, 1995.

de l'École des Mines de Paris⁷⁵, de H. Small sur les co-citations dans la lecture scientifique⁷⁶.

LA QUESTION DES ÉCHELLES, DISSÉMINATION DES ARCHIVES ET FORMES COURTES

De l'intérêt qu'il pourrait y avoir à inclure ce type de technologies dans les infrastructures de l'Édition numérique et à la décliner et à la disséminer aux niveaux d'échelle qui conviennent et aux points stratégiques où les savoirs se fabriquent, se déplacent, se transforment et s'affrontent⁷⁷. C'est ce que la réflexion engagée autour des « Aolab⁷⁸ » (archive ouverte de laboratoire) tentait d'exprimer.

Au cœur des processus d'émergence intellectuels formes courtes et longues ont besoin d'un milieu rendant possible une plasticité des rapports de vitesse et de lenteur dans la co-construction plus ou moins conflictuelle en tout cas controversé des savoirs. Les formes courtes jouent un rôle essentiel. En effet elles ne sont jamais centrées sur elles-mêmes, elles ne cessent de porter et d'exprimer qu'elles sont toujours en attente d'une autre forme, d'un autre texte, d'un autre argument, d'un autre récit d'un autre événement de pensée ? Cette capacité subversive et productive d'instabilités, d'anaphores intellectuelles peut être forte.

Et ce d'autant que les modes d'inscription et de répétition-altération, d'expérimentation, de simulation sont porteurs d'une activité de lecture plus fébrile, labile instable. Au milieu de phénomènes de reconfiguration des processus associatifs, analogiques que l'acte de penser implique. Tout mode d'écriture et tout dispositif qui favorisent

75 <http://gabriel.gallezot.free.fr/Solaris/d02/2polanco1.html>

76 E. Garfield, *Citation indexing*, New York, Wiley, 1979. J.-P. Courtial, M. Callon, F. Laville, « Co-words analysis as a tool for describing the networks of interaction between basic and technological researches : the case of polymer chemistry », *Scientometrics*, 22, 1, 1991. H. Small, « Co-citation in the scientific literature : a new measure of the relationship between two documents », *Journal of the American Society for Information Science*, 24, 1973.

77 Brian Kleiner, Isabelle Renschler, Boris WERNLI, Peter Farago & Dominique Joye, *Pour une réflexion sur les infrastructures de recherches en sciences sociales*, Zurich, 2013.

78 Aolab, Université de Paris 7, Laboratoire Crics.

pour suivre Douglas Hofstadter l'établissement de connexion par la bande, d'expériences de pensée etc. sans rien devoir à la causalité ou à des schèmes convenus qui rendent plus manifeste pour chaque énoncé, chaque proposition formulée, chaque narration, le fait d'être entouré d'un amas, d'un halo de variantes elles-mêmes qui peuvent être suggérées par le glissement (la traduction) d'un quelconque des innombrables traits qui les caractérisent⁷⁹. Ces dispositifs se rapprochent du souhait énoncé par M. Foucault.

Mais d'une manière plus générale il s'agit de pas laisser l'invention dans les dispositifs de l'édition scientifique dans les mains des grands ensembles molaires de l'édition y compris les infrastructures Open. L'articulation de ces archives ouvertes dites de laboratoire avec les autres modes éditoriaux se faisant sous des conditions de filtrages spécifiques afin de maintenir la circulation vers le dehors de la recherche, les va-et-vient documentaires, tout en préservant toutefois la possibilité toujours renégociable par le collectif d'avoir un espace clos de travail, de réserve documentaire, réserve de données en devenir, où l'on peut parler, écrire, expérimenter « barbare » voire « hérétique ». Liberté inconditionnée pour l'exercice de la recherche, espace qui est en quelque sorte le lieu de la mixité entre modes de recherche et antichambre des modes d'exposition de la recherche.

Cela est d'autant plus important que la production (collective aux échelles prés, y compris dans le cas des individus toujours déjà pris dans leur milieu « cognitif productif », des œuvres et des travaux, repose sur des rapports variables entre transparence et opacité, pratiques de la rétention calculée, entre procédures de construction de la preuve, de la testabilité et puissance et statut des communautés interprétatives.

Les possibilités de traiter des données non-structurées ou semi structurées (petit ou gros corpus) de manière émergente (bottom up) à partir de traitements statistiques relativement simples, mais aussi à présent à partir d'algorithmes issus du Deep learning et des travaux de l'IA, afin de faire apparaître des structures, des formes, permettent d'envisager à terme une prolifération des formes sémiotiques associées, des blocs sémantiques comme graphes pour des orientations dans les espaces

79 Hofstadter D. *Ma Thémagie*, Paris, Interéditions, 1988 et Gödel, Escher, Bach – Les brins d'une guirlande éternelle, 2008.

des savoirs. Ce que les approches industrielles et celles des systèmes d'information d'entreprise ont peu à peu mis en place pourrait être source d'inspiration pour les plateformes de l'édition numérique en vue de renforcer l'arsenal hybride pour la « transversalité »

Peut-être serait-il donc intéressant de réfléchir à des manières de faire co-habiter divers types de solutions, de technologies, de normes, de fonctions lors de la construction des infrastructures dans le domaine des sciences en général et des SHS en particulier. On peut imaginer que selon le niveau d'échelle auquel on considère les pratiques et les processus de différenciation des champs de recherche, les fonctionnalités spécifiques qui y sont privilégiées ainsi que leurs domaines d'interventions, les modes d'écriture / lecture activées, il n'est pas nécessaire de mettre en place des configurations de normes, de metadata par trop intrusives et coercitives. De même est-il possible et même souhaitable de créer des milieux d'intelligence troués, sorte d'archipels où sont utilisés (et sont développés) selon les singularités des recherches, leur hétérogénéité, des technologies de datamining et de mapping orientées « bottom up » et qui prennent en charge les données non-structurées ou semi-structurées.

Cela est d'autant plus important que l'on pourrait être tenté à partir les mémoires hypertextuelles en réseaux des sciences, des textes SHS et pourquoi pas à partir du monde des formes courtes des réseaux (Twitter, Facebook) une « explicitation » des « communautés interprétatives » qui enveloppent et traversent les mondes scientifiques, les mondes des SHS en particulier⁸⁰. Cette explicitation consistant à une cartographie (toujours en droit comme en fait, évolutive) des agencements interprétatifs constitutifs de ces communautés.

Les devenirs des infrastructures de l'édition numérique scientifiques sont encore loin de ces développements et les intelligences collectives, de milieux relativement sophistiquées. Cela devrait engager encore plus fortement le mouvement de l'Open Archive à la pointe de ses immenses responsabilités politico-cognitives afin d'occuper le cœur partout présent

80 Stanley Fish, *Your P's and B's : The Digital Humanities and Interpretation*, January 23, 2012 Stanley Fish, *Quand lire c'est faire. L'autorité des communautés interprétatives*, trad. de l'anglais (américain) par Étienne Dobenesque, préface d'Yves Citton, postface inédite de Stanley Fish, Paris, 2007 ; et Stanley Fish, <https://opinionator.blogs.nytimes.com/2011/12/26/the-old-order-changeth/>

des milieux de l'intelligence afin d'offrir de manière inconditionnée, des puissances inédites pour le libre exercice des facultés. Marche vers le commencement de l'infini ?

Jean-Max NOYER
Professeur Émérite des Universités
Université de Toulon
Consultant