

Les concepts information et document à l'épreuve du fait numérique : approche épistémologique

Délivré par École Nationale de Formation Agronomique de Toulouse

Diplôme : Master 2 « Enseignement et Formation pour l'Enseignement Agricole et le Développement Rural »

Option : Formation aux métiers de la Recherche et de l'Expertise en éducation (RE)

Présenté et soutenu par Isabelle Couturier
Le 30 juin 2015

Directrice de mémoire

Mme Cécile Gardiès
Maître de conférences (HDR) en
Sciences de l'information et de la
communication

Membres du jury

Mme Viviane Couzinet
Professeure en Sciences de l'information et de la
communication

Mme Isabelle Fabre
Maître de conférence en Sciences de l'information
et de la communication

Remerciements

Merci à Cécile Gardiès, pour sa bienveillance, son enthousiasme et son exigence, trois combustibles idéaux pour susciter et soutenir une motivation ; qu'elle m'a insufflée dès mon entrée dans le métier de professeure-documentaliste, renouvelée par l'entremise de l'aventure passionnante « gapedocienne », et ravivée par l'invitation à la fréquentation stimulante de la communauté scientifique SIC.

Merci à Viviane Couzinet et à Isabelle Fabre qui ont accepté de participer au jury, pour les précieux conseils qu'elles m'ont donnés, et pour leur encouragement à poursuivre sur ma lancée.

Merci à Cécile S., Sylvie S. et Laurent E., qui m'ont précédée dans l'aventure et qui m'ont accompagnée et soutenue lors de conversations très animées (merci à Cécile pour son intérêt constant nonobstant le combat personnel qu'elle mène courageusement, merci à Sylvie, « bébé chercheuse » en train de devenir grande et merci à Laurent pour le prêt prolongé de la bible des Sciences de l'information-documentation !).

Merci aux professionnel-les de l'information de la BU de l'Université du Maine, de la Médiathèque Louis Aragon du Mans, du CDI de l'ENFA (merci à Chantal Grenier-Gire pour sa disponibilité et sa diligence) et de la BU de Laval, sans l'aide desquels ce travail n'aurait pu être mené à bien.

Merci à mes ami-es, à ma belle-mère, à mes frères, mes sœurs et mes parents qui se sont intéressé-es à mon sujet et m'ont accordé un soutien sans faille.

À Fabrice et Abel qui ont supporté, avec une patience infinie, mes découragements et... mes emballements !

Sommaire

Introduction.....	5
Chapitre 1 Le concept information à l'ère numérique : la tentation pléonastique.....	9
1.1 Les relations ambiguës entre information et numérique.....	9
1.1.1 Qu'est-ce que "le numérique" ?.....	9
1.1.2 "Le numérique" : de la technique au « concept ».....	12
1.2 La théorie mathématique de l'information : un pavé dans la mare des SIC.....	16
1.2.1 La théorie mathématique de l'information : l'information comme donnée mesurable.....	16
1.2.2 La théorie mathématique de l'information : influences contemporaines en SIC....	23
1.3 L'information-connaissance en SIC : priorité à la question du sens.....	26
1.3.1 Les théoriciens fondateurs des sciences de l'information.....	26
1.3.2 Une information humaine, sociale, vivante avant d'être numérique : les héritiers des fondateurs des sciences de l'information.....	30
Chapitre 2 Le fait numérique : un perturbateur conceptuel pour le document en SIC ? Où la théorie mathématique de l'information resurgit.....	36
2.1 Les bases conceptuelles du document en SIC.....	36
2.1.1 Les précurseurs : la notion d'extensivité et de preuve.....	36
2.1.2 Le lien indissoluble entre le document et l'information.....	38
2.2 Les caractéristiques controversées du concept de document à l'heure numérique.....	46
2.2.1 Les travaux du RTP-DOC : le document numérique : un nouveau concept ?.....	46
2.2.2 De l'objet stable à l'entité disparate.....	49
2.2.3 De l'unité documentaire à l'unité informationnelle.....	59
2.3 Le document numérique, révélateur de caractéristiques originelles du concept document.....	71
2.3.1 Le document, objet informationnel avant d'être numérique.....	71
2.3.2 La double partition.....	74
2.3.3 La valeur sociale du document.....	78

Chapitre 3 Analyse épistémologique des concepts en SIC : éléments d'interprétation pour distinguer le savoir scientifique du « fait numérique »	81
3.1 Point sur l'analyse épistémologique.....	81
3.1.1 L'épistémologie : étude critique des sciences.....	81
3.1.2 L'épistémologie des sciences de l'information.....	82
3.2 Grille de lecture épistémologique.....	83
3.2.1 Méthodologie de construction de la grille.....	83
3.2.2 Présentation de la grille d'analyse épistémologique.....	85
3.3 Analyse des limites et de la validité de la grille épistémologique.....	100
3.3.1 Limites de la grille.....	100
3.3.2 Validité et perspectives.....	101
Conclusion	106
Liste bibliographique.....	110
Table des matières.....	116
Index lexical.....	118
Index des auteurs.....	119
Répertoire des sigles.....	120
Sommaire des annexes.....	121

Introduction

L'enseignement de l'information-documentation dans l'enseignement secondaire, dans la pratique, n'est pas chose aisée dans un contexte qui devrait paradoxalement aider les professeurs-documentalistes. En effet, le discours, notamment médiatique, met l'accent sur l'information. L'expression « société de l'information »¹ s'est largement répandue avec le développement de l'internet et l'accès généralisé à l'information dans nos sociétés contemporaines. La maîtrise de celle-ci est devenue un enjeu majeur au niveau international² et national. Le législateur s'est emparé de la question et, dans le domaine de l'éducation, la *Loi sur la refondation de l'école de la République* du 8 juillet 2013 met en avant l'éducation aux médias et à l'information (EMI) et crée le service public du numérique éducatif. Cette mission relève de l'intérêt général car « *nos sociétés sont profondément transformées par le numérique. La société de l'information ouvre des perspectives nouvelles en matière d'accès à la connaissance et à la formation*³ ». Désormais quand on parle d'information, il s'agit de fait essentiellement d'information numérique. En effet, l'information à laquelle sont confrontés les jeunes en situation d'apprentissage scolaire est numérique (celle qui n'est pas numérique à l'origine est numérisée ou en voie de numérisation). Ceux que Marc Prensky avait qualifiés, de façon controversée, de « *natifs du numérique* » sembleraient *a priori* constituer un public particulièrement réceptif à la formation à la culture informationnelle.

Or, il est encore plus difficile de légitimer l'intervention des professionnels de l'information à l'ère numérique car la position de médiateur, qui était souvent incontournable dans une bibliothèque pour accéder aux documents primaires, tend à s'effacer sur la toile. L'un des changements très importants accompli par les réseaux interconnectés est l'accès direct (sans passer par les documents secondaires) à l'information.

L'information-documentation est une discipline inégalement implantée dans l'Enseignement secondaire en France. La documentation est inscrite dans les référentiels de

1 On verra plus loin pourquoi cette expression ne fait pas l'unanimité dans le champ des SIC

2 Déclaration de Prague en 2003 : « *vers une société compétente dans l'usage de l'information* » ; Proclamation d'Alexandrie en 2005 : « *permettre aux gens sur tous les chemins de la vie de chercher, d'évaluer, d'utiliser, de créer l'information pour des objectifs personnels, sociaux, professionnels et éducationnels* »

3 [LOI n° 2013-595 du 8 juillet 2013 d'orientation et de programmation pour la refondation de l'école de la République](#)

l'Enseignement agricole pour les formations Bac professionnel agricole et Brevet de technicien supérieur agricole et des pratiques régulières avec les enseignants d'autres disciplines en situation de pluridisciplinarité sont relativement anciennes à l'Éducation nationale comme dans l'enseignement agricole. Les apports des professeur-es-documentalistes ont évolué, comme l'ont montré différents auteurs (Gardiès, 2011), d'un aspect méthodologique pur (apprentissage du fonctionnement des logiciels de recherche documentaire, connaissance de l'organisation du centre de documentation et d'information) vers une approche plus disciplinaire (développement d'une culture informationnelle) en étant passés par un positionnement transversal (apprendre à apprendre).

L'enseignement n'allait pas de soi pour la discipline scolaire documentation dans la mesure où la majorité des professeur-es documentalistes, s'ils-elles ont bien reçu une formation pédagogique, ils-elles n'ont, pour la plupart, pas de formation universitaire dans ce domaine spécifique. Cela est lié à l'histoire de la discipline universitaire par elle-même, très récente⁴, au faible nombre de formations (licence, master) dans les Sciences humaines et sociales qui proposaient la mention Information-Communication et aux épreuves du concours non centrées sur les Sciences de l'information et de la communication (SIC) jusqu'à la session de 2001. Les candidats aux concours viennent d'options variées tels le français, l'histoire ou encore les langues. Cette position particulière par rapport aux autres collègues de disciplines plus reconnues était, depuis notre entrée dans le métier, une préoccupation constante et elle a pris un tournant nouveau au fil des travaux engagés par le GAP⁵ documentation auquel nous appartenons depuis sa création par l'ENFA⁶. Ce groupe de réflexion a été mis en place suite à une refonte des programmes de l'Enseignement agricole qui a été à l'origine d'un changement profond dans la manière d'aborder l'information-documentation avec un ancrage scientifique délibéré. Dans un premier temps, il s'agissait pour les membres du GAP, professeur-es documentalistes, guidé-es par des chercheur-es en SIC, en Sciences de l'éducation et par l'Inspection, de faire un travail d'explicitation des nouveaux textes réglementaires à destination des pairs. Les notions identifiées pour l'enseignement de la documentation ont été décryptées en lien avec les sciences de l'information. Très tôt, au sein du GAP et lors des

4 Création du CAPESA en 1990 et du PLPA en 1991

5 Groupe d'animation et de professionnalisation : plusieurs GAP existent au sein de l'enseignement agricole. Ils sont régis par la note de service DGER/MSSI/N2013-2039 du 20 mars 2013

http://www.formco.agriculture.gouv.fr/fileadmin/user_upload/plus_info/Documents_de_references/DGERN2013_2039.pdf

6 École nationale de formation agronomique

formations organisées en région auprès des professeur-es documentalistes, a émergé la question du manque de maîtrise des savoirs scientifiques et donc la nécessité de réaliser un travail approfondi autour des concepts fondamentaux des Sciences de l'information et de la communication. C'est ainsi qu'en observant les caractéristiques relevées par les théoriciens fondateurs de l'information, nous nous sommes rendue compte non seulement que le concept document était riche de dimensions, insoupçonnées jusqu'alors, mais aussi que ces dernières semblaient difficiles à accorder avec les pratiques informationnelles en ligne. Partant du constat évoqué plus haut que malgré les discours ambiants valorisant l'information, celle-ci était de plus en plus insaisissable, difficile à maîtriser, nous nous sommes demandée si le savoir scientifique sur le document pouvait permettre d'éclairer d'une façon différente les questions soulevées par « le numérique » au sujet de l'information.

Les concepts fondamentaux des sciences de l'information permettent-ils de comprendre les raisons du hiatus entre les membres de ladite société de l'information et les professionnels de l'information ? D'un point de vue scientifique, « le numérique » reconfigure-t-il les caractéristiques des notions clés que sont l'information et le document ? L'évolution technologique impose-t-elle donc de repenser les concepts ? Ou bien au contraire, les apports scientifiques sur ces concepts permettent-ils de baliser la réflexion et de stabiliser le savoir dans un contexte technologique évolutif ? Une approche épistémologique peut-elle éclairer la question de l'information à l'heure numérique ? Une analyse des discours des chercheurs en sciences de l'information permet-elle d'appréhender l'influence « du numérique » sur les concepts ? Les concepts tels qu'ils avaient été définis restent-ils valides, offrent-ils une résistance au fait numérique ? Afin de traiter ces questions, nous nous sommes rendue compte progressivement qu'une approche épistémologique des deux concepts information et document devenait incontournable.

Dans un premier temps, après avoir cerné ce que recouvre le terme « numérique », il s'agira d'analyser, le concept « information », à partir d'un éclairage scientifique issu des SIC. Nous ferons un détour nécessaire par la théorie mathématique de l'information qui, nous l'avons constaté très rapidement dans nos recherches, joue un rôle essentiel – d'un point de vue épistémologique - dans la façon d'aborder les concepts information et document à l'heure numérique. Nous verrons que de nos jours certains chercheurs en SIC sont très influencés par

cette théorie mathématique puis nous en viendrons au concept tel qu'il a été défini par les premiers théoriciens de l'information dans le champ pluridisciplinaire des Sciences de l'information et de la communication. Nous terminerons par le regard des héritiers contemporains de la vision « meyratiennne » de l'information.

La question de la validité des concepts scientifiques s'est déjà posée et se pose encore à diverses équipes de recherche. À l'âge numérique, plus que le concept « information », c'est le concept « document » qui est sujet à controverse. Nous ferons donc, dans un deuxième temps, le point sur les bases conceptuelles du document qui ont été posées par les théoriciens en Sciences de l'information, puis nous analyserons les nouvelles théories élaborées à ce sujet, et nous ferons ressortir les éléments mouvants pour en venir aux dimensions permanentes qui émergent mettant en exergue la notion d'information avant tout.

À la lumière des éléments théoriques relevés sur les deux concepts majeurs des sciences de l'information, nous avons essayé de dégager les traits caractéristiques permettant de positionner les discours, dans l'environnement numérique, dans une veine tantôt mécaniste, tantôt cognitiviste. Nous tenterons donc, dans un troisième temps, après un rappel sur ce qu'est l'épistémologie et une explicitation de ce choix méthodologique, de proposer un modèle d'analyse des concepts fondamentaux à l'aide de repères épistémologiques. Nous analyserons pour finir les limites et la validité de ce modèle afin d'envisager les utilisations éventuelles qui pourraient en être faites.

Chapitre 1

Le concept information à l'ère numérique : la tentation pléonastique

Le rapport entre information et numérique apparaît d'emblée lorsque l'on se penche sur la signification de ce terme. Cette relation entre les deux items cependant dénote comme nous allons le voir maintenant une acception particulière de la notion d'information.

1.1 Les relations ambiguës entre information et numérique

1.1.1 Qu'est-ce que "le numérique" ?

Ce terme employé très couramment n'est pas simple à définir précisément et sert dans les discours de synonyme à des mots aussi variés que « virtuel », « multimédia », « Internet » ou « Net », « réseau », « toile », « web », « électronique », « hypermédia », « cyber », « digital », "informatique", "technologies de l'information et de la communication"...

Étymologiquement, "numérique" se rapporte au "nombre" : « *adjectif [qui] est dérivé savamment (1616) du latin numerus avec le suffixe -ique. Le mot signifie qui a rapport au nombre, qui appartient aux nombres est essentiellement employé en mathématiques (par exemple dans extraction numérique, 1756 ; tableaux numériques, 1868) et plus tard en informatique. En sont dérivés deux termes informatiques récents, numériser, v tr (V. 1970) et numérisation n. f. (1974), francisations de digitaliser et digitalisation* ». (Rey, 2010).

L'acception actuelle du terme se rapporte à la notion d'information dans un sens particulier : « *numérique : se dit de la représentation de données d'information ou de grandeurs physiques au moyen de caractères, généralement une suite de chiffres, ce qui permet de les traiter ou de les stocker par des procédés informatiques ; système, dispositif employant ce mode de représentation (opposé à analogique). Le numérique : nom masculin, l'ensemble des techniques de télécommunications utilisant des signaux numériques* » (Robert, 2006). Les "données d'information" renvoient ici à la théorie mathématique de l'information⁷, celle-ci étant prise comme une donnée quantifiable.

Selon S. Cacaly, un « *signal numérique est un signal électronique qui varie de façon discontinue en fonction du temps et ne peut prendre qu'un nombre fini de valeurs. Les circuits électroniques des équipements informatiques fonctionnant exclusivement avec une logique à*

7 Voir partie 1.2.1

deux états, dite binaire, le signal utilisé ne peut prendre que deux valeurs qui correspondent à une alternative, tout ou rien. On appelle bit (abréviation de binary digit) les impulsions élémentaires correspondant à l'existence ou à l'absence d'un courant électrique. ». L'adjectif est pris dans le sens technique et il n'y a pas d'entrée pour « document numérique » ou « information numérique » dans ce dictionnaire mais une entrée pour « document électronique » (Cacaly, 2004 : 167).

Pour A. Boulogne, numérique « qualifie un dispositif, système ou procédé délivrant un signal qui décrit les phénomènes étudiés de façon discontinue par rapport au temps aux moyens de codes binaires (0 et 1) (Boulogne, 2004 : 176). L'information est déterminée (...) pour une présentation numérique par un ensemble de bits. » (Boulogne, 2004 : 128). On peut constater à nouveau que ce terme comporte une acception de l'information qui, comme nous le verrons ultérieurement, s'oppose à celle qui a été théorisée dans le champ des Sciences de l'information et de la communication par ses fondateurs (Meyriat, Escarpit).

Ce lien direct avec l'information -toujours au sens de donnée quantifiable- est détaillé par Y-F. Le Coadic qui précise que « la numérisation (ou digitalization en anglais) de l'information consiste à transformer les valeurs qui décrivent l'intensité des sons et des couleurs, perceptibles en mode analogique, en mesures qui sont codées en suites de 0 et de 1 pour être reconnues, traitées et restituées, au final sous forme analogique – seule forme reconnaissable par les appareils visuels et auditifs humains-, par les machines numériques. La numérisation de l'information présente des avantages importants. En effet les données numériques sont plus faciles à transformer, à transporter, à diffuser et à éditer. Elles sont disponibles sur tout support : papier, cédérom, CD-Audio, etc. La numérisation permet des traitements plus rapides, ce qui permet des traitements de données plus volumineuses. La numérisation permet surtout dans certains cas de réaliser une compression de données, (la compression numérique) élargissant de nouvelles perspectives en termes de volume d'informations transmissible ». (in Lamizet, 1997 : 414). Cette fois encore, l'information, ici, est synonyme de « donnée » et le sens n'entre pas en ligne de compte dans l'explication du procédé dont l'auteur vante les mérites d'un point de vue purement quantitatif.

Dans la littérature scientifique, c'est d'abord le terme "électronique" qui est privilégié pour étudier cette nouvelle technique. Ainsi H. Fondin pose la question du "traitement documentaire de l'écrit à l'heure du tout-électronique" utilisant cette expression comme

synonyme de "numérique" (Fondin, 1998 : 7). Il précise que "*l'enjeu documentaire est désormais dans le document électronique*", enjeu fondamental pour les « *informatistes*⁸ » qui doivent « *regarder favorablement les nouveaux moyens de communication des savoirs, et afficher face à la société leurs compétences documentaires, celles sur la documentation, celles sur le document, celles sur l'information contenue dans le document* ». (Fondin, 1998 : 9-10).

Le terme numérique s'est imposé progressivement et a maintenant presque exclusivement remplacé le terme "électronique". Puis un autre glissement sémantique d'importance s'est opéré : de l'adjectif (qui qualifie donc à la base une technique), on est passé au substantif, élargissant le terme vers une conception plus "culturelle" que technique.

8 H. Fondin reprend le terme proposé par Meyriat en précisant qu'il désigne les personnels spécialistes du traitement de l'information documentaire et culturelle ; archivistes, bibliothécaires et documentalistes

1.1.2 "Le numérique" : de la technique au « concept »

A. Moatti dénonce cet usage impropre en constatant que «*la création récente par les pouvoirs publics français d'un Conseil national du numérique consacre quasi officiellement un mésusage du mot « numérique ». Ce mot est un adjectif, mais il est le plus souvent utilisé comme substantif : le*⁹ numérique, presque avec un N majuscule*» (Moatti, 2012 : 133). L'auteur explique ainsi que cet emploi dénote une vision quasi idéologique du support : «*On voit néanmoins apparaître là un premier élément d'interprétation possible. C'est cette convergence, dans des champs d'usage très différents (l'informatique, les télécommunications, la télévision, la photographie ...), vers le support numérique qui exalte l'importance de ce dernier. Le numérique, substantivé, presque personnifié, envahit tout. Le concept est à ce point invasif qu'il en arrive à faire oublier ce à quoi il se rapporte (image photographique, enregistrement musical, traitement de l'information ...) : tout converge dans le* numérique. L'adjectif numérique* abolit le substantif qu'il est censé qualifier : raccourci symbole d'un monde pressé qui en vient à oublier le substantif, voire à le refouler, pour désigner plus rapidement le fait du* numérique, quel qu'en soit le support. Un monde pressé par le numérique lui-même et la célérité de propagation et de vie qu'il induit.* » (Moatti, 2012 : 134)

En qualifiant le numérique de "concept", cet ingénieur, actif dans les domaines de la diffusion de la culture scientifique et des bibliothèques numériques, va plus loin dans l'analyse sémantique. Un concept, représentation mentale générale et abstraite d'un objet, est ce qui rend compte du réel. Cela suppose des théoriciens, un domaine de validité, or, il n'y a pas de théoriciens du "numérique" car il s'agit d'un moyen technique. Cependant en arrière-plan du terme, ce n'est pas le procédé technique mais bien une vision du réel qui émerge.

Ainsi B. Bachimont estime que le numérique a profondément changé notre rapport au monde. «*Le numérique est un nouveau support d'inscription qui a révolutionné nos sociétés en permettant d'aborder et de traiter tous les types de contenus. La massification et la spatialisation des contenus ouvrent de nouvelles possibilités analytiques et de nouveaux objets de recherche et de connaissance. Pour discerner plusieurs registres de possibilités, il faut distinguer deux niveaux dans notre rapport au numérique : le computationnel (l'information abstraite et son calcul) et le numérique proprement dit (la diversité de ses réalisations effectives). Comment penser cette tension entre l'universalité du computationnel et la contingence de ses réalisations ?* » (Bachimont, 2012). «*L'information abstraite et son*

9 Les termes et expressions signalés par un astérisque sont en italique dans le texte original

calcul » renvoie ici encore à la donnée numérisée mais au lieu d'évoquer seulement la représentation numérique de l'information, c'est la relation à la connaissance qui est modifiée, selon le Docteur en informatique, par la migration de tous les contenus en une seule forme par la technologie numérique et les nouvelles modalités d'usage rendues possibles.

C. Genin essaie quant à lui de définir ce que l'on peut qualifier de "culture numérique" en dépassant les querelles entre défenseurs invétérés « du numérique » et opposants radicaux (tel Alain Finkielkraut) : *« pour d'aucuns parler de culture numérique est un oxymore. Comment articuler culture et nombre ? Comment le nombre peut-il faire culture, être culture ? L'expression de l'esprit serait-elle quantifiable, réductible à un calcul, qui plus est binaire ? Après tout, ce ne serait pas si idiot. Pythagore, selon lequel « tout est nombre », avait montré que les mathématiques ordonnaient l'univers des dieux et des hommes. Boole parvint à exprimer les opérations de l'esprit logique par une algèbre. Aujourd'hui on appelle ainsi « numérique » ce procédé qui consiste à convertir du qualitatif en quantitatif, à traduire des quales en quanta, par opposition à la traduction dite analogique »* (Genin, 2004). Il ne s'agit pas là de transformer les messages en données mathématiques pour pouvoir les transporter via un canal mais bien de changer la valeur même des messages au cours de l'opération. Selon ce philosophe, *« la culture numérique n'est donc pas du côté de l'objet (hardware), indubitablement un bien de consommation (ware). Elle est dans l'esprit et l'usage qui font fonctionner cet automate. Nous devons donc parler du numérique comme un mode de penser¹⁰ qui avance une thèse sur le monde : tout ce qui existe est susceptible de relever d'un programme, lui-même susceptible d'être exprimé par un calcul binaire, lui-même susceptible de permettre toutes sortes de simulations et d'actions prédéterminées »* (Genin, 2004). Il s'agit bien de dépasser l'aspect technique de l'outil numérique jusqu'à le considérer comme un « mode de penser ». Le philosophe fait sienne, en citant Wiener, l'approche mathématique de l'information définie comme *« une suite continue ou discontinue d'événements mesurables, distribués dans le temps »* (Genin, 2004). Revenant sur les critiques sociologiques et philosophiques (notamment celle de Heidegger, 1968) de la cybernétique, en tant que danger pour la liberté ontologique, il montre que les craintes face au « numérique » procèdent du même ressort. *« La critique métaphysique s'imagine un monde composé d'irresponsables, «jouets» inconscients d'un projet planétaire qui les dépasserait »* (Genin, 2004). S'inscrivant en faux contre cette approche négative, l'auteur insiste sur les bouleversements induits par « le

10 sic

numérique » : « *Par cette dissociation de l'information et du support naturel habituel, la numérisation est analogue à la chimie capable de produire des arômes artificiels. En quelque sorte le virtuel serait une information artificielle, une information de synthèse* » (Genin, 2004). Cette approche de la numérisation lui confère des vertus quasi magiques en considérant que telle l'alchimie qui transforme le métal en or précieux, le codage des données induit une métamorphose de l'information naturelle en information synthétique. Cette mutation serait due à la séparation de l'information et du support autrement dit la séparation du fond de la forme.

Les discours autour « du numérique » pris comme concept renvoient donc clairement à l'acception mécaniste de l'information¹¹ et non au « *contenu cognitif d'une communication réalisée ou possible* » (Meyriat, 1983 : 66).

Ce procédé technique a certes des conséquences notables en terme d'inscription de l'information et de diffusion, mais pour les tenants de « *l'information humaine* » ou « *vivante* », qui s'intéressent avant tout au sens, « le numérique » tue l'information car « au moment où une culture informationnelle commence enfin à imprégner le public scolaire et même le grand public, le focus qu'elle pose sur le « numérique » en vient de fait à occulter la question même de l'information* » (Baltz, 2013 : 6). Le paradoxe évoqué dans la problématique tient donc en partie à cet « *impérialisme numérique* » qui « *nous force à reconsidérer l'information là où elle se manifeste dans toute sa « nature* » » (Baltz, 2013 : 9).

Cerner ce que recouvre le terme "numérique" revient à évoquer la notion de l'information : de la technique de codage de l'information, on est passé au "concept" sans qu'il y ait pour autant de véritables bases théoriques. Ainsi l'information numérique dans certains esprits est un pléonasme puisque « le numérique », c'est l'information. Y. Jeanneret, à propos de ce qui est communément appelé « *la société de l'information* », pointe le raccourci qui est fait entre l'essor des technologies de la communication et la circulation des savoirs. Il déplore le manque d'une « *initiation raisonnée et synthétique aux grands questionnements à partir desquels le rapport entre dispositifs techniques et socialisation des savoirs peut être formulé* ». (Jeanneret, 2000 : 12). Cette critique peut s'appliquer de la même façon à ce que d'aucuns appelle « l'âge numérique » et qui pourrait ressembler à une « *société du numérique* ».

11 Voir partie 1.2.1

Dorénavant l'information est numérique ou en voie de numérisation et, dans le cadre scolaire, c'est à celle-ci que sont confrontés les sujets apprenants. Si une approche typologique facilite la plupart du temps la compréhension d'un concept ou constitue à tout le moins un moyen de l'aborder, dans le cas de l'information numérique (ou électronique) et du document numérique (ou électronique), il semblerait que cela soit plus complexe. Parler d'information numérique semble entraîner la confusion au lieu de simplement caractériser l'information à l'instar de l'information écrite, sonore, audiovisuelle...

Quand on évoque l'information écrite, audiovisuelle ou encore sonore, on fait référence au mode d'inscription de l'information mais plus que le procédé technique, c'est la forme obtenue qui est définie. Ainsi on ne « lit » pas de la même façon une image, un texte, un contenu audiovisuel. En revanche, mentionner l'information numérique – cette expression est souvent remplacée par le vocable « numérique » ou au contraire le raccourci « information » revient dans le discours à parler de l'information numérique – englobe toute forme finale de restitution du contenu. Cela tient à la particularité de cette technique qui revient à fusionner les différents types d'information en un unique procédé. L'information écrite peut se numériser (est numérisée), de même pour l'iconique, l'audio-visuelle, etc. Qu'elle soit écrite, image, son, l'information peut être numérique et l'adjectif numérique accolée au concept « information » ne permettrait donc en rien de caractériser celui-ci et il faudrait en revenir aux catégories créées avant l'émergence de ce vocable aux contours imprécis, d'un point de vue scientifique, à savoir : « image », « son », « audiovisuel », pour qualifier l'influence de la forme de l'information sur le fond.

L'information numérique peut-elle être entendue comme un type d'information comme un autre ne remettant pas en cause le concept « information » tel qu'il a été défini par les théoriciens fondateurs en Sciences de l'information et de la communication ? « Le numérique », comme on vient de le voir est un terme vague qui, qualifiant à la base un procédé technique, électronique pour être précis, s'est étendu à une acception conceptuelle menant pour certains à une véritable acculturation reposant sur un élément clé : « l'information ». Aborder l'information et le document numériques à travers l'analyse épistémologique des deux concepts « information » et « document » peut-elle permettre de lever ces ambiguïtés sémantiques ?

1.2 La théorie mathématique de l'information : un pavé dans la mare des SIC

1.2.1 La théorie mathématique de l'information : l'information comme donnée mesurable

Les Sciences de l'information et de la communication (SIC), en France, sont une discipline assez récente (création de la 52^e section du Comité consultatif des universités en 1975¹²) et il n'est donc pas surprenant qu'un de ses objets d'étude principal, l'information, soit influencé par une théorie née au début des années 1950. Claude E. Shannon, comme le montre E. Dion, en publiant sa *Théorie mathématique de la communication* (autrement appelée *Théorie mathématique de l'information*) a provoqué pendant plusieurs décennies des controverses passionnées autour des outils et concepts développés dans son ouvrage (Dion, 1997 : 34-51). Les résultats de ces recherches étaient pourtant destinés au départ à un public restreint (les ingénieurs de la société de télécommunications américaine Bell) et n'avaient pas de prétention autre que celle de proposer une transmission, la plus efficace et la plus rapide possible, de messages codés sous forme de signaux.

Ainsi, Shannon précise que le « *problème fondamental de la communication est de reproduire en un point, soit exactement, soit approximativement, un message recueilli en un autre point. Fréquemment, les messages ont une signification*, c'est-à-dire qu'ils se réfèrent ou sont corrélés à un autre système par certaines entités physiques ou conceptuelles. Ces aspects sémantiques de la communication sont sans rapport avec les problèmes techniques* ». (Shannon, 1975 : 65-66). Ce sont bien ceux-ci que le professeur du MIT¹³ propose de résoudre en raisonnant de la façon suivante : « *L'aspect important est que le message actuel soit choisi dans un ensemble* de messages possibles. Le système doit fonctionner pour chaque sélection possible et non seulement pour celle qui sera choisie, étant donné qu'on ignore cette sélection au départ* ». (Shannon, 1975, 66). C'est pourquoi, il va se servir de la mesure logarithmique d'une part, et s'appuyer d'autre part, sur les découvertes en linguistique pour élaborer sa théorie. Le calcul des probabilités qu'une lettre suive une autre, la connaissance de l'occurrence des lettres dans une langue donnée (l'anglais notamment) permettent de déterminer l'ensemble des choix possibles afin que le choix optimal se fasse en fonction du canal utilisé. « *La redondance de l'anglais ordinaire, si l'on exclut les structures statistiques*

12 Devenue la 71^e section du Conseil supérieur des universités en 1985

13 Massachusetts Institute of Technology

supérieures à huit lettres environ, est de 50 %. Ce qui signifie que, dans l'anglais écrit, la moitié de ce qui s'écrit est déterminée par la structure du langage, et l'autre moitié est choisie librement. » (Shannon, 1975 : 96).

Cette approche mathématique permet d'utiliser la construction de la langue pour réduire les incertitudes liées au fait que l'on ne peut connaître à l'avance le message qui va être émis et ce, toujours indépendamment de ce que le message contient cognitivement. Shannon invente une méthode permettant de déterminer les diverses probabilités. Il calcule ainsi l'« entropie » de la source d'information, qui correspond à la « forme H » : les « quantités de la forme H (...) jouent un rôle central dans la théorie de l'information » (Shannon, 1975 : 87). Il définit ainsi le calcul de l'entropie « comme mesure de l'information, du choix et de l'incertitude » (Shannon, 1975 : 87) ou encore « comme une mesure du hasard » (Shannon, 1975 : 142). Ses démonstrations montrent bien que les impératifs liés à la technique (les problèmes matériels) s'éloignent de la réalité objective puisque lorsqu'il met au point la formule calculant la capacité d'un canal, il remarque que « la nature se paie en exigeant précisément beaucoup d'incertitude, de sorte qu'il est impossible d'obtenir réellement un débit supérieur à C » (Shannon, 1975 : 116), la « capacité C d'un canal avec bruit [étant] le débit maximal possible de transmission, c'est-à-dire le débit quand la source est correctement adaptée au canal » (Shannon, 1975 : 115).

À aucun moment, dans ses articles, Shannon n'aborde la question de la signification des messages. S'il évoque la fidélité, l'évaluation, c'est au sujet de l'intégrité du message reçu par rapport à celui qui a été émis. Il se sert des formules et théories déterminées en statistique linguistique pour calculer au mieux les choix successifs possibles de construction de « séquences » qui vont permettre de les transmettre le plus rapidement et le plus efficacement en fonction des canaux utilisés. Ainsi, selon lui, « si une source peut produire seulement un message particulier, son entropie est nulle, et un canal n'est pas nécessaire. Par exemple, une machine à calculer faite pour calculer les digits successifs de π produit une séquence donnée sans élément aléatoire. Aucun canal n'est nécessaire pour « transmettre » cela à un autre point. On pourrait construire une seconde machine pour calculer, à ce point, la même séquence » (Shannon, 1975 : 104-105). Dans ce passage, la source (d'information) est prise

dans une acception très restrictive en fonction des paramètres de transmission. En développant cette approche nouvelle de la communication, Shannon a eu besoin de nommer le phénomène observé pour mieux le mesurer. Ce faisant, il a « créé », selon Abraham Moles, le mot « information » car le « *chercheur scientifique* » est un « *créateur linguistique. Tant que les mots restent dans le cadre même qui leur a donné naissance, il n'y a là aucun inconvénient : c'est un processus nécessaire, celui du langage scientifique comme parallèle au langage ordinaire ; les mathématiciens, entre autres, ne se privent pas de prélever des mots du langage le plus simple (filtre, anneau, catastrophe, coupure, etc.) et de leur donner un sens extrêmement particulier qui provoque la perplexité du public profane* » (Shannon, 1975 : 18). Shannon aurait pu utiliser un autre terme du langage courant et il avait d'ailleurs employé le terme « communication » auquel on a substitué celui d'« information » par la suite. L'expression « *quantité d'information* » est en fait la vulgarisation de la formule mathématique exacte « *logarithme du maximum de vraisemblance d'une distribution multinomiale* » (Dion, 1997 : 36).

Ce qui explique un tel retentissement de cette théorie, d'après Emmanuel Dion, c'est qu'elle « *est la seule à avoir proposé une approche véritablement quantifiée du phénomène de communication* » (Dion, 1997 : 15). De ce fait, elle ne s'intéresse qu'aux « *aspects formels de la communication en négligeant tout ce qui [peut] concerner le contenu des messages. Elle lui a donc interdit de s'appliquer à tous les cas où les problèmes de sens, de finesse, d'allusion, de valeur, de vérité, d'exclusivité, d'histoire, d'objectif, etc., sont en jeu* » (Dion, 1997 : 15-16). Ces aspects d'ordre cognitif sont précisément ceux qui sont retenus par les chercheurs en SIC et qui, comme nous le verrons plus loin, rendent les deux approches inconciliables pour les théoriciens fondateurs.

La théorie mathématique permet en revanche d'apporter des « *solutions pratiques à des problèmes matériels et concrets, comme l'amélioration de la quantité de communications qu'il est possible de véhiculer (...)* » (Dion, 1997 : 16). C'est à ce titre qu'elle a fasciné et fascine encore de nombreux chercheurs dans des disciplines aussi variées que l'informatique, la statistique, la biologie, la psychologie, etc. Le fait de disposer de formules (algorithmiques) représentant le processus de communication arrive à point nommé pour renforcer la

légitimation des sciences de l'information. Il est d'ailleurs notable de remarquer qu'alors qu'E. Dion -dans une volonté de réhabiliter l'importance de la théorie mathématique de l'information- ne cesse de préciser au long de son exposé que celle-ci est circonscrite (en premier lieu par son auteur) à des opérations concrètes particulières et n'a pas vocation à être extrapolée dans d'autres contextes, il se fait parallèlement le relais de réflexions de biologistes, psychologues, linguistes, et même d'artistes qui y ont vu (en vain souvent) les moyens de résoudre des problèmes conceptuels (Dion, 1997 : 129-143).

Il rapproche de façon plus convaincante cette théorie de celle des jeux (Dion, 1997 : 142). Pour la théorie des jeux, « l'information » est « parfaite » lorsque chaque joueur connaît et se souvient de tous les coups joués par tous et « incomplète » quand au moins un des éléments fait défaut. Il souligne que les outils de la théorie mathématique sont applicables dans un univers limité dont il est possible de réduire l'incertitude en procédant par choix successifs entre deux alternatives : cette théorie « *ne pouvant en effet en toute rigueur être appliquée que dans des univers probabilisés, elle pose indirectement le problème de la distinction entre ce qui est probabilisable et ce qui ne l'est pas* » (Dion, 1997 : 45). Du point de vue mathématique, une situation de double choix représente une unité d'information mesurée en bit sur la base du logarithme binaire. En langage informatique, les deux états sont codés 0 et 1. Ainsi selon ce principe, deux bits correspond à quatre états (2×2), trois bits à huit états différents ($2 \times 2 \times 2$), etc. Plus les choix possibles sont nombreux, plus il y a d'informations (au sens mathématique).

Cependant, il existe, selon Escarpit, cofondateur du Comité des sciences de l'information et de la communication¹⁴, une seule sorte de « *machine* » qui est capable, lorsqu'on lui laisse le choix entre deux propositions, d'opter pour « *la troisième* » : c'est le cerveau humain (Escarpit, 1976 : 42). Dans son essai paru un an après la publication de la traduction française de la théorie mathématique, celui-ci analyse la portée de la vision de Shannon dans le champ des SIC, avant l'usage généralisé des ordinateurs et bien avant le développement de l'internet. Il propose une lecture historique dans une perspective marxiste de la théorie mathématique en montrant que ces travaux marquent l'« *aboutissement* » des

14 Qui deviendra en 1974 la Société française des Sciences de l'information et de la communication (SFSIC)

recherches sur le rendement dans la communication (Escarpit, 1976 : 7). S'inspirant du principe de « l'entropie », phénomène de la thermodynamique, qui représente la perte de l'énergie thermique (la chaleur) non transformée en énergie mécanique (le travail), les théoriciens de l'information ont mis au point un modèle de quantification de l'information similaire à celui de l'énergie. Plus l'entropie est élevée, plus les éléments du système sont désordonnés et plus la part de l'énergie inutilisable pour l'obtention d'un travail est grande.

Appliquée au domaine du transport d'information, cette formule sert à obtenir le contre processus de l'entropie, la « *nég-entropie* » : l'information est donc un signe négatif qui vient éliminer un phénomène lui-même négatif. Escarpit précise qu'il est « *important, pour bien comprendre l'attitude intellectuelle de la théorie de l'information, de ne pas assimiler l'information à des acquisitions comme le savoir ou la connaissance, mais de lui maintenir son double signe négatif* » (Escarpit, 1976 : 21). Selon cette vision, « *on peut imaginer une source « froide », chargée d'entropie qui fait passer du « froid » (c'est à dire de l'indétermination) sur une source « chaude » et produit au passage de l'information, le processus étant irréversible. La mesure de l'information potentielle contenue dans la source froide est son imprévisibilité statistique* » (Escarpit, 1976 : 21). Il montre bien les limites de ce qu'il appelle le « *modèle mécaniste* » qui en cherchant à prévenir l'imprévisible s'oppose à l'esprit humain dont la liberté s'exprime par l'imprévisibilité de ses choix (Escarpit, 1976 : 71). La réflexion d'Escarpit s'attarde surtout sur la dimension temporelle en reprenant la notion d'entropie mais dans une perspective diachronique : « *l'information est quantité négative apportée à l'esprit par l'événement quand il passe du futur au passé. Le savoir est une construction assez stable pour freiner l'écoulement du temps et rester disponible à l'échelle d'une vie humaine, assez mouvante pour accueillir sans cesse de nouvelles réponses et les insérer dans des structures lisibles* » (Escarpit, 1976 : 59-60). Escarpit pointe les « *problèmes de compatibilité** » entre « *les exigences du canal* » et « *les exigences de la pensée humaine* » dont la caractéristique fondamentale est de « *pouvoir produire à tout moment de nouveaux énoncés* » (Escarpit, 1976 : 43).

Weaver¹⁵ explique que si la mesure de l'entropie est de 0, alors il n'y a pas « information » car pas d'incertitude, pas de liberté de choix. Si l'incertitude provient de la liberté de l'émetteur (choix multiples), il y a plus d'information et c'est une « bonne » information, si l'incertitude

15 Warren Weaver : mathématicien américain, collaborateur de Shannon, à l'origine de la publication à destination du grand public de la *Théorie mathématique de l'information*

provient d'erreurs ou de bruit dans le canal, l'information est plus grande mais c'est une « *mauvaise* » information. Cette double connotation se rapporte à un point de vue quantitatif (ce qui peut se mesurer pour améliorer l'efficacité de transmission du message) et non d'ordre qualitatif (Shannon, 1975 : 50-51).

Si ce calcul était transposable dans le domaine de l'information-connaissance, pour répondre à une recherche d'information, il suffirait d'appliquer les formules shannoniennes et d'éliminer progressivement de façon dichotomique les réponses jusqu'à parvenir à l'unique information pertinente – ce qui correspond dans la formule à une entropie nulle. Or, dans ce cas l'information pertinente n'est *pas* une information au sens shannonien, puisque l'entropie étant nulle, il n'y a pas d'information car c'est ce qui est imprévisible qui fait l'information. En d'autres termes, plus il y a de choix, plus l'incertitude augmente, plus l'information est grande (au sens « donnée mesurable »).

Les procédés mathématiques permettent la résolution de problèmes et recherchent l'efficacité et la rapidité mais, comme le reconnaît (en le déplorant) E. Dion, « *il y a (...) un grand achoppement pour l'esprit à éliminer, dans toutes les formes de communication, la redondance, l'évidence, l'automatisme, le mécanisme* »(Dion, 1997 : 99). Les contraintes qu'avaient à régler les ingénieurs, afin que le maximum de signaux puissent être transportés dans un canal tout en conservant l'intégrité du message, ne sont donc pas les mêmes que celles d'individus qui souhaitent que le contenu du message soit bien compris par leurs interlocuteurs. E. Dion, en 1997, regrettait que la théorie mathématique ne soit pas plus connue dans une société où le terme « information » était omniprésent. Au moment où, comme nous l'avons vu en partie 1.1, c'est le vocable « numérique » qui prend le dessus, cette théorie qui s'appuie sur le codage binaire surgit, de façon consciente ou non, dans les réflexions actuelles sur ce qu'Y. Jeanneret appelle « *les médias informatisés* » (Jeanneret, 2000 : 65).

Et, comme nous le verrons dans la partie suivante, des chercheurs en SIC abordent l'information sous cet angle. Dès 1949, dans sa note introductive à la théorie mathématique, c'est Weaver lui-même qui suggère, et autorise donc, toutes les extrapolations. Bien que précisant que « *le mot information*, dans cette théorie, est utilisé dans un sens spécial, à ne*

pas confondre avec son langage courant. En particulier, information ne doit pas être confondu avec signification » (Shannon, 1975 : 37), il contredit cette précaution sémantique en démontrant quelques lignes plus loin qu'il est « *indésirable* » de séparer les « *trois niveaux des problèmes de communication* » (Shannon, 1975 : 57-58). Ces trois niveaux sont le *problème technique*, niveau d'exactitude de la transmission des symboles de la communication ; le *problème sémantique*, niveau de précision de la signification véhiculée lors de la transmission ; et le *problème de l'efficacité*, niveau d'efficacité avec laquelle la signification reçue dans la communication influence la conduite dans le sens désiré (Shannon, 1975 : 32). Weaver pousse le raisonnement au bout et appelle de ses vœux « *une théorie réelle de la signification* » (Shannon, 1975 : 60).

En ces temps de connexion permanente, où l'internet met « *hors-circuit la diachronie*¹⁶ », cet espoir est comblé par les recherches sur le web sémantique qui reviennent à compléter, comme le proposait Weaver, le schéma de Shannon¹⁷ en introduisant la boîte « *bruit sémantique* » entre la source d'information et l'émetteur (le « bruit » du schéma d'origine devenant le « *bruit technique* ») et à insérer une boîte « *récepteur sémantique* » entre le récepteur (technique) et le destinataire (Shannon, 1975 : 59).

1.2.2 La théorie mathématique de l'information : influences contemporaines en SIC

La confusion – la fusion – information-numérique peut expliquer l'approche quantitative actuelle de certains courants des SIC. P-D. Pomart, ancien président de l'ADBS¹⁸ fait ressortir « *l'aspect entièrement volatil* » que peut prendre l'information dont, précise-t-il, la forme peut être séparée du sens du fait du traitement électronique (Cacaly, 2004 : 122). Une innovation technologique serait ainsi à l'origine de la transformation de la nature même de l'information. S'inscrivant dans la droite ligne de la conception mathématique, il définit l'information comme « *un facteur de réduction de l'incertitude. La nouveauté d'un message (et donc la puissance de son caractère informatif) résulte de l'originalité de l'assemblage et de la combinaison des unités élémentaires d'information qui, elles, sont connues* » (Cacaly, 2004 : 122). Le procédé technique permet de réaliser ces appariements entre les différents fragments d'informations disponibles, connus.

16 Cf. CRARY, Jonathan. 24/7. *Le capitalisme à l'assaut du sommeil*. Zones. 2014

17 Voir schéma en annexe 1

18 Association des professionnels de l'information et de la documentation

L'individualité n'est pas prise en compte car « *au contraire de la connaissance et, a fortiori, de la compétence, l'information est extérieure à un individu ou à un groupe d'individus, elle vit une existence autonome et de ce fait peut être acquise facilement sur un marché où se rencontrent l'offre et la demande, dès lors qu'existent les outils de guidage vers cette offre* » (Cacaly, 2004 : 122). On comprend ici que l'on peut considérer l'information comme une donnée brute et lui attribuer une valeur chiffrée. Or, il y a amalgame entre la propriété de l'information, qui est immatérielle, dans le sens où on ne peut en déposséder l'individu qui la transmet, et le caractère « immatériel » de la circulation réticulaire qui n'est pas liée à un support matériel (comme pouvait l'être le livre). Pourtant ce n'est pas de la disparition du support dont il s'agit mais de la migration des différents supports vers un seul mode d'inscription : le support numérique. L'information était immatérielle avant « le numérique », elle l'est avec « le numérique » et le restera après. Mais c'est devenu un enjeu économique et de pouvoir accru, du fait de la généralisation de l'électronique et de l'évolution du capitalisme, qui a organisé la mutation de la société industrielle, productrice de biens matériels vers la « *société d'information* » autrement qualifiée de « *société de haute productivité* »¹⁹ qui « produit » des biens immatériels (les services) reposant sur le bien immatériel par excellence : l'information.

Y-F. Le Coadic précise que la théorie mathématique est « *relative à la mesure et à la transmission de signaux électriques* » et que ce concept n'a rien à voir « *avec la cognition et la communication humaine* » (Le Coadic, 2004 : 5-6). En SIC, selon lui, l'information est une « *connaissance inscrite (enregistrée) sous forme écrite (imprimée ou numérisée), orale ou audiovisuelle sur un support spatio-temporel* » (Le Coadic, 2004 : 6). Le but de l'information, c'est la connaissance. Bien que dénonçant les dérives liées à l'utilisation de la théorie mathématique dans le champ des Sciences de l'information et de la communication, il reprend à son compte les « *équations fondamentales de la science de la communication* », telle celle de Brookes :

$$C + \Delta C = C'$$



où l'on passe d'un état de connaissance C à un nouvel état de connaissance C' par l'apport

¹⁹ Cf. Rapport de Simon Nora et Alain Minc. *L'informatisation de la société*. 1978

d'une connaissance ΔC extraite d'une information (ΔI), ΔC étant l'effet de la modification (Le Coadic, 2004 : 11). Il s'agit bien là d'une approche quantitative de l'information. L'information venant combler un « *état anormal de connaissance* » (Le Coadic, 2004 : 11). Le professeur en sciences de l'information au CNAM²⁰ détaille ainsi une série de mesures des activités de construction, de communication et d'usage de l'information (infométrie) et revient sur les lois bibliométriques et sur la modélisation shannonienne des processus de communication. À propos de la numérisation, il est prudent face à l'assertion suivante : « *toute information peut désormais être traitée, codée, stockée, transmise par les mêmes procédures et les mêmes réseaux* » (Le Coadic : 2004 : 92). Cette « *idée erronée de l'information* » dans le domaine informatique « *évacue* » selon lui « *toute idée de sens, de signification* » revenant à dire qu'est « *information tout ce qui peut faire l'objet d'un traitement numérique* » (Le Coadic, 2004 : 93). Nous constatons à nouveau²¹ que l'expression « information numérique » serait alors pléonastique. Il y a confusion entre le traitement technique de l'information et la nature de cette même information. Y-F. Le Coadic met en garde contre les abus sémantiques : « *virtuel, cyber, hyper, numérique mais avant tout électronique ! Cessons de qualifier ces dispositifs informationnels et Internet de “virtuels” et autre “cyber” pour les considérer sous leur angle bien réel, technique, électronique et informatique* » (Le Coadic, 2004 : 101). Avec cette évolution technique, on est passé, constate-t-il « *d'une emphase documentaire à une emphase informationnelle, d'une orientation système à une orientation usager* », ce qui suppose « *un nouveau cadre conceptuel* ». (Le Coadic, 2004 : 112). Il semble ardu pour ces auteurs précités de s'éloigner de la conception mathématique de l'information au moment où les principes de Shannon sont plus que jamais en corrélation avec les moyens de diffusion actuels de l'information.

Nous l'avons vu, la théorie mathématique précède les recherches fondamentales en sciences de l'information, quoique des précurseurs comme Otlet, fondateur de l'Institut international de la documentation, dans les années 1930, ou encore Briet, première directrice des études de l'INTD²², dans les années 1950, aient toujours mis l'accent, à propos de la documentation, sur le travail intellectuel nécessaire. Pourtant des chercheurs contemporains en SIC convoquent cette théorie qui semble pertinente au regard des évolutions techniques.

20 Conservatoire national des arts et métiers

21 Voir partie 1.1

22 Institut national des techniques de la documentation

Les théoriciens fondateurs des sciences de l'information en France se sont néanmoins immédiatement différenciés de la théorie mathématique. Nous allons aborder le concept information du point de vue de ce que Meyriat, cofondateur et premier président de la SFSIC, avait proposé de nommer « *l'informatologie* » afin d'étudier en quoi l'environnement numérique requestionne le concept information tel que l'ont défini les théoriciens fondateurs de l'information-documentation. (Estivals, 1993 : 331)

1.3 L'information-connaissance en SIC : priorité à la question du sens

1.3.1 Les théoriciens fondateurs des sciences de l'information

Selon Meyriat, l'information, « *ressource fondamentale pour nos sociétés contemporaines (...), joue un rôle aussi décisif dans l'ordre de l'esprit que l'énergie dans l'ordre de la matière* » (Meyriat, 1983 : 65). L'analogie entre les deux termes « information » et « énergie » est poussée très loin par Shannon qui va utiliser un phénomène thermodynamique, l'entropie, pour l'appliquer à l'information. Mais pour Meyriat, la comparaison s'arrête là car il place la question du sens, du « *contenu cognitif* » au centre de ses écrits sur le concept. Pour lui, comme pour Shannon, le « *concept d'information n'est pas isolable de celui de communication* » mais au contraire de ce dernier, ce n'est pas l'aspect formel de la communication qui intéresse Meyriat mais son contenu ; ses dimensions psychologique, sociologique, politique, économique, juridique, etc. (Meyriat, 1983 : 65).

Le terme information, comme nous l'avons souligné précédemment, est équivoque puisqu'il s'agit dans l'approche mécaniste d'une donnée quantifiable dans le processus de communication. Meyriat s'inscrit en faux contre cette acception car « *ne peut être information pour nous que ce qui a un sens. Notre information est donc tout à fait différente de ce que les informaticiens, les généticiens ou les mathématiciens, pour ne parler que d'eux, désignent par le même terme* » (Meyriat, 1983 : 67). D'emblée, Meyriat prend ses distances par rapport à la théorie mathématique, respectant d'ailleurs la pensée de son fondateur, pour qui la question du sens n'entraîne pas en ligne de compte. L'information en SIC est « *le contenu cognitif d'une communication réalisée ou possible* » (Meyriat, 1983 : 66). Est information ce qui peut être communiqué : dire qu'un individu possède une information mais qu'il ne la communique pas, c'est affirmer qu'elle est précisément communicable. La question de la prévisibilité n'est pas en cause dans cette acception, c'est la relation entre individus dans la communication qui fait l'information. C'est la forme donnée par l'esprit au contenu de la communication, le sens reconnu car il « *n'y a de communication réalisée (réussie) que si tous les acteurs donnent le même sens à ce qui est communiqué* » (Meyriat, 1983 : 66-67). Le fait d'attribuer un sens, c'est faire « *acte de connaissance* » (Meyriat, 1983 : 67). Tant que cet acte n'est pas réalisé, il n'y a pas d'information car celle-ci n'existe pas en tant que telle et c'est pourquoi elle ne peut être mesurée. « *C'est par le sens attribué qu'une information se distingue d'une donnée ou d'une série de données* ». Le sens « *n'est pas une propriété intrinsèque de l'objet connu* » (Meyriat,

1983 : 67). Si quelqu'un prend connaissance d'un événement, cela apporte du sens pour lui et il s'agit donc d'une information mais tant qu'elle n'est pas activée dans un processus de communication, Meyriat la qualifie « *d'information latente* » (Meyriat, 1983 : 67). La Science de l'information a pour objet d'étudier ce que « *devient le sens attribué au contenu de la communication dans le processus même de la communication* » (Meyriat, 1983 : 69).

Meyriat analyse ce qu'est « *l'attribution de sens* » et précise que l'écart peut être grand entre le sens apparent d'un message et celui qu'il recouvre réellement. L'information, qui résulte d'une attribution de sens et, est de ce fait un acte de connaissance, signifie une « *modification (par ajout ou transformation) de l'état de connaissance de celui qui la reçoit* » (Meyriat, 1985 in Couzinet, 2001 : 202-203). Quand ce qui est visé par l'émetteur est la modification du comportement du destinataire, c'est l'information de type commerciale qui est en cause. En effet, selon Meyriat, cette attitude est celle qui « *est à l'origine de la démarche publicitaire* ». « *L'enjeu* » étant de « *faire accueillir et accepter un sens donné à l'avance* » (Meyriat, 1983 : 72). Jusqu'à vouloir faire admettre, dans le cas de la propagande, un sens erroné et aboutir à ce que Meyriat nomme la « *contre-information* ». (Meyriat, 1983 : 72). Dans tous les cas, il se produit une activité mentale à chaque extrémité du processus de communication : du côté de l'émetteur (« *volonté de celui qui l'émet* ») et du récepteur (modification du « *savoir implicite ou explicite* » pour « *l'esprit qui la reçoit* ») (Meyriat, 1985 in Couzinet, 2001 : 203). De plus, il précise que « *toute connaissance dépend de la culture de celui qui connaît* » (Meyriat, 1983 : 71). Il insiste sur l'importance de l'environnement social, culturel, économique des individus qui exerce nécessairement une influence sur leur capacité à attribuer un sens et implique une interprétation. C'est la différence fondamentale avec « l'information » des informaticiens pour qui « *l'information n'a pas de sens ; elle est neutre, c'est seulement un support, un objet mental qui se prête à des opérations et à des calculs indépendants de toute valeur sémantique* », suivant en cela la « *notion désincarnée de l'information* » théorisée par les mathématiciens (Meyriat, 1985 in Couzinet, 2001 : 204-205).

Dans l'environnement numérique, cette approche fondée sur l'attribution de sens est d'autant plus difficile à conserver si l'on s'attarde sur les moyens techniques qui « accélèrent »

la diffusion et « effacent » les distances. Meyriat constate que « *les demandes nouvelles font en effet appel à des mémoires non-écrites (audiovisuelles, électroniques, optiques...)*. Souvent aussi elles prennent de vitesse toute mémoire et sollicitent une réponse immédiate, sans délais. D'où l'importance prise par les outils qui ont rassemblé d'avance de très nombreux éléments de réponse à toutes sortes de questions possibles » (Meyriat, 1991, in Couzinet, 2001 : 285). Il pose également la question du moyen de transport choisi et de l'influence de ce dernier sur le sens attribué par rapport à l'information qu'il transmet. Ainsi il se demande si « *la transmission elle-même* » est « *neutre* », comment elle intervient « *sur le contenu transmis* » et si « *la nature des procédés techniques utilisés (mécaniques, électriques, électroniques...)* » est une « *variable secondaire ou principale au regard du sens, donc de l'information* » (Meyriat, 1983 : 70). En posant ces questions, il fait le lien, comme nous le verrons dans la partie 2 avec le document, en tant qu'« *objet auquel a été confiée la fonction de transmettre de l'information* » (Meyriat, 1983 : 73). Au sujet de la transmission, il reconnaît que les professionnels de l'information sont « *dépendants des techniques de la communication** » (Meyriat, 1985 in Couzinet, 2001 : 212). Au milieu de la décennie 1980, il note que le développement des techniques « *offre des possibilités jusqu'ici insoupçonnées à une meilleure diffusion de l'information* » (Meyriat, 1985 in Couzinet, 2001 : 212). Il emploie à nouveau le terme « *dépendance* » à l'égard des outils, soulignant par-là l'importance de leurs rapports avec l'information mais mettant l'accent, une fois encore, sur leurs capacités ou non à « *véhiculer du sens* » (Meyriat, 1985 in Couzinet, 2001 : 212). Ainsi, pour lui, il ne peut y avoir de « *producteurs d'informations* », mais des « *producteurs de connaissances* » et des « *transmetteurs d'informations* ». C'est la relation qui s'établit entre les individus qui importe et l'activité mentale du sujet sur lui-même. (Meyriat, 1985 in Couzinet, 2001 : 215).

Escarpit met également l'accent sur le lien entre communication et information : « *nous disons que la communication est un acte et que l'information est son produit* » (Escarpit, 1976 : 100). D'un point de vue étymologique, informer, donner une forme, correspond à l'acception shannonienne. Ainsi « *un objet « informé » serait un objet entièrement connu et prévisible, qui aurait perdu toute son entropie* » (Escarpit, 1976 : 109). Cependant dans sa théorie diachronique de l'information, Escarpit fait figurer au rang de « *machines* » particulières dans le processus de communication, les « *M6* » ou « *machines à langages (animal hominisé)* » qui sont seules capables d'abstraction. Ainsi ces « *machines* »

dans le processus de communication peuvent non seulement « *informer* » mais encore « *s'informer* », c'est à dire qu'elles « *peuvent influencer par feedback les sources de stimuli qui les atteignent* » et « *reconnaître les limites d'influence qui définissent leur identité et leur champ de conscience* » (Escarpit, 1976 : 110). Dans ce processus itératif, les êtres humains sont « *producteurs* » d'information (au sens mathématique) et « *recréent l'entropie au fur et à mesure qu'[ils] l'épuisent* », en contradiction avec l'approche shannonienne (Escarpit, 1976 : 110). Cette approche « *colorée de marxisme* » (Escarpit, 1976 : 3) liée au sujet, aux groupes et sous-groupes et aux rapports qu'ils entretiennent les uns avec les autres est fondamentale et met en avant la notion de contexte dans le processus de communication : « *comme l'individu, le groupe a besoin de se constituer en identité stable, de s'objectiver en institutions, mais il ne peut le faire que par une stratégie de communication qui lui permette à la fois de dominer les identités des individus ou des groupes qui la composent et de résister à la domination que les groupes dans lesquels il est inclus tentent d'exercer sur son identité. Le processus de communication est ici étroitement lié au fonctionnement de l'appareil économique, social et politique* » (Escarpit, 1976 : 113).

1.3.2 Une information humaine, sociale, vivante avant d'être numérique : les héritiers des fondateurs des sciences de l'information

Y. Jeanneret, pour lever les ambiguïtés du terme, propose de nommer *information₁*, la donnée mesurable et *information₂*, « *l'information sociale* » qui correspond à la « *connaissance communiquée ou communicable* » de Meyriat (Jeanneret, 2000 : 42). Rejetant l'expression « société de l'information », qui véhicule une idéologie particulière et dans laquelle on serait entré - « *comme si une société pouvait entrer en elle-même* », note-t-il avec ironie (Jeanneret, 2000 : 11) -, il fait le constat que les dites « technologies de l'information et de la communication » font la part belle à *l'information₁* en occultant *l'information₂*. L'information que l'on nomme désormais « numérique » est à rapprocher de la troisième appellation qu'il propose : *l'information₃* (*l'information₁₊₂*) correspond à la « *dimension d'information₂ d'une « technologie de l'information » comme celle que nous connaissons* » (Jeanneret, 2000 : 51). Il reconnaît cependant que ce qu'il nomme, « *l'écrit d'écran* » (Jeanneret, 2000 : 75) pose des questions supplémentaires mais il interroge toujours *l'information₃* à travers la question de la connaissance, du partage des savoirs et de l'opération mentale qui s'opère dans l'esprit du sujet : « *savons-nous contrôler ce que nous confions à la machine et dont nous nous dépossédons (ce que nous acceptons de ne plus pouvoir interpréter, afin d'accéder à d'autres messages) ?* » (Jeanneret, 2000 : 51). Il propose de distinguer les trois notions information, connaissance et savoir qui sont interdépendantes mais qui ne sont pas équivalentes. Ainsi l'information, c'est « *la relation entre le document et le regard porté sur lui* ». La connaissance, c'est « *le travail productif des sujets sur eux-mêmes pour s'approprier des idées ou des méthodes* ». Le savoir, ce sont « *les formes de connaissances qui sont reconnues par une société* ». (Jeanneret, 2000 : 85).

B. Juanals souligne, quant à elle, la situation paradoxale de cette société dite de l'information dont l'« *un des enjeux majeurs* » est « *la question de l'utilisation et de la maîtrise de l'information* » (Juanals, 2003 : 15). Elle pose la question de savoir si les « *mutations dans les accès à l'information observées à l'heure actuelle* », c'est-à-dire au moment de l'explosion de l'information numérique connectée, vont « *réellement dans le sens d'une démocratisation des connaissances* » (Juanals, 2003 : 14). Elle note aussi le flou sémantique qui entoure la notion d'information qui recouvre selon elle trois aspects distincts :

le sens lié au développement de la presse sous la III^e République, « *information que l'on porte à la connaissance d'un public* », le sens contemporain « *ensemble de connaissances* », « *action d'informer un public* » et le sens singulier issu de la théorie mathématique et lié à la transmission de signaux (Juanals, 2003 : 15). L'unique terme information désigne ces trois sens à la fois. Mais pour elle, la question de la signification, du contenu des messages est essentielle. Ainsi elle ne souscrit pas au discours ambiant selon lequel « *la technique est créditée du pouvoir considérable d'accroître le potentiel intellectuel de ses utilisateurs* » (Juanals, 2003 : 17). Elle note que certains auteurs au sujet de l'information literacy (Bawden et Robinson) accordent une attention particulière au sens : la culture de l'information doit « *englober les habiletés techniques mais ne pas s'y limiter, comme à une technologie ou à un ensemble de technologies particulières ; la compréhension, la signification et le contexte doivent occuper une position centrale* » (Juanals, 2003 : 20). Elle affirme également au sujet des dispositifs de communication que « *sens et savoirs sont toujours à construire par le lecteur en fonction des finalités qui lui sont propres. Ils ne peuvent donc y être présents à un état préexistant* » (Juanals, 2003 : 30).

Les dispositifs informatisés d'accès à l'information permettent une « *mise en relation entre apprenants potentiels et données à transformer en connaissances* » (Juanals, 2003 : 30). C'est le passage des unes aux autres qui constitue l'information, car ce passage s'effectue par l'attribution de sens... Elle propose une relation entre les termes information, connaissance et savoir, différente de celle proposée par Y. Jeanneret que nous avons reprise plus haut : ainsi selon elle, les « *informations deviennent un savoir lorsque le sujet fait intervenir la notion de sens, ce qui suppose l'analyse et l'appropriation de l'information dans son système de connaissances* » (Juanals, 2003 : 134). Mais le point commun dans leur approche respective est la notion de sens. Elle précise plus loin : « *le sujet recherche des données qui deviennent pour lui des informations dans la mesure où elles font sens par rapport à son projet et à ses finalités* » (Juanals, 2003 : 134). L'espoir illusoire que la facilitation d'accès à des contenus (qui ne sont que des données numérisées) est synonyme d'acquisition accrue de connaissances génère déception et frustration dans la mesure où « *cette réduction du processus éducatif à une question d'accès aux données, fondamentalement différente de l'accès au savoir, implique une représentation utopique d'un sujet sans intériorité, sans histoire, sans contexte socioculturel, sans spécificités psychologiques et affectives, sans désir...* » (Juanals, 2003 :

135). Le terme « données » renvoie à la théorie mathématique (l'auteure a précisé que l'information recouvrait désormais les trois acceptions) mais elle se rapproche de Meyriat en précisant que les données par elles-mêmes ne suffisent pas pour être qualifiées d'informations, car « *l'informatique manipule des données abstraites qui ne sont pas encore de l'information* » (Meyriat, 1986 in Couzinet, 2001 : 259). Les informaticiens se servent du terme information pour désigner ces « *données, éléments neutres manipulables et calculables, mais dénués par eux-mêmes de sens, et qui doivent en recevoir un pour devenir porteurs d'information* » (Meyriat, 1986 in Couzinet, 2001 : 251). L'internet ne contient que des données et leur « *processus de transformation en informations et en connaissances relève d'une opération mentale à construire par le sujet* » (Juanals, 2003 : 135). L'illusion d'accès direct à l'information est très largement entretenue alors que le développement des applications et outils numériques génère un profit croissant. B. Juanals conclut en notant que « *la recherche d'informations et la construction de connaissances ont toujours constitué des activités complexes, dont la difficulté s'est encore accrue avec le recours à la recherche documentaire informatisée* » (Juanals, 2003 : 171) car « *la technologie interactive ne peut remplacer en aucune manière l'activité intellectuelle du lecteur* » . (Juanals, 2003 : 175).

Cette assertion est corroborée par B. Simonnot qui précise que des études montrent que « *l'usage de l'internet et des moteurs commerciaux n'améliore pas les compétences documentaires des internautes, notamment celles des étudiants* » (Simonnot, 2012 : 153). La stratégie des moteurs de recherche commerciaux (objets de son étude) consiste à maintenir l'utilisateur dans une « *forme d'émerveillement* » (Simonnot, 2012 : 150). C'est pour cette raison qu'elle milite pour une connaissance approfondie du fonctionnement des moteurs de recherche commerciaux (souvent les seuls dispositifs d'accès à l'information en ligne utilisés) afin de « *lutter contre la fascination que suscitent ces dispositifs* » (Simonnot, 2012 : 213). Cet état de fait ayant été fabriqué de toute part, suivant l'analyse d'Y. Jeanneret : « *l'entre-réseau (...) n'est devenu l'internet (l'objet imaginaire fascinant et l'espace d'expérimentation de certains groupes sociaux) que parce qu'il a été montré et raconté par les médias qui le précèdent* ». (Jeanneret²³, 2000 : 56). B. Simonnot aspire également à une vision plus distancée car le critère économique ne saurait être le critère « *principal qui arbitre les*

23 Yves Jeanneret constate que cette vision a la vie dure puisque plus de dix ans après l'avoir critiquée, il persiste et signe dans la réédition de son ouvrage (Jeanneret, 2011)

possibilités et les conditions d'accès à l'information dans nos sociétés » (Simonnot, 2012 : 218). Revenant sur les discours superlatifs au sujet du volume illimité de données disponibles sur l'internet, elle rappelle que « *la capacité de lecture des individus, elle, reste limitée* » (Simonnot, 2012 : 49). Or l'information nécessite un « *travail d'interprétation et de production de sens* » (Simonnot, 2012 : 25) : « (...) *un programme informatique fournit seulement des données ou des résultats de calculs. Il peut aussi permettre d'accéder à des textes, des images ou autres. C'est l'attention qui y est portée qui en font ou non des informations effectives* » (Simonnot, 2012 : 27). Elle relaie une des « *hypothèses fondamentales* » de « *l'approche sense making* » de B. Dervin qui considère que « *l'information n'est pas une chose qui existe de manière indépendante ou externe aux individus mais un produit de l'observation humaine* » (Simonnot, 2012 : 90). Ceci en lien avec l'environnement socio-culturel de l'intéressé.

Ainsi B. Dervin, dans sa théorie du « *sense-making* » fait ressortir la difficulté d'émettre une question qui corresponde réellement au besoin d'information. À travers un travail exploratoire, elle recense plusieurs facteurs qui expliquent pourquoi la question de départ d'usagers ne fournit pas au bibliothécaire la bonne indication correspondant à leur besoin réel. (« *There are many reasons why patrons' initial questions do not provide the librarian with good guidance to their real needs* » - Dervin, 1986 : 6). Elle a donc mis au point des techniques d'interview permettant de guider par des questions neutres les requérants afin de mieux cerner leurs besoins. C'est l'utilisateur qui, seul, a la clé de la pièce à l'intérieur de laquelle sa requête réelle est cachée. (« *the user's real query as hidden inside a room to which the user has the only key* » - Dervin, 1986 : 6). En fait, il s'agit de démêler en quelque sorte les méandres de la pensée car il s'agit bien de l'élaboration des connaissances qui résulte d'un procédé cognitif complexe. Cependant en prenant garde aux trois aspects suivant : la prise en compte de la *situation* de l'utilisateur (son vécu personnel, ses qualifications au moment de la recherche qui sont à l'origine du manque), - la lacune ou le *fossé*, d'un point de vue cognitif, ce qui manque pour que l'utilisateur puisse réaliser son projet et qui sera comblée ou dépassé grâce au pont construit par l'information reçue et - les *usages* de l'information prévus (ce que l'utilisateur espère faire une fois franchi le pont) (« *situation, gaps, and expected uses* »), il est possible d'atteindre progressivement l'information pertinente (Dervin, 1986 : 7).

Meyriat distingue donc clairement l'information-donnée qui appartient aux informaticiens de l'« *information-connaissance* » qui intéresse les sciences de l'information. Il précise toutefois à propos des données que « *ces éléments discrets ne sont pas fournis spontanément par les réalités ou phénomènes que l'on étudie, ils en sont extraits par une opération de l'esprit par une observation attentive, par un appareil de mesure imaginé à cet effet. Ce sont en réalité des « construits ». Mais ils n'ont pas de sens en eux-mêmes (...). Ils peuvent recevoir un sens, mais c'est l'esprit de l'opérateur qui le leur attribue : c'est à ce moment qu'ils supportent une information* » (Meyriat, 1985 in Couzinet, 2001 : 205). S'opposant de même à la conception mathématique de l'information, B. Simonnot montre bien que l'information ne procède pas automatiquement à une « *diminution de l'incertitude* » mais peut au contraire faire émerger des doutes ou des questionnements là où il n'y en avait pas auparavant (Simonnot, 2012 : 159). Elle explique également que « *l'accessibilité est une condition nécessaire mais non suffisante pour l'appropriation des dispositifs* ». Et cet obstacle est irréductible puisqu'il est d'ordre cognitif propre à l'individualité de chaque sujet. (Simonnot, 2012 : 163).

La confusion sémantique au sujet de l'information en ligne est telle que, selon C. Baltz, ce terme en vient à être remplacé par le vocable « numérique ». Pourtant l'information est certes « *la matière première du numérique* » (Baltz, 2013 : 4) mais « *l'information numérisée vient d'« ailleurs »** », elle n'est pas « *donnée telle quelle sous la main* » (Baltz, 2013 : 6-7). Le fait que tout converge vers « le numérique » donne « *l'impression que l'essentiel du système est en place et n'aurait plus qu'à fonctionner* » (Baltz, 2013 : 9). Ainsi il constate que « *l'impérialisme du numérique nous force à reconsidérer « l'information » là où elle se manifeste dans toute sa « nature »* » (Baltz, 2013 : 9). Il est nécessaire de redonner sa place à « *l'information humaine* », « *vivante* » dans les formats numériques. (Baltz, 2013 : 10).

Les arguments avancés pour justifier ce changement de la nature même de l'information liée à la technologie numérique sont principalement de deux ordres : celui du temps d'abord (la circulation accélérée de l'information réticulaire donnant l'illusion d'un présent permanent) et celui de l'espace ensuite (l'effacement des distances conduisant à une prétendue ubiquité). Nous l'avons vu précédemment, l'approche mathématique de

l'information a été clairement rejetée par les théoriciens fondateurs des SIC ; le nœud central de leur réflexion tournant autour du sens, de la signification des messages. Cette préoccupation cognitive est partagée par l'ensemble des chercheurs en sciences de l'information mais de nos jours alors que les uns, héritiers de la pensée de Meyriat ou d'Escarpit se détournent de la théorie mathématique au motif qu'elle n'étudie pas le même objet qu'eux, les autres (bien que certains s'en défendent) reprennent à leur compte les formules shannonniennes dans un environnement où les algorithmes font loi. Les divergences de points de vue s'accroissent quand on s'intéresse au deuxième concept majeur des sciences de l'information, celui de document. Le paramètre numérique est l'élément déclencheur à l'origine de la réflexion scientifique.

Chapitre 2

Le fait numérique : un perturbateur conceptuel pour le document en SIC ? Où la théorie mathématique de l'information resurgit...

La formation des concepts « *constitue une construction abstraite qui vise à rendre compte du réel* ». Cette construction « *ne retient pas tous les aspects de la réalité concernée* ». (Quivi, 2011 : 122). Appréhender la complexité de la réalité nécessite la mise en relation de plusieurs concepts. Ainsi le concept de document est-il étroitement lié à celui d'information, ce dernier, comme nous l'avons vu plus haut, étant indissociable de celui de communication. De plus, le terme document, comme celui d'information, est un mot très couramment employé et sa conceptualisation n'est pas l'apanage des sciences de l'information. Issu à l'origine de la sphère juridique, le document est devenu objet d'études pour les historiens et les anthropologues, notamment. Cependant il a été théorisé dans le domaine de la documentation dès la fin du XIX^e siècle.

2.1 Les bases conceptuelles du document en SIC

2.1.1 Les précurseurs : la notion d'extensivité et de preuve

Otlet a ouvert la voie dès le début des années 1930 en définissant les documents comme des « *éléments de nature matérielle* » qui permettent « *l'enregistrement de la pensée humaine et de la réalité extérieure* » (Otlet, 1934 : 10). Cette « pensée » étant représentée par les signes tracés. L'expansion de la masse documentaire au tout début du vingtième siècle l'a amené à se pencher sur les moyens de traiter efficacement ces objets et donc de les étudier finement. Il explique « *combien le document aide à réfléchir, à faire prendre claire conscience des idées* » (Otlet, 1934 : 248). Il est en outre à l'origine de la notion d'extensivité considérant qu'au-delà d'un livre, un document, « *c'est une inscription (...) une simple figure conventionnelle de la signalisation (rond, triangle, barrière fermée). C'est même moins, c'est le signe que le boy scout trace à la craie, sur les arbres ou les rochers...* » (Otlet, 1934 : 43). Dans sa définition du document, il précise ainsi que « *l'unité physique, matière du document, est marquée soit par la continuité matérielle de sa surface (ex. : la surface d'une lettre, d'un journal), soit par un lien matériel entre plusieurs surfaces (ex. : les feuilles reliées d'un livre),*

soit par un lien immatériel (ex. : les divers tomes d'un même ouvrage) ». (Otlet, 1934 : 45-46). Le caractère immatériel de l'information supportée par le document n'est donc pas le propre de la toile.

Briet est allée plus loin encore en incluant les êtres vivants et en précisant que c'est l'attention que l'on porte à un objet, quel qu'il soit qui en fait un document. Une opération volontaire au sujet de l'objet en question est nécessaire pour en faire un document. Elle précise sa pensée par une série poétique de questions : *« une étoile est-elle un document ? Un galet roulé par un torrent est-il un document ? Un animal vivant est-il un document ? Non. Mais sont des documents les photographies et les catalogues d'étoiles, les pierres d'un musée de minéralogie, les animaux catalogués et exposés dans un Zoo. »* (Briet, 1951 : 7).

Ainsi elle définit le document comme étant *« tout indice concret ou symbolique, conservé ou enregistré, aux fins de représenter, de reconstituer ou de prouver un phénomène ou physique ou intellectuel »* (Briet, 1951 : 7). La volonté de l'émetteur de *« l'indice »* est ici privilégiée et ce sont les fonctions de mémorisation et de preuve qui sont mises en exergue. M. Buckland souligne l'importance de ce choix du terme *« indice »* car c'est le fait de considérer un objet aux fins de preuve qui en fait un document. (*« Briet's use of the word « indice » is important, that it is indexicality – the quality of having been placed in an organised, meaningful relationship with other evidence – that gives an object its documentary status »* (Buckland, 1997 : 806)). Meyriat illustre cette notion d'indice utilisé comme preuve significative lorsqu'il donne comme exemple *« Sherlock Holmes [qui] passe son temps à convertir en informations des indices fournis par des interlocuteurs qui n'en ont aucune connaissance »* (Meyriat, 1983 : 66).

Dans l'immédiat après-guerre où la question des échanges d'informations, notamment stratégiques, est un enjeu majeur, Briet, a bien conscience de l'importance de la technique pour aider les professionnels à être les plus efficaces et pertinents possibles. Elle soupçonne déjà la marge d'évolution possible sur ce plan et se félicite de ce qu'*« un épais dossier se glisse (...) dans la poche d'un veston »* ou de ce qu'*« une bibliothèque entière est renfermée dans un sac à main »* (Briet, 1951 : 9) : un annonceur ne saurait trouver mieux pour vanter les mérites d'une tablette numérique... Mais ne perdant jamais de vue la fonction du document,

Briet met l'accent sur l'importance du travail de l'esprit. Dans son optique, visant à faire comprendre les intérêts d'une documentation scientifique de qualité, elle insiste sur le fait que les facultés de « *l'homo documentator* » (Briet, 1951 : 29), que sont l'ordre, la signalisation et la sélection, correspondent à « *trois démarches essentielles aux opérations intellectuelles* ». Elle faisait déjà le constat il y a plus de soixante ans que les « *facteurs espace et temps* » donnaient à « *l'effort intellectuel un rythme différent** » (Briet, 1951 : 11). Comme elle le présentait le rythme s'est encore accru et la difficulté à sélectionner les informations pertinentes s'est renforcée.

2.1.2 Le lien indissoluble entre le document et l'information

2.1.2.1 Une relation non équivalente entre les deux concepts

Meyriat remarque que le concept « document » est particulièrement difficile à définir car il ne peut s'envisager sans recourir à ses dérivés, documentation et documentologie. Or, la logique scientifique requiert de partir du radical, en l'occurrence, « document », pour pouvoir cerner les termes qui en découlent (Meyriat, 1981 in Couzinet : 143) D'une façon générale, il n'apprécie pas cette racine étymologique, notamment pour qualifier les professionnels de l'information que sont les « documentalistes » car on n'y retrouve pas la « *relation plus fondamentale qui est établie avec l'information à transmettre et avec l'utilisateur de cette information* » (Meyriat, 1978 in Couzinet, 2001 : 113). S'agissant du document, il ajoute que du point de vue du « documentaliste », « *nous n'avons aucun respect pour lui, aucune considération pour ses qualités esthétiques ou autres. Nous avons besoin de connaître, d'identifier ses origines, ses conditions de production, sa structure matérielle, mais dans la seule mesure où cela nous permet d'évaluer l'information qu'il contient. Après quoi, nous nous intéressons à l'information elle-même en faisant abstraction du document* » (Meyriat, 1978 in Couzinet, 2001 : 118). Ce terme, « *dans le langage commun, évoque toujours l'idée d'écrit et de papier* », image quelque peu dévalorisante « *à une époque où l'on cultive avec prédilection le multimédia* » (Meyriat, 1993 in Couzinet, 2001 : 342).

Cette dernière assertion résonne singulièrement à l'heure où le concept est remis en cause du fait de la technologie numérique comme nous le verrons dans la partie suivante.

Quoiqu'il en soit, Meyriat insiste tout particulièrement sur la relation entre les deux notions document et information. Ces dernières « *sont inséparables l'une de l'autre, et leur conjonction est essentielle dans [la] définition* » (Meyriat, 1981 in Couzinet, 2001 : 144). Ainsi un document est « *un objet qui supporte de l'information, qui sert à la communiquer, et qui est durable (la communication peut donc être répétée)* » (Meyriat, 1981 in Couzinet, 2001 : 143). Il y a deux aspects, l'un matériel (l'objet), l'autre conceptuel (l'information). Ce dernier élément est fondamental. « *Il n'y a document que s'il y a signification* » (Meyriat, 1981 in Couzinet, 2001 : 161). C'est pourquoi l'on ne peut « *définir un document sans tenir compte du signifié du message qu'il a fonction de transmettre* » (Meyriat, 1981 in Couzinet, 2001 : 144). La relation entre les deux notions information et document n'est pas équivalente et dans le débat terminologique pour nommer la spécificité scientifique, plutôt que « documentologie », Meyriat qui avait un temps proposé « informatologie », soumet l'expression « *science de l'information* » qui a le mérite selon lui de « *mettre l'accent sur l'essentiel, c'est-à-dire l'information, et d'attribuer au document un rôle subordonné* » (Meyriat, 1978, in Couzinet, 2001 : 114). La conception de Meyriat positionne de façon hiérarchique le rapport entre les deux concepts ; n'est considéré comme document que ce qui apporte une information. Autrement dit si l'information portée par un document n'est pas activée, l'objet porteur n'est *pas* un document au sens des sciences de l'information. « *Un objet produit intentionnellement pour être un document peut cesser de l'être s'il n'est pas reçu comme tel* » (Meyriat, 1978, in Couzinet, 2001 : 115). Meyriat le qualifie alors de « *document virtuel* » (Meyriat, 1978, in Couzinet, 2001 : 163). En corollaire, dans le cadre des sciences de l'information, « *la signification de l'information n'est pas séparable de son support. Les conditions dans lesquelles ce support est conçu, produit et distribué, les contraintes qu'il subit à ces différentes étapes de son existence influent évidemment sur sa capacité informative* » (Meyriat, 1981, in Couzinet, 2001 : 152).

2.1.2.2 La bipartition documents par intention – document par attribution

Reprenant la notion d'extensivité élaborée par Otlet et précisée par Briet, Meyriat opère une distinction entre ce qu'il appelle les documents « *par intention* » (un auteur a une volonté d'informer) et les documents « *par attribution* », pour lequel la volonté de l'émetteur peut être toute autre que celle du récepteur (Meyriat, 1978, in Couzinet, 2001 : 115-116). Tout objet peut donc devenir document à partir du moment où il y a « *volonté d'obtenir une*

information » (Meyriat, 1981, in Couzinet, 2001 : 145). Dans tous les cas, le document est « *le produit d'une volonté, celle d'informer ou de s'informer (la seconde étant toujours nécessaire)* » (Meyriat, 1981, in Couzinet, 2001 : 147). « *L'utilisateur fait* le document* » (Meyriat, 1981, in Couzinet, 2001 : 147). D'autre part, le dessein de l'acte d'informer ou de s'informer, par le moyen du document, est conditionné par l'usage prévu de l'information, de son utilité en somme. Ainsi Meyriat distingue plusieurs types d'information selon leur genre en opérant le croisement entre la fonction de l'information (« *utilité explicite* » ou « *diffuse* ») et la durée de vie de cette information (« *instantanée* » ou « *durable* »). Les sciences de l'information s'intéressent au « *document susceptible d'entrer dans le système secondaire de transfert de l'information : c'est à ce moment que l'information qu'il supporte entre dans [leur champ]* » (Meyriat, 1981, in Couzinet, 2001 : 157). Cette information est de type « *spécialisée (scientifique et technique...)* ». Le fait de fixer l'information sur un support permet de la conserver et de pouvoir s'y référer en temps utiles. Elle est ainsi soumise au « *processus documentaire* » (Meyriat, 1981, in Couzinet, 2001 : 157). « *L'information peut être donnée et reçue par des voies diverses, soit par contact direct (visuel, oral, éventuellement tactile...) entre deux acteurs d'une communication, soit indirectement en utilisant un support* », le document (Meyriat, 1981, in Couzinet, 2001 : 162). Si les interlocuteurs ne sont pas dans le même espace physique et ne peuvent donc communiquer par la voix et/ou le geste, l'information ne peut circuler que par l'intermédiaire d'un moyen matériel. Les supports peuvent être de nature très variée (CD-Rom, cassettes vidéo, DVD-Rom, etc.) mais ce qui fait leur « *vertu documentaire* », c'est leur capacité à fixer l'information, à être « *stable et durable, aussi longtemps que tout accident matériel, climatique ou autre, ne vient pas compromettre [leur] intégrité* » (Meyriat, 1978 in Couzinet : 117). Cette notion d'intégrité du document est importante pour cerner une caractéristique essentielle qui consiste à permettre des « *exploitations répétées aussi nombreuses qu'on le veut, et à des vérifications sans limite dans le temps* » (Meyriat, 1978 in Couzinet : 117-118).

Concernant les images animées ou les messages sonores, Escarpit crée la catégorie particulière que sont les semi-documents. Contrairement au document, « *le semi-document ne peut être l'objet d'une lecture* » (Escarpit, 1976 : 137), dans la mesure où il n'est ni stable, ni disponible pour être exploré ultérieurement. En revanche, à l'instar du document, il est caractérisé par une « *absence l'un à l'autre de la source et du destinataire* » (Escarpit, 1976 :

138). Il distingue deux sortes de semi-documents : les *semi-documents immédiats*, une pièce de théâtre, un morceau de musique qui comportent un support écrit et des interprètes et les *semi-documents différés*, un film ou un enregistrement sonore qui nécessitent un support mécanique et un décodeur commandé éventuellement par un opérateur (Escarpit, 1976 : 138). « *La technologie moderne* » en constituant ces semi-documents « *à partir du message phonique ou du message visuel animé* » n'a cependant « *pas modifié les contraintes qui pèsent sur la constitution du document* » (Escarpit, 1976 : 123). Celui qui « *fabrique le semi-document* », note-t-il, prend une place plus importante que « *la source du message* » (Escarpit, 1976 : 138) : il donne l'exemple du metteur en scène d'une pièce de théâtre par rapport à l'écrivain. Il souligne toutefois que la sollicitation du canal auditif est préférable car « *ayant une capacité inférieure au canal visuel [il] permet une perception maximale de l'information transmise* », et « *une concentration au moins égale à celle de la lecture du texte et certainement supérieure à celle de la lecture de l'image* » (Escarpit, 1976 : 141). En effet, le son, dont « *la linéarité et le débit* » sont compatibles avec « *la capacité de perception humaine* », est « *le meilleur vecteur du semi-document* » (Escarpit, 1976 : 144). Quoiqu'il en soit, « *le semi-document, adapté pour des apprentissages opératoires* », ne peut remplacer, selon Escarpit, le document « *pour l'acquisition d'un savoir* » (Escarpit, 1976 : 144). Le document a ceci de particulier qu'il permet ainsi que nous allons le voir ci-dessous une indépendance par rapport au temps.

2.1.2.3 L'indépendance par rapport au temps

Escarpit détaille, en recourant aux concepts linguistiques²⁴, cette autre caractéristique majeure du document, « *objet informationnel visible ou touchable et doué d'une double indépendance par rapport au temps* » (Escarpit, 1976 : 120).

D'une part, le document permet de s'affranchir de la diachronie. Il précise que le système humain dispose de deux canaux d'émission (phonique et moteur) et de trois canaux de réception naturels qui correspondent à trois des cinq sens (l'ouïe, le toucher et la vue), à savoir le canal auditif, le canal tactile et le canal visuel. Ainsi, le message oral reçu par le canal auditif ne peut être dispensé que par étape successive de façon linéaire. Ce caractère diachronique est effacé lorsque ce message, transposé sous forme d'inscription, sollicite le canal visuel. Escarpit considère donc que le message radiodiffusé est « *une information à distance sans la médiation d'un document* » (Escarpit, 1991 : 119). Le canal acoustique n'admet en effet, à distance ou pas, que des messages « *linéairement inscrits dans le temps comme des séquences* d'événements* » (Escarpit, 1976 : 120). Pour que ce message radiodiffusé perde cette caractéristique diachronique, il s'agirait de le transcrire de façon scripturaire et d'en faire ainsi un document. En revanche, l'enregistrement de ce même message radiophonique répondrait à la deuxième propriété du document²⁵ en permettant une ré-écoute dans un autre contexte que sa diffusion initiale.

Car le document permet, d'autre part, de s'affranchir de la chronologie : il devient possible de prendre connaissance en un autre temps et en un autre lieu d'un événement passé qui s'est produit à un instant T, le document étant « *un support matériel de la trace qui peut être conservé, transporté, reproduit* » (Escarpit, 1976 : 120). Cela donne pour le message de façon générale la possibilité d'être communiqué en dehors du moment de sa production.

Ces deux propriétés du document sont la « *synchronie* » et la « *stabilité* » et permettent à l'information qu'il porte d'être délivrée de la dépendance à l'écoulement du temps, contingence à laquelle est soumise toute vie et a fortiori toute production humaine comme l'exprime si bien le poète²⁶. Cette indépendance du document par rapport au temps est relative, ainsi que le souligne Meyriat car celui-ci est lui-même, comme tout objet, soumis à

24 S'appuyant notamment sur les travaux de Ferdinand de Saussure (1857-1913), édités en 1957 dans *Le cours de linguistique générale*

25 Dans ce cas, il s'agit comme nous l'avons vu précédemment de ce qu'Escarpit appelle un semi-document

26 « L'homme n'a point de port, le temps n'a point de rive. Il coule, et nous passons ! » Lamartine, 1820 (Le Lac, in *Méditations poétiques*)

l'écoulement du temps et aux dégradations qu'il peut subir, intentionnellement ou non (Meyriat, 1978 in Couzinet : 117). Mais sans cette inscription sur un support matériel, l'information ne peut être mémorisée ou vérifiée aux fins de servir de preuve, au sens juridique du terme mais aussi et surtout au sens rhétorique permettant de justifier, d'argumenter, de consolider l'état des connaissances. Escarpit précise qu'il s'agit de « *compenser l'effet temporel* » et non de « *l'inverser* ». Lors de l'utilisation du document, le temps est « *réintroduit* » par la lecture de l'information qu'il véhicule (Escarpit, 1976 : 122). La qualité de lecture est productrice de sens : en fonction de la façon dont sont perçus les signes, le degré de compréhension sera plus ou moins grand. Il est nécessaire pour que celui-ci soit maximal que le « *lecteur* » ait connaissance de « *la « mémoire externe » contenue dans le texte. Cette mémoire est ce qu'on appelle le contexte*²⁷, produit de la synchronie du texte* » (Escarpit, 1976 : 125). Escarpit met l'accent sur le caractère libre et indépendant de l'esprit qui analyse le texte selon ses propres choix. Cette dimension textuelle du document est une condition nécessaire pour définir ce concept car le texte, « *produit de l'écriture (...) remplit simultanément trois fonctions* » liées aux trois formes de communication possibles : la trace, la parole et l'écriture. Ces fonctions sont la « *fonction iconique* », trace stabilisée dans un objet, la « *fonction discursive* » - la « *linéarité temporelle du discours* », produit de la parole, étant « *figurée (...) par la linéarité spatiale de l'écriture* » - et la « *fonction documentaire* », « *stabilisation de l'ensemble du message mixte (trace/icône et parole/discours) sur un support qui le rend indépendant du temps et synchroniquement disponible* » (Escarpit, 1976 : 121). Le fait qu'il y ait une inscription suppose une composante visuelle (« *visible* », au sens strict, faisant appel à la vue). Pour être diffusable, ce texte est inscrit sur un support matériel (« *touchable* », là aussi au sens strict du terme, faisant appel au toucher, à la matérialité). L'ensemble constituant « *un objet informationnel* » (Escarpit, 1976 : 120). Le document est un « *anti-événement* », il représente « *une cumulation de traces* » (permanentes, dans le cas du bouche à oreille ou fixes, dans le cas du texte). Cela signifie que « *ces traces restent disponibles pour une lecture, c'est-à-dire pour une exploration libre de toute contrainte événementielle ou chronologique, en fonction du projet et de la stratégie destinée à le réaliser* » (Escarpit, 1976 : 59). Le document dans ces conditions permet la « *constitution d'un savoir* » en cohérence avec ses racines étymologiques, (*doceo*, s'instruire) et c'est ce qui fait, selon Escarpit, « *toute la différence entre la notion d'information et la notion de savoir* »

27 Escarpit reprend dans ce passage la terminologie développée par le linguiste Noam Chomsky, professeur émérite au MIT, notamment dans *Deep Structure, Surface Structure and Semantic Interpretation*, 1968

(Escarpit, 1976 : 59).

Selon son approche diachronique de la théorie de l'information, comme nous l'avons souligné en première partie, il constate que « *l'information est une quantité négative apportée à l'esprit par l'événement quand il passe du futur au passé* », tandis que « *le savoir est une construction assez stable pour freiner l'écoulement du temps et rester disponible à l'échelle d'une vie humaine, assez mouvante pour accueillir sans cesse de nouvelles réponses et les insérer dans des structures lisibles* » (Escarpit, 1976 : 59-60). La différence avec la conception de Meyriat, réside dans la notion de connaissance que ce dernier introduit entre l'information et le savoir. Dans une perspective suivant plus l'approche mathématique, Escarpit rapproche l'information d'une rupture dans la communication. Ainsi il définit l'événement comme « *un produit informationnel du temps* » (Escarpit, 1976 : 58). Le document permet non seulement « *une réactivation de l'événement* » mais aussi « *une production informationnelle nouvelle* » (Escarpit, 1976 : 123). Cette activité ne se fait pas au niveau du canal (codage et décodage) mais bien « *au niveau de la source et du destinataire qui en ont chacun de son côté l'initiative* » (Escarpit, 1976 : 123), c'est à dire au niveau des interlocuteurs humains dans le cadre du processus de communication. Les termes « *volontaire* » ou encore « *initiative* » employés par Escarpit mettent l'accent sur le travail intellectuel et le choix qui incombent aux individus et qui ne peuvent être effectués par des « machines » car « *la lecture du texte* » est « *une lutte, une conquête de la signification* » (Escarpit, 1976 : 143). Nous reconnaissons dans cette expression une conception similaire à celle de Meyriat qui s'intéresse à l'aspect cognitif des notions des sciences de l'information.

À l'approche du XXI^e siècle, constatant les évolutions technologiques croissantes, Escarpit²⁸ se demande cependant si « *la clef de notre avenir informationnel* » ne repose pas sur cette capacité à relier « *tout ce qui est périssable dans l'information* » et qui ne peut pas être « *constitué en texte ou en icône* », c'est-à-dire les paroles et les traces ou gestes, qui par essence ne sont pas fixés, stables (Escarpit, 1976 : 145). Le stockage des paroles, gestes, traces sous forme de « *semi-documents* », la faculté à reproduire « *l'événement « à la demande »* » (Escarpit, 1976 : 145) offriraient ainsi le contre-point à la constitution des

28 Il ne reniait rien de ses analyses, publiées en 1976, dans la réédition de sa *Théorie générale* en 1991.

« *anti-événements* » que sont les documents (Escarpit, 1976 : 59).

La technologie numérique et plus encore la connectivité permettant la circulation réticulaire des informations sur l'internet y parvient-elle et entraîne-t-elle naturellement la « mort » du document ? Ce qui est en mouvement dans les échanges en ligne représente-t-il effectivement un moyen de pérenniser la « *part périssable de l'information* » ou au contraire le fonctionnement de la toile renforce-t-il la nécessité de constituer des documents afin de satisfaire aux exigences de l'élaboration des connaissances ? Les temps modernes signent-ils la disparition « *du vieux et stérile débat entre le fixe et le mouvant, entre ce qui reste et qui s'envole, entre l'héritage et l'environnement culturel* » ? (Escarpit, 1976 : 145). Nous pouvons, à cet égard, introduire une nuance dans les approches respectives d'Escarpit et de Meyriat, pour qui les « *informations « périssables »* » sont celles qui sont d'utilité immédiate : « *heure de départ du prochain tram (...) prévisions météorologiques* », ceci en corrélation avec le contexte de celui qui reçoit l'information. Car les exemples cités peuvent être, dans le cadre d'un travail historique, des informations d'utilité durable (Meyriat, 1981 in Couzinet, 2001 : 150).

En réalité de nos jours, des chercheurs remettent en question ce concept de document tel qu'il avait été théorisé en avançant un concept nouveau, celui de « document numérique ». Les évolutions technologiques impliquent un questionnement, voire un repositionnement de certains concepts. C'est d'ailleurs le fondement de la réflexion médiologique initiée par Régis Debray qui s'interroge sur la façon dont « *une rupture de nos méthodes de transmission et de transport suscite une mutation dans les mentalités et les comportements* » et à l'inverse sur la manière dont « *une tradition culturelle suscite, assimile ou modifie une innovation technologique* » (Debray, 2014). Nous allons à présent poser les termes du débat actuel dans la communauté scientifique des SIC autour du concept de document à l'ère dite numérique.

2.2 Les caractéristiques controversées du concept de document à l'heure numérique

2.2.1 Les travaux du RTP-DOC : le document numérique : un nouveau concept ?

2.2.1.1 Qui est Roger Pedauque ?

En choisissant, comme le souligne B. Simonnot, de faire abstraction de toute la littérature produite auparavant sur ce concept (Simonnot, 2012 : 37), la mutation étant telle que les outils intellectuels précédents n'auraient pas de validité opératoire, le Réseau thématique pluridisciplinaire « Documents et contenu : création, indexation, navigation » (RTP-DOC) du CNRS définit un nouveau concept, celui de « document numérique ». L'adjectif venant (comme pour l'information numérique²⁹) non qualifier le substantif mais changer sa nature même.

Ce réseau considérant que « *le contenu n'a de valeur que par rapport à un contexte* » se demandait à l'aube de ses réflexions « *comment rendre compte du statut du texte, des relations avec les ensembles auxquels il se réfère, de la confiance que le lecteur peut lui accorder, de l'articulation avec les informations orales, de sa globalité de sens, sans référence à la notion de document ?* » (Salaün, 2003 : 7). D'un « *objet matériel manipulable* » dans la conception « *traditionnelle* », le document numérique est devenu un « *puzzle dont les morceaux sont agencés à la demande du lecteur* » (Salaün, 2003 : 4). Le document n'aurait de « *forme à proprement parler qu'à deux moments : celui de sa conception par son auteur (...) et celui de sa reconstruction par un lecteur* » (Salaün, 2003 : 4). Nous pourrions voir ici (sans qu'il y soit fait allusion) une référence à la partition élaborée par Meyriat comme nous l'avons vu dans la partie précédente entre document par intention et document par attribution. Au fil des travaux de l'équipe pluridisciplinaire cependant, c'est une vision informaticienne, guidée par un souci d'applicabilité technique, qui émerge sachant que les « *chercheurs en sciences sociales ont été moins représentés dans la rédaction collective* ». (Salaün, 2004 : 16). Menés à travers une démarche plurielle nécessaire car « *la théorie du document (...) engage trop de savoirs spécialisés pour pouvoir être le produit d'un seul individu* ». (Salaün, 2004 : 15), les recherches du RTP-DOC accordent aux considérations techniques, voire pratiques une place prépondérante. Le résultat des réflexions collectives a paru dans un ouvrage signé du

29 Voir partie 1.2.2

pseudonyme Roger T Pedauque, avatar éponyme de l'équipe du CNRS.

2.2.1.2 La démarche et les conclusions du réseau thématique RTP-DOC

Constatant les effets du « *nouveau processus de modernisation* » en matière documentaire, l'ambition de l'équipe de recherche, qui a rassemblé des chercheurs dans les domaines du traitement de l'image, de l'édition numérique, du W3C³⁰, de l'anthropologie, de l'histoire du livre, des sciences du langage, des sciences de l'information, de l'ingénierie des connaissances, des Télécoms, des sciences de la communication, de la gestion et des sciences sociales, est, selon eux, du même niveau que celle de Paul Otlet dans sa « *tentative de fonder une « documentologie » au début du XX^e siècle (...) emblématique du besoin de rationaliser la prolifération documentaire qui s'est installée* » (Pedauque, 2007 : 15-16). Le passage du « *document-papier imprimé* » au document numérique consacre un phénomène constitutif de nos sociétés contemporaines, que les chercheurs appellent la « *redocumentarisation* », « *réingénierie documentaire initiée par le Web* ». (Salaün, 2012 : 12). Le rôle joué par l'internet et plus précisément par une de ses applications, la toile, qui permet de consulter les pages accessibles sur des sites via un navigateur, est fondamental dans cette approche du document. Ce n'est donc pas tant les caractéristiques du document liées à la technique numérique qui sont observées que celles liées aux interconnexions sur le réseau. Plus que du « document numérique », il s'agit là plutôt du « document en ligne ». La multiplicité des disciplines concernées par les études ouvre des chantiers tout aussi variés et pour donner une cohérence à l'ensemble, les recherches ont concerné trois aspects : la « *forme* » (ou le « *signe* ») : « *vu* », reliée au « *support* », à la « *structure* » et aux « *formats* » - le « *texte* » (ou le « *contenu* ») : « *lu* », relié au « *sens* », au « *méta* », aux « *moteurs* » et aux « *ontologies* » - le « *medium* » (ou la « *relation* ») : « *su* », relié à la « *collection* », à l'« *hyper* », aux « *portails, collaboratifs, sites* » (Pedauque, 2007 : 20). Le support, le contenu et la visée communicationnelle du document ont donc été délibérément abordés de façon distincte en opposition à la conception traditionnelle, à savoir le caractère indissociable du document et de l'information, comme nous l'avons souligné précédemment. Mais la nouveauté du format et de l'usage réticulaire légitime cette approche selon l'équipe qui précise toutefois que les trois aspects sont interdépendants. (Pedauque, 2007 : 17)

30 World Wide Web Consortium : organisme, créé en 1994, chargé de normaliser les différents formats de l'internet (html, xml, png, etc.), il regroupe plusieurs centaines d'entreprises partenaires (dont Adobe Systems Inc., Alcatel-Lucent, Google Inc., Orange, etc.).

Le document est désormais défini comme une « *représentation d'une vérité partagée au-delà du chaos (le silence et le bruit), de la cacophonie (la confusion et le sensible) et de l'oubli (l'intime et l'éphémère)* ». Ce qui caractérise donc le document, c'est la faculté à être « vu » (reconnaissance des signes, lisibilité), « lu » (compréhension, assimilation) et « su » (sélection, transmission) » (Pedauque, 2007 : 17). Derrière le slogan facilement mémorisable, « vu », « lu », « su », nous pouvons voir apparaître dans cette définition du « nouveau » concept, l'influence de la théorie mathématique : les termes « chaos », « cacophonie » renvoient à la notion de désordre qu'il s'agit de combattre, ce qui caractérise une approche négative du phénomène informationnel. Il est intéressant de noter que ce sont les notions de « compréhension » et d'« assimilation » qui sont mises en avant, présentant le rapport à l'information dans son résultat : appréhension par la connaissance ou appropriation des connaissances acquises (ce terme de « connaissance » n'est d'ailleurs pas utilisé) au détriment de la notion de « sens », de signification. L'acte de connaissance s'effectue quand l'esprit donne forme³¹ à « *l'objet dont il prend possession* », il le reconnaît, lui attribue un sens (Meyriat, 1983 : 66). De même le processus de communication, sans lequel, rappelons-le, d'après Meyriat, il n'y a *pas* information, est réduit dans la nouvelle définition pré-citée à la « sélection » et à la « transmission », opérations que l'on peut aisément voir réalisées par des outils dans cette approche mathématique.

Ces travaux du CNRS font référence dans la quasi-totalité des productions contemporaines puisque les différents auteurs que nous avons lus évoquent, avec certes plus ou moins de distance critique, ces études quand ils abordent le document numérique, observant que plusieurs attributs du document changent en raison de la technologie numérique, et plus particulièrement de la toile, parmi lesquels : la matérialité et la stabilité en lien avec la question du support - la temporalité - l'unité documentaire – la fonction documentaire. Les chercheur-es en SIC ne s'accordent toutefois pas, comme nous allons le voir maintenant sur les conséquences conceptuelles des éléments mouvants relevés.

31 Meyriat fait référence aux racines étymologiques du mot information : informo, donner une forme

2.2.2 De l'objet stable à l'entité disparate

2.2.2.1 Dissociation support-contenu ?

Selon l'équipe du CNRS, le « *numérique déplace la question du support du document, qui en assurait la stabilité grâce à la fixité de l'information, vers la problématique de sa structure* » (Pedauque, 2007 : 17). Le groupe de réflexion souligne la possibilité « *inédite* » de réaliser des « *traitements formels différents d'un même contenu* » par le format XML (Pedauque, 2007 : 18). Cependant, les recherches tournent autour de considérations techniques soulevées par la migration des documents vers le support numérique, tels les « *outils de lecture automatique des documents* » ou « *outils d'OCR* » (Pedauque, 2007 : 32) ; par la problématique de l'homogénéisation des procédés employés comme « *la question de l'interopérabilité des plateformes de numérisation* » (Pedauque, 2007 : 36) ; par le fonctionnement des outils de « *gestion de document déjà au format numérique (...) au format PDF (...) ou issus de logiciels de Conception Assistée par Ordinateur* » pour n'en citer que quelques uns (Pedauque, 2007 : 43). La problématique des chercheurs n'est, là encore, pas d'évaluer les conséquences de la nouvelle forme du support sur la nature de l'information mais de régler les problèmes de compatibilités entre les différents outils.

Si l'on aborde le traitement documentaire, le « *tout-électronique* » induit nécessairement des pratiques nouvelles selon H. Fondin. Ce dernier rappelle les liens entre le support et l'information qu'il transporte : « *le texte, véhicule linguistique de l'information documentaire, représente toujours un enjeu documentaire parce qu'il est porteur de sens explicite et implicite, donc source de renseignement pour celui qui l'interroge ; accessible après un traitement plus ou moins lourd du fait d'une faiblesse, voire d'une absence de structure de repérage des informations ; couché sur un support, ce qui oblige à considérer à la fois un contenant et un contenu* » (Fondin, 1998 : 9). Plutôt que d'évoquer une dissociation support-contenu, il précise que les deux sont à prendre en considération simultanément.

Le format numérique impose le traitement de l'information, au sens mécaniste du terme, c'est à dire son codage. Il faut donc que « *l'élément de connaissance* » à transmettre soit « *susceptible d'être représenté à l'aide de conventions pour être communiqué* » (Fondin, 1998 : 129). L'opération mathématique qui transforme l'information en « *donnée* » s'appuie sur deux éléments « *une rubrique – ou variable – identifiant la nature de l'information* » et

« une valeur – ou contenu – représentation de la connaissance » (Fondin, 1998 : 129). Se plaçant du point de vue de la constitution d'une base de données par un professionnel, H. Fondin s'éloigne de la conception de Meyriat pour qui l'information n'existe qu'au moment où une opération de l'esprit s'effectue. H. Fondin explique néanmoins que l'« informatiste » s'intéresse à la « dimension informationnelle » du texte, « le cadre dans lequel [ce] texte est inséré par rapport à l'ensemble des savoirs (...) sa contribution à l'avancement du savoir de l'humanité » (Fondin, 1998 : 34). La difficulté avec le document électronique, c'est que « pour pouvoir reconnaître les informations, il faut pouvoir reconnaître les règles de codage et de décodage des chaînes de caractères échangées ainsi que celles, qui sont aussi complexes que les informations transmises, de tous les éléments qui les entourent pour assurer leur mise en forme et leur mise en page » (Fondin, 1998 : 35). D'où la nécessité de constituer un « modèle normalisé ». L'intention d'Otlet lors de l'élaboration de son *Traité de documentation* était de même nature. Mais « sur le plan éditorial, jusqu'alors, le document était un assemblage élaboré par un auteur, une composition originale. Désormais il ressemble à une entité composite, à un ensemblier » (Fondin, 1998 : 47). Il est le « résultat de nombreuses interventions », construit par « chaque destinataire, et non par un auteur », « susceptible d'être modifié à tout moment, échangé (...) récupéré, transformé » (Fondin, 1998 : 47). Ces caractéristiques nouvelles justifient, selon lui, de considérer le « document électronique » comme « un document en tant que tel » (Fondin, 1998 : 47).

Le lien indéfectible du document avec l'information n'est pas remis en cause : « l'information, c'est à dire le contenu, ne peut être dissociée de l'objet qui la supporte, le contenant, le document. L'un et l'autre forme un tout, une unité » (Fondin, 1998 : 71). Mais l'information est prise au sens de la théorie mathématique. Le fonctionnement du format numérique conduit à une approche usager loin des approches systèmes des « -thèques » où un document avait une place unique. C'est le « flux d'informations » plutôt que le « stock de documents » qui est apparent pour les utilisateurs. Le professionnel de l'information a une place plus discrète « en amont du système documentaire » et ne se trouve plus être un « intermédiaire obligé » (Fondin, 1998 : 117-118). Reprenant à son compte cet état de fait, il occulte les qualités spécifiques de l'« homo documentator », l'« ordre », la « signalisation » et la « sélection » corroborant l'idée répandue que le fonctionnement de la toile suffit à la réalisation de ces « occupations intellectuelles » (Briet, 1950 : 11). Confortant d'autre part le cliché habituel, il constate que les activités sur l'internet « se dématérialisent » et s'effectuent

à travers un seul et même outil (certes unique, mais l'équipement indispensable pour accéder au contenu est tangible) : cette « *documentation sans papier* » ne va pas sans poser des problèmes d'ordre technique, juridique, économique, social, culturel et même ergonomique (Fondin, 1998 : 178-179). Cette expression montre que la « dématérialisation » concerne en réalité la disparition (ou supposée telle) du papier, support de prédilection jusqu'alors. Il concède toutefois qu'« *il y a autre chose que les machines ont encore beaucoup de mal à saisir, à prendre en compte, c'est le sens, l'information véhiculée par un langage au fonctionnement complexe, car il associe l'élément et le contexte, l'explicite et l'implicite, l'apprentissage et l'expérience, l'évidence et le caché* » (Fondin, 1998 : 312). Cette assertion vient finalement contredire la possible transposition de « *l'élément de connaissance* » par le traitement informatique, évoquée plus haut, car la dimension humaine est irréductible à la machine.

Cette prétendue dématérialisation, permise par « *le numérique* », et tant vantée, participe, selon D. Cotte, des « *utopies de la communication* » qui resurgissent à chaque innovation technologique (Cotte, 2004 : 34). La fusion de tous les supports en un même format serait capable d'homogénéiser les éléments qui participent à l'élaboration des connaissances. L'incantation forte qui somme de « *passer au numérique* » pousse à s'interroger sur les bouleversements auxquels est soumis le document (Cotte, 2004 : 33). Nonobstant ces changements de statut, les documents restent des « *objets qui permettent la relation entre les hommes* » (Cotte, 2004 : 34). L'auteur, ici, met l'accent sur le processus de communication. Il note qu'auparavant il fallait qualifier la forme autre qu'écrit en précisant « *sonore* », « *audiovisuel* » mais que « *le numérique* » ne préjuge pas de « *la forme finale de réception* » du fait de la migration des différentes formes possibles dans ce format unique (Cotte, 2004 : 34). Le « *caractère numérique* » du document est un « *état transitoire* » qui nécessite un « *appareillage* » pour la lecture. Cette notion constitue une des caractéristiques du « *concept de document numérique* » (Cotte, 2004 : 34). La « *volonté d'élargissement* » du document (jusqu'à l'animal chez Briet) en vient selon lui à « *tuer le concept lui-même* » (Cotte, 2004 : 35). L'Antilope, en tant qu'élément naturel, ne peut rien dire, et n'est donc pas un objet informatif. Ce n'est qu'en produisant un document, un discours sur elle que sa capacité informative se révèle. Le simple fait de la placer dans un zoo ne suffit pas. La notion d'extensivité qui n'a cessé de progresser au fil de la conceptualisation du document semble se

heurter à la technologie numérique qui en fusionnant les divers supports en un seul tend à l'unification du document. La notion de destinataire est cruciale pour D. Cotte, c'est par la réalisation d'acte volontaire qu'est produit un document, « *pour un public et un usage donné* » (Cotte, 2004 : 36). D'une manière générale, il considère que « *le document, l'information n'existent pas à l'état naturel, mais dans certaines circonstances, selon une certaine volonté* » (Cotte, 2004 : 31). Un objet ne peut donc être un document « *en soi* » mais en étant « *documenté* », il est à l'origine du document. Ce faisant, il ne reprend pas la partition élaborée par Briet et développée par Meyriat entre « *document originaux* » et « *documents dérivés* ». Le sens est attribué par l'individu lors de l'observation de l'« objet » d'étude, quel qu'il soit, et le document qui en émane entre dans le « *système secondaire de transfert de l'information* ». L'épithète « *secondaire* », comme le précise Meyriat, n'est pas dépréciative mais indique qu'il se rapporte au document primaire et « *s'appuie sur lui* » (Meyriat, 1981 in Couzinet, 2001 : 167). En revanche, D. Cotte rappelle, à juste titre, que la plupart des documents pour être confectionnés font appel à l'électricité (téléphone, phonographe, radio, télévision...) et pourraient donc être qualifiés de « *documents électriques* » mais que cela n'apporte rien à la définition (Cotte, 2004 : 36). D'un point de vue conceptuel, il estime donc que la dimension technique du document numérique ne suffit pas à le distinguer en tant que tel.

Ce spécialiste de la gestion numérique de l'information rappelle que le « *document numérique* » ne diffère pas systématiquement du « document » s'il peut être restitué sous une forme analogique (Cotte, 2004 : 39). En revanche, il est significatif de qualifier ceux qui « *subsistent sous une forme consultable au moyen d'un appareillage électronique* » de « *documents numériques* », tels les « *cd-roms* » et « *sites web* » (Cotte, 2004 : 39). Le document numérique ainsi délimité comporte plusieurs « *caractéristiques* » propres : l'« *ouverture* » car il est « *susceptible de subir toutes les altérations, voulues ou non* » ; l'« *automaticité* » car il est « *paramétrable* » puisque constitué de données qui peuvent être rééditées à volonté ; la « *plasticité* », ou la « *recomposition* », qui confère la possibilité de transposer un même contenu dans des formats divers mais n'est pas, selon lui, contrairement à l'idée reçue, un affranchissement pur et simple du support car le « *passage d'un média à un autre n'est pas si fluide que cela* » mais requiert souvent une restructuration (Cotte, 2004 : 39-40). L'interface technique est nécessaire mais pas suffisante pour définir le document numérique. Les formats classiques subsistent et l'on imprime les documents numériques qui

redeviennent ainsi en quelque sorte des documents traditionnels. Il conclut que « *l'expression « document numérique » est « mal formée »* car elle met l'accent sur la « *dimension technique* » et dénote une homogénéité trompeuse ne signifiant au final pas grand-chose car « *tout document au XXI^e siècle est, à un moment donné de son cycle, numérique* » (Cotte, 2004 : 41). En prenant garde de ne pas mêler aspect technique et caractéristique conceptuelle du document, cet auteur opère quand même une dichotomie support-contenu car les nouvelles dimensions (« ouverture », « automaticité », « plasticité ») du document consultable par le truchement d'un « *appareillage* » sont opérationnelles si l'on considère que ce qui est supporté, ce sont des « données », au sens mathématique du terme. Ainsi ces observations ressortent d'une vision partielle du concept d'information car le travail de l'esprit nécessaire, « *le regard porté sur* » le document, pour reprendre l'expression d'Y. Jeanneret, est évacué (Jeanneret, 2000 : 85).

G. Lallich-Boidin, documentaliste spécialisée en informatique de formation, et J-P. Metzger, contributeurs de RTP-DOC en leur temps, constatent, pour leur part, que les innovations technologiques font émerger des « *objets qui n'ont pas de dénomination courante* », tel le « *document numérique* » (Metzger, 2004 : 12). Ils définissent le document comme « *une relation quaternaire asynchrone entre un auteur, un discours, un support et un lecteur* » (Metzger, 2004 : 12). Plutôt que le terme information, ou signe, les auteurs préfèrent le terme « *discours* ». Sans inscription sur un support, les discours ne peuvent devenir des documents. Un document est « *donc nécessairement un objet matériel* » (Metzger, 2004 : 12). Que l'auteur l'ait ou non construit intentionnellement, c'est « *le lecteur qui dote cet artefact du statut de document* » (Metzger, 2004 : 12). Suivant en cela Meyriat, les auteurs affirment qu'« *un document sans lecteur n'est pas un document* » (Metzger, 2004 : 12). La particularité du document numérique est d'être « *sur support électronique, d'être perceptible via la technologie numérique* » (Metzger, 2004 : 12). Ce support « *interdit l'enregistrement d'objets du monde et se limite à l'enregistrement de la représentation de ces objets* » (Metzger, 2004 : 12). De là à estimer que la représentation suffit à faire saisir le monde réel, il n'y a qu'un pas que nous ne franchirons pas. Cette constatation ne signifie pas, contrairement à l'assertion de D. Cotte rapportée plus haut, que les objets de la vie réelle ne sont pas porteurs d'information. Comme le souligne R. Casati, malgré toutes les possibilités encensées par les thuriféraires de la technologie numérique, cette dernière ne permet pas la circulation d'objets concrets :

« toute représentation (...) peut migrer vers le numérique [mais] toute chose n'est pas numérisable ». Concrètement, le philosophe prend l'exemple du sandwich³² dont il est possible de numériser une représentation mais qu'il est impossible de manger « numériquement » (Casati, 2013 : 157). Ces objets porteurs d'information peuvent donner naissance à des documents secondaires, numériques ou non. Dans cette approche, nous constatons que le fait de ne pas associer de façon systématique le concept d'information (au sens information-connaissance) au concept de document occulte de fait l'importance de celle-ci dans notre être au monde. Comme le rappelait Meyriat, l'information « est partout, dès lors qu'il y a relation entre êtres humains » (Meyriat, 1986 in Couzinet, 2001 : 250).

La question du support ne saurait être remise en question, même à l'épreuve du fait numérique puisque l'espace documentaire, est le « lieu où s'organisent les ensembles de documents », il est « le support matériel de médiation entre les deux pôles de la communication », autorisant de ce fait « la désynchronisation entre les processus d'énonciation et de réception, d'interprétation ». Cette matérialité « lui permet de traverser le temps » (Metzger, 2004 : 13). Nous en venons donc au deuxième aspect changeant du document à l'ère numérique, relatif à la temporalité.

32 La trivialité de l'exemple choisi est intéressante car il semble parfois nécessaire de rappeler des évidences aussi élémentaires quand d'aucuns estiment que l'internet « peut tout faire »

2.2.2.2 La nouvelle temporalité du document

Ainsi, « *les textes sur écran non seulement peuvent être modifiés après leur diffusion ou construits collaborativement par exemple, mais ils ont une durée de vie qui peut s'avérer très éphémère* » (Pedauque, 2007 : 51). Le document numérique pose donc, selon les domaines de spécialités, des questions relatives au temps. Les « *informaticiens (...) s'interrogent sur la synchronisation des différents modes (sons, images animées avec le texte par exemple), et sur les effets sur cette synchronisation des interactions avec les utilisateurs* » (Pedauque, 2007 : 86). Les historiens « *dont le métier consiste à reproduire une période à partir de documents, se confrontent à leur authentification et à leur contextualisation* » (Pedauque, 2007 : 87). Ce travail de contextualisation et d'authentification, est en effet constitutif de la démarche historique, comme l'a montré V. Couzinet³³, mais s'il peut s'avérer complexe sur la toile, ce travail est tout aussi difficile dans d'autres contextes. Ces dernières préoccupations de validité rejoignent, selon l'équipe du CNRS, celles des professionnels de l'information. Les auteurs raisonnent la question de la temporalité en proposant une partition en trois « univers » : « *l'univers socio-historique* » (événements et périodes), « *l'univers du discours* » (construction) et « *l'univers documentaire* » constitué par les documents, numériques en l'occurrence. Le deuxième univers permet, à travers l'énoncé, la reconstruction du premier univers (contextualisation). Quant à l'univers documentaire, par sa matérialité, il est le « *support de la mémoire* » (Pedauque, 2007 : 87).

L'enregistrement d'« *énoncés* » sur un support matériel leur permet de s'affranchir du temps social. L'univers documentaire contribue à la construction de l'univers social mais aucun des documents qui le compose ne peut être interprété sans être inscrit dans le temps social. En lisant le contenu du document, le récepteur doit reconstruire l'univers social de l'auteur à partir du discours produit. « *Toute unité documentaire est considérée comme un événement surgissant dans un champ d'interprétation* » (Metzger, 2004 : 17). Il s'agit bien d'attribuer un sens à ce qui est lu. Le temps de lecture fait partie intégrante du temps documentaire. Le temps social, est le temps dans lequel s'inscrit une société humaine, ce n'est pas une construction intellectuelle. C'est l'écoulement du temps auquel rien ni personne n'échappe. Selon les disciplines, on évoque les « *ères géologiques, périodes historiques, âges de la vie* » (Metzger, 2004 : 18). Cependant certains documents influencent par eux-mêmes

33 Voir partie 2.3

ce temps social. Une annonce médiatique d'une « *mesure gouvernementale peut déclencher des mouvements sociaux* » qui s'inscrivent dans l'histoire de la société (Metzger, 2004 : 18). Le temps documentaire, lui, est soumis de façon interne au temps social puisque le temps de l'auteur précède toujours celui du lecteur. « *Le support numérique des documents ne bouleverse pas cette distribution en trois univers* » mais pose des questions pratiques : effacement de la « *distance temporelle* » du fait de la vitesse de diffusion des fichiers numériques ; « *délinéarisation* » accrue par la multiplicité des parcours possibles avec l'activation des liens ; absence de « *norme de nommage* » des fichiers (un même document peut porter des noms différents selon le lieu où il réside, ce que le DOI³⁴ pourrait résoudre) (Metzger, 2004 : 19).

Les trois univers ont en commun « *l'essence temporelle* ». « *Les modalités actuelles de l'information et de la transmission (spatiale et temporelle) poussent à s'interroger sur les temporalités : une des clés d'accès à la signification et au sens. Tout dispositif numérique d'information et de communication ne peut aujourd'hui négliger le temps et doit l'appréhender dans toute sa complexité* » (Metzger, 2004 : 21). Cette question relative au temps, prégnante sur la toile, n'est cependant pas nouvelle d'un point de vue général comme l'analyse d'Escarpit l'a pointée³⁵. Les auteurs insistent néanmoins sur la nécessaire prise en compte de ces trois dimensions temporelles représentées dans les trois univers au sein des applications informatiques, notamment pour les archives audiovisuelles (Pedauque, 2007 : 88). Reprenant les travaux de M-C. Luesebrink³⁶, les chercheur-es font la distinction entre le « *temps de la machine* » et le « *temps cognitif* ».

Le « *temps de la machine* » comprend trois temps : « *le temps mécanique (...) de réalisation des opérations par la machine et le réseau* » ; « *le temps de la lecture (...) pendant laquelle le lecteur est physiquement et mentalement engagé dans l'interaction avec le document (lire, regarder, écouter)* » et le « *temps d'interaction (...) pendant laquelle le lecteur engage une interaction avec le document (...), sélectionne un document ou construit une requête d'accès à ces éléments* » (Pedauque, 2007 : 89).

Le « *temps cognitif* » est également triple : le « *temps réel* » est celui « *du déroulement des événements relatés dans le document* » ; le « *temps narratif* » est celui de « *l'exposé des*

34 Digital Object Identifier ou identifiant numérique d'objet : norme ISO permettant d'identifier une entité sur l'internet.

35 Voir partie 2.1.2.3

36 Marjorie Coverley Luesebrink, auteure américaine de fiction hypertextuelle

événements » et le « *temps mythique (...) donne des informations sur le contexte des faits relatés* », contexte « *social, culturel, historique* » d'interprétation (Pedauque, 2007 : 90). Autant de dimensions à prendre en compte pour représenter le temps par les auteurs de documents multimédias. Ainsi, « *l'affichage à l'écran d'un document annihile tout indice de repérage contextuel et notamment le temps* » (Pedauque, 2007 : 94). Les chercheurs proposent donc de rendre techniquement visible l'effet temporel en jaunissant par exemple le fond d'écran pour représenter l'ancienneté ou bien en jouant sur les polices de caractère employées (Pedauque, 2007 : 95). Ces diverses recommandations visent à surmonter des obstacles d'ordre technique, qui finalement nuisent aux possibilités d'interprétation de l'information, mais ne garantissent pas que l'effet escompté se produise au niveau de la réception.

Pour D. Cotte, la « *différence fondamentale* » à l'ère numérique provient de la temporalité dans la mesure où la « *fabrication/conception de l'écrit en tant qu'objet, et ses moments de réception/appréhension peuvent être parfaitement disjoints dans le temps et dans l'espace* » (Cotte, 2004 : 37). Nous pouvons remarquer la résurgence (sans qu'elle soit notifiée par l'auteur) de la partition de Meyriat, détaillée précédemment, entre document par intention et document par attribution.

Poursuivant la réflexion temporelle, J-P. Metzger estime, sur le mode existentialiste, qu'« *un document ne naît pas document, il le devient* », sous l'impulsion du lecteur, chercheur, récepteur d'information (in Papy, 2008 : 87). D'emblée, il relie les deux notions fondamentales, le document et l'information. Il attache une importance particulière au contexte de production et à ce qu'il nomme « *la temporalité de la documentation exploitée* » (in Papy, 2008 : 87). Cette question du temps est étroitement liée à la mémoire. Reliant les notions de temps, de mémoire et de document, il développe trois phases successives : la première phase, documentaire, est celle qui permet « *l'établissement de la preuve documentaire* ». Cette dimension correspond à la première fonction originelle du document. La deuxième phase, explicative ou compréhensive permet « *l'établissement d'une réponse à la question « pourquoi ? »* ». Nous reconnaissons le lien avec l'information-connaissance qui contribue à l'enrichissement cognitif. La dernière phase, représentative est « *la mise en forme du discours porté à la connaissance du lecteur* ». Cet aspect reprend les attributs du processus de communication (in Papy, 2008 : 93). Le document est perçu comme support permettant

d'apporter des réponses à un questionnement. Il précise que « *rien n'est en tant que tel un document, même si tout résidu du passé est potentiellement trace. Le document n'est pas simplement donné, il est cherché et trouvé* » (in Papy, 2008 : 100). Il rejoint la vision de Meyriat pour qui c'est « *l'utilisateur qui fait* le document* » (Meyriat, 1981 in Couzinet, 2001 : 147). Ainsi « *tout ce qui peut être interrogé par une personne en situation de recherche d'information* » peut devenir document (in Papy, 2008 : 100). De ces considérations édifiantes, nous parvenons pourtant difficilement à reconnaître des problématiques nouvelles liées au fait numérique.

Se rapprochant de la définition du document, donnée par Escarpit, comme étant à la fois un « texte » et un « anti-événement », J-P. Metzger évoque l'« énoncé » que « *ni la langue, ni le sens ne peuvent tout à fait épuiser* » (in Papy, 2008 : 102). Ce dernier est « *une modalité d'existence propre à un ensemble construit de signes graphiques, phoniques ou autres (...) [qui] lui permet d'être en rapport avec un référentiel, (...) d'être enfin doté d'une matérialité répétable* ». Ce référentiel est « *le lieu, la condition, le champ d'émergence, l'instance de différenciation des individus ou des objets, des états de choses et des relations qui sont mises en jeu par l'énoncé lui-même* » (Metzger in Papy, 2008 : 102-103). Un énoncé est « *rémanent : il est conservé grâce à des supports ou techniques matériels (dont le livre bien entendu n'est qu'un exemple), selon certains types d'institutions (la bibliothèque entre autres), et avec certaines modalités statutaires (qui ne sont pas les mêmes selon qu'il s'agit d'un texte religieux, d'un règlement de droit ou d'une vérité scientifique)* » (in Papy, 2008 : 106). L'énoncé, le référentiel sont des éléments qui ont trait au contexte global dans lequel s'insère l'information et qui est fondamental pour l'interprétation, la signification. Le « *temps documentaire* » précisément est celui « *de la vérité scientifique* », « *socialement établie* », constitutif du savoir (in Papy, 2008 : 107). Ces réflexions sur la vertu documentaire à l'heure numérique ne diffèrent en rien de la conception d'Escarpit, selon laquelle le document permet la « *constitution d'un savoir* » (Escarpit, 1976 : 59) .

2.2.3 De l'unité documentaire à l'unité informationnelle

2.2.3.1 La fragmentation des contenus

L'aspect morcelé de l'« information » présente sur la toile est soulignée par B. Juanals qui constate quelle est « *dispersée car multilocalisée, hétérogène, difficilement validable et incontrôlable* » (Juanals, 2004 : 64). L'accent est mis sur le caractère non linéaire du « discours » sur ce support, offrant plusieurs parcours de lecture possible. Cette caractéristique existe dans les encyclopédies et certains artistes se sont emparés de cette forme d'écriture comme le rappellent B. Juanals ainsi que le RTP-DOC. Le mouvement littéraire Oulipo avait à cœur de donner au lecteur « *une importance toute particulière au travers de ses parcours de lecture* ». (Juanals, 2004 : 59). Ce point de vue est à nuancer si l'on veut éviter l'écueil de l'approche mathématique en précisant que la dispersion, l'hétérogénéité sur la toile concerne les données. L'attribution de sens qui les transforme alors, et seulement alors, en informations est un autre sujet.

B. Simonnot constate également que « *l'unité documentaire bien définie dans le monde de l'imprimé a été, semble-t-il mise à mal par le passage au numérique* ». Elle évoque une « *fragmentation des contenus* », une « *granularité* » des « *objets informationnels* » (Simonnot, 2012 : 14). Ainsi « *aux débuts du web grand public, la notion de document est devenue floue : on consulte un site, une page ou une rubrique sans pour autant avoir l'impression de consulter un document* » (Simonnot, 2012 : 19). Cela entraîne un déficit d'organisation des documents, une perte de repère au sein d'une collection, et même un manque d'organisation des unités informationnelles entre elles, ce qui participait à la cohérence d'ensemble. Elle pointe la difficulté à « *restituer* » le « *contexte de publication du document* » par « *les nouveaux services de recherche d'information* », parmi lesquels les moteurs de recherche, en comparaison avec les « *services documentaires* » traditionnels soumis à un traitement spécifique : signalétique, indexation (Simonnot, 2012 : 36). Or, cette connaissance est « *indispensable pour pouvoir en apprécier le contenu* » (Simonnot, 2012 : 36). La possibilité offerte de « *recomposition diverse de fragments de textes, images, sons* », le fait de pouvoir redécouper, extraire, rassembler, commenter, annoter, traduire des contenus impose de reconsidérer « *la notion de document « original » [qui] se perd dans la multitude des reproductions, reprises, rediffusions, commentaires* » sur la toile (Simonnot, 2012 : 37). Dans ces conditions, précise-t-elle, « *il est difficile de reconstituer la manière dont tel ou tel document a été composé, d'en dresser une signalétique et une généalogie qui participent aux*

éléments que nous prenons en compte pour l'interpréter ». (Simonnot, 2012 : 38) La notion aurait donc « *tendance à s'effacer au profit d'autres : information, ressource* » (Simonnot, 2012 : 38). Cependant l'emploi du conditionnel dénote, chez l'auteure, une certaine prudence car pour aborder les entités informationnelles qui circulent sur l'internet, elle emploie régulièrement ce mot « document », faute de mieux, ce terme étant comme le soulignait Meyriat, très connoté « *papier* » (Meyriat, 1993 in Couzinet, 2001 : 342). Mais si le terme rendant compte du concept est insuffisant, ou mal trouvé, cela justifie-t-il pour autant la dénégation sa valeur symbolique ?

B. Simonnot revient à la notion d'unité documentaire, en tant qu'« *entité autonome* » déjà entamée, comme le notait Briet, avec « *la multiplicité des procédés de reproduction* » (Simonnot, 2012 : 38). Dépossédé de ses attributs de preuve et de conservation, le document en ligne serait devenu un prestataire de services, la notion de « *ressource documentaire* » étant « *fortement connotée par l'utilité du document pour la personne ou la communauté qui la consulte* » (Simonnot, 2012 : 38). La profusion de « *documents* » en ligne, leurs variétés et leur capacité à être construits en temps réel apportent des avantages : la possibilité de « *consultations simultanées* » sans dégradation ; « *les capacités de stockage des collections* » (qui se retrouvent à l'abri de « *destructions accidentelles* ») ; une réduction des « *coûts de reproduction et de diffusion* », l'essentiel des coûts étant porté sur le « *processus de création* » (Simonnot, 2012 : 48). Mais cette économie est relative car elle ne prend pas en compte les coûts de matériels et de maintenance des réseaux et autres serveurs informatiques (Simonnot, 2012 : 48). Nous pouvons observer que cette acception est partielle puisque seule la dimension économique est prise en compte.

De plus, au sujet de la conservation, B. Juanals met en garde contre « *la confusion* », s'agissant de la conservation des données, « *entre la capacité de stockage de données numérisées multimédias rendues possibles par les supports électroniques (...) et la conservation à long terme de ces données* », celle-ci pouvant être raccourcie du fait d'un accident matériel. Il est nécessaire, selon elle, de mettre en place une politique de conservation spécifique qui a un coût : la solution technique la plus durable à long terme étant, selon elle, « *l'émulation* », sauvegarde sous forme de métadonnées du logiciel ayant servi à créer le fichier et du système d'exploitation de l'ordinateur (Juanals, 2004 : 136). Le

terme donnée, nous écarte de la notion qui nous intéresse, l'information. Un des changements fondamentaux induit par le « web », selon B. Simonnot, et plus précisément par les dispositifs d'accès à l'information que sont les moteurs de recherches commerciaux, objet de son étude, est le fait que les « *objets potentiellement accessibles (...) ne constituent pas une collection organisée, contrairement aux documents traités par les SRI³⁷ traditionnels* » (Simonnot, 2012 : 55). Le biais introduit par ces moteurs de recherche dans la relation entre l'information et les utilisateurs est la volonté de dispenser les internautes de la consultation des « *documents primaires* » en faisant comme si l'information était directement fournie par leur propre service (Simonnot, 2012 : 75). Ce qui est pour partie vraie puisque la plus grande opacité (sous prétexte de protection contre la concurrence) règne sur la manière dont les moteurs « *procède à l'appariement entre requêtes et documents* » (Simonnot, 2012 : 72). Ce faisant, l'auteure fait référence à la partition conceptuelle du document en deux niveaux et invalide, comme nous le remarquions précédemment, la prétendue « disparition » de la notion du « document » sur la toile.

Pour permettre une meilleure orientation dans le gisement d'« informations », l'équipe pluridisciplinaire RTP-DOC s'est intéressée aux études lancées par Tim Berners-Lee dont la « *volonté de développer un « Web sémantique » joue un rôle comparable à celui de Paul Otlet un siècle plus tôt* ». (Pedauque, 2007 : 16) Il s'agit de la création de nouveaux modèles de classification des connaissances, les « ontologies ». Ce travail de conceptualisation, établissant des relations entre des termes, prolonge les systèmes de classifications antérieures mais avec une orientation usager plutôt qu'une orientation système. L'ambition est de pouvoir libérer les internautes « *d'une bonne partie de leurs tâches de recherche, de construction et de combinaison des résultats ainsi que d'exploitation des contenus des ressources grâce aux capacités accrues des machines à accéder à ces contenus* et à effectuer des raisonnements* sur ceux-ci* » (Pedauque, 2007 : 100). Nous retrouvons ici le débat ancien sur les capacités « cognitives » des machines et autres robots et ces recherches ne font que poursuivre le « *rêve cybernétique* » que les créateurs de machines espèrent réaliser en « *demand[ant] à l'ordinateur de s'attaquer à ce qu'il y a de plus spécifiquement humain dans les performances de notre cerveau (...)* » (Escarpit, 1976 : 76).

Les chercheur-es de l'équipe RTP-DOC rappellent les diverses réalisations de

37 Système de recherche d'information

systèmes de représentation de l'organisation des connaissances, permettant l'exploitation de corpus documentaires, depuis les « *taxinomies utilisées en sciences naturelles au XVII^e et XVIII^e siècles en passant par la classification décimale de Dewey (1876), les réseaux sémantiques de Quilian (1968) ou les ontologies de l'ingénierie des connaissances (Gruber, 1993)* », l'informatisation ayant permis, selon l'équipe, une « *accélération dans la réflexion sur la terminologie* » (Pedauque, 2007 : 132). L'accès direct à l'information sur l'internet fait que le travail de documentation est désormais perçu comme un « *parasitage lorsqu'il vient s'intercaler entre l'usager et le corpus informatisé* » (Pedauque, 2007 : 134). Cependant les auteur-es soulignent que « *l'expérience de l'utilisation réelle d'internet a montré les limites d'un accès médiatisé par un intermédiaire commercial* », Google notamment, dont les visées sont lucratives (Pedauque, 2007 : 135). Il s'agit donc de tendre vers une normalisation afin de faciliter l'échange des connaissances mais en « *restant proche de l'usage* » car « *normaliser, c'est imposer une vision du monde* » (Pedauque, 2007 : 135). Ce souci d'« objectivité » est une préoccupation constante pour l'équipe qui laisse supposer ainsi que la machine offre une « neutralité » permettant de limiter la subjectivité intrinsèque à toute forme de réalisation humaine, « oubliant » de ce fait que les programmations sont toujours réalisées par des hommes et des femmes...

Ainsi B. Simonnot réfute la prétendue neutralité du « Web » et spécifiquement des moteurs de recherche commerciaux, qui par leur mode de fonctionnement et l'agencement des résultats (liés, entre autres, aux « *préférences* » manifestées à leur insu ou non par les internautes (Simonnot, 2012 : 109)) produisent un « *message en tant que tel* » correspondant de fait à « *une ligne éditoriale* » à l'instar de celle établie mais de façon plus visible et plus transparente par les maisons d'édition (Simonnot, 2012 : 139). La transparence en ligne est donc tout autant sujette à caution que la neutralité, dans la mesure où elle est la « *qualité d'informer de manière claire et complète sur son fonctionnement et ses pratiques* » (Simonnot, 2012 : 144), qualité loin d'être appliquée par les moteurs de recherches commerciaux, dont le principal, Google, est l'un des trois dispositifs correspondant au « *lu* » de la règle des trois participes de J-M. Salaün (Salaün, 2012). B. Simonnot constate que « *savoir sélectionner les documents, les trier, les classer de manière à les retrouver [sont] autant de compétences informationnelles qui sont peut être les moins partagées. Et le partage en ligne ne résout pas ce problème* » (Simonnot, 2012 : 162). Si la notion de « *notoriété des sources et des auteurs* » vient à remplacer celle de preuve ou d'autorité, c'est « *faute de réelles connaissances ou dans*

l'impossibilité d'analyser finement le contenu intrinsèque par des raisonnements » (Simonnot, 2012 : 181). Ainsi les dispositifs d'accès à l'information « *ne participent pas à donner du sens aux documents qu'ils proposent et ils n'en respectent pas toujours l'intégrité dans la présentation qu'ils en font* » (Simonnot, 2012 : 201) De plus, ils ont pour objectif premier de récolter des données sur les internautes pour pouvoir en faire commerce et, en contraignant les requérants à se conformer aux « *contraintes réelles ou supposées* » imposées pour la formulation de leurs recherches, entament « *une tentative de clôture de l'espace informationnel de l'internaute* » (Simonnot, 2012 : 210). Procédant de manière opaque, les moteurs de recherche commerciaux encouragent « *une déresponsabilisation dans les pratiques* » (Simonnot, 2012 : 223). Elle plaide donc pour « *une formation solide du plus grand nombre d'individus qui vont être sollicités davantage pour des pratiques informationnelles qui ne s'arrêtent pas à l'accès* ». (Simonnot, 2012 : 225) Pour déjouer ces plans liés aux intérêts économiques, R. Casati invite, lui aussi, à « *former des citoyens technologiquement responsables* » (Casati, 2013 : 194) et à refuser toute « *normativité automatique* ». (Casati, 2013 : 199)

Certain-es chercheur-es de l'équipe du CNRS s'appuient pourtant sur un modèle, qui « *peut contenir des caractéristiques sur les connaissances (ce que l'utilisateur connaît ou ce qu'un groupe connaît), l'expérience (savoir-faire) et les compétences, les préférences (par exemple si un utilisateur préfère un siège de couloir dans un avion), les objectifs, les centres d'intérêts, etc.* », pour montrer la spécificité des documents numériques (Pedauque, 2007 : 111). Le « *document virtuel (...) possède les 2 caractéristiques suivantes : a) il est construit via la réutilisation de différentes sources d'information et b), il est généré dynamiquement à la demande en fonction de l'utilisateur (Milosavljevic et al., 2001). On parlera de document virtuel personnalisable (DVP)* » (Pedauque, 2007 : 112). La personnalisation est assortie de l'adaptabilité, c'est-à-dire que les réseaux devancent les besoins en mettant à disposition de nouveaux services : « *les services Web sémantiques* ». La définition « *usuelle* » d'un service est qu'il « *permet à un utilisateur, non seulement d'obtenir de l'information, mais aussi d'effectuer des changements sur l'état du monde. Le commerce électronique est, bien-sûr, un exemple privilégié de ces approches, mais l'exposition de tout processus fonctionnel sur le Web, comme la souscription d'une police d'assurance, la réservation de voyages en ligne, les*

services bancaires ou des fonctions entières d'une entreprise comme la mise en œuvre de la gestion de la chaîne logistique, relèvent également de cette problématique. Cette notion de services est aujourd'hui au cœur des stratégies des principaux acteurs du monde du logiciel » (Pedauque, 2007 : 112). Nous mesurons, à nouveau, à travers ses précisions, le caractère corporatif qui explique le souci d'applicabilité guidant les chercheur-es³⁸.

Ces deux caractéristiques, virtualité et personnalisation, sont constitutives du concept théorisé par Meyriat pour qui le récepteur produit le document en activant l'information qu'il supporte. Avant celle-ci le document est « virtuel » au sens propre du terme : il ne devient réel que sous les yeux de celui ou celle qui y cherche une réponse. Tandis que la notion de virtualité évoquée plus haut est en fait une référence à la possibilité technique offerte par l'outil numérique, et « virtuel » est synonyme, dans cette acception, de « numérique ». D'autre part, le lien avec la personne est d'ordre cognitif dans la conception de Meyriat, car c'est l'individu qui érige un objet, quel qu'il soit, en document, en attribuant un sens à l'information qu'il porte. Alors que la « personnalisation » est, dans l'approche technique, la génération dynamique à la demande, ce qui signifie prosaïquement l'activation de liens mais l'on ne se pose pas la question de savoir s'il y a attribution de sens chez le récepteur.

Le résultat des recherches pluridisciplinaires sur « *la construction de RTO³⁹ à partir de corpus (...) s'inscrit dans des perspectives où les besoins, les modes d'utilisation, les modes de rédaction sont situés, c'est-à-dire anticipés (...) »* et « *l'accélération numérique dont on ne connaît pas encore les conséquences n'affectera pas les acquis de cette interdisciplinarité et les perspectives qu'elle a permis de dégager »* (Pedauque, 2007 : 145). L'orientation de ces recherches cependant apparaît idéologique puisqu'elle considère qu'il sera possible de prévoir ce que les requérants recherchent et veulent savoir, refusant par là l'imprévisibilité ontologique et ignorant que « *c'est une duperie d'établir une frontière entre le canal et ce qui vient avant et après lui (...) ce qui fait l'originalité irremplaçable et le prix inestimable de la pensée humaine, c'est le pouvoir de non-pertinence, l'imprévisibilité vraie de l'énonciation et non plus l'imprévisibilité domptée à l'intérieur d'un système probabiliste »* (Escarpit, 1976 : 41-42). La notion de prévisibilité sous-jacente à celle d'anticipation nous renvoie, en effet, à nouveau à la théorie mathématique dont toute l'édification repose sur les probabilités.

38 Les travaux effectués répondant de fait à des appels à projet d'entreprises

39 Ressources terminologiques ou ontologiques

La difficulté d'un point de vue technique pour les développeurs du « *Web sémantique* » est d'intégrer « *la diversité des sources d'information distribuées et leur hétérogénéité* » (Pedauque, 2007 : 108). L'enjeu étant de pouvoir traiter plusieurs « *sources d'information* », « *base de données rationnelle* » (par exemple une « *liste de films* »), base contenant « *des données représentées en XML* (par exemple « *pour chaque film, les salles (nom, adresse) où le film peut être vu* ») afin de répondre au mieux à une requête (Pedauque, 2007 : 109). Se posent donc des questions de compatibilité des formats ou structures utilisés et des questions « *sémantiques* » en lien avec les ontologies. L'idée est d'homogénéiser l'ensemble car il faut « *s'attendre à une explosion du nombre d'ontologies utilisées* » (Pedauque, 2007 : 110). Cette « *explosion* » redoutée est finalement un retour au phénomène entropique que les chercheur-es veulent endiguer à tout prix (au détriment de la charge significative de l'information). Un des atouts principaux des documents en ligne, d'après les chercheurs et comme nous l'avons évoqué précédemment, est de permettre la personnalisation des réponses car « *tous les utilisateurs ne sont pas intéressés par les mêmes informations, n'ont pas les mêmes attentes, connaissances, compétences, centres d'intérêts et cherchent à satisfaire leurs besoins le plus facilement possible* » (Pedauque, 2007 : 111). Cet état de fait n'est en réalité pas nouveau : avant l'ère dite numérique, les attentes de tout un chacun n'était pas les mêmes mais la technologie numérique serait à même d'opérer une sélection inédite de « *l'information* ». L'enjeu face « *aux problèmes de données surabondantes* » est de « *guider l'accès à l'information en fonction des besoins* » afin « *d'effectuer un filtrage des ressources* ». (Pedauque, 2007 : 111). Les traces laissées par les recherches et les mouvements effectués par l'internaute auparavant sont un moyen de réaliser cette personnalisation. Or, cette caractéristique suppose un choix opéré par la machine (les robots « *intelligents* ») à la place du chercheur d'information, une sélection parmi les « *informations* » disponibles, de celles jugées pertinentes parce qu'elles correspondraient au profil numérique du requérant. Le risque pointé par R. Casati est de fermer l'horizon des possibles puisque les moteurs, tel Google, « *[analysent] vos recherches passées pour construire votre moi en ligne, et vous proposer des résultats dont on prévoit qu'ils seront pertinents pour vous* », restreignant ainsi toute sérendipité, notion qui s'est développée aux premiers temps du « *Web* » quand la pratique du « *surf* », de la navigation au hasard, était courante (Casati, 2013 : 177). Désormais, au contraire, les données reçues apportent confirmation plutôt que vérification et les pratiques ne laissent pas place au recoupage, à l'élargissement (Casati, 2013 : 178). Situation paradoxale en

un lieu où il y a des données à foison, où le gisement d'informations potentielles est croissant.

2.2.3.2 La fonction documentaire : le document, un prestataire de service ?

Au sujet de la médiation réalisée par le truchement du document, les chercheur-es avancent l'hypothèse d'un changement radical du rapport à la connaissance (au niveau de celui qui a marqué le passage du *volumen* au livre) résultant, là encore, de la « dissociation entre support et contenu induite par le passage au numérique » (Pedauque, 2007 : 167). Selon l'équipe du CNRS, les nouveaux usages du document ne sont pas liés à une évolution de « la nature des informations transportées » mais à « la diversification des supports et des techniques d'édition » car « chaque acteur du réseau de circulation de l'information est à la fois producteur et consommateur de contenus numériques et non plus seulement de documents » (Pedauque, 2007 : 168). Cette remarque concerne non pas seulement le document numérique mais bien le document qui se situe « sur le réseau ». Ce dernier n'est « plus un objet pérenne » mais un « conteneur de données » ou un « actant prenant en charge un contenu » (Pedauque, 2007 : 173). Dans cette conception, le document est une entité autonome, interchangeable, un service. À nouveau, l'action humaine indispensable au regard de l'approche des théoriciens fondateurs, dans la mesure où le lien fondamental est celui qui s'effectue avec l'information supportée, n'entre pas en ligne de compte.

Ce lien avec les individus est présent quand les chercheur-es évoquent la dimension dynamique du document numérique qui en fait « une action dans le processus relationnel... ». Mais plutôt que de lui assigner la place d'intermédiaire, de moyen matériel, ils lui attribuent un rôle actif par lui-même. L'« action » en question ne se réduit-elle pas pourtant dans cette conception à la manipulation de la souris ? L'équipe souligne « le rôle prédominant de l'usage du document » et voit en lui « un objet doté d'un statut social » (Pedauque, 2007 : 177). Ne faisant pas le lien systématique avec la notion fondamentale d'information, les chercheur-es caractérisent le document dans ce cas comme étant une « inscription sur un support informatique se présentant sous forme de fichiers ou d'un ensemble de fichiers liés » (Pedauque, 2007 : 184). Ce qui est inscrit en l'occurrence n'est pas précisé. Sur l'internet, « les documents perdent leurs auteurs, les auteurs perdent leurs documents » (Pedauque, 2007 : 193). Ainsi la difficulté est de parvenir à apparier un document avec celui qui l'a émis. Une

ambiguïté s'installe entre les notions de document et d'information (cette dernière étant, au fil des travaux présentés, très peu nommée) car la difficulté posée par les objets en ligne consiste à retrouver la paternité du contenu par lui-même et non du support.

Ainsi, dans le domaine de la publication, « *l'édition ne repose plus simplement sur un modèle de choix a priori par un groupe-filtre mais aussi sur un modèle de validation a posteriori ou une publication distribuée (Chartron & Rebillard, 2004)* » (Pedauque, 2007 : 197). Cette problématique de « *l'auctoritativité (...) propension pour les auteurs à s'affirmer auteur en dehors des autorités établies* » provient de la possibilité nouvelle induite par les pratiques et usages sur la toile. La « *reproductibilité à l'identique des documents, caractéristique héritée de la culture de l'imprimerie fait place dans l'univers numérique aux spécifications nécessaires à la restitution des manifestations physiques des documents : ce sont les conditions de lisibilité et de structure temporelle des documents qui seront décrites pour s'adapter aux appareils de lecture* » (Pedauque, 2007 : 198). RT Pedauque souligne que « *l'univers du web se compose d'une myriade d'outils remodelant les rôles de la chaîne éditoriale, transformant la relation auteur/lecteur et faisant apparaître l'étroite dépendance des conditions de légitimation des informations avec la façon dont elles circulent* » (Pedauque, 2007 : 199). Considérer la toile comme un catalyseur de cette relation entre information et modalité d'évaluation de celle-ci peut s'envisager mais de là à estimer que c'est la toile qui la révèle, le pas est si profond, que nous nous garderons de le franchir. Ces questions de légitimation et de validation sont bien antérieures à ce format et il apparaît réducteur de penser qu'elles se réglèrent facilement avec les autres formes de médias. Le choix opéré par l'éditeur qui a publié *L'effroyable imposture*⁴⁰, sous format papier, soulève des questions intellectuelles et éthiques qui n'ont rien à envier à celles auxquelles nous devons faire face dans l'édition ou l'auto-édition en ligne.

Plusieurs modèles éditoriaux, non encore stabilisés, en effet, émergent sur le Web. Le phénomène de l'auto-publication liée à une faible visibilité des travaux individuels ouvre l'espace à une sélection de contenus relevant d'une évaluation collective. Cette évolution est à mettre cette fois encore au crédit de « *l'Internet [qui] donne un coup d'accélérateur à la dissociation entre contenus et supports* » (Pedauque, 2007 : 206). Cette prétendue dissociation est mise en avant pour remettre en cause les caractéristiques du document. Il s'agit de facto de

40 Ouvrage de Thierry Meyssian dans lequel il développait la théorie du complot à propos des attentats du 11 septembre 2001

l'affranchissement du format papier et « *donc des contraintes physiques de la reproduction en de multiples exemplaires [qui] s'apparente à un basculement vers le domaine des services* » mais la question de « *la sélection intellectuelle et [de] la mise en visibilité des œuvres demeure un problème crucial* » (Pedauque, 2007 : 207), dans le domaine littéraire notamment. L'équipe du CNRS ne va pas plus loin dans ses investigations à ce sujet et il est permis de se demander, à la lumière d'une conception sociale de l'information, si ces obstacles ne seraient pas dus à l'appréhension par trop partielle du rôle de l'information dans le processus de communication.

J-M. Salaün, s'est essayé à formaliser les nouvelles formes d'éditions qui se profilent sur la toile. S'appuyant sur les travaux issus des réflexions de l'équipe RTP-DOC du CNRS qu'il a dirigée, il théorise le document numérique à travers un comparatif entre la bibliothèque et le « *Web* », ce dernier figurant une « *nouvelle étape* » dans l'histoire de la bibliothèque (Salaün, 2012 : 31). Il reconnaît qu'il n'y a « *pas de consensus actuel autour d'une théorie opérationnelle* » (Salaün, 2012 : 33). Retournant sur le terrain des sciences sociales⁴¹ et reprenant les définitions établies par Otlet et Briet, il estime que le « *document désormais numérique n'est plus qu'un ensemble de signaux dans un vaste flux (...) un ensemble de pièces éparses réagencées à la volée* » (Salaün, 2012 : 48). « *Les données sont les plus petites unités documentaires représentant le monde réel. Ces unités documentaires sont « données », c'est à dire prises comme telles, non discutées, objectives en soi, neutres* » (Salaün, 2012 : 88), ce qui correspond à la définition canonique de l'information du point de vue de la théorie mathématique.

Renversant la relation originelle entre l'individu et le document (le document permettant d'extérioriser la pensée humaine), J-M. Salaün utilise la célèbre formule rimbaldisienne pour l'appliquer à l'univers documentaire dans lequel il s'agit de gérer notre « double » numérique : « *« Je » est un document* » (Salaün, 2012 : 89). L'individu documenté se fait le « *relais de la circulation de l'information* » (Salaün, 2012 : 122). Ainsi l'individu devient le média dans le processus de communication. Le lien souligné par Meyriat, et par Y. Jeanneret après lui, entre information et connaissance, est ignoré. Or, l'attribution de sens, étant acte de connaissance, effectuée par l'individu sur lui-même en fait plus qu'un simple passeur de données.

41 Nous avons noté plus haut que les recherches du RTP-DOC étaient axées prioritairement vers les sciences de l'ingénieur

Les buts de la documentation selon Otlet sont caduques et en lieu et place de la véracité, de la preuve et des faits, l'heure est à « *l'efficacité* », « *au croisement des opinions* » et au « *consensus* » (Salaün, 2012 : 87). Ainsi dans une approche résolument marketing, il redéfinit le document. « *Le document (...) n'est pas un bien stable mais un service labile* » (Salaün, 2012 : 91) (...) *le marché de biens a détaché le document de son support (...) l'accès est devenu un enjeu commercial* » (Salaün, 2012 : 104). Le document est « *une promesse* », un « *bien d'expérience* » dont la valeur varie en fonction de ce que l'acquéreur y trouvera car « *lire un document en épuise la valeur pour son consommateur* » (Salaün, 2012 : 116). S'opposant en cela à la conception de Meyriat, pour qui « *la capacité informative d'un document n'est jamais épuisée par les utilisations déjà faites d'informations qu'il contient* » (Meyriat, 1978 in Couzinet : 117). L'information n'est pas un bien comme un autre car le fait de transmettre une information n'en dépossède pas celui qui la communique. Or, dans l'approche commerciale opérée par J-M. Salaün, les trois dimensions du document numérique correspondent chacune à trois marchés dominés par trois groupes sur le Web. Ainsi le « *néo-document* », ou document numérique, sert les stratégies commerciales des composantes de ce que l'on pourrait appeler le « Big Three » (Google, Apple, Facebook)⁴² : Google se chargeant du contenu « *lu* », Apple du design « *vu* » et Facebook de la transmission « *su* », sachant que la séparation entre ces trois attributs tend à s'effacer car les intérêts économiques des firmes les poussent à investir les domaines de leurs concurrents (Salaün, 2012 : 143). La formule « *vu* », « *lu* », « *su* » s'applique de plus parfaitement au sein de chaque firme. Ainsi Facebook relie un ensemble de « *documents* » que sont les abonnés : l'identité (signe de reconnaissance, papier d'identité) correspond au « *vu* », les différentes caractéristiques postées (le contenu) sont le « *lu* » et les relations (les amis regroupés autour de centres d'intérêts communs) représentent le « *su* » (Salaün, 2012 : 140). L'auteur concède toutefois que les « *techniques du Web documentaire [servant] à indexer les individus* » pose des questions éthiques et (mises sur le même plan) de rentabilité économique (Salaün, 2012 : 143). L'enjeu pour les professionnels de l'information est donc, selon J-M. Salaün, de se former tant en informatique qu'en Sciences de l'information afin que ceux qu'il baptise les « *architectes de l'information* » sur l'internet soient capables de gérer les « *néo-documents* », les documents numériques (Salaün, 2012 : 145-146).

42 Par référence à ce qu'on appelle le « Big Four », ou encore GAFA (regroupement des trois entreprises et d'Amazon, firmes dominantes sur le Web commercial) : le fait de s'en tenir à « GAF » tient-il aux impératifs liés la « triade » qui n'accepte que trois référents ?...

Cette nécessité de tenir compte de l'environnement technique n'est pas nouvelle, les professionnels de l'information, théoriciens et/ou praticiens (les deux étant souvent liés), ont toujours été attentifs aux évolutions technologiques (dès Otlet) et les implications du format numérique n'ont pas échappé aux chercheurs contemporains. Mais en gardant une vue générale, et non partielle (technique, économique voire commerciale), comme nous allons le voir dans la partie suivante, les caractéristiques originelles du concept document présentent de fait une stabilité certaine à l'ère numérique.

2.3 Le document numérique, révélateur de caractéristiques originelles du concept document

Nous venons de voir dans les parties précédentes que si les chercheurs s'accordent à trouver dans le document numérique des spécificités, il n'y a pas consensus au sujet de la nature de celles-ci au sein des SIC. L'approche technique, privilégiée par les uns, est rejetée par les autres qui considèrent que cette dimension n'est pas la condition principale de la réinterrogation du concept. Nous allons aborder dans cette dernière partie théorique les travaux qui, à l'instar des théoriciens fondateurs du document, replacent de façon systématique l'information, notion fondamentale et supérieure d'un point de vue hiérarchique, au cœur de leurs réflexions sur le document numérique et s'appuient sur les bases théoriques toujours fondées de nos jours.

2.3.1 Le document, objet informationnel avant d'être numérique

Le recours aux bases conceptuelles pour repenser le document à l'heure numérique est d'autant plus justifié, selon V. Couzinet, C. Courbières et G. Regimbeau, en raison de l'« *anticipation terminologique et encore étonnante* » des précurseurs. Ainsi Otlet, dans les années 30, « *en vient à formuler la notion d'hyperdocumentation* », cette étape venant clore le cycle de la documentation (Couzinet, 2001 : 469). À ce stade ultime, « *le document type de l'avenir devra donc comporter en lui toutes les possibilités héritées, jointes à celles qu'il sera possible d'acquérir encore ; il sera une synthèse des caractéristiques acquises les plus efficaces, jointes aux finalités qui doivent lui être assignées* » (Otlet, 1934 : 429). Briet, quant à elle, tout en développant l'idée d'extension du document jusqu'à inclure dans cette catégorie les animaux vivants, avait prévu, dès 1951, le morcellement allant croissant des éléments permettant de réaliser des études et d'accéder à la connaissance. En avançant les notions d'« *unité documentaire* », d'« *idée élémentaire* » et d'« *unité de pensée* » (Briet, 1951 : 10), elle exprime le « *processus qui a ouvert le passage d'abord au professionnel du document, puis au professionnel de l'information : de l'unité du support à l'unité informationnelle* » (Couzinet, 2001 : 471).

En replaçant le concept document dans son rapport « *subordonné* » à l'information-connaissance, il semble que la distance entre technique et science soit plus nette et que l'étude

épistémologique puisse s'amorcer. Les interrogations dans cette optique ne sont pas celles qui préoccupent les ingénieurs et autres informaticiens (compatibilité de formats, durée de vie, conservation, etc.) mais celles relatives au contenu, au sens, à la validité et in fine à l'élaboration des connaissances. Rappelant les réserves de Meyriat à l'égard du terme, V. Couzinet met en exergue l'importance qu'il accordait avant tout à l'information, désignant « *l'information comme l'objet spécifique de la documentation et [plaçant] le document en position secondaire* » (Couzinet, 2001 : 472). Il précise ainsi que « *le document est défini comme un objet matériel support d'information : ce support est subordonné à l'information, il peut avoir des formes et même des natures différentes (support papier, support magnétique, optique, etc.), il peut être transformé sans que l'information soit modifiée pour autant, il peut être éliminé après que l'information ait été transférée sur un autre support : c'est elle qui est important, c'est elle d'ailleurs qui intéresse l'utilisateur* » (Meyriat, 1994 in Couzinet, 2001 : 396).

Y. Jeanneret insiste sur cette « *solidarité lexico-sémantique* » entre les deux notions (Gardiès, 2012 : 51) car « *ce qui constitue l'information, c'est l'interprétation d'un document, en fonction de sa forme matérielle, par un sujet qui le perçoit et en comprend le sens, en vertu de la culture des formes sur laquelle repose toute légitimation et toute transmission culturelles* » (Jeanneret, 2000 : 80).

À propos du caractère hypertextuel du document numérique qui focalise l'attention des chercheurs (avec la particularité d'activer ses liens en interconnexion), Y. Jeanneret et J. Davallon mettent en garde contre des raccourcis trompeurs dus aux « *formes complexes dont est fait tout écrit d'écran [qui] sont écrasées pour être finalement réduites à une combinaison graphique de parcours. Le reste [étant] purement et simplement négligé* » (Davallon, 2004 : 47). Or c'est ce résidu qui est crucial comme ils le précisent en insistant sur la nature du « lien » qui « *n'est pas censé relier des objets, mais une réalité plus immatérielle*, l'information, qui occupe la place des points d'intersection des arcs* dans le graphe* » (Davallon, 2004 : 47). Ils démontrent que la transformation de l'information en donnée numérique ne se fait qu'au prix d'une réduction sémantique dans la mesure où elle « *engage une coïncidence entre contenu et expression, qui place en équivalence des séries d'« unités » sémantiques isolées, interchangeable et insécables* » (Davallon, 2004 : 47). Dans leur

réflexion sur le « document numérique », à travers leur approche hypertextuelle, les chercheurs s'interrogent sur le concept supérieur, l'information. Ainsi ils constatent que « *le texte informatisé n'est pas étudié comme un objet à manipuler mais comme une structure régie par le code* » (Davallon, 2004 : 49). Cette conception ne tient pas compte du contexte, ne laisse pas place à l'intention et à la diversité d'interprétations liées à la variété des individus engagés dans le processus de communication. C'est précisément pour s'écarter de cette vision « informaticienne » qu'ils proposent de nommer « *signe passeur* » ce que d'autres appellent « *lien hypertexte* » (Davallon, 2004 : 50). Ainsi ils mettent l'accent sur ce qui, du point de vue des SIC, est fondamental : le sens ; et mettent de côté l'aspect technique, informatique qui permet l'activation d'un mot, ou d'un ensemble de mots par l'intermédiaire de la souris. Par cette appellation, ils entendent « *prendre au sérieux ce que signifie une nouvelle forme de lecture gestualisée, refuser l'isolement artificiel des signes pour les intégrer à une construction, à un contexte, indépendamment desquels ils n'ont aucun sens* » (Davallon, 2004 : 50). Ils s'inscrivent en faux contre la « *conception (...) euphorisant[e] qui fait du réseau un support de circulation et de l'usager, l'opérateur qui assure à l'ensemble sa dynamique* » (Davallon, 2004 : 51). Tout ce que l'internaute met en œuvre (activation des « *signes passeurs* », notes produites, copiés-collés, etc.) constitue en soi un « *espace répondant à une organisation textuelle* » (Davallon, 2004 : 53). Dans leur volonté de mettre la signification au cœur de leur approche, les chercheurs refusent une conception désincarnée et interchangeable de ce qui se joue sur les réseaux. La situation dans laquelle se trouve l'usager est primordiale : il y a d'une part le rapport entre ce qui apparaît à l'écran et le code informatique qui permet l'activation, il y a d'autre part la réalité des espaces qui sont reliés par ce « *signe passeur* » et il y a surtout le « *programme d'activité* » dans lequel est engagé l'internaute. Ces différents contextes sont essentiels pour que ce qui est lu produise du sens. Les auteurs proposent de nommer ce document particulier au triple contexte, le « *textiel* » (Davallon, 2004 : 54).

2.3.2 La double partition

Après avoir posé les bases de la relation hiérarchique entre le concept information et le concept document, Meyriat introduit une catégorisation essentielle, peu reprise par les chercheur-es en SIC ; celle des « *documents par intention* » et « *documents par attribution* » (Meyriat, 1978, in Couzinet, 2001 : 115-116). V. Couzinet a creusé cette notion, nous le verrons plus loin, au regard des méthodes de recherches des historiens. Deux autres domaines reposent également sur l'étude de documents : l'ethnologie et la muséologie.

L'enquêteur ethnologue est un « *chercheur de documents humains* » (Griaule, 1975). Les gestes, les postures, la communication non verbale exprimés par les êtres humains observés par l'ethnologue lui procurent des renseignements et sont à ce titre des « documents » (Couzinet, 2001 : 480). Mais contrairement à l'approche de J-M. Salaün, qui fait de l'internaute un document fait homme, c'est ici parce que l'individu est l'objet d'une étude, donc dans le sens de l'attribution et non dans le sens intentionnel. Le muséologue illustre quotidiennement par ces choix l'extensivité de la notion et l'attention particulière accordée aux objets jusqu'aux plus monumentaux. La notion s'élargit et sort des murs des musées en valorisant les éléments du patrimoine que sont les paysages remarquables, les monuments et en documentant des formes d'art, non stabilisées par essence puisque volontairement éphémères, telles les œuvres du Land art, par le biais de photographies ou de films (Couzinet, 2001 : 485). Il s'agit là encore d'après V. Couzinet d'une approche selon l'angle attributif puisque ce n'est pas seulement la réalité matérielle, créée intentionnellement pour instruire qui intéresse le muséologue mais le sens donné à l'objet de son observation, embrassant de ce fait l'univers entier.

Les chercheur-es invoquent parallèlement une autre partition élaborée par Meyriat, et par Briet avant lui, entre document primaire et document secondaire (document sur le document). Le document primaire est toujours à l'origine du document secondaire. Dans le cas de la fameuse antilope de Briet, le fait de placer l'animal dans un zoo en fait un document. Les articles sur cet animal sont des documents secondaires : « *l'antilope cataloguée est un document initial et les autres documents sont des documents seconds ou dérivés* » (Briet, 1951 : 8). Ceux-ci à leur tour peuvent « *passer dans la catégorie des documents primaires en raison, par exemple, de leur valeur historique* » (Couzinet, 2001 :

484) desquels peuvent à leur tour émaner des documents secondaires (fiches signalétiques, résumés, autres), etc. V. Couzinet précise que « *le décalage des degrés s'établit selon le contexte de production et d'utilisation de l'information* » (Couzinet, 2001 : 484). La définition du document se fait toujours, et quelle que soit l'échelle de précision, à l'aune de l'information. Au sujet des œuvres artistiques érigées en documents, en soulignant « *le paradoxe du musée dont la mission est de représenter des pièces, qui à l'origine, ne lui sont pas destinées* » (Couzinet, 2001 : 485), le propre de l'art étant de porter un regard critique sur la société dont les musées sont précisément le reflet, V. Couzinet aborde la notion de temporalité dans la mesure où « *en conférant intentionnellement la qualité de document au produit artistique, on a mis à l'écart les notions d'œuvre et de beauté intemporelle pour faire place à une temporalité sociale revendiquée* » (Couzinet, 2001 : 484). Les choix et construction opérés par les conservateurs de musée, les ethnologues ou encore les historiens en attribuant à des objets une certaine valeur documentaire constituent ce que V. Couzinet nomme la « *société documentaire* » à travers laquelle émergent les « *fonctions réelles et symboliques* » du document (Couzinet, 2001 : 486).

Les auteur-es constatent que la distinction faite par Meyriat entre le « document par intention » et le « document par attribution » (...) est *particulièrement opératoire pour rendre compte du caractère spécifique des documents sur support électronique* » (Couzinet, 2001 : 488). Le document hypertexte exacerbe la séparation entre l'intention d'un émetteur, d'un producteur et l'attribution, la (re)-construction par un récepteur. Les modalités de lecture et d'écriture propres à l'environnement numérique sont singulières car l'organisation spatiale sollicite l'œil dans des directions différentes et parce qu'une « *des règles de bases (...) sur le Web est de faire court* » (Couzinet, 2001 : 489). Le document numérique nécessite de la part du récepteur une activité accrue (ne serait-ce que technique) et participe de ce fait par « *sa structure même* » à faciliter « *la prise d'information* » (Couzinet, 2001 : 490). Car les auteur-es soulignent la condition essentielle de « *l'activation à réception* » qui définit le document selon ses théoriciens pionniers, ainsi selon Estivals : « *dans la chaîne primaire de création tant que le document est créé mais n'est pas encore utilisé par un récepteur, il reste un document virtuel. Dès lors qu'il est employé et décodé par un récepteur, le document devient réel : il a trouvé son utilisateur* » (in Couzinet, 2001 : 490).

Le problème particulier soulevé, d'un point de vue scientifique et conceptuel, par les chercheur-es dans le cadre de la réflexion sur les évolutions du document ne concerne pas tant le document numérique par lui-même - « *l'introduction de l'information sur supports électroniques n'a pas véritablement changé la donne quand il s'agit de documents stabilisés, gravés* » (Couzinet, 2001 : 491) - que le document sur la toile. Il s'agit des questions liées aux « *informations en ligne dont la nature est relativement volatile et transformable* » (Couzinet, 2001 : 491). Les auteur-es relèvent deux conséquences à cette « *instabilité de l'information* » : la possibilité pour l'émetteur de « *revenir sur son contenu comme « on revient sur sa parole* » (...) *rétractation d'autant plus facile qu'elle n'est plus appuyée sur la preuve tangible du document* » (Couzinet, 2001 : 491). Et son corollaire, le fait de « *prêter des propos à un auteur sans que ce dernier puisse réellement prouver qu'ils n'ont pas été tenus puisqu'on sait que techniquement il lui était loisible de les transformer* » (Couzinet, 2001 : 491).

Pour les documents en ligne, se pose aussi le problème du référencement : en effet « *la référence d'une information en ligne correspond à la mention d'un chemin pour y accéder à un instant donné mais n'est plus la garantie d'une preuve matérielle* » (Couzinet, 2001 : 492). L'accès aléatoire à l'information se retrouve dans le phénomène des liens brisés. L'archivage est un moyen de pallier ces risques de pertes informationnelles mais génère des obstacles. Il s'agit d'effectuer une « *documentarisation* » de l'information » (Couzinet, 2001 : 492), c'est-à-dire de constituer des documents là où il n'y en avait pas pour se préserver du risque de ne pas retrouver l'information. Cependant, se pose alors la question du statut de ces nouveaux documents : « *il faudra s'en remettre à des substituts qui serviront de témoin, comme on pourrait le concevoir avec une bande magnétique conservant des paroles émises à la radio. La capture d'écran, la mise en mémoire de fichiers, le tirage sur imprimante peuvent constituer des doubles de l'information* ». Les chercheur-es s'interrogent alors sur la valeur de ce type de référence, à savoir, s'il sera « *possible pour autant de proposer une référence portant sur un fac-similé et non plus sur la source réelle* » (Couzinet, 2001 : 493).

Nous en revenons à la graduation opérée entre les documents, notamment ceux qu'Escarpit nommait « *les documents-relais* » (Escarpit, 1976 : 158). La règle admise consistant à indiquer « *la « date de visite* » (...) *prépare le récepteur à l'éventualité d'une*

source indisponible ou transformée » (Couzinet, 2001 : 493). Dans le cadre de la « documentarisation », le document vers lequel on peut se tourner pour vérification « *met en avant le simulacre au détriment de l'original* » (Couzinet, 2001 : 493). Les modifications et transformations successives des documents en ligne réinterrogent par conséquent « *le contrat de confiance* » entre auteurs et récepteurs » (Couzinet, 2001 : 493).

Les principes de construction et d'évaluation des documents scientifiques illustrent bien l'existence de ce contrat implicite dans la mesure où « *les contrôles effectués par le comité de rédaction et par le comité de lecture garantissent le respect des contraintes en même temps qu'ils garantissent le niveau des contenus* » (Couzinet, 2001 : 496) or « *la multiplication possible des évaluateurs [et] les conséquences que cette multiplication induit sur les pratiques scientifiques et plus généralement sur la construction des connaissances* » modifient l'approche du concept (Couzinet, 2001 : 495). Nous pouvons remarquer à ce sujet qu'un retour à ces formes de validation s'est opéré progressivement sur la toile, un site collaboratif comme Wikipedia ayant un comité de relecture qui se charge de valider ou non les articles proposés. La transparence n'est pas pour autant au rendez-vous car les membres de ce comité et son fonctionnement ne sont pas présentés de façon claire aux utilisateurs. Les « *attentes* » réciproques entre les auteurs, les destinataires et les lecteurs ne sont plus du même ordre et V. Couzinet fait le parallèle avec les attentes mutuelles entre les émetteurs et les récepteurs du document (Couzinet, 2001 : 496). Le document qui est le moyen grâce auquel la communication de l'information se réalise participe du contrat tacite, car « *la notion de communication porte en elle la nécessité de l'accord, ou du pacte implicite, qui permet l'établissement de l'échange* » (Couzinet, 2001 : 497). Ce moyen n'est pas anodin puisqu'il « *contraint le contenu et ses usages* » (Couzinet, 2001 : 497). Sa construction d'une part, et l'utilisation qui en est faite d'autre part relèvent de codes mis en place et admis par la société. Le document, « *produit par intention ou ayant fonction attribuée, (...) est aussi le résultat d'une construction sociale* » (Couzinet, 2001 : 499). L'enjeu concernant le document sur la toile est de déterminer les termes de ce « *construit social* » (Couzinet, 2001 : 499).

2.3.3 La valeur sociale du document

Comme nous l'avons vu, en raison des transformations variées permises par le réseau, la notion de stabilité est mise à mal et c'est la « *notion de mémoire [qui] paraît centrale pour l'analyse de l'hyperdocument, objet qui semblerait davantage du côté du flux que de la fixité* » (Couzinet, 2001 : 500). La durée de vie des objets informationnels sur la toile étant aléatoire, un document est donc défini « *comme un objet faisant fonction de mémoire pour une instance réceptrice* » (Couzinet, 2001 : 500). L'irruption des réseaux dans les usages entraîne bien une transformation du document du point de vue des SIC car « *l'utilisation du support électronique et d'internet modifie le statut du document par lequel les connaissances circulent et se mettent en forme sur les réseaux au sein d'un champ donné* » (Couzinet, 2001 : 500). Cependant cette constatation ne remet pas en cause les traits dégagés par Meyriat et tend même à les renforcer. La « *partition* » sur la toile est à l'œuvre en ce que le document y est « *le résultat d'une double construction produit par l'usage d'un destinataire qui le reconnaît en tant que vecteur de réponse et par les interactions qui ont présidé à sa naissance* » (Couzinet, 2001 : 500). S'appuyant sur une analyse de l'approche historique des documents, Viviane Couzinet confirme la pertinence de la scission en deux catégories pour cerner les enjeux de la circulation de l'information de nos jours. Elle compare l'approche informationnelle du document avec la conception des historiens dont le document est « *la matière première* » (Couzinet, 2004 : 21). Elle repère, en croisant les deux démarches historique et documentaire, quatre critères délimitant le concept de document : le contexte, le contenu, l'identification et la valeur qu'on peut lui attribuer. Le document est « *inscrit dans une séquence d'événements qui l'a précédé, dans une suite qui a préparé sa naissance, dans un temps et un lieu* » (Couzinet, 2004 : 24). Il n'est « *pas une entité indépendante* » (Couzinet, 2004 : 24). L'auteur de la source (synonyme de document pour les historiens) son parcours, sa légitimité, les influences qu'il a pu subir font l'objet d'une attention toute particulière ; ces éléments permettent de déterminer la valeur intrinsèque du document. Elle souligne que l'approche historique montre bien que le document est « *un construit social* », fruit de son contexte d'élaboration, des interactions sociales, économiques, techniques (Couzinet, 2004 : 25). À propos des caractéristiques relevées par les SIC, elle s'attarde sur le support et le contenu. La notion d'objet d'étude dans sa dimension matérielle est très importante pour pouvoir saisir « *son utilité, son statut social, parfois son prix* » (Couzinet, 2004 : 26). Elle précise que « *la recherche du sens ne peut s'envisager sans la prise en compte du matériau*

qui la compose », et ce quel que soit le support. Il ne saurait donc y avoir de dissociation du support et de son contenu (Couzinet, 2004 : 28). Ce matériau, « *entendu comme média (...) est le croisement de son utilité, des intentions de l'auteur et de son destinataire chacune pouvant être détournées par le récepteur* » (Couzinet, 2004 : 28). Tous les documents « *ne se ressemblent pas* » et il est nécessaire de prendre en compte le support pour adapter son questionnement vis à vis du document (Couzinet, 2004 : 27). Cette attitude est cruciale pour trouver le sens de l'information portée par le document. Et la multiplicité des récepteurs potentiels avec leur singularité fait qu'il « *n'y a pas de documents morts ou de documents vivants, mais il y a des documents latents en attente de questionnements* » (Couzinet, 2004 : 29). La source d'information de ces documents putatifs est inextinguible. En tant qu'objet d'étude, le document est ainsi « *défini comme concept à quatre dimensions. Il est possible de l'analyser comme ayant une valeur, étant un « construit social », porteur d'un contenu et détenteur d'une identité propre* » (Couzinet, 2004 : 29).

C. Courbières, pour sa part, tranche le débat de façon délibérément « *provocatrice* » en affirmant qu'« *un document, ça n'existe pas* » (Courbières, 2012 : 103). Elle précise sa pensée au sujet de la « *virtualité* » du document en poussant au bout la logique de Meyriat pour qui l'important, c'est l'information. Dans une approche sémiotique, elle indique que « *l'information en tant que donnée accède au statut d'information en tant qu'élément de connaissance à partir du moment où elle est saisie comme signe* » (Courbières, 2012 : 103). Elle revient également sur la distinction élaborée par Meyriat et reprise par Escarpit entre « *document par intention* » (l'objet élaboré en vue d'informer) et « *document par attribution* » (l'objet élaboré dans la recherche d'information). Selon elle, le document n'est réel qu'à deux endroits : là où il est émis et là où il est reçu. Nous constatons que cette caractéristique ne s'applique pas seulement au support numérique, ou à la spécificité de la toile, mais à la nature même de ce qu'est un document. La notion de « *virtualité* » avancée par C. Courbières ne renvoie pas aux réseaux et à l'agrégation d'éléments disparates mais à la vertu qui consiste en la possibilité à un moment ou à un autre de servir de moyen de s'informer (Courbières, 2012 : 104). « *Un objet documentaire n'existe donc qu'à partir du moment où il signifie pour quelqu'un* » (Courbières, 2012 : 104). Se référant à la philosophie de Hans Robert Jaus⁴³, elle définit le document comme « *un horizon documentaire* » (Courbières, 2012 : 104). Elle

43 Hans Robert Jaus : philosophe allemand (1921-1997), théoricien de la réception

précise que « *l'horizon documentaire remet en question la stabilité informationnelle de l'objet documentaire, non pas en raison de la forme ou du support de ce dernier, mais parce que le document se construit à partir de son contexte de réception* » (Courbières, 2012 : 104). La « virtualité » intrinsèque du document, le fait qu'il soit toujours potentiel lui fait « *proclamer son impossible définition* » (Courbières, 2012 : 115).

Cette difficulté à appréhender le concept de document et à le définir provient sans doute en partie de l'origine étymologique et de la dénotation insatisfaisante de ce terme, maintes fois relevées par Meyriat. Cependant en l'assujettissant au concept information dans le sens information-connaissance, nous prenons conscience de l'importance de ce concept dans l'approche sociale de l'information en tant que vecteur des connaissances reconnues par une société à un moment donné, fût-elle à l'âge dit « numérique ». La tentative de reconstruction voire de détournement⁴⁴ auquel ce concept est soumis de nos jours révèle comme nous l'avons montré précédemment un retour en force de la théorie mathématique de l'information qui évacue la signification, rendant ce faisant service aux intérêts marchands de ceux qui détiennent les canaux de la communication sur la toile et ont pour objectif de faire admettre un sens convenu à l'avance.

Pour tenter de lever les ambiguïté sémantiques et conceptuelles et donner des clés afin de repérer les résurgences de l'approche mathématique de l'information, dans le débat autour du fait numérique, nous avons effectué une ébauche d'analyse épistémologique qui nous a conduite à dégager une série d'indicateurs axés autour des visions respectivement sociale et mécaniste des deux concepts des sciences de l'information. Nous allons dans la partie suivante présenter la méthodologie utilisée, les résultats obtenus et une analyse de ces derniers.

44 Après le terme d'« information », d'« informaticien » et de « réseau », les sciences de l'information vont-elles se voir confisquer le terme « document » ?

Chapitre 3

Analyse épistémologique des concepts en SIC : éléments d'interprétation pour distinguer le savoir scientifique du « fait numérique »

Après une présentation de la méthodologie appliquée dans l'analyse épistémologique des deux concepts centraux des sciences de l'information, nous allons expliciter la grille épistémologique élaborée, et en aborder les limites et la portée éventuelle.

3.1 Point sur l'analyse épistémologique

3.1.1 L'épistémologie : étude critique des sciences

L'épistémologie est le nom qui a été donné à ce qui s'appelait auparavant la philosophie des sciences. Selon H. Barreau, un des traits caractéristiques de l'épistémologie est ce rapport entre la philosophie et la science : « *en Occident, la philosophie depuis Platon a accompagné le développement de la science* » dans sa volonté de dépasser la conception « vulgaire », commune du monde (Barreau, 2002 : 7). L'approche épistémologique a pour objectif d'opérer la distinction entre ce qui relève de la science et ce qui n'en relève et d'étudier la nature de la connaissance scientifique : les aspects descriptifs et les aspects normatifs. Suivant l'épistémologie historique d'H. Barreau, peu de disciplines peuvent être qualifiées de sciences théoriques : la discipline reine et mère étant les mathématiques. Les phénomènes observés dans le monde ont fait l'objet de recherches et le but est de dégager des lois et des concepts pour apporter une explication rationnelle aux éléments dont chacun éprouve les effets au quotidien. Pour le reste ce sont les techniques qui en découlent : ingénierie, informatique. Les autres disciplines qui sont des sciences pratiques, selon l'approche du philosophe des sciences, sont la psychologie, la linguistique et l'économie. L'histoire n'est pas une science théorique et les autres disciplines telle la sociologie ne sont pas mentionnées.

Selon Descartes, les mathématiques donnent le pouvoir aux hommes de connaître les lois que Dieu a imposé à la nature. (Kremer, 2006 : 8). L'épistémologie « *contemporaine (...) pose la question de savoir comment les conjectures, hypothèses, et théories scientifiques rendent compte de ce qu'est la nature* » (Kremer, 2006 : 9).

L'épistémologie est parfois rapprochée de la théorie de la connaissance, partie de la

philosophie qui étudie le phénomène de la connaissance. Le lien est étroit avec les sciences de l'information dont l'un des principaux objets d'étude est en relation avec la connaissance.

3.1.2 L'épistémologie des sciences de l'information

Meyriat, en son temps, avait posé les « *fondements d'une science de l'information* » et les « *jalons pour son épistémologie* » (Meyriat, 1983 : 67). Il soulignait que « *la liaison entre information et connaissance est (...) essentielle* ». Celle-ci étant entendue dans « *l'acception la plus large* » du « *mot « connaissance », qui désigne l'acte de l'esprit s'appliquant à un objet quel qu'il soit. Il peut s'agir de la connaissance que l'on appelle « scientifique », et qui constitue un savoir structuré, aussi bien que celle du fait ou de l'événement le plus trivial* » (Meyriat, 1985 in Couzinet, 2001 : 203). Meyriat, comme l'a montré V. Couzinet, dans un ouvrage hommage (Couzinet, 2001), a grandement contribué avec Escarpit, notamment, à délimiter les objets de cette science. Ceux-ci ne font cependant toujours pas consensus actuellement mais le fait de s'interroger sur les concepts est une des caractéristiques scientifiques (Quivi, 2011).

Le travail épistémologique s'est donc poursuivi et se poursuit encore à travers de nombreux travaux de recherche de la « documentation », dont C. Gardiès rend compte à partir d'une analyse « *d'un point de vue épistémologique [des] concepts, [de] leur étendue et [de] leur juxtaposition* » afin « *d'offrir un cadre de lecture* » (Gardiès, 2012 : 155). Se plaçant dans la continuité des conceptions d'auteurs, tels que Meyriat, Escarpit, G Losfeld, R. Miège, J-P. Metzger, H. Fondin, V. Couzinet, C. Courbières ou encore Y. Jeanneret, elle qualifie « *les fondements théoriques de la documentation au sein des SIC comme une branche scientifique spécifique (...) : l'information-documentation. Cette branche peut permettre d'étudier le partage des savoirs au travers de la construction, circulation et transformation d'objets informationnels. Ce partage se situe dans des processus de médiation, d'énonciation et d'appropriation du sens échangé dans des situations de communication. Le travail conceptuel autour de cette branche scientifique a cherché à formaliser cet espace théorique particulier* » (Gardiès, 2012 : 197-202). Elle souligne avec V. Couzinet « *la nécessité d'ancrer les pratiques professionnelles des praticiens de la documentation sur les Sciences de l'information et de la communication, car la méconnaissance par les acteurs eux-mêmes des enjeux de cohésion de la discipline, de culture commune ou de positionnement des acteurs, perceptible dans les publications des chercheurs de la discipline constitue « un frein au développement de supports de médiation correspondant à la spécificité française des SIC »* »

(Gardiès, 2012 : 201).

Cette ignorance, nous l'avons éprouvée nous-même et c'est consciente de cette exigence impérieuse que nous sommes allée à la rencontre des théoriciens des sciences de l'information, ce qui a été l'occasion pour nous de nous initier à la démarche de recherche. Ainsi, nous avons ébauché un modèle épistémologique susceptible d'aider les néophytes à s'orienter dans les courants des SIC en France, que nous détaillons maintenant.

3.2 Grille de lecture épistémologique

3.2.1 Méthodologie de construction de la grille

Le choix du sujet et la volonté de comprendre le paradoxe entre difficulté à enseigner l'information-documentation à ladite ère numérique et volonté de la société de mettre en avant l'information nous a conduite à appréhender les rapports entre les concepts information et document et « le numérique ». Très vite au cours des lectures les ambiguïtés, confusions et amalgames entre savoir scientifique et considérations techniques sont apparues. Démêler les différents aspects relevait de l'analyse épistémologique.

D'où le choix d'une méthodologie axée autour du travail épistémologique effectué. Cette piste a permis d'orienter les lectures, d'être attentive aux termes et expressions utilisés et d'engager une analyse du discours scientifique des auteurs qui écrivent sur les concepts fondamentaux d'« information » et de « document » en SIC en prenant la dimension numérique en compte. L'ambition étant de trouver une possibilité de situer les chercheur-es en fonction de la manière dont ils abordent les notions et concepts clés.

Une observation du point de vue adopté par les auteur-es a révélé que la différenciation d'approche du concept d'information se cristallise autour de la théorie mathématique de l'information. Les fondateurs des SIC en France ont d'ailleurs été quasi contraints de construire la discipline scientifique contre cette théorie qui a étendu son influence bien au-delà du domaine de l'ingénierie. Le lien entre les deux concepts d'information et de document, comme nous l'avons vu précédemment, est tel que cette différenciation entre approche sociale et mathématique, se retrouve dans les controverses, bien que de façon plus subtile, autour du document. Il a été intéressant de repérer également les mentions faites (ou pas) des travaux des autres chercheur-es de la discipline.

C'est au fil des lectures et de la compréhension des concepts que des constantes sont

apparues entre les discours prônant une vision sociale, spirituelle, significative du concept d'information et par conséquent de celui de document, dans la filiation de Meyriat et, les propos s'inspirant d'une vision plus arithmétique, concrète, insensible, de ces concepts, dans la lignée de Shannon ou Weaver.

Nous avons été incitée pour mieux cerner les enjeux, à intégrer la dimension historique dans l'étude de l'évolution des concepts « information » et « document ». De même, la prise en compte du parcours personnel des chercheur-es s'est avérée utile pour donner des clés de lecture pertinentes sur les rapports de force qui se jouent au sein des SIC entre les tenants d'une vision mathématique de l'information et ceux qui mettent la question du sens au cœur de leurs préoccupations. Le mot « numérique » n'est pas une référence SIC ou SHS⁴⁵ mais est plus en lien avec les sciences dures. Beaucoup d'auteur-es SIC sont d'abord des professionnel-les (archivistes ou gestionnaires de documents) qui ne sont pas confronté-es aux méthodes de la recherche. Se dégage ainsi une conception « utilitariste », menant à des applications concrètes.

45 Sciences humaines et sociales

3.2.2 Présentation de la grille d'analyse épistémologique

Le tableau ci-dessous reprend les indicateurs issus de l'analyse épistémologique des deux concepts. La colonne de gauche présente les critères des tenants de l'information-connaissance ; celle de droite, ceux qui considèrent l'information comme une donnée.

Indicateurs	Vision sociale exclusion de la théorie mathématique	Vision mécaniste intégration de la théorie mathématique
1	Proximité des deux termes « document » / « information » lien permanent entre document et information	Éloignement des deux termes (ou occurrence moindre) lien distendu entre document et information
2	Occurrence des termes « sens » / « signification » / « esprit »... subjectivité	Occurrence des termes « donnée » / « flux » / « flot » objectivité
3	Prégnance accordée à l'esprit humain	Importance accordée aux chiffres, formules et autres équations
4	Distinction claire des notions ex. ce qui relève de la technologie et ce qui n'en relève pas	Confusion entre les notions Outils/esprits mis sur le même plan
5	Veiller au sens des mots définition claires, simples, concises des concepts avec un emploi correct des mots	Jouer avec les mots au détriment de la justesse confusion langage courant / langage scientifique complexité des définitions, emploi « détourné » des mots
6	Idée de sens à apporter, d'ordre à trouver (positive) neg-entropie	Idée de désordre à régler, de flux à démêler (négative) entropie

Indicateurs	Vision sociale exclusion de la théorie mathématique	Vision mécaniste intégration de la théorie mathématique
7	Accroissement des connaissances Ouverture du potentiel informationnel quête de sens infinie	Réduction d'incertitude Clôture du potentiel informationnel finitude
8	Information n'est pas une donnée pré-construite pas d'existence autonome	Information = donnée pré- construite existence autonome
9	Information opération cognitive – pas un bien comme un autre notion humaine	Information / document = bien de consommation notion de marché
10	Prise de distance avec l'objet de recherche attitude « scientifique » Espace de validité clairement établi et circonstancié SIC	Emphase, fascination non déguisée pour l'objet de recherche superlatifs attitude « non scientifique » Élargissement obligatoire des concepts du fait de la technique pas de choix dans la réflexion, conclusions dictées par « le numérique »
11	Importance du contexte (social, économique, politique) de l'individu en lien avec l'information	Décontextualisation Assemblage, agrégation d'éléments épars pour fabriquer l'information
12	Références explicites aux théoriciens fondateurs Appui sur recherches précédentes	Pas, peu de références explicites Références implicites (éloignement) Peu d'appui sur recherches précédentes

Pour chacun des douze critères relevés, nous proposons ci-dessous une explicitation en reprenant les éléments de l'analyse théorique présentée dans les parties précédentes, le rang de présentation est indicatif et il n'y a pas de notre point de vue d'ordre de priorité entre les différents indicateurs, lesquels sont parfois (souvent) corrélatifs.

Indicateur numéro 1

Nous avons souligné le caractère indissociable des deux concepts document et information, ou, pour les théoriciens les plus reculés comme Otlet ou Briet, le lien entre le document et l'esprit, la pensée, les connaissances. Briet qualifie ainsi la documentation de « *moyen puissant de collectivisation des connaissances et des idées* » (Briet, 1951 : 30). Meyriat considère que les deux notions sont « *inséparables* » dans la définition (Meyriat, 1981 in Couzinet, 2001 : 144). Escarpit s'éloigne quelque peu de la conception de Meyriat dans la mesure où il reprend pour partie l'approche shannonienne de l'information mais en l'inscrivant dans une perspective diachronique : l'information étant « *un produit informationnel du temps* » : un événement (Escarpit, 1976 : 68) et le passage au document permettant la production d'un « *anti-événement* » et donc la « *constitution d'un savoir* » (Escarpit, 1976 : 59). Cependant comme nous l'avons évoqué plus haut il marque, au sujet de « *la signification* », les « *limites du modèle mécaniste de la théorie de l'information qui, en ce cas, n'est qu'une théorie de la communication* » (Escarpit, 1976 : 27) car « *la notion d'information est liée à toute forme de pensée (...) mais la pensée de type humain (...) a un pouvoir particulier d'invention* » (Escarpit, 1976 : 39) et « *le schéma mécaniste ignore délibérément les aspects psychologiques et sociologiques de l'avant-canal et de l'après-canal* » (Escarpit, 1976 : 40). D'un point de vue cognitif les concepts information et document ne peuvent donc être isolés. Les héritiers de cette conception perpétuent cette approche, tel Y. Jeanneret qui voit dans l'acception nouvelle du document, comme étant indépendant de la notion information, une marque idéologique « *où les chantres de la commotion culturelle négligent la matérialité documentaire (...) [or,] s'il faut accorder une importance aux objets documentaires, c'est précisément parce que l'information n'est pas un objet, mais une relation posée face à l'objet* » (Jeanneret, 2000 : 80).

V. Couzinet rappelle cette priorité absolue à l'information quand il s'agit du document, rapport distendu si ce n'est négligé, quand on se penche sur les travaux du RTP-DOC ou de J-

M. Salaün, lesquels, pris par des considérations techniques de formats, de faisabilité ou par des considérations économiques, de rentabilité commerciale, n'évoquent pas ce lien établi scientifiquement par les théoriciens précurseurs. L'occurrence du terme information dans les travaux de l'équipe du CNRS est d'ailleurs très faible et l'on retrouve plus souvent le terme « contenu », « signe », « texte ». La cooccurrence des deux termes ne signifie d'ailleurs pas pour autant toujours une vision sociale de l'information et donc du document car la relation entre eux peut l'être au sens arithmétique comme c'est le cas dans l'approche pratique faite par H. Fondin, guidé par des soucis d'applicabilité. Dans le cadre du travail professionnel, cette acception est valide mais d'un point de vue épistémologique, cette vision est incomplète pour rendre compte du concept scientifique.

Indicateur numéro 2

La question du sens, nous l'avons montré dans la première partie, est la ligne de démarcation entre les tenants de la vision meyratienne et ceux de la vision shannonienne. Et, rappelons-le, c'est Shannon lui-même qui trace cette ligne de façon très nette. L'emploi du terme « donnée », renvoyant à la théorie mathématique et à une certaine neutralité, objectivité est en opposition avec les termes « sens », « signification ». Dans le premier cas, le rapport est désincarné, « objectif », « neutre » et donc facilement quantifiable, délimitable (l'accent est mis sur le caractère « discret », dénombrable de l'« information ») tandis que dans le deuxième cas, l'« attribution de sens », l'« opération mentale » du sujet sur lui-même rend l'information nécessairement subjective. Il est intéressant de noter que cette remarque sur la terminologie (et la connotation idéologique qu'elle véhicule) est valable pour les écrits sur le document. De la même façon, nous avons pu constater la présence bien moindre des éléments ayant trait à la signification, à la compréhension, au regard de ceux relevant du champ lexical de la « mesure » dans certains travaux intégrant la théorie mathématique et la relation inversement proportionnelle dans d'autres récusant cette même théorie. Ainsi Meyriat revient toujours au sens, au contenu cognitif, à la signification à plus forte raison au sujet du document, « objet » qui sert à « communiquer » l'information, qui ne l'intéresse que par sa capacité « à véhiculer du sens » tandis que J-M. Salaün considère que « les données sont les plus petites unités documentaires représentant le monde réel. Ces unités documentaires sont « données », c'est à dire prises comme telles, non discutées, objectives en soi, neutres »

(Salaün, 2012 : 88). L'approche sociale de l'information en fait un élément subjectif puisque lié à la connaissance, qui est individuelle tandis que l'approche mathématique considère la donnée comme une abstraction, déliée de toute attache et donc objective. C'est le point de vue de la machine, pour qui en effet ce ne peut être autre chose, qui est adopté. C'est une indication d'une vision qu'Escarpit qualifie de « mécaniste ».

Indicateur numéro 3

En cohérence avec ce qui précède, nous observons une prégnance accordée à l'humain, à l'individu en tant qu'être pensant dans les travaux d'Otlet, de Briet, de Meyriat, d'Escarpit, d'Y. Jeanneret ou encore de V. Couzinet tandis que les considérations concrètes, l'importance accordée aux chiffres, aux formules et autres équations ressortent particulièrement des travaux d'H. Fondin ou du RTP-DOC, qui sont gouvernés par un souci d'applicabilité.

Abraham Moles, pour exalter la valeur du travail de Shannon reprend à son compte en la nuancant l'expression du physicien William Thomson : « *c'est une bien pauvre science, disait lord Kelvin, celle qui n'est pas capable de s'exprimer dans des mesures* » (in Shannon, 1975 : 14). Derrière l'utilisation immodérée des chiffres et autre lois bibliométriques, se loge sans doute la volonté de légitimer une science qui selon Meyriat en 1986, est « *toute récente, et qui sort à peine des incertitudes de sa petite enfance* » (Meyriat, 1986 in Couzinet, 2001 : 251). Presque trente ans après, nous pourrions estimer que sans avoir atteint l'âge adulte, elle est adolescente mais l'obsession « comptable » de certain-es chercheur-es les conduit à perdre de vue les objets majeurs que sont les concepts d'information et de document et à être happé-es par les soucis (légitimes eux-aussi) d'applicabilité technique. Escarpit, en réinterprétant la théorie mathématique, montre que la réduction de la communication au transport de signaux est improductive. Les êtres humains face à l'impossibilité de se faire comprendre par le biais des canaux à leur disposition, si ceux-ci n'ont pas la capacité suffisante, s'en affranchissent et font la révolution (Escarpit, 1976 : 42). Il y a, au contraire, dans une certaine vision contemporaine, une soumission aux outils au détriment de l'idée d'émancipation humaine : ainsi, les travaux sur le « *Web sémantique* » s'éloignent du principe de précaution posée par Briet, qui incitait à « *commander, toutes facultés en éveil, aux robots de demain* » car la « *machine vaudra ce que vaut le servant* » (Briet, 1951 : 29).

L'information est le « *contenu cognitif d'une communication* », processus « *par lequel*

s'établit une relation mentale entre deux ou plusieurs êtres humains ». Cette relation mentale « *relève de l'esprit et non pas de la matière* » (Meyriat, 1983 : 65-66). Cette subjectivité intrinsèque au phénomène informationnel est mise en évidence par les liens qu'Y. Jeanneret établit entre l'information, le document et la connaissance : l'information, c'est « *la relation entre le document et le regard porté sur lui* ». La connaissance, c'est « *le travail productif des sujets sur eux-mêmes pour s'approprier des idées ou des méthodes* » (Jeanneret, 2000 : 85).

Indicateur numéro 4

Le débat autour des deux concepts que nous analysons est très complexe à l'heure dite numérique car il est particulièrement ardu de démêler ce qui ressort de la technologie (application concrète) et ce qui ressort de la science (abstraction, conceptualisation). C'est un des aspects qui nous a semblé représenter également la ligne de démarcation entre tenants de la vision sociale de l'information et tenants de la vision mécaniste. Y. Jeanneret a consacré un ouvrage complet à ce sujet en insistant sur les confusions induites par des expressions telles que la « *société de l'information* » et les « *nouvelles technologie de l'information et de la communication* » (Jeanneret, 2000 : 52). Le brouillage s'accroît encore avec un vocable unique mais très équivoque, comme nous l'avons vu auparavant, « le numérique ». Plus qu'une querelle sémantique, il s'agit bien d'une conception idéologique et d'une différenciation épistémologique. La vigilance est de mise concernant la distinction entre ce qui relève des outils, des considérations techniques et ce qui relève du concept, de la représentation de la réalité. Ainsi si l'on reprend la définition de l'information d'Y-F Le Coadic : « *connaissance inscrite (enregistrée) sous forme écrite (imprimée ou numérisée), orale ou audiovisuelle sur un support spatio-temporel* », on voit typiquement le mélange des genres entre considération technique et approche conceptuelle, où il n'est pas précisé que les « connaissances » orale ou audio-visuelle peuvent, elles aussi, être numérisées (Le Coadic, 2004 : 6). Le rapport entre le concept d'information et les implications techniques est posé différemment par Meyriat lorsqu'il se demande si « *la nature des procédés techniques utilisés (mécaniques, électriques, électroniques...)* » est une « *variable secondaire ou principale au regard du sens, donc de l'information* » (Meyriat, 1983 : 70). L'amalgame approche technique-approche conceptuelle est presque permanent dans les résultats des réflexions du réseau thématique du CNRS, ceci étant lié au mélange des genres entre « chercheur-es » de disciplines très diverses, certain-es étant avant tout des ingénieur-es. Le propos n'est pas ici de déprécier le travail de ces derniers

mais d'un point de vue épistémologique, il ne concerne pas les mêmes objets d'étude. Ainsi au sujet de la « *numérisation massive* », l'équipe rend compte des projets nationaux et internationaux qui mettent en place des stratégies pour répondre aux appels à projets pour réussir de façon optimale, informatiquement, la migration de collections documentaires au format numérique. Les conséquences sur le contenu et la réception sont évoquées en matière de conditions d'accès technique, d'interopérabilité, de conservation des données et de gestion des droits à travers une vision commerciale et non en terme d'attribution de sens.

Indicateur numéro 5

Le mélange des genres entraîne une rigueur moindre dans l'emploi des mots. Nous pouvons soulever l'extrême exactitude observée par des auteur-es comme Meyriat, Escarpit, V. Couzinet, ou encore C. Courbières qui respectent le sens littéral des mots utilisés. En revanche, les approximations, voire les détournements de sens sont plus réguliers pour le RTP-DOC ou J-M. Salaün. La notion de « virtualité » en est un exemple édifiant. Le terme « virtuel » est utilisé par Otlet ou Meyriat mais il est devenu un synonyme pour évoquer ce qui circule sur la toile. Certains chercheur-es reprennent cet usage ordinaire alors que nous avons vu que c'est un des éléments de la rigueur scientifique que de déterminer précisément le langage employé (un mot courant peut être employé de façon plus spécifique mais dans le cas du terme « virtuel », c'est l'inverse qui se produit). Chez Otlet, et chez Meyriat ensuite, l'acception est liée à la potentialité au sens littéraire. Tandis que ce terme, développé avec l'internet, tend à être un marqueur de son fonctionnement (la nouveauté de la réticularité étant figurée dans son caractère virtuel) et le sens « fictif » est prédominant, au sens mécanique du terme, véhiculant de plus une idée de magie. C. Courbières en avançant la notion de « *virtualité* » à propos du document numérique (ou pas) réhabilite le sens premier en faisant ressortir la potentialité « *consubstantielle* » du document.

Une autre illustration d'utilisation approximative avec le terme « fixité » est observable dans la phrase suivante : « *le numérique déplace la question du support du document, qui en assurait la stabilité grâce à la fixité de l'information, vers la problématique de sa structure* » (Pedauque, 2007 : 17). Cela signifierait qu'il n'est pas possible de fixer « l'information » numériquement or sur un CD-Rom par exemple, cela est techniquement faisable. Le problème est (volontairement ou non) mal posé. Car la stabilité ressort de la « *fixation* » de

l'information : « *l'information est fixée sur un support ; cette fixation, par écriture ou enregistrement, est la première phase de l'émission, puisqu'elle est réalisée en vue d'une communication : un document, au moins virtuel, a été créé* » (Meyriat, 1981 in Couzinet, 2001 : 164) : nuance de taille qui fait appel à une action volontaire contrairement au substantif « fixité » qui renvoie à un état immuable, fixe, invariable qui est relatif au document et non à l'information (encore un exemple du manque de distinction claire entre les notions).

Nous avons eu également l'occasion plus haut⁴⁶ de revenir sur les qualificatifs de « transparence » et de « neutralité » décernés hâtivement au fonctionnement de la toile.

Indicateur numéro 6

Escarpit pointait l'attitude intellectuelle « *négative* » à l'égard de l'information qui prévalait dans la conception mécaniste et cette dimension est une constante pour les tenants contemporains de cette vision. Nous sommes selon eux submergés par le « *flot* », les « *flux* » d'information et il faut y trouver un sens et régler le désordre. Le « *néo-document* » ou document numérique serait le moyen de surmonter le « *chaos* », la « *cacophonie* », l'« *oubli* » (Pedauque, 2007). L'appui apportées par les mesures et autres lois étant dans cette conception une aide précieuse. C'est le caractère entropique à combattre qui est sous-jacent. À l'inverse, une vision plus positive consiste à discerner dans l'information, une connaissance communiquée, un « *sens partagé* » venant contribuer à la compréhension de l'environnement et enrichir de façon cumulative le savoir des individus (Meyriat, 1986 in Couzinet, 2001 : 256). La multitude n'est pas inquiétante car elle est synonyme de richesse intellectuelle et la mise en ordre se fait dans l'esprit de celui ou celle qui reçoit l'information. Il s'agit d'un processus néguentropique. L'individu a la maîtrise de cette information puisqu'elle est connaissance communiquée ou communicable et il ne peut donc être dépassé par elle. Tout dépend du niveau auquel on se place : si l'on fait face à la profusion de données, on risque d'être submergé par le flot mais si l'on considère l'information-connaissance, on réalise un acte constructif d'attribution de sens en cohérence avec les connaissances déjà acquises. Cette assimilation de l'information à un « *flux* » est une « *idée d'inspiration économique* » où « *la connaissance et le savoir sont assimilés à un stock, à du capital qui résulte de ce flux d'information* » (Lamizet, 1997 : 410). La surabondance des informations évoquée comme

46 Voir partie 2.2.3.1

corrélative à l'arrivée de la toile est à nuancer selon l'approche sociale, puisqu'elle est alors perçue comme « *une ressource fondamentale et constitutive de toute société* » par Meyriat (in Couzinet, 2001 : 164) : la toile accélère la circulation des données mais cette capacité se situe au niveau du canal. La mise en avant de cette profusion est le moyen de sommer les citoyens de s'emparer de la technologie sans que l'aspect cognitif soit valorisé. Il y a une forme de fatalité dans les discours qui contredit la vision sociale selon laquelle l'information est un élément constitutif de la communication entre êtres humains.

Indicateur numéro 7

L'acquisition et l'accroissement des connaissances dans la vision sociale n'est pas en corrélation automatique avec une réduction de l'incertitude. Cette notion d'incertitude relative au besoin d'information ressort de l'attitude quelque peu négative, voire pessimiste, d'une certaine relation avec l'information, que nous venons d'évoquer.

L'extension des connaissances appelle de nouvelles connaissances et peut de se fait en quelque sorte accroître l'incertitude⁴⁷. Ce terme connoté péjorativement ne rend pas compte du processus cognitif en jeu permettant la relation entre êtres humains et la compréhension mutuelle nécessaire à toute communication humaine. La quête de la « *signification* » est infinie car elle est liée à l'interprétation de chaque individu et à son histoire personnelle tandis que la tentative des chercheurs du « *Web sémantique* » laisse croire à une possibilité d'information totale, personnalisée qui se constitue d'elle-même à partir des traces laissées par l'internaute : « *les réponses ont déjà été données avant même que les questions ne soient posées* » (Salaün, 2012 : 7). Il est notable qu'une évolution technologique soit créditée du pouvoir de rassembler toutes les connaissances humaines, comme s'il ne pouvait plus y avoir de découverte scientifique. Il suffit dans cette vision d'apparier la connaissance avec celui qui en a besoin sur le modèle des moteurs de recherche qui apparient résultats et requêtes. C'est faire l'économie un peu rapide de l'étape cognitive nécessaire et c'est faire abstraction comme le soulignent Y. Jeanneret, et V. Couzinet du contexte dans lequel le-la requérant-e se trouve. Cette tentative de « *clôture de l'espace informationnel* » est en cohérence avec l'approche mathématique qui ne peut se faire que dans des « univers limités », probabilisables : « *en théorie de l'information (...) le monde idéal des mathématiques traite de la modélisation de*

⁴⁷ Socrate exprimait cette contradiction par la célèbre maxime : « la seule chose que je sais, c'est que je ne sais rien »

phénomènes qui n'existent pas dans le monde réel : la signification, la volonté, la preuve » (Dion, 1997 : 49-50). Comme nous l'avons pointé dans la première partie, il y a une velléité actuelle de s'appuyer sur les performances techniques des réseaux pour faire croire que ces éléments spirituels sont désormais calculables, prévisibles. À l'inverse, pour la discipline information-documentation, le potentiel informatif de tout document est inépuisable et l'on peut ré-exploiter à l'infini un même document : « *la capacité informative d'un document n'est jamais épuisée par les utilisations déjà faites d'informations qu'il contient* » (Meyriat, 1978 in Couzinet : 117) . L'approche mathématique de l'information est tout autre car « *l'information est un concept qui n'a de sens que dans un contexte bien précis : celui d'un ensemble fini et probabilisé, caractérisé par un niveau d'incertitude donné* » (Dion, 1997 : 88). Nous pouvons constater à nouveau à quel point le cadre dans lequel les mots sont utilisés induit une interprétation complètement différente : « *sens* » et « *contexte* » ici renvoient à l'« *ensemble fini* » et ne sont pas employés pour mettre en relation « *signification* » et « *champ d'interprétation* » du sujet qui fait acte de connaissance. L'illusion de contrôle du temps par le biais de la circulation « en temps réel » représente une conception particulière du rapport au temps faisant croire à une maîtrise du temps et réfutant par conséquent la fuite inexorable de celui-ci qui participe à l'illusion magique démasquée par Y. Jeanneret dès 2000, au commencement de la démocratisation de l'internet, où « *Bill Gates [était le] modèle de tous les consultants-thaumaturges* » (Jeanneret, 2000 : 9).

Indicateur numéro 8

La relation capitale avec la communication fait que par essence l'information n'est pas une donnée pré-construite avec une existence autonome. C'est ce qu'exprime C. Baltz en notant que l'information « *n'est pas donnée sous la main* », qu'elle « *vient d'ailleurs* ». A contrario les tenants de la vision mécaniste considèrent que l'information est une donnée pré-construite qui existe par elle-même : les éléments sont « *connus* », présents, « *prévisibles* » : il ne reste plus qu'à les agréger. Pour Meyriat, « *l'information n'est pas un acquis, un objet constitué mais une modification de l'état de connaissance de celui qui la reçoit* » (in Couzinet, 2001 : 202-203). Pourtant, la représentation numérique de celle-ci lui fait vivre « *une existence autonome* » selon P-D. Pomart, omettant de préciser que dans ce cas, pour les sciences de l'information, il ne s'agit encore que de « *données à transformer en connaissances* » (Juanals, 2003 : 30). B. Juanals insiste en indiquant que « *sens et savoirs (...) ne peuvent donc (...) être présents à un état préexistant* » (Juanals, 2003 : 30). Cette distinction entre objet indépendant et matière interdépendante dénote l'intégration ou l'exclusion de la théorie mathématique puisque selon Meyriat, ce que manipulent les informaticiens, ce sont des « *données abstraites qui ne sont pas encore de l'information* » (Meyriat, 1986 in Couzinet, 2001 : 259) tandis que pour J-M. Salaün les « *données sont les plus petites unités documentaires représentant le monde réel* » (Salaün, 2012 : 88). Le souci d'opérer la distinction entre donnée pré-construite et information (qui suppose modification et ne peut donc être un état figé) est révélateur de la prise en compte de ce qu'Escarpit appelait « *l'avant-canal et l'après-canal* » (Escarpit, 1976 : 40). Autrement dit, le fait de garder (ou pas) à l'esprit la nécessaire intervention humaine. Ainsi B. Simonnot rappelle que « *c'est l'attention* » portée aux « *données* » qui en font ou non « *des informations effectives* » (Simonnot, 2012 : 27). Elle relaie une des « *hypothèses fondamentales* » de « *l'approche sense making* » de Brenda Dervin qui considère que « *l'information n'est pas une chose qui existe de manière indépendante ou externe aux individus mais un produit de l'observation humaine* » (Simonnot, 2012 : 90).

Indicateur numéro 9

L'information n'est pas un bien, un matériau qui peut s'échanger comme un autre. C'est pour cela que Meyriat souligne que l'on ne peut au sens des sciences de l'information « produire » l'information mais que l'on peut la transmettre. Le fait de considérer l'information comme pouvant être « autonome » permet de l'envisager comme pouvant « être acquise facilement sur un marché où se rencontrent l'offre et la demande, dès lors qu'existent les outils de guidage vers cette offre » (Cacaly, 2004 : 122). Si l'information est considérée comme une donnée, objective, neutre, entité autonome, alors elle peut se monnayer, et être considérée à l'instar de l'objet qui la supporte comme un bien marchand qui fluctue en fonction des besoins. Cet aspect économique fait l'objet d'un ouvrage complet de J-M. Salaün. D'un point de vue économique, les constats effectués sont corrects mais d'un point de vue épistémologique, il s'agit d'une réduction considérable du concept puisque le focus est mis sur une utilisation particulière et non sur les propriétés intrinsèques de l'objet d'étude. C'est cette conception que montre B. Simonnot en remarquant que l'objectif des « dispositifs d'accès à l'information », moteurs de recherche ou réseaux sociaux, « semble davantage de susciter le désir plutôt que de répondre à des besoins informationnels » (Simonnot, 2012 : 136). D'un point de vue épistémologique, nous abordons les notions dans une acception la plus complète possible et c'est dans ce sens que l'approche du RTP-DOC ou de J-M. Salaün au sujet du document nous semble véhiculer une conception partielle, de l'information qu'il supporte, dans sa seule valeur économique. Le document est envisagé comme un prestataire de service qui génère du profit pour les concepteurs de logiciels, pour les éditeurs de livres numériques ou encore pour les entreprises de « management de la relation client » (Pedauque, 2007 : 109). Cet aspect économique de l'information et du document qui la supporte n'est pas lié à la technologie et Meyriat l'inclut dans la typologie des informations mais l'analyse épistémologique nous incite à ne pas en faire un élément primordial de la définition des concepts.

Indicateur numéro 10

La posture adoptée par les chercheur-es est révélatrice également du courant dans lequel ils-elles se placent. Le fait numérique entraîne une dimension technique incontournable et s'agissant du concept d'information, l'approche mathématique incite naturellement à adopter un point de vue « *techno-centré* » pour reprendre l'expression d'Y. Jeanneret. Les chercheur-es en SIC qui suivent la route tracée par les fondateurs des sciences de l'information-documentation ne sont pas pour autant luddistes et sont au contraire conscients de l'importance des enjeux techniques puisqu'ils se placent dans la branche mère des sciences de la communication et que comme l'a montré C. Gardiès, ils entretiennent de plus un lien étroit avec la pratique. Les évolutions des canaux par lesquels circulent l'information sont à prendre en compte mais la remarque formulée par Meyriat reste d'actualité : « *des outils nouveaux sont le plus souvent utilisés d'abord par les techniciens et les industriels, qui en profitent pour imposer leurs exigences et leurs conditions à ceux qui veulent communiquer* » (Meyriat, 1981 in Couzinet, 2001 : 132).

La tonalité du discours est significative également de la prise de recul plus ou moins grande avec les objets d'études, cette distance étant symptomatique de l'attitude scientifique. Nous retrouvons chez les laudateurs de la vision mécaniste une certaine emphase, si ce n'est fascination envers les performances des outils numériques. Tandis que ses contempteurs tendent à rester modéré-es dans leur propos et à réserver une place idoine aux machines loin de « *l'imagerie publicitaire qui vante inlassablement le même modèle télépathique de communication* » (Merzeau, 2012). Ainsi à propos de l'internet et des discours dithyrambiques qu'il génère, Y. Jeanneret propose une « *piste pour cultiver le discernement : ne pas nous laisser prendre à l'effet d'abstraction de ceux qui nous demandent de considérer l'objet comme une essence, un « en-soi »* » (Jeanneret, 2000 : 57). Nous pouvons a contrario relever dans certains discours un goût immodéré pour les superlatifs au sujet de ce qui se passe sur les réseaux et un champ lexical relevant de l'exploit : « *spectaculaire* », « *inédit* », « *sans précédent* », « *succès* », « *explosif* », « *vitesse* », « *accélération* », « *transformation radicale* », etc. Le tout assorti parallèlement d'une certaine vision pessimiste car la faculté de l'être humain à avoir prise sur ces phénomènes présentés comme extra-ordinaires semble bien mince.

Indicateur numéro 11

La prise en compte du contexte (social, économique, politique) de l'individu qui participe à l'attribution de sens s'oppose dans les approches respectives à une certaine décontextualisation du réseau qui permettrait un assemblage, une agrégation d'éléments épars pour « fabriquer » l'information. C'est ce qu'Y. Jeanneret et J. Davallon qualifient d'« *unités* » *sémantiques isolées, interchangeableables et insécables* » (Davallon, 2004). Cette remarque est à mettre en relation avec la prétendue « *neutralité* » du réseau qui livre des données « *neutres* », « *objectives* », selon J-M. Salaün, tandis que Meyriat insiste sur la dimension culturelle de la connaissance car « *toute connaissance dépend de la culture de celui qui connaît* ». Nous avons rappelé combien les travaux de Brenda Dervin ont montré l'importance de l'environnement du chercheur d'information dans son rapport à celle-ci. Y. Jeanneret et J. Davallon insistent sur l'intégration de toute démarche de recherche dans « *un programme d'activité* » et J-P. Metzger précise que l'unité documentaire est un « *événement surgissant dans un champ d'interprétation* ». L'interprétation est liée au contexte, élément crucial dans l'analyse d'Escarpit mais aussi de C. Courbières, pour qui « *le document se construit à partir de son contexte de réception* » (Courbières, 2012 : 39). V. Couzinet opère sur ce point, comme nous l'avons vu, un rapprochement avec la méthodologie d'étude du document en histoire : « *tous les auteurs recommandent de le réinsérer dans le contexte historique et idéologique qui permet de comprendre ses tenants et ses aboutissants* » (Couzinet, 2004 : 23). B. Juanals, constatant que le débat porte sur l'accès technique aux données qui circulent sur la toile, s'inscrit en faux contre « *une représentation utopique d'un sujet sans intériorité, sans histoire, sans contexte socioculturel, sans spécificités psychologiques et affectives, sans désir...* » (Juanals, 2003 : 135). H. Fondin reconnaît que « *les machines ont encore beaucoup de mal (...) à prendre en compte (...) le sens (...) car il associe l'élément et le contexte* » (Fondin, 1998 : 312).

Indicateur numéro 12

Dans une perspective épistémologique, il est tout naturel de veiller aux références présentes dans les travaux scientifiques car une des conditions de la science est de s'appuyer sur les recherches précédentes ou en cours. Le fonctionnement de validation par les pairs subsiste même si les modalités et le rythme des échanges ont évolué. Cet indicateur est donc un indicateur « transversal » à toute entreprise d'analyse épistémologique mais il n'est apparu qu'en dernier dans les éléments que nous avons relevés progressivement car il nous a fallu dans un premier temps prendre connaissance des différents travaux et ce n'est qu'en cumulant les lectures que les affiliations sont apparues plus nettement. Nous avons désormais pris l'habitude de consulter très rapidement la liste bibliographique des ouvrages et articles que nous sélectionnons. À cet égard, il est intéressant de noter qu'un ouvrage comme celui de J-M. Salaün n'en possède pas (les références utilisées sont distillées au gré des notes de bas de page mais ne sont pas recensées à la fin de l'essai). De plus, comme nous l'avons rapporté, B. Simonnot s'étonne de ce que le texte collectif du RTP-DOC « *ne propose aucune référence bibliographique pour approfondir les différentes notions, se coupant ainsi de l'ensemble des travaux dont il souhaite faire la synthèse* » (Simonnot, 2012 : 37). Elle avance l'hypothèse de la difficulté à trouver un terrain d'entente au sein des multiples disciplines représentées dans l'équipe. Il apparaît en effet chimérique d'espérer s'accorder sur les propriétés d'un concept selon des angles de vue différents. Les concepts ont un domaine de validité relativement défini et si les controverses scientifiques sont inévitables parce que constitutives de la démarche scientifique par elle-même, la raison veut qu'elles soient circonscrites. Meyriat pense ainsi que « *les objets (sociaux) n'appartiennent pas de droit à une science plutôt qu'à une autre. Une science ne se définit que partiellement par son objet, mais davantage par ce qu'elle cherche à expliquer dans les objets qu'elle étudie, par les modèles explicatifs qu'elle en propose, on pourrait dire ses paradigmes* » (Meyriat, 1983 : 68).

Nous avons découvert à travers notre approche épistémologique des deux concepts information et document que certains travaux contemporains, se positionnant pourtant en SIC, n'avaient que de lointains rapports avec les « *paradigmes* » des sciences de l'information. Immersée dans les discours ambiants sur « le numérique », nous avons pu émettre dans un premier temps l'hypothèse que le « fait numérique » était l'élément perturbateur remettant en

cause les « *modèles explicatifs* » mais notre analyse épistémologique nous a conduite à repositionner la dispute au regard de la théorie mathématique de l'information et du « brouillage » conceptuel qu'elle induit.

3.3 Analyse des limites et de la validité de la grille épistémologique

3.3.1 Limites de la grille

Nous avons tenté une approche épistémologique qui reste incomplète en raison du nombre restreint de lectures, principalement françaises. Les critères relevés peuvent servir à situer un texte, un propos mais rien n'indique que dans d'autres expressions d'un-e même auteur-e les mêmes termes soient employés et qu'une démarche similaire soit suivie. Il ne s'agit pas de généraliser trop hâtivement les caractéristiques relevées et de vouloir placer dans des cases les chercheur-es.

Concernant l'occurrence de termes, il eût été utile d'appliquer sur les textes lus une méthodologie reposant par exemple sur un outil d'analyse sémantique afin de calculer de façon précise et comparative la présence, l'occurrence ou la cooccurrence des termes relevés. Cependant, le simple relevé de termes, de champs lexicaux, ne saurait suffire en soi puisque les termes peuvent être employés pour mieux s'en démarquer ou être utilisés avec un détournement (volontaire ou non) de sens. Les indicateurs proposés reposent ainsi plus sur une impression d'ensemble, un ressenti que sur des preuves formelles d'occurrence. Le regard porté sur les textes et les discours est nécessairement subjectif et le risque d'interprétation erronée n'est pas exclu.

Nous avons réalisé un test de validité à partir de deux articles écrits par des chercheur-es en Sciences de l'information et de la communication au sujet de la culture informationnelle et de la complexité documentaire : deux approches qui ont pour objet les deux notions fondamentales des Sciences de l'information : le concept d'information et le concept de document. Les résultats de l'analyse montre à ce stade de nos recherches une certaine validité et laisse entrevoir des perspectives que nous allons maintenant exposer.

3.3.2 Validité et perspectives

3.3.2.1 Test de la grille

Le premier article⁴⁸ « épluché » est un texte d'Alexandre Serres, Maître de conférences en Sciences de l'Information et de la Communication, co-responsable de l'URFIST⁴⁹ de Rennes. Il s'intitule : *Une certaine vision de la culture informationnelle*. Le second article⁵⁰ « décortiqué » a été écrit par Caroline Courbières, Professeure en Sciences de l'information et de la communication du LERASS⁵¹ de Toulouse. Il est titré : *Virtualité, représentation, signification : approche de la complexité documentaire*.

Afin d'éviter de renouveler l'énumération des douze indicateurs explicités précédemment qui serait par trop fastidieuse, nous proposons une synthèse croisée des éléments relevés dans les deux articles faisant apparaître deux visions nettement opposées : l'une étant orientée vers les outils et l'autre mettant la question du sens et de la signification au-dessus de toute autre considération.

Nous avons mis en évidence par des indications typographiques certains termes et expressions dans les textes sélectionnés en correspondance avec les indicateurs répertoriés. Consciente que plus les interventions typographiques sur les textes étaient nombreuses, plus la lecture de ceux-ci étaient rendue difficile, nous nous sommes arrêtée à quatre traits distinctifs. La volonté de rester lisible en Noir&Blanc a aussi guidé notre décision de ne pas multiplier les repères typographiques. Les éléments privilégiés par conséquent ont été ceux qui pouvaient être révélateurs d'un maximum des douze indicateurs. Si l'ensemble de ceux-ci ne font pas l'objet de commentaires spécifiques dans les textes testés, huit des douze indicateurs semblent être opérationnels dans l'étude de cas en question.

Ainsi, nous pouvons faire une première remarque d'ensemble qui apparaît visuellement en parcourant les deux articles : la part du discours sur les outils, la technique, « le numérique » est dominante chez A. Serres tandis que pour C. Courbières, ce sont les notions de sens, de signification qui sont prépondérantes. L'un et l'autre l'annoncent d'emblée : A. Serres précise : « *Je centrerai le propos ici sur la première de ces questions, celle de la technique, en développant ma propre approche personnelle* » et C. Courbières postule que

48 Annexe 2

49 Unité Régionale de Formation à l'Information Scientifique et Technique

50 Annexe 3

51 Laboratoire d'Études et de Recherches Appliquées en Sciences Sociales

l'on ne peut « *définir l'objet documentaire en faisant abstraction de la charge signifiante de son contenu informationnel* ».

D'un point de vue épistémologique cependant, les deux auteur-es se positionnant dans le champ des SIC et plus précisément dans celui de l'information-documentation, il est opportun de s'interroger sur la place qu'ils accordent aux concepts fondamentaux information et document. Or, le premier discours gravite autour des notions techniques tandis que le second développe largement les diverses propriétés des deux concepts en question. Même si A. Serres se targue d'adopter une « *réflexion critique contre les mythes de la calculabilité généralisée* », qui n'est autre que l'approche mathématique de l'information, il refuse « *la vision idéelle des processus cognitifs* », qui correspond à l'approche sociale de l'information. Les indicateurs d'occurrence ou de non-occurrence des termes « information » et « document » (n°1), de la prise en compte plus ou moins grande des outils et de l'esprit (n°4) se vérifient ici. L'importance accordée aux notions de « sens », de « signification » est prégnante chez C. Courbières tandis qu'elle est bien moindre chez A. Serres. Le terme central de « *sens* » dans son article a une seule occurrence (dans l'acception d'idée intelligible) et quasiment toutes les autres références à la compréhension, l'appropriation, à la cognition sont en lien direct avec des outils ou techniques et non avec des êtres pensants. On ne retrouve pas systématiquement les termes « données », « flux » ou « flot » chez cet auteur (indicateur n°2) : nous avons relevé une seule occurrence du mot « métadonnées », présentée toutefois avec « *l'indexation* » comme une des seules « *notions stables, pérennes* » susceptibles d'être enseignées dans le cadre de la discipline d'information-documentation. L'insistance portée sur la « *logique interne* », « *profonde* » des outils nous semble aller dans le sens du calcul et de la vision mécaniste (indicateur n°3).

L'indicateur n°3 oppose l'attention accordée à l'esprit humain à la prédilection pour les chiffres et autres formules. La référence aux individus est omniprésente chez C. Courbières qui a le souci constant du « *sujet récepteur* ». Ce « *quelqu'un* » qui se saisit de l'objet documentaire, est susceptible de lui « *attribuer* » une « *fonction informative* » qui est toujours associée à une action d'ordre cognitif : « *regard prêt à interprété* », « *action institutionnalisante du récepteur* », « *la com-préhension* », etc. A. Serres, de son côté, avance une « *formule* » qu'il juge lui-même commode « *les trois R* » mais quand il évoque les acteurs de la communication, il ne les place pas dans une attitude volontaire : ce sont « les jeunes », les « élèves », les « étudiants ». Au contraire, ce sont les Technologies de l'information et de la

communication qui sont dotées d'une « *dimension cognitive, théorique, abstraite* » et il incite à « *penser le numérique* », les « *supports* » et autres « *techniques de l'intelligence* ». Nous trouvons ici l'amalgame entre ce qui relève de la technologie et ce qui relève de l'esprit (indicateur n°4).

L'indicateur suivant (n°5) est particulièrement opératoire pour distinguer l'attitude mécaniste de l'attitude sociale. Les mots employés par C. Courbières sont soigneusement pesés et employés dans un sens précis. Cette exactitude étant à relier à la définition très rigoureuse qu'elle présente des concepts des sciences de l'information : le document est un « *objet signifiant* » qui « *se construit à partir de son contexte de réception* ». C'est la « *fonction informative* » qui est portée par le document, qui ne devient réel qu'à partir du moment où un récepteur lui attribue un sens. Nous avons commenté plus haut l'acception littérale qu'elle fait de la notion de « *virtualité* », ne reprenant pas l'usage impropre qui en fait, comme nous le soulignons en première partie de ce travail, une qualité intrinsèque au « numérique » ou à la toile. Face à cette justesse d'expression (qui est corrélée avec la justesse conceptuelle), A. Serres commet des approximations : il évoque « *l'évaluation des sites webs* » et plus loin « *l'évaluation de l'information trouvée sur le web* ». D'un point de vue épistémologique, les implications conceptuelles de ces deux expressions sont très différentes. De même il « détourne » quelque peu l'usage du terme « *résistance* » qu'il emploie tantôt au sens littéral comme une expression de l'opposition (connotée négativement dans le cadre du refus de la technique), tantôt dans le sens du « *recul critique* » (connoté positivement, dans le cadre de la distance prise avec la technique).

Au sujet du concept d'information, il est intéressant de noter qu'A. Serres, en décalage total avec Meyriat, déplore « *la confusion entre information et connaissance, qui fait des ravages dans les têtes* ». Nous sommes tentée de rapprocher cette remarque de celle d'Escarpit formulée presque 40 ans en arrière : « *Il y a un emploi métaphorique de la théorie mathématique de l'information qui a fait des ravages* » car elle permet de souligner à quel point les deux approches, sociale et mécaniste, sont antinomiques.

Les tonalités respectives des deux articles tendent à corroborer les indicateurs n°6 et n°10 : le propos d'A. Serres est parfois incantatoire voire virulent tout en se défendant d'être idéologique ; tandis que C. Courbières n'éprouve, semble-t-il, pas le besoin d'appuyer ses dires par des considérations péremptoires, voire pessimistes. Ainsi, A. Serres insiste sur la

« *crise* » de l'éducation, le déficit d'autorité des enseignants, le « *retard* » accumulé par la France eu égard à l'usage des technologies. Le « web 2.0. » vient faire « *exploser le monopole du savoir de l'école* » et le nouveau modèle du savoir « *d'une immanence radicale, heurte de plein fouet (...) le fonctionnement des institutions éducatives, mais plus profondément peut-être tout système, tout projet éducatif* ». La crainte du désordre, le refus de l'entropie apparaissent ici de façon sous-jacente et le caractère inéluctable semble donner raison à ceux qui se soumettent en quelque sorte aux outils. Au contraire, l'approche philosophique de C. Courbières qui considère le document comme un « *horizon d'attente* » laisse toute latitude à l'imprévisibilité et ouvre le champ des possibles.

Nous terminerons par ce qui est le plus aisé à relever : les références (citations dans le texte et liste bibliographique), cet indicateur (n°12), comme nous l'expliquions plus haut ne nous est venu à l'esprit qu'à la fin de nos lectures préparatoires à la rédaction de la partie théorique de notre mémoire. Il est difficile en effet d'avoir une idée de la diversité des points de vue épistémologiques, a priori. En revanche, cet indicateur prend tout son sens en dernière analyse car C. Courbières se situe dans la filiation de J. Meyriat, de R. Escarpit, d'Y. Jeanneret ou encore de V. Couzinet. Tandis qu'A. Serres ne cite pas de chercheur-es en SIC mais des philosophes tels B. Stiegler dont la réflexion actuelle tourne autour des technologies numériques ou R. Debray, fondateur de la médiologie.

Certes, les articles n'ont pas été publiés avec les mêmes visées communicationnelles, ni dans des contextes similaires mais les auteur-es s'expriment tout deux en tant que membres des SIC et abordent pour l'un la culture informationnelle et pour l'autre la complexité documentaire. Aussi, dans une approche épistémologique, telle que la nôtre, il a paru légitime de leur faire passer le test de la grille que nous avons élaborée. Tous les indicateurs ne se sont pas révélés du même niveau de pertinence mais plus des deux tiers ont pu être vérifiés sur les deux articles. En observant la concordance des différents indicateurs, il est loisible de déterminer des deux visions, laquelle est prédominante. Dans l'essai présent, les huit indicateurs convergeaient et si nous n'avons pu vérifier les indicateurs n°7, n°8 et n°9 et 11 ; notre analyse ne nous a pas conduite à invalider ces derniers. Ainsi C. Courbières présente une approche sociale des concepts information et document qui s'oppose à la vision mécaniste d'A. Serres.

3.3.2.2 Perspectives d'utilisation de la grille et de prolongement du travail de recherche

La grille a pour ambition de donner quelques clés de lecture afin de repérer quel chemin théorique les différents auteur-es empruntent de façon délibérée ou involontaire. La pensée scientifique n'est pas détachée du contexte socio-historique, culturel dans lequel elle s'inscrit et elle évolue tout naturellement. Il serait intéressant de tester de façon approfondie la grille sur plusieurs textes scientifiques et de prendre plusieurs extraits d'époques variées pour un-e même auteur-e. Nous avançons l'hypothèse que plus le pourcentage de critères relevant de l'une des deux colonnes est élevé plus les remarques épistémologiques sont pertinentes pour évaluer le courant auquel le-la chercheur-e appartient.

Dans notre démarche personnelle, le fait d'appréhender les concepts des sciences de l'information en gardant à l'esprit cette scission entre vision mécaniste (l'information est une donnée, neutre, autonome, qui se prête à des calculs de probabilité) et vision sociale (l'information est une connaissance communiquée ou communicable dans le cadre du processus communicationnel qui suppose à la fois échange entre êtres humains et attribution de sens) nous a été d'une aide précieuse. Nous avons pu ainsi nous détacher du discours ambiant sur la rupture radicale induite par la technologie numérique et adopter une attitude raisonnée envers les concepts. À travers les controverses épistémologiques, cependant, ce sont finalement deux visions du monde qui émergent.

Le paradoxe soulevé à l'aube de notre réflexion sur la difficulté à enseigner l'information-documentation dans un cadre de surabondance informationnelle n'est pas réglé mais les termes du débat semblent mieux définis. Nous pourrions également envisager un travail praxéologique auprès des professeur-es-documentalistes afin de construire un outil susceptible d'aider à forger une culture des SIC, qui reste encore lacunaire (comme nous le rappelions en introduction). La délimitation des concepts, leur mise en relation et leurs caractéristiques respectives sont autant de connaissances indispensables à posséder pour effectuer la transposition didactique, le passage du savoir savant au savoir enseigné.

Conclusion

L'enseignement des notions d'information et de document est particulièrement difficile en ces temps de mise en avant de l'information numérique. Pourtant l'information semble être une valeur essentielle de la société contemporaine puisqu'elle sert à la qualifier. Ceux qui n'en sont (toujours) pas convaincus n'ont que trop tardé et l'école, au premier chef, est sommée d'entrer dans « l'ère du numérique » à travers le plan numérique pour l'éducation. Pour quelles raisons, dans ces conditions, la liaison s'effectue-t-elle si imparfaitement entre les professionnels de l'information et les jeunes qui sont nés à cette époque nouvelle ? Les concepts des sciences de l'information n'auraient-ils plus de validité opératoire du fait de la mutation technologique en cours ? Ou bien au contraire ces concepts pourraient-ils permettre une meilleure compréhension des enjeux sociétaux de la période ?

La première intention de notre démarche de recherche a été d'appréhender sous l'angle « numérique », les deux concepts fondamentaux des sciences de l'information : information et document. Ce dernier étant particulièrement remis en cause dans les controverses scientifiques ayant pour arrière-plan la dimension numérique. Notre participation au GAP documentation, par le biais de la recherche-action articulée autour de l'analyse de ces concepts élaborés par les théoriciens fondateurs des sciences de l'information et de la communication en France, a nourri notre réflexion et amené cette question de départ consistant à interroger les problématiques actuelles du concept information à l'heure numérique à la lumière des caractéristiques originelles du concept document.

Il nous est apparu très rapidement qu'il était nécessaire avant toute observation de cerner précisément ce qui se cache derrière l'expression « le numérique ». Les premières ambiguïtés ont émergé car ce vocable est particulièrement équivoque et laisse place, nonobstant les discours officiels et médiatiques qui font comme si tout le monde s'accordait naturellement sur ce qu'il implique, à toutes les divagations. La critique opérée par Yves Jeanneret il y a quinze ans à propos de l'expression « nouvelles technologies de l'information et de la communication », voile sémantique masquant une idéologie bien réelle guidée par les intérêts économiques, est particulièrement édifiante pour ce qui ressemble de plus en plus au mythe d'une « société du numérique ». Nous avons montré que ce terme flou est récusé par plusieurs chercheur-es et, pour ce qui concerne notre recherche, nous renchérissons sur l'emploi à mauvais escient de ce terme qui au sens littéral concerne un procédé technique qui

n'est « vraiment » plus nouveau (par opposition à l'analogique). Pourtant ce qui entraîne les débats passionnés depuis quelques décennies c'est, dans les pays occidentaux du moins, l'utilisation massive du réseau, ou de la toile, et non pas, stricto sensu, la technologie numérique.

Le phénomène numérique, objet de toutes les attentions, dissimule donc plutôt la question du réseau. Ainsi la question s'est renouvelée en ces termes : les concepts fondamentaux information et document offrent-ils une résistance face au « fait numérique » ou bien l'évolution de la société nécessite-t-elle une refondation des théories des sciences de l'information ? En faisant le point sur l'acception du terme numérique, nous nous sommes heurtée à une nouvelle ambiguïté sémantique de taille concernant le concept information. Le « numérique » signifiant numérisation des informations. Ces dernières étant prises au sens de données, mesurables, calculables.

Or, en posant les bases conceptuelles des sciences de l'information, Meyriat (un des fondateurs des sciences de l'information) a regretté cette appropriation par les mathématiciens, et par les informaticiens après eux, de ce terme qui intéresse les professionnels de l'information mais au sens de l'information-connaissance (Meyriat, 1978). L'orientation de notre travail a pris dès lors un tour épistémologique et la décision a été prise d'effectuer une approche épistémologique des concepts en intégrant cette dimension numérique.

Un des obstacles majeurs dès l'origine pour appréhender les concepts étudiés dans ce mémoire est donc dû à la théorie mathématique de l'information, élaborée par des ingénieurs au milieu du XX^e siècle. Ces recherches menées sur la capacité optimale d'un canal de transmission avaient pour objectif l'amélioration de la rapidité et de l'efficacité des techniques de communication (Shannon, 1975). Elles ont abouti à la création d'une unité de mesure, baptisée « information » qui a été l'objet de toutes les extrapolations dans les sciences humaines, entre autres. Cette théorie de surcroît est en adéquation totale avec les techniques de numérisation et les techniques de communication dont la rapidité et la capacité volumique ne cessent de croître conformément à ce que Robert Escarpit nommait la recherche du rendement dans la communication (Escarpit, 1976).

Nous nous sommes donc penchée sur cette théorie et sur l'influence qu'elle exerce actuellement au sein des chercheur-es en SIC. Nous avons présenté ensuite les propriétés du concept central pour les sciences de l'information qu'est celui d'information dont le nœud

fondamental tourne autour de la signification et de l'échange entre êtres humains. Vision sociale du concept toujours accréditée par certain-es chercheur-es contemporain-es à l'heure dite numérique.

La césure épistémologique est plus nette, comme nous l'avons pointé, au sujet du second concept analysé, celui de document. Nous sommes revenue sur les bases théoriques posées dès le début du XX^e siècle par Paul Otlet et développées par Suzanne Briet après la deuxième guerre mondiale. Elles ont été reprises à partir de la fin des années 70 par les fondateurs des sciences de l'information, Jean Meyriat et Robert Escarpit, notamment, qui ont placé ce concept en position subordonnée par rapport au concept prioritaire d'information. Nous avons ensuite rendu compte des controverses actuelles au sujet du document numérique, considéré comme un nouveau concept par plusieurs chercheur-es. Des éléments mouvants apparaissent qui remettent en cause certains traits du concept de document, telles la stabilité, la temporalité, l'unité documentaire, la fonction documentaire. Cependant les arguments avancés pour apporter une nouvelle définition du concept font, d'un point de vue épistémologique, insuffisamment la part entre ce qui relève du technologique et ce qui relève du scientifique. Les tentatives de nouvelle théorisation du concept document apparaissent comme un reflet de l'influence grandissante de la technique dans l'approche scientifique de l'information. La rigueur théorique conduit d'ailleurs certain-es chercheur-es contemporain-es à recourir aux bases conceptuelles qui, selon eux, sont particulièrement opératoires pour appréhender le concept de document à l'heure numérique, notamment la bipartition entre « document par intention » et « document par attribution » (Couzinet, 2001).

Au fil des lectures, nous avons tenté de lister les indicateurs susceptibles de lever les ambiguïtés sémantiques et conceptuelles liées aux deux concepts étudiés. Nous avons ainsi remarqué que les divergences d'idées portaient sur l'exclusion ou l'intégration de la théorie mathématique de l'information. Nous avons donc proposé une grille de lecture épistémologique organisée autour des éléments distinguant la vision sociale (information-connaissance) de la vision mécaniste (information-donnée) des deux concepts avec une douzaine d'indicateurs dénotant l'influence, consciente ou non, de la théorie incriminée. La grille ainsi confectionnée présente certaines limites de validité que nous avons répertoriées, dont les principales sont le risque de simplification dichotomique et la tentation de placer les

chercheur-es dans des cases. Mais en tenant compte de ces précautions d'usage, le test que nous avons effectué, en passant deux articles au crible des indicateurs, nous donne l'espoir que ce modèle épistémologique puisse contribuer à donner des clés de compréhension aux professionnel-les de l'information, tel-les les professeur-es-documentalistes qui souhaiteraient pouvoir comprendre les ressorts épistémologiques des débats actuels autour de ces concepts et envisager d'un point de vue plus pratique son éventuel usage pour confectionner les textes de savoir à enseigner.

En prolongement de ce travail, nous pensons intéressant d'approfondir l'analyse épistémologique en affinant la grille et en la testant sur un corpus de textes scientifiques dans la perspective de contribuer, à notre niveau, à la poursuite du cadrage théorique des sciences de l'information.

Liste bibliographique

- BACHIMONT, Bruno. 2012. *Pour une critique phénoménologique de la raison computationnelle*. [Consulté le 20 novembre 2014]. Disponible sur : <http://www.ina-expert.com/e-dossier-de-l-audiovisuel-l-education-aux-cultures-de-l-information/pour-une-critique-phenomenologique-de-la-raison-computationnelle.html>
- BALTZ, Claude. 2013. Numérique : « Meurtre » de l'information ?... In SALEH Imad, ZACKLAD Manuel, LELEU-MERVIEL Sylvie et alii. *Pratiques et usages du numérique : Acte de H2PTM'13*. Paris: Lavoisier, Hermès. p. 3-12..
- BALTZ, Claude. 2011. *Cyber, informationnelle, numérique...A nouvelles économies, nouvelles cultures ?* Décembre 2011. [consulté le 24 novembre 2014]. Disponible sur : <http://www.ina-expert.com/e-dossier-de-l-audiovisuel-l-education-aux-cultures-de-l-information/cyber-informationnelle-numerique-a-nouvelles-economies-nouvelles-cultures.html>
- BARREAU, Hervé. 2002. *L'épistémologie*. Paris : Presses universitaires de France, 5^e édition. 127 p. (Que sais-je ? 1475)
- BOULOGNE, Arlette. 2004. *Vocabulaire de la documentation*. Paris : ADBS éditions. 334 p.
- BRIET, Suzanne. 1951. *Qu'est-ce que la documentation ?* Paris : ÉDIT (Éditions documentaires industrielles et techniques). 48 p.
- BUCKLAND, Mickael. 1997. What is a document ? *Journal of the American Society for Information Science* (1986-1998), 09/1997, p.804. Disponible sur : <http://www.columbia.edu/cu/libraries/inside/units/bibcontrol/osmc/bucklandwhat.pdf>
- CACALY, Serge (dir). 1997. *Dictionnaire encyclopédique de l'information et de la documentation*. Paris : ADBS ; Nathan. 634 p. (Collection « ref »).
- CAPURRO, Rafael. 2006. *Ethik der Informationsgesellschaft : ein interkultureller Versuch*. Beitrag zur Tagung [Shapes of Things to Come](#), Humboldt-Universität, Berlin, 15.-17. Februar. [consulté le 7 octobre 2014]. Disponible sur : <http://www.capurro.de/parrhesia.html>

- CASATI, Roberto. 2013. *Contre le colonialisme numérique : manifeste pour continuer à lire*. Paris : Albin Michel. 208 p. (Bibliothèques Idées).
- COTTE, Dominique. 2004. Le concept de document numérique. *Communication & langages*, juin 2004, n°140, p. 31-41. Disponible sur :
http://www.persee.fr/web/revues/home/prescript/article/colan_0336-1500_2004_num_140_1_3265
- COURBIÈRES, Caroline ; FRAYSSE, PATRICK ; GARDIÈS, Cécile. 2007. Distance et immédiateté : incidences du document numérique sur le traitement de l'information. *Études de communication : langages, information, médiations*, n°30, p. 71-81. Disponible sur :
<http://edc.revues.org/478>
- COURBIÈRES, Caroline. 2012. Virtualité, représentation, signification : approche de la complexité documentaire. *Études de communication : langages, information, médiations*, n°39, p. 103-115. Disponible sur :
<http://www.cairn.info/revue-etudes-de-communication-2012-2-page-103.htm>
- COUZINET, Viviane. 2001. *Jean Meyriat, théoricien et praticien de l'information-documentation*. Paris : ADBS éditions. 512 p.
- COUZINET, Viviane. 2004. Le document : leçon d'histoire, leçon de méthode. *Communication & langages*, juin 2004, n°140, p. 19-29. Disponible sur :
http://www.persee.fr/web/revues/home/prescript/article/colan_0336-1500_2004_num_140_1_3264
- DAVALLON, Jean ; JEANNERET, Yves. 2004. La fausse évidence du lien hypertexte. *Communication & langages*, juin 2004, n°140, p.43-54. Disponible sur :
http://www.persee.fr/web/revues/home/prescript/article/colan_0336-1500_2004_num_140_1_3266
- DERVIN, Brenda ; DEWDNEY, Patricia. 1986. Neural questioning : a new approach to the reference interview. *Research Quarterly*, vol 25, n°4, p. 506-513. Disponible sur :
<http://www.jstor.org/discover/10.2307/25827718?uid=3738016&uid=2&uid=4&sid=21106617145843>

- DION, Emmanuel. *Invitation à la théorie de l'information*. Paris : éditions du Seuil. 1997, 156 p. (Points)
- ESCARPIT, Robert. (1976) *Théorie générale des sciences de l'information et de la communication*. Paris : Hachette. 218 p. (Hachette université).
- ESTIVALS, Robert (dir.). 1983. *Les sciences de l'écrit : encyclopédie internationale de bibliologie*. Paris : Éditions Retz. (Références).
- FONDIN, Hubert. 1998. *Le traitement numérique des documents*. Paris : Hermès. 382 p.
- GARDIÈS, Cécile. 2006. *De la mouvance identitaire à l'ancrage scientifique des professionnels de l'information-documentation dans l'Enseignement agricole*. Université Toulouse III Paul Sabatier. 374 p.
- GARDIÈS, Cécile (dir.). 2011. *Approche de l'information-documentation : concepts fondateurs*. Toulouse : Cépaduès éditions. 232 p.
- GARDIÈS, Cécile. 2012. *Dispositifs info-communicationnels de médiation des savoirs : cadre d'analyse pour l'information-documentation*. [HDR] École nationale de formation agronomique.
- GENIN, Christophe. 2004. "Culture numérique": une contradiction dans les termes ? In *Patrimoine et culture numérique*. Berlin 31 août-02 septembre. [consulté le 20 novembre 2014]. Disponible sur : http://www.archimuse.com/publishing/ichim04/0461_Genin.pdf
- JEANNERET, Yves. 2000. *Y a-t-il (vraiment) des technologies de l'information ?* Villeneuve-d'Ascq : Presses universitaires du Septentrion. 134 p. (Les Savoirs mieux,).
- JUANALS, Brigitte. 2003. *La culture de l'information : du livre au numérique*. Paris : Hermès science publications. 243 p.
- KREMER-MARIETTI, Angèle ; DHOMBRES, Jean. 2006. *L'épistémologie : état des lieux et positions*. Paris : Ellipses. 111p.
- LAMIZET, Bernard ; SILEM, Ahmed. 1997. *Dictionnaire encyclopédique des sciences de l'information et de la communication*. Paris : Ellipses. 590 p.
- LE COADIC, Yves-François. 2004. *La science de l'information*. Paris : Presses universitaires de France 3e édition. 128 p. (Que sais-je ?, 2873)

- MERZEAU, Louise. *Internet et fraternité*. [consulté le 20 septembre 2014]. Disponible sur : <http://mediologie.org/tribune/internet-fraternite/merzeau.html>
- METZGER, Jean-Paul ; LALLICH-BOIDIN, Geneviève. 2004. Temps et documents numériques. *Document numérique*, avril, vol 8, p.11-21. [consulté le 13/10/2014]. Disponible sur : <http://www.cairn.info/revue-document-numerique-2004-4-page-11.htm>
- METZGER, Jean-Paul. 2008. Temps, mémoire et document. In PAPY, Fabrice. *Problématiques émergentes dans les sciences de l'information*. Paris : Lavoisier. p. 87-109.
- MEYRIAT, Jean. 1978 . De l'écrit à l'information : la notion de document et la méthodologie de l'analyse du document. In *Inforcom 78*, Société française des sciences de l'information et de la communication, premier Congrès, Compiègne. Paris : SFSIC, Vol 1. p. 23-32.
- MEYRIAT, Jean. 1981 . Document, documentation, documentologie. *Schéma et schématisation*, 2ème trimestre, n°14, p. 51-63.
- MEYRIAT, Jean. 1981 . L'informatologie, science sœur de la bibliologie.. *Schéma et schématisation*, 2ème trimestre, n°15, p. 9-19.
- MEYRIAT, Jean. 1983 . De la science de l'information aux métiers de l'information. *Schéma et schématisation*, n°19, p.65-74.
- MEYRIAT, Jean. 1985 . Information vs communication ?. In LAULAN, A-M. *L'espace social de la communication : concepts et théories*. Paris : Retz-CNRS. p. 63-89 *Schéma et schématisation*, 2ème trimestre, n°14, p. 51-63.
- MEYRIAT, Jean. 1986 . La science de l'information. *Revue des sciences morales et politiques*, n°2, p. 255-269.
- MEYRIAT, Jean. 1991 . Vers une troisième civilisation de l'écrit. In *Le livre dans la Caraïbe : systèmes de production et de diffusion*, actes du XX^e congrès de l'Association des bibliothèques universitaires, de recherche et institutionnelles de la Caraïbe (ACURIL), Caracas, Vénézuéla, 19-25 mai 1991, p. 5-14.
- MEYRIAT, Jean. 1993 . Un siècle de documentation : la chose et le mot. *Revue de bibliologie : schéma et schématisation*, n°40, p. 39-45.
- MEYRIAT, Jean. 1994 . Y a-t-il une place pour une théorie de la documentation ?. *Documentaliste – Sciences de l'information*, juillet-octobre, vol. 30, n°4-5, p. 192-198.

- MOATTI, Alexandre. 2012. Le numérique, adjectif substantivé. *Le Débat*, mai-juin, n°170, p. 133-137.
- OTLET, Paul. 1934 *Traité de documentation : le livre sur le livre, théorie et pratique*. Bruxelles : Ed. Mundaneum. 431 p.
- PAPY, Fabrice (dir.). 2008. *Problématiques émergentes dans les sciences de l'information*. Paris : Lavoisier. 351 p.
- PEDAUQUE, R.T. 2007. *La redocumentarisation du monde*. Toulouse : Cépaduès éditions. 213 p.
- *Qu'est-ce que la médiologie ?* [consulté le 20 septembre 2014]. Disponible sur : <http://mediologie.org/presentation/>
- QUIVY, Raymond ; VAN CAMPENOUGHT, Luc. 2011. *Manuel de recherche en sciences sociales*. Paris : Dunod. 262 p. (4^e édition)
- REY, Alain. 2010. *Dictionnaire historique de la langue française*. Paris : Éditions Le Robert. 2617 p.
- SALAÛN, Jean-Michel. 2004. Chronique inachevée d'une réflexion collective sur le document. *Communication & langages*, juin 2004, n°140, p.9-17. Disponible sur : http://www.persee.fr/web/revues/home/prescript/article/colan_0336-1500_2004_num_140_1_3263
- SALAÛN, Jean-Michel. 2003. *Documents et numérique*. Version 2 , 1^{er} décembre. [consulté le 13/10/2014]. Disponible sur : <http://archivesic.ccsd.cnrs.fr/docs/00/06/23/26/HTML/index.html>
- SALAÛN, Jean-Michel. 2012. *Vu, lu, su : les architectes de l'information face à l'oligopole du Web*. Paris : La Découverte. 151 p. (Cahiers libres)
- SHANNON, Claude E. ; WEAVER, Warren. 1975. *Théorie mathématique de la communication*. Paris : Éditions Retz et C.E.P.L. 188 p. (Les Classiques des Sciences Humaines)
- SIMONNOT, Brigitte. 2007. Évaluer l'information. *Documentaliste-Sciences de l'Information*, 2007/3, Vol. 44, p. 210-216. Disponible sur : <http://www.cairn.info/revue->

documentaliste-sciences-de-l-information-2007-3-page-210.htm

- SIMONNOT, Brigitte. 2012. *L'accès à l'information en ligne : moteurs, dispositifs et médiations*. Cachan : Hermès science publications. 249 p.

Table des matières

Introduction.....	5
Chapitre 1 Le concept information à l'ère numérique : la tentation pléonastique.....	9
1.1 Les relations ambiguës entre information et numérique.....	9
1.1.1 Qu'est-ce que "le numérique" ?.....	9
1.1.2 "Le numérique" : de la technique au « concept ».....	12
1.2 La théorie mathématique de l'information : un pavé dans la mare des SIC.....	16
1.2.1 La théorie mathématique de l'information : l'information comme donnée mesurable.....	16
1.2.2 La théorie mathématique de l'information : influences contemporaines en SIC....	23
1.3 L'information-connaissance en SIC : priorité à la question du sens.....	26
1.3.1 Les théoriciens fondateurs des sciences de l'information.....	26
1.3.2 Une information humaine, sociale, vivante avant d'être numérique : les héritiers des fondateurs des sciences de l'information.....	30
Chapitre 2 Le fait numérique : un perturbateur conceptuel pour le document en SIC ? Où la théorie mathématique de l'information resurgit.....	36
2.1 Les bases conceptuelles du document en SIC.....	36
2.1.1 Les précurseurs : la notion d'extensivité et de preuve.....	36
2.1.2 Le lien indissoluble entre le document et l'information.....	38
2.1.2.1 Une relation non équivalente entre les deux concepts.....	38
2.1.2.2 La bipartition documents par intention – document par attribution.....	40
2.1.2.3 L'indépendance par rapport au temps.....	42
2.2 Les caractéristiques controversées du concept de document à l'heure numérique.....	46
2.2.1 Les travaux du RTP-DOC : le document numérique : un nouveau concept ?.....	46
2.2.1.1 Qui est Roger Pedauque ?.....	46
2.2.1.2 La démarche et les conclusions du réseau thématique RTP-DOC.....	47
2.2.2 De l'objet stable à l'entité disparate.....	49
2.2.2.1 Dissociation support-contenu ?.....	49
2.2.2.2 La nouvelle temporalité du document.....	55
2.2.3 De l'unité documentaire à l'unité informationnelle.....	59
2.2.3.1 La fragmentation des contenus.....	59
2.2.3.2 La fonction documentaire : le document, un prestataire de service ?.....	66
2.3 Le document numérique, révélateur de caractéristiques originelles du concept document.....	71
2.3.1 Le document, objet informationnel avant d'être numérique.....	71
2.3.2 La double partition.....	74
2.3.3 La valeur sociale du document.....	78

Chapitre 3 Analyse épistémologique des concepts en SIC : éléments d'interprétation pour distinguer le savoir scientifique du « fait numérique ».....	81
3.1 Point sur l'analyse épistémologique.....	81
3.1.1 L'épistémologie : étude critique des sciences.....	81
3.1.2 L'épistémologie des sciences de l'information.....	82
3.2 Grille de lecture épistémologique.....	83
3.2.1 Méthodologie de construction de la grille.....	83
3.2.2 Présentation de la grille d'analyse épistémologique.....	85
3.3 Analyse des limites et de la validité de la grille épistémologique.....	100
3.3.1 Limites de la grille.....	100
3.3.2 Validité et perspectives.....	101
3.3.2.1 Test de la grille.....	101
3.3.2.2 Perspectives d'utilisation de la grille et de prolongement du travail de recherche.....	105
Conclusion.....	106
Liste bibliographique.....	109
Index lexical.....	118
Index des auteurs.....	119
Répertoire des sigles.....	120
Sommaire des annexes.....	121
Annexe 1 Schéma de Shannon.....	122
Annexe 2 Test de la grille épistémologique : une vision mécaniste (intégration de la théorie mathématique de l'information).....	123
Annexe 3 Test de la grille épistémologique : une vision sociale (exclusion de la théorie mathématique de l'information).....	136

Index lexical

- communication. .10, 13, 15, 20, 34, 37., 41, 49, 52, 54, 71, 74, 77, 79, 84, 86, 89, 94, 99, 104
- connaissance. 4, 11, 19, 22, 25, 28, 35, 40, 46, 55, 61, 66, 68, 74, 76, 83, 86, 89, 94, 100, 102
- document numérique.....9, 14, 43, 49, 63, 72, 88, 105
- document par attribution.....44, 55, 72, 76, 105
- document par intention.....44, 55, 72, 76, 105
- document primaire.....5, 50, 61, 71, 75
- document secondaire.....5, 54, 71, 74
- donnée.....8, 12, 20, 22, 24, 28, 47, 50, 56, 58, 60, 62, 66, 69, 76, 82, 85, 88, 94, 99, 102, 104
- électronique.....8, 14, 21, 26, 47, 50, 58, 61, 72, 74, 87
- immatériel.....22, 34, 69
- information numérique.....4, 9, 13, 23, 28, 44, 103
- information-connaissance.....20, 24, 31, 52, 55, 68, 77, 82, 89, 104
- information-donnée.....31, 105
- matérialité.....41, 46, 52, 56, 84
- matériel.....16, 22, 34, 36, 40, 44, 51, 56, 58, 64, 69, 71, 73, 75
- processus de communication.....17, 23, 27, 42, 46, 49, 55, 65, 70
- savoir.....10, 13, 19, 25, 28, 39, 41, 48, 56, 79, 84, 89, 92, 101
- sens.....9, 13, 17, 21, 23, 29, 40, 44, 49, 53, 60, 66, 69, 75, 77, 79, 81, 85, 87, 92, 94, 98, 102
- signification.....15, 21, 23, 29, 33, 37, 42, 46, 54, 56, 70, 77, 82, 84, 90, 98, 105, 108
- stabilité.....40, 46, 67, 74, 76, 88, 105
- support.....9, 11, 13, 22, 26, 37, 45, 49, 55, 58, 63, 68, 72, 74, 87, 100
- virtuel.....8, 13, 23, 37, 61, 72, 88

Index des auteurs

Bachimont.....	11
Baltz.....	13, 32, 91
Barreau.....	78
Boulogne.....	9
Gardiès.....	5
Rey.....	8
Cacaly.....	8, 9
Briet.....	35, 36, 48, 71, 84, 86
Buckland.....	35
Cacaly.....	21, 22, 92
Casati.....	51, 52, 60, 63
Cotte.....	49, 50, 51, 55
Courbières.....	68, 76, 88, 94, 98, 100, 101
Couzinet.....	52, 68, 69, 70, 71, 72, 73, 74, 75, 76, 79, 84, 86, 94, 95, 99
Davallon.....	69, 70, 94
Dervin.....	31, 92, 94
Dion.....	15, 17, 18, 20, 91
Escarpit.....	18, 19, 27, 38, 39, 40, 41, 42, 59, 62, 73, 84, 92, 100
Estivals.....	24, 72
Fondin.....	9, 10, 47, 48, 49, 95
Gardiès.....	69, 79, 80
Genin.....	12, 13
Jeanneret.....	13, 20, 28, 29, 30, 51, 69, 84, 87, 91, 94
Juanals.....	28, 29, 30, 56, 58, 92, 95
Kremer.....	78
Lallich-Boidin.....	51
Lamizet.....	9, 89
Le Coadic.....	9, 22, 23, 87
Merzeau.....	94
Metzger.....	51, 52, 53, 54, 55, 56, 94
Meyriat.....	13, 24, 25, 26, 28, 30, 31, 32, 33, 35, 36, 37, 38, 40, 42, 46, 50, 52, 55, 66, 69, 79, 86, 87, 89, 90, 91, 92, 93, 94, 95
Moatti.....	11
Moles.....	17, 86
Otlet.....	34, 68
Papy.....	55, 56
Pedauque.....	44, 45, 46, 47, 52, 53, 54, 59, 61, 62, 63, 64, 89, 93
Pomart.....	21, 91
Quivi.....	34
Regimbeau.....	68
Salaün.....	44, 45, 60, 65, 66, 67, 71, 85, 86, 88, 90, 92, 93, 94, 95
Serres.....	98, 99, 100, 101
Shannon.....	15, 16, 20, 21
Simonnot.....	30, 31, 32, 44, 56, 57, 58, 60, 92, 93, 95
Weaver.....	19, 20, 21

Répertoire des sigles

ADBS : Association des professionnels de l'information et de la documentation

BU : Bibliothèque universitaire

CDI : Centre de documentation et d'information

CNAM : Conservatoire nationale des arts et métiers

CNRS : Conseil national de la recherche scientifique

DOI : Digital object identifier (identifiant numérique d'objet)

EMI : Éducation aux médias et à l'information

ENFA : École nationale de formation agronomique

GAP : Groupe d'animation et de professionnalisation

INTD : Institut national des techniques de la documentation

LERASS : Laboratoire d'études et de recherches appliquées en sciences sociales

MIT : Massachusetts Institute of Technology

OCR : Optical character recognition (reconnaissance optique de caractères)

PDF : Portable Document Format (format de document multiplateforme)

RTO : Ressources terminologiques ou ontologiques

RTP-DOC : Réseau thématique pluridisciplinaire « Documents et contenu : création, indexation, navigation »

SFSIC : Société française des Sciences de l'information et de la communication

SHS : Sciences humaines et sociales

SIC : Sciences de l'information et de la communication

SRI : Système de recherche d'information

URFIST : Unité régionale de formation à l'information scientifique et technique

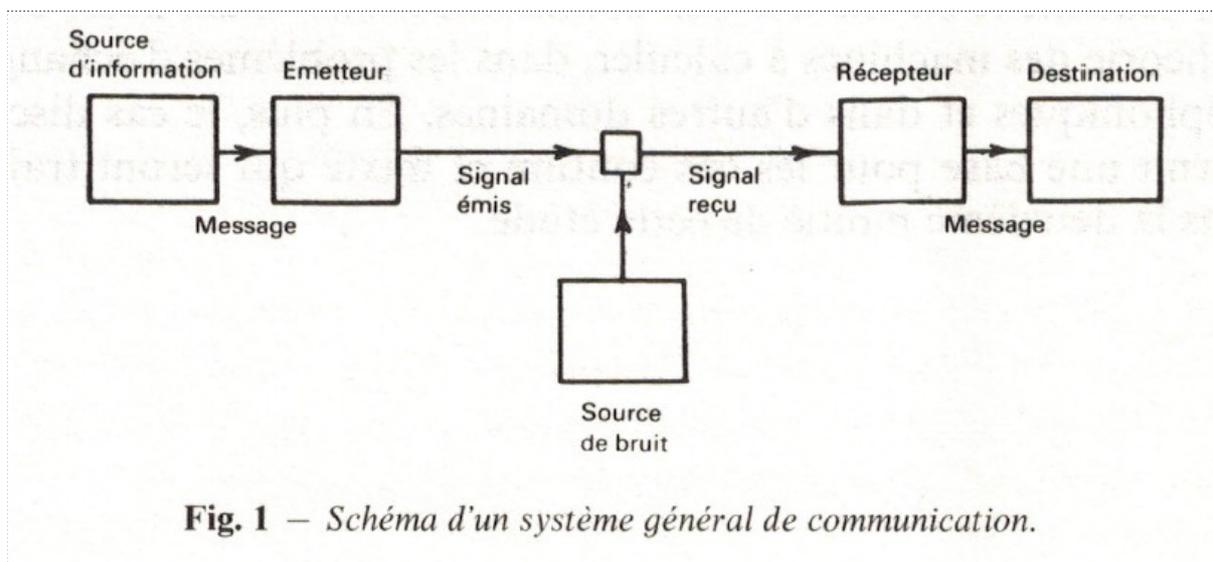
XML : Extensible Markup Language (langage à balise extensible)

Sommaire des annexes

Annexe 1 : schéma de Shannon.....	122
Annexe 2 : test de la grille épistémologique : une vision mécaniste (intégration de la théorie mathématique de l'information).....	123
Annexe 3 : test de la grille épistémologique : une vision sociale (exclusion de la théorie mathématique de l'information).....	136

Annexe 1

Schéma de Shannon



in SHANNON, Claude. E. ; WEAVER, Warren. *Théorie mathématique de la communication*. Paris : Retz, 1975. 188 p. Fig.1 – Schéma d'un système général de communication, p.69.

Annexe 2

Test de la grille épistémologique : une vision mécaniste (intégration de la théorie mathématique de l'information)

Mise en évidence des indicateurs de la grille épistémologique dans le texte ci-dessous :

- Termes grisés : occurrences du mot information
- Termes encadrés : occurrences du mot document
- Termes et expressions en gras : éléments se rapportant au sens, à l'esprit, à la signification, aux opérations cognitives
- Termes et expressions soulignés : éléments se rapportant à la technique, aux outils, à la technologie

Une certaine vision de la culture informationnelle

Alexandre Serres,

Maître de conférences en Sciences de l'Information et de la Communication

Ce texte reprend et développe une intervention faite lors du séminaire « *Enseignement et médias* », organisé le 16 mai 2009 à Paris par Ars Industrialis, le CIEM et la revue Skhole.fr.

Avant de présenter ma propre approche de la culture informationnelle, il convient peut-être de donner quelques repères sur cette expression ambiguë, qui connaît une certaine vogue actuellement et fait l'objet de débats et de recherches de plus en plus nombreuses.

Pendant plusieurs années, « culture informationnelle » pouvait être considérée comme l'une des traductions possibles en français de l'expression anglo-saxonne, presque intraduisible, d'*information literacy*. Certes, en France « maîtrise de l'information » a été longtemps préférée et reste la traduction la plus répandue, mais au Canada francophone, les Québécois utilisent « culture de l'information » comme équivalent de *information literacy*. Parler de culture informationnelle implique donc d'évoquer, au moins brièvement, cette notion matricielle d'*information literacy*, sur laquelle deux observations peuvent être faites, au regard de sa courte histoire un peu plus que trentenaire.

Si l'idée de « maîtriser l'information » est apparue en 1974 aux États-Unis dans le monde de l'entreprise, pour désigner les nouvelles compétences informationnelles nécessaires aux futurs travailleurs du savoir, elle a été immédiatement développée dans deux autres sphères : le monde des bibliothèques, devenu son principal terreau de croissance, et le monde politique, institutionnel, qui lui a conféré une dimension citoyenne. Trois origines peuvent donc être relevées, donnant à cette notion vague de « maîtrise de l'information », une triple dimension,

à la fois économique, bibliothéconomique et citoyenne, qui la caractérise toujours aujourd'hui[1].

La deuxième observation sur l'histoire de cette notion concerne son institutionnalisation progressive, à travers les nombreuses et diverses déclarations, textes, recommandations d'instances internationales et politiques (UNESCO, IFLA, Commission Européenne, ministères de l'éducation, etc.), allant souvent de pair avec une « idéologisation » croissante : au début simple ensemble de compétences méthodologiques et documentaires, la maîtrise de l'**information** est devenue rapidement « *l'une des « cinq habiletés essentielles » pour pouvoir intégrer le marché du travail dans l'avenir* »[2], puis rien moins que la « *condition de survie et de succès dans la société de l'information* »[3], notamment par son assimilation à « la formation tout au long de la vie », et enfin elle a été considérée comme l'un des nouveaux « *droits humains de base dans un monde numérique* »[4].

Autrement dit, ce qui caractérise aujourd'hui le discours dominant sur *l'information literacy* et sa version française de « maîtrise de l'**information** » est, entre autres, la dimension idéologique d'un discours *d'adaptation* à la société dite « de l'**information** », avec un inévitable effet de tautologie : « il faut former à l'**information** pour former à la société de l'**information** ». Il est intéressant de relever la convergence, voire la congruence, entre les deux expressions (de maîtrise et de société de l'**information**), dont il conviendrait de poursuivre la généalogie et la critique.

Aujourd'hui, l'émergence récente de la notion de culture informationnelle peut être appréhendée à travers les catégories de continuité et de rupture : continuité avec cette généalogie, puisque la culture informationnelle est issue de ce mouvement de *l'information literacy*, dont elle garde encore de nombreuses marques (notamment la dimension info-documentaire, majoritairement prégnante, la persistance du modèle des compétences, l'accent mis sur les compétences procédurales dans la maîtrise des TIC, etc.) ; mais aussi rupture avec cette origine, ou du moins évolution profonde, peut-être à terme changement de paradigme, avec le passage de la *maîtrise* à la *culture* de l'**information**.

Car au-delà des questions terminologiques encore en suspens, l'emploi désormais répandu de l'expression de « culture informationnelle » (ou de culture de l'**information**[5]) exprime plusieurs évolutions assez nettes :

- au plan éducatif, avec la remise en cause des approches purement procédurales des formations documentaires, longtemps dominantes et l'évolution vers une véritable éducation à l'**information**, mettant les savoirs, les connaissances, et le recul critique au premier plan[6] ;
- au plan théorique, la culture informationnelle, notion infiniment plus complexe et plus riche que celle de maîtrise de l'**information**, n'en finit pas de susciter des tentatives de définition, des débats sur son périmètre, ses contenus, etc. ; après plusieurs travaux pionniers dans les années 90 et au début des années 2000, elle donne lieu actuellement

à l'émergence d'un nouveau champ de recherche, en voie de structuration, comme en témoignent par exemple les travaux de l'ERTé « *Culture informationnelle et curriculum documentaire* »[7]. Parmi les différents signes de cette activité de recherche, signalons notamment la tenue à Lille, en octobre 2008, du premier colloque scientifique international sur « *L'Education à la culture informationnelle* », qui a rencontré un net succès[8] et dont les actes sont en cours de publication[9], la création et les travaux du GRCDI[10], la floraison de nombreuses publications[11]...

- enfin, aux plans idéologique ou axiologique, la culture informationnelle permet de sortir des « discours de l'adaptation » et de poser, dans toute sa complexité, la question des finalités éducatives.

La culture informationnelle peut être définie de plusieurs manières, notamment selon le sens donné au mot-valise de culture, et elle fait l'objet d'une grande multiplicité d'approches, de conceptions, de visions, actuellement en débat. Cette multiplicité d'approches peut être, selon nous, résumée de manière très schématique en deux grands « pôles » : d'une part, l'approche « sociologisante » de la culture informationnelle, appréhendée comme l'ensemble des usages, des pratiques, des représentations, des savoirs spontanés... d'une population donnée, concernant l'**information** (la culture informationnelle des étudiants, des lycéens, etc.) ; d'autre part, l'approche éducative, socio-politique, selon laquelle la culture informationnelle constitue un projet, un objectif à atteindre, mais aussi un ensemble de contenus, de savoirs, de notions. A l'évidence, ces deux approches, ces deux pôles (« sociologique et éducatif », pour dire vite) sont étroitement complémentaires, car pour bâtir un projet éducatif sur l'**information**, il est indispensable de partir de l'observation des usages et des représentations et de prendre en compte les nouvelles pratiques informationnelles des jeunes.

Précisons qu'en ce qui concerne le GRCDI, petit groupe de recherche fondé par l'URFIST de Rennes et regroupant une douzaine de personnes, nous nous situons plutôt dans le « pôle éducatif » de la culture informationnelle, du côté des savoirs, des fondements et des finalités, ce qui n'invalide nullement l'autre pôle, celui des usages, des pratiques et des représentations. Et dans ce pôle éducatif, trois grandes questions, selon moi, permettent de poser la problématique de la culture informationnelle dans toute sa complexité :

- tout d'abord, penser la culture informationnelle à partir de (et dans le cadre de) la question de la technique, des supports et des *hypomnémata* : le projet éducatif de la culture informationnelle ne peut s'élaborer que sur la base d'une véritable pensée des supports et des techniques de l'intelligence ;

- ensuite, développer une approche globale, ouverte, non exclusive, de cette culture informationnelle, une nouvelle culture qui engloberait toutes les facettes de l'**information** (depuis la recherche d'**information** sur internet jusqu'au décryptage des médias en passant par la maîtrise conceptuelle des notions et des savoir-faire informatiques de base) ; autrement dit,

une culture trans-informationnelle (une *transliteracy*) réunissant les trois cultures longtemps séparées et bien distinctes : culture des médias, culture documentaire et culture informatique.

- enfin, poser et discuter sans cesse la question des finalités sociopolitiques de l'éducation à l'**information** : si l'on veut sortir du discours stérile, tautologique, de l'adaptation, si le projet de développer une culture informationnelle à l'école et à l'université, depuis l'école primaire jusqu'au doctorat, est bien un projet éducatif, au sens fort du terme, alors la question des finalités de ce projet ne peut être esquivée : à quoi veut-on former, et pour faire quoi ? La question des finalités peut sembler relever de l'évidence : tout le monde sera d'accord pour dire qu'il faut former des citoyens critiques, des **esprits libres**, etc., ou bien des salariés bien adaptés au marché de l'emploi de l'économie de la connaissance. Mais il s'agit souvent de fausses évidences et la conception de la formation à l'**information** est sous-tendue par un certain nombre de présupposés, de valeurs, vaguement partagées, qui sont autant « d'allants de soi » à questionner.

A ces grandes questions s'ajoute une quatrième, également essentielle à nos yeux : celle des contenus didactiques de cette culture informationnelle, sur laquelle travaillent plusieurs membres du GRCDI. Toutes ces questions sont étroitement reliées : en effet, la détermination et l'explicitation des finalités sociopolitiques, éducatives, d'une culture informationnelle dépendra en bonne part de l'approche préalable de la question technique et de la pensée des supports ; de même, de l'assignation de telle ou telle finalité dépendra le choix des contenus didactiques à promouvoir, ainsi que le périmètre, le territoire de cette culture informationnelle.

Je centrerai le propos ici sur la première de ces questions, celle de la technique, en développant ma propre approche personnelle. Donner la primauté, aux plans théorique et épistémique, aux relations entre la culture informationnelle et la question de la technique, celle-ci débordant celle-là de toutes parts, nécessite, selon moi, de réfléchir à trois types de questionnements imbriqués : penser les supports, penser le numérique, penser les contradictions entre l'école et le numérique.

Vouloir penser la culture informationnelle à partir de la « question de la technique » implique tout d'abord de prendre véritablement en compte le rôle des traces et des outils dans l'élaboration des savoirs. Dans la lignée des penseurs et des chercheurs de la technique, depuis Leroi-Gourhan, Gilbert Simondon jusqu'à Bernard Stiegler, en passant par Régis Debray, Bruno Latour, Pierre Lévy, Jean-Max Noyer, etc., il faut sortir de l'occultation, de l'impensé, de l'oubli, voire du mépris des techniques, dans lesquels trop de penseurs, d'enseignants, de responsables tiennent encore la *tekhne*. Vieille antienne, diront certains, tout ou presque ayant été écrit sur la question, notamment dans les années 90 qui ont vu fleurir de nombreux travaux fondamentaux sur la technique et les TIC. Pourtant, même si l'hybridation sociotechnique ne cesse de progresser dans nos sociétés « hyper-matérielles », même si nos prothèses les plus diverses (et notamment **cognitives**) ne cessent de brouiller les frontières du naturel et de l'artificiel, la prégnance des discours

anti-, pro- ou tout simplement *a*-techniques (*i.e.* faisant l’impasse sur le rôle des supports et des outils) reste toujours assez forte, notamment dans le monde éducatif et universitaire. Et concernant plus précisément le champ de la culture informationnelle, de nombreux « dragons entre la technique et nous », pour reprendre l’expression de Régis Debray[12], sont toujours là : la **vision idéale des processus cognitifs**, la conception instrumentale des techniques, la croissance en la prétendue neutralité de l’outil, voire l’occultation, la dénégation des techniques, etc.

Pour prendre l’exemple de l’université, cet « impensé de la technique » se manifeste, selon nous, à travers des phénomènes aussi différents que le mépris, diffus et persistant dans les filières de sciences humaines, pour les formations pratiques aux TIC, l’état souvent problématique des matériels (ou de l’accès à ceux-ci), le retard de l’université, pointé par le Rapport Isaac[13], à intégrer réellement les TICE dans les pratiques pédagogiques, l’incapacité des universités à développer et mettre à la disposition de la communauté universitaire des outils aussi performants et intuitifs que ceux développés par les entreprises ou les communautés de « l’open source »[14], mais aussi dans les « discours enchantés » et naïvement technophiles sur les TIC... Un fil invisible ne relie-t-il pas ces différents phénomènes, hétérogènes dans leur manifestation et leurs causes (un ensemble de facteurs que l’on pourrait mettre sur le compte du fameux « retard français » en matière de TIC) et la difficulté théorique, épistémologique, à penser les supports et les techniques, à prendre véritablement au sérieux la question des outils informationnels ? Nous ne parvenons toujours pas à penser les « hybrides sociotechniques », qui entrelacent discours, société et technique[15] et nous préférons continuer d’opposer, dans des dichotomies stériles, culture et technique, discours et outils, social et nature. Pourtant, sur un autre plan, une partie de la crise de l’école ne vient-elle pas, comme le montre Bernard Stiegler[16], de l’incapacité de celle-ci à penser ses propres supports de transmission, les plus anciens comme les plus nouveaux ?

L’une des conséquences majeures, au plan de la culture informationnelle et de l’éducation en général, de cette « pensée des supports » qu’il faut développer, concerne la formation initiale des enseignants. En attendant de rendre obligatoire, dans la préparation des CAPES, une formation à l’histoire des hypomnemata, comme le propose Bernard Stiegler dans son récent entretien pour Mediadoc[17], l’école (au sens très large, incluant l’université) ne pourra réellement intégrer, et surtout penser, les TIC qu’à partir d’une formation théorique des enseignants sur ces questions, d’une véritable réflexion sur le rôle des supports et des outils dans la constitution et l’évolution de chaque discipline, sur les processus d’extériorisation de la mémoire collective, en bref d’une réhabilitation de la culture technique, que d’autres, ici même, appellent de leurs vœux[18].

Car un autre axe de réflexion, induit par cette pensée des supports, concerne la culture technique des élèves (mais aussi des enseignants), notoirement insuffisante dans le domaine des outils informatiques, contrairement aux apparences de l’explosion des usages des TIC. Dans le souci de développer surtout l’utilisation, la maîtrise pratique des outils, les institutions

éducatives ont laissé de côté toute la **dimension cognitive, théorique, abstraite, des TIC**. Or la **culture technique repose avant tout sur la compréhension des schèmes techniques**, comme l'a montré Gilbert Simondon[19]. Ainsi, pour prendre le seul domaine de la recherche d'**information**, nous ne cessons de plaider avec d'autres, pour développer, dans la formation documentaire et informatique des élèves et des étudiants, **la compréhension d'internet et des outils de recherche**, la « déconstruction » de Google par exemple, la maîtrise des notions de la recherche d'**information** mais aussi de l'informatique, maîtrise conceptuelle loin d'être assurée aujourd'hui chez les « digital natives ».[20]

Penser le numérique, ensuite. Là encore, un vaste programme, qui déborde la culture informationnelle et qui est entrepris depuis longtemps dans différents champs. Mais il nous paraît évident que la problématique de l'éducation à l'**information** doit se fonder sur la **compréhension et l'analyse** de ce que l'on peut appeler, en faisant référence à la médiologie, la « **logique d'outil** » du numérique et de l'**immense processus historique de numérisation des traces et des signes** : à savoir les nouvelles conditions de production, de circulation, de diffusion et d'appropriation des énoncés, sous les conditions du numérique, les nouvelles modalités de la lecture, de l'apprentissage, de l'enseignement, permises par les supports et outils du numérique, etc. L'une des difficultés résiderait ici, me semble-t-il, dans la différenciation à faire, dans l'observation des usages du numérique, entre la part *contingente* des usages, autrement dit ce qui relèverait des modalités d'**appropriation spontanée**, des compétences enseignées, mais aussi des modes passagères, et la part *irréductible*, la **logique interne, profonde, du numérique**. Les deux nous paraissent souvent confondues dans de nombreuses analyses, comme en témoigne une certaine réification des usages d'internet chez les jeunes. Par exemple, si l'explosion de la pratique abusive du copier-coller chez les étudiants relève de l'usage, voire du « mésusage », qui peut être combattu et diminué, en revanche, la « manipulation » de la trame des textes et des images relève bien de la **nature profonde, de la logique même du numérique**, qui le différencie en cela radicalement de l'imprimé. Bruno Bachimont[21] l'a très bien montré en ce qui concerne les images, et il nous semble que nous sommes loin d'avoir tiré toutes les leçons de cette *plasticité* du numérique, notamment dans le domaine éducatif. Là encore, la culture informationnelle devrait revisiter davantage les intuitions de la médiologie sur la logique des supports et des outils.

Enfin, cette pensée du numérique prend tout son sens avec le troisième questionnement, qui touche précisément à l'analyse des contradictions entre les TIC et les nouveaux médias d'une part, et l'école, le système éducatif d'autre part ; contradictions, oppositions, ruptures qui ont déjà été également analysées à maintes reprises et qui expliquent d'ailleurs une large part de la crise profonde des systèmes éducatifs et universitaires : crises des temporalités, de l'attention, de l'autorité, des modes de transmission et d'apprentissage, etc. Toute la question ici est peut-être de déterminer, comme pour les usages, la part irréductible de la part contingente de ces contradictions. Autrement dit, si

l'école et l'université doivent s'appropriier, intégrer, utiliser, développer, etc., les TIC, elles ne peuvent le faire en masquant un certain nombre de contradictions, dont certaines peuvent paraître parfois insolubles[22] : comment, par exemple, combiner les nouveaux modes de lecture en vigueur chez les jeunes (la lecture zapping, l'hyper attention, etc.) et la **lecture profonde, la *deep attention***, sans lesquelles il ne peut y avoir de transmission possible, voire **d'auto-construction des sujets, d'individuation psychique et collective**, pour reprendre l'expression de Simondon chère à Bernard Stiegler ? « *La lecture profonde, qui était auparavant naturelle, est devenue une lutte.* », dit Nicolas Carr, dans un article incisif, « *Est-ce que Google nous rend idiots ?* »[23]

Une véritable éducation à l'**information** pourrait constituer ici l'une des pistes éducatives possibles, pour réconcilier les deux logiques disjointes de l'imprimé et du numérique, notamment autour de la formation des élèves à une meilleure évaluation des sites web, des **documents** et des images ; pour réhabiliter l'analyse, la lecture soutenue, une certaine **rigueur intellectuelle**, ne faudrait-il pas développer une nouvelle herméneutique et une nouvelle philologie, à partir de la redocumentarisation des traces, des images, des **documents** ? Pouvoir remonter la trace des diverses manipulations d'une image, établir les différentes versions d'un texte, mais aussi rendre visibles les réseaux d'acteurs, mobilisés dans les processus d'écriture, ou plus simplement pouvoir identifier un site web... : tous ces apprentissages informationnels permettraient peut-être de re-créeer de la « *deep attention* » sur le web. L'éloge de la lenteur devient ici une urgence pédagogique.

Sur un autre plan et à une autre échelle, l'une des logiques profondes d'internet, et notamment du « web 2.0 », est constituée par cette démocratisation sans précédent de la production de l'**information**, ce modèle horizontal, non hiérarchique, des savoirs, cette distribution des expertises, dont Wikipedia est la figure emblématique. Comment ne pas voir à quel point ce modèle du savoir, d'une immanence radicale, heurte de plein fouet, non seulement le fonctionnement des institutions éducatives mais plus profondément peut-être tout système, tout projet éducatif, qui ne peut exister sans un point de référence transcendantal, une certaine « verticalité »[24] ? Une « *école 2.0* », au-delà de l'intérêt à court terme du slogan, est sans doute l'un des plus beaux oxymorons, qui ait fleuri récemment sur internet : car jusqu'à quel point une institution, dont la mission première reste, envers et contre tout, la transmission d'un héritage commun, peut-elle se dissoudre dans l'apprentissage en réseaux et ce nouvel égalitarisme entre élèves et enseignants ? Les propos de Hannah Arendt sur la crise de l'éducation prennent ici une nouvelle résonance, à l'heure des réseaux et de la crise de l'autorité : « *Dans le monde moderne, le problème de l'éducation tient au fait que par sa nature même, l'éducation ne peut faire fi de l'autorité, ni de la tradition, et qu'elle doit cependant s'exercer dans un monde qui n'est pas structuré par l'autorité ni retenu par la tradition.* »[25]. Certes, ce n'est pas Internet ni Wikipedia qui sont à l'origine de la profonde crise de légitimité et d'autorité qui frappe la plupart des systèmes éducatifs des pays développés. Mais les nouveaux « pharmakon » du web 2.0 aggravent incontestablement cette

crise, en faisant exploser le monopole du savoir de l'école, en **remettant en cause les expertises cognitives**, et notamment celle des enseignants, en disqualifiant la notion même d'autorité, de plus en plus remplacée par la popularité... Comment **refonder l'autorité cognitive des enseignants** dans les nouveaux environnements numériques, est la grande question qui est devant nous. Si nous nous éloignons un peu de la culture informationnelle proprement dite, celle-ci, qui est souvent à l'avant-garde de l'introduction des TIC, ne peut faire l'impasse sur ces contradictions profondes, dont l'âpreté se constate tous les jours dans la difficulté du métier d'enseignant.

Ces questions de la technique, des supports, du numérique, de l'école devraient donc constituer, selon moi, le premier grand axe de réflexion pour le champ de la culture informationnelle, à l'intérieur du champ plus vaste de l'éducation. Les deux autres axes, la *transliteracy* et la question des finalités, seront abordés ici très brièvement.

La *transliteracy* est une expression forgée par des chercheurs américains pour désigner l'ensemble des compétences, des *littéracies* mises en œuvre dans l'usage des TIC : compétences informatiques, documentaires, informationnelles, communicationnelles, médiatiques, etc. Il s'agit là de promouvoir une formation globale à l'ensemble de ces compétences et cultures, longtemps distinctes et dont le web a signé l'acte de mariage. S'informer sur internet fait appel aujourd'hui à différents savoirs et savoir faire, qui relèvent de champs autrefois séparés, par exemple la connaissance de l'ordinateur, de la recherche d'information, des médias, tous domaines aujourd'hui étroitement imbriqués sur les réseaux. La formation aux nouveaux médias ne doit-elle pas prendre acte de cette hybridation des pratiques et des cultures, tout en gardant bien au clair les différences épistémiques profondes et irréductibles, notamment entre les trois cultures : celles de l'information, des médias et de l'informatique ?[\[26\]](#)

L'une des voies possibles de rapprochement et de convergence entre ces trois cultures serait le travail didactique sur les contenus d'enseignement, notamment sur les notions communes, les savoirs partagés ; ainsi la question, centrale entre toutes, de la formation des élèves et des étudiants à l'évaluation de l'information trouvée sur le web, thématique qui intéresse autant les acteurs de l'éducation aux médias que ceux de l'information-documentation (qui sont souvent les mêmes d'ailleurs) constitue-t-elle un territoire partagé, commun, sur lequel un grand travail didactique reste à mener. Pour résumer ce point, je dirai donc que la culture informationnelle, selon moi, devrait, sinon englober, du moins mieux articuler les trois cultures et les trois formations, info-documentaire, informatique et d'éducation aux médias.

Enfin la dernière grande question est celle des finalités, autrement dit de la **réflexion collective**, à la fois pédagogique et politique, sur la détermination des finalités et des objectifs de cette formation à l'information. Nous sommes de plus en plus nombreux à penser et à dire qu'il faut sortir résolument du discours « adaptationniste », *i.e.* du modèle d'une maîtrise de l'information visant seulement et avant tout l'employabilité des futurs travailleurs du savoir, l'adaptation des élèves et des étudiants aux TIC et à Internet. Mais

une fois proclamé ce refus, tout reste à faire dans la réflexion sur les finalités de la culture informationnelle, réflexion qui s'inscrit dans le cadre plus vaste des finalités de l'enseignement.

Pour conclure, cette approche personnelle de la culture informationnelle peut être résumée par une formule que j'ai souvent employée : celle des trois R, Réaliser, Réfléchir et Résister, formule commode qui a le mérite de condenser une approche se voulant équilibrée.

Le premier R de Réaliser englobe toute la culture et la pratique concrète des outils et des techniques numériques : à l'université, les « fractures numériques » sont multiples, au regard des pratiques, et il existe encore une grande partie des étudiants qui maîtrisent mal Internet et les outils informatiques en général. Et quant à ceux qui les maîtrisent, ou qui les utilisent, ils manquent généralement d'une culture technique et informatique de base. Mais le défi, à l'école comme à l'université, n'est plus seulement d'utiliser les outils ou internet, ni de former les élèves à cette utilisation : le véritable enjeu consisterait plutôt à élever les niveaux d'usage, à développer la créativité, à stimuler les pratiques sophistiquées des outils au service des apprentissages, des savoirs et de la connaissance. Et sur ce plan, convenons que beaucoup reste à faire. Le développement de la « réalisation » pratique, de l'appropriation technique sous toutes ses formes, en s'emparant concrètement des outils, l'encouragement de l'inventivité pédagogique dans les usages du numérique, constitue, selon moi, un axe essentiel de la culture informationnelle

Dans le même temps, une éducation à la culture informationnelle se doit de viser la **réflexion**, autrement dit la **compréhension, la connaissance, le savoir sur les outils et les réseaux de l'information**. Ce deuxième R du triptyque ouvre sur un champ immense, celui des sciences et des savoirs de l'information, des médias, de la communication, des techniques informatiques et des supports. Pour nous en tenir au seul domaine de l'information-documentation, nous sommes plusieurs à penser, notamment au GRCDI, qu'il y a ici matière à transmission, à enseignement, à didactisation, peut-être demain à discipline. Sans entrer dans les détails de ce débat et de cette controverse d'une nouvelle discipline, disons simplement qu'il s'agit, partout où cela est possible, de mettre l'accent sur les savoirs et sur les notions fondamentales. Les formations aux nouveaux médias, à internet et à la recherche d'information sont confrontées, ici plus qu'ailleurs, au défi de l'innovation permanente, qui rend très vite caducs les compétences techniques et les savoirs acquis. D'où l'importance cruciale de développer l'enseignement de notions stables, pérennes, de longue durée : par exemple, si les modalités de l'indexation vont certainement encore changer dans les années à venir, la notion, essentielle, d'indexation ou de métadonnées n'a pratiquement pas changé depuis Callimaque. L'enjeu de ce deuxième R de la **Réflexion** concerne donc les contenus didactiques, les savoirs, les concepts fondamentaux de l'information, qu'il faut délimiter, définir, didactiser,

transmettre et **faire construire**... C'est tout l'objet du chantier didactique de l'**information**-documentation, sur lequel sont activement mobilisés plusieurs membres du GRCDI[27].

Enfin avec le troisième R, celui de la résistance ou du recul, se profile toute la **dimension citoyenne, critique de l'éducation à l'information, mais aussi auto-réflexive**. Mais résister à quoi ? Aux outils, à l'information et au numérique.

Apprendre aux élèves à maîtriser les traces qu'ils laissent, à se méfier des techniques du profilage marketing, à être vigilants par rapport au *ranking* des moteurs de recherche, à ne pas se laisser entièrement « googliser », etc. : le champ d'une formation critique aux outils informationnels est très vaste.

Former à l'**information** suppose aussi de former *contre* l'**information**, pas seulement contre ses infopollutions, mais aussi contre ses mythes, ses illusions, notamment celle de la confusion entre **information** et connaissance, qui fait des ravages dans les têtes. Enfin le recul face au numérique est une variante de la **réflexion critique contre les mythes de la calculabilité généralisée**, les risques du temps réel, les effets pervers du numérique, etc.

Il faudrait ajouter à tout cela la **dimension auto-critique, auto-réflexive, sur ses propres comportements informationnels, sur les phénomènes d'addiction à la connexion permanente**.

On le voit, le champ de ce troisième axe du recul critique est immense et demande bien entendu des éclaircissements, des approfondissements, que nous ne pouvons développer ici. Notons seulement que cette dimension critique (et auto-réflexive) de la culture informationnelle serait sans doute celle qui pourrait conférer un véritable sens à cette nouvelle formation, en incarnant fortement ces finalités éducatives et politiques à construire et permettant d'échapper aux discours de l'adaptation.

Appropriation des outils et inventivité des usages, réflexion théorique et acquisition de connaissances, recul critique sur les outils : pris séparément, chacun de ces trois axes suscite généralement l'adhésion, parfois des consensus locaux. Certains types de « discours » sur les TIC peuvent même se référer à l'un ou l'autre des trois axes : pensons par exemple au discours des « geeks », enthousiastes du web 2.0 et des outils numériques, ou, à un niveau plus modéré, aux discours officiels sur la « société de l'**information** », qui ne mettent généralement l'accent que sur le premier R de la réalisation pratique, de la prise en mains, de la formation technique, au mieux de l'inventivité. A l'inverse, de nombreux discours, d'intellectuels et d'enseignants notamment, sont frappés au coin d'une certaine méfiance, d'une critique parfois systématique d'internet, voire d'une position assumée de « technophobie », toutes positions pouvant se référer au dernier R de la résistance et du recul. Quant à l'axe de la réflexion, des savoirs et des notions à transmettre, notons qu'il correspond plutôt aux préoccupations d'un certain nombre d'enseignants-documentalistes, d'enseignants-chercheurs en Sciences de l'**information** et de la communication et de professionnels des

bibliothèques et de l'**information**.

Pour nous, tout l'enjeu, mais aussi toute la difficulté, est de concevoir et de mettre en œuvre un nouveau projet éducatif d'une formation à l'**information**, qui puisse combiner ces trois axes, à parts égales. Il faut, à la fois, former *à, par, sur* et *contre* l'**information** et ses nouvelles médiations...

[1] Sur les questions d'origine, de traduction et de sens de la notion *d'information literacy*, voir notamment les travaux d'Olivier Le Deuff, notamment : Le Deuff, Olivier. « La culture de l'information : Quelles « littératies » pour quelles conceptions de l'information ? », In *VI.ème Colloque ISKO-France'2007*

7 et 8 juin 2007, à Toulouse, IUT de l'Université Paul Sabatier : IUT. Disp. sur : http://archivesic.ccsd.cnrs.fr/sic_00286184/fr/

[2] Information Literacy Committee, « *Information Literacy* », Ocotillo Report '94, Arizona, USA, Maricopa Center for Learning and Instruction (MCLI), 1995. Disp. sur : <http://hakatai.mcli.dist.maricopa.edu/ocotillo/report94/rep7.html>. (trad. disponible sur <http://www.ebsi.umontreal.ca/formanet/culture.html>).

[3] Grand dictionnaire terminologique de l'Office Québécois de la Langue Française (OQLF), 2002.

[4] IFLA (International Federation of Library Associations and Institutions), *La proclamation d'Alexandrie sur la maîtrise de l'information et l'apprentissage tout au long de la vie*, novembre 2005. Disp. sur : <http://www.ifla.org/III/wsis/BeaconInfSoc-fr.html>.

[5] Entre les deux expressions existent des nuances de sens, voire des conceptions différentes, qui font d'ailleurs aujourd'hui l'objet de réflexions chez les acteurs de ce champ de recherche, nuances ou différences qu'il serait trop long de développer ici.

[6] Voir notamment : Ballarini-Santonocito, Ivana, Duplessis, Pascal. « *Petit dictionnaire des concepts info-documentaires : approche didactique à l'usage des enseignants documentalistes* ». SavoirsCDI, 2007 [en ligne]. Disp. sur :

<http://savoircdi.cndp.fr/culturepro/actualisation/Duplessis/dicoduplessis.htm#auteur> ;

Duplessis, Pascal, Serres, Alexandre. « Une nouvelle étape vers une didactique de l'information ? ». FADBEN. Les savoirs scolaires en information-documentation : 7 notions organisatrices, *Médiadoc*, mars 2007. p. 5-9 ; Duplessis, Pascal. « Les référentiels de compétences info-documentaires : enjeux, intérêts et limites. » In Devis-Duclos, Sylvie et al. *Le Livre bleu des professeurs documentalistes*. Scéren, CRDP d'Orléans-Tours, 2006. p. 133-146

[7] Voir le site de l'ERTé : http://geriico.recherche.univ-lille3.fr/erte_information/

[8] Voir le programme et les textes du colloque sur le blog créé à l'occasion : <http://ertecolloque.wordpress.com/>, ainsi que l'ensemble des communications des intervenants sur ArchiveSic : <http://archivesic.ccsd.cnrs.fr/ERTE>

[9] Ils seront publiés prochainement aux Presses de l'ENSSIB.

[10] Groupe de Recherche sur la Culture et la Didactique de l'Information ; voir le site <http://culturedel.info/grcdi/>

[11] Voir l'étude bibliométrique sur les travaux concernant la culture informationnelle que nous avons menée en janvier 2008 : Serres, Alexandre. « *Quelques observations bibliométriques sur la culture informationnelle*. ». Pre-publication, février 2008. Disp. sur : http://www.uhb.fr/urfist/files/Etude_bibliometrie_culture_informationnelle.A%20Serres.doc

[12] R. Debray, *Cours de médiologie générale*, Paris, Gallimard, 1994

[13] Henri Isaac, *L'Université numérique. Rapport à Madame Valérie Pécresse, Ministre de l'Enseignement Supérieur et de la Recherche*, Paris, Université Paris Dauphine, 2007, p. 37 et 45-47. Disp. sur : http://www.nouvelleuniversite.gouv.fr/IMG/pdf/Rapport_univ_num.pdf

[14] Pourquoi, par exemple, les enseignants et les étudiants qui veulent créer un blog, utiliser un wiki, une plate-forme de partage de signets, un agenda collaboratif, des outils bureautiques partagés, etc., utilisent-ils avant tout les plates-formes et outils du web 2.0, gérés par Google, Yahoo et d'autres entreprises ?

[15] B. Latour, *Nous n'avons jamais été modernes. Essai d'anthropologie symétrique*, Paris, La Découverte, 1994

[16] Voir notamment B. Stiegler, *Prendre soin. 1. De la jeunesse et des générations*, Flammarion, 2008.

[17] Stiegler, Bernard. Entretien avec... Bernard Stiegler (Philosophe, Directeur de l'IRI, Professeur à l'université de technologie de Compiègne). Interview par Serres, Alexandre et Ballarini-Santonocito, Ivana. *Revue Mediadoc*, 10 mai 2009. Disponible sur : <http://www.fadben.asso.fr/spip.php?article78>

[18] Voir sur Skhole.fr, le texte de Jean-Hugues Barthélémy, *Penser l'école avec Gilbert Simondon* : <http://skhole.fr/penser-l-%C3%A9cole-avec-gilbert-simondon-par-jean-hugues-barth%C3%A9my>

[19] Gilbert Simondon, *Du mode d'existence des objets techniques*, Paris, Aubier, 1989

[20] Voir sur ce point les travaux de Bruillard et Fluckiger, les textes de l'association EPI (par exemple un texte récent de Jacques Baudé : <http://www.epi.asso.fr/revue/articles/a0905d.htm>)...

[21] Chercheur à l'INA

[22] Nous avons tenté de développer l'analyse de ces contradictions entre l'école et les TIC dans un texte récent : Serres, Alexandre. « L'école au défi de la culture informationnelle ». In Dinet, Jérôme (sous la dir. de). *Usages, usagers et compétences informationnelles au 21^e siècle*. Paris : Hermès Lavoisier, 2008. Chap. 2, pp. 41-70

[23] <http://www.internetactu.net/2009/01/23/nicolas-carr-est-ce-que-google-nous-rend-idiot/>
« Lorsque nous lisons en ligne, nous avons tendance à devenir de “simples décodeurs de l'information”. Notre capacité à interpréter le texte, à réaliser les riches connexions mentales qui se produisent lorsque nous lisons profondément et sans distraction, reste largement inutilisée. »

[24] Nous renvoyons ici à un très beau texte de Philippe Meirieu, prononcé lors d'une conférence donnée à l'Université de Barcelone, en juillet 2008 : « *Le maître, serviteur public. Sur quoi fonder l'autorité des enseignants dans nos sociétés démocratiques ?* ».
Disp. sur : http://www.meirieu.com/ARTICLES/maitre_serviteur_public_version2.pdf

[25] Hannah Arendt - *L'éducation moderne*, disp. sur <http://skhole.fr/hannah-arendt-%C3%A9ducation-moderne>

[26] Serres Alexandre. « Educations aux médias, à l'information et aux TIC : « ce qui nous unit est ce qui nous sépare » », *Colloque international de l'ERTé, « L'éducation à la culture informationnelle »*, Lille, 16-17-18 octobre 2008. Intervention à la table ronde « Acteurs et territoires de l'éducation à l'information ».

[27] Voir par exemple les travaux de Pascal Duplessis (<http://esmeree.fr/lestroiscouronnes/>) et Ivana Ballarini-Santonocito en Pays de la Loire (http://www.pedagogie.ac-nantes.fr/64788248/0/fiche_ressourcepedagogique/&RH=1159790386593), d'Agnès Montaigne à l'IUFM de Rouen (<http://documentaliste.ac-rouen.fr/spip/spip.php?article193>), de Nicole Clouet à l'IUFM de Caen...

Annexe 3

Test de la grille épistémologique : une vision sociale (exclusion de la théorie mathématique de l'information)

Mise en évidence des indicateurs de la grille épistémologique dans le texte ci-dessous :

- Termes grisés : occurrences du mot information
- Termes encadrés : occurrence du mot document
- Termes et expressions en gras : éléments se rapportant au sens, à l'esprit, à la signification, aux opérations cognitives
- Termes et expressions soulignés : éléments se rapportant à la technique, aux outils, à la technologie

Virtualité, représentation, signification : approche de la complexité documentaire

Par Caroline Courbières, LERASS – Université de Toulouse

Introduction

Le concept de document fonde la documentation en tant qu'ensemble de pratiques, dispositifs et usages. Il justifie, tout en dépendant d'elle, l'activité de tous les acteurs que l'on peut désigner sous l'appellation de « professionnels de l'information ». Pourtant, un document, ça n'existe pas... Cette entrée en matière n'est pas aussi provocatrice qu'elle semble l'être si l'on daigne se poser la question que lançait Jean Meyriat dans le cadre de sa réflexion sur le territoire des Sciences de l'information et de la communication : « ET LA SÉMIOLOGIE ? » (Meyriat, 1993a). Comment définir l'objet documentaire en faisant abstraction de la **charge signifiante de son contenu** informationnel ? Comment expliquer la possibilité d'une représentation documentaire déconnectée d'une éventuelle **signification** documentaire ? Plus largement, comment envisager la documentation, discipline improbable au regard des Sciences de l'information et de la communication ? Ces quelques interrogations jalonnent l'itinéraire que suit cet article entre virtualisation et actualisation.

L'impossible définition du document

L'évocation des étapes fondatrices de la réflexion sur le concept de document nous invite tout d'abord à confronter les notions de signe, d'information et de savoir – défini en tant que formes de connaissance reconnues par une société (Jeanneret, 2007) –, selon lesquelles se déploie sa dimension épistémologique. Dans les années trente, Paul Otlet propose une

première définition de l'objet documentaire caractérisé par un support et un signe (Otlet, 1934, 13). À l'aube des années cinquante, Suzanne Briet reprend ces deux caractéristiques fondamentales qui identifient le **document** en tant qu'**objet signifiant**. Elle intègre dans sa propre typologie documentaire l'animal et l'objet du monde naturel tout en introduisant la notion d'unité documentaire comme objet de la quête scientifique (Briet, 1951, 9) ; ce faisant, elle pose tout à la fois la nature polysémiotique de l'objet documentaire et le nécessaire processus de transformation de l'objet en objet documentaire. Conservant la notion de support posée par Paul Otlet et la nature sémiotique de l'objet avancée par Suzanne Briet, Jean Meyriat introduit celle d'**information** pour définir ce qu'est un **document** : « un objet sur lequel sont enregistrés des signes représentant de l'**information**. Ces signes n'appartiennent pas nécessairement au langage écrit ; et s'ils lui appartiennent, l'objet peut être fait de bien d'autres matières que du papier » (Meyriat, 1993b, 152).

En reliant à son tour la notion d'**information** (documentaire) à celle de signe, Jean Meyriat souligne la nuance fondamentale posée en son temps par Robert Escarpit entre la notion d'**information** et la notion de savoir (Escarpit, 1991, 203) – nuance qui doit être entendue à la lumière de la différence essentielle de nature entre l'**information** mathématique et l'**information** sociale largement explicitée par Yves Jeanneret (Jeanneret, 2007). La notion d'**information** évoquée par Robert Escarpit renvoie à l'**information** mathématique, la notion de savoir s'apparente à l'**information** sociale et son inscription documentaire ; si la première se saisit dans une logique de transmission linéaire, la seconde implique un travail interprétatif qui déborde le modèle traditionnel de la communication et prouve « que le savoir et le **sens** ne sont jamais simplement donnés, mais qu'ils sont élaborés » (Jeanneret, 2008, 38). L'**information** en tant que donnée accède au statut d'**information** en tant qu'élément de connaissance à partir du moment où elle est **saisie comme signe**.

La référence au signe comme caractéristique de l'**information** documentaire sous-tend de la même manière la définition autorisée de l'**information** dans le Vocabulaire de la documentation : « élément de connaissance susceptible d'être représenté, à l'aide de conventions pour être conservé, traité ou communiqué. Toute **information** se caractérise par un contenu (**signifiant**), un **signifié** et une forme » (Boulogne, INTD-ER, 2004). L'éventualité d'une représentation de l'**information** documentaire inscrite dans sa définition même préfigure la virtualité de l'objet qui ne devient documentaire qu'à partir du moment où il participe d'une stratégie informationnelle située. C'est ainsi que Robert Escarpit a défini le **document** comme un « anti-événement » :

Au cours des derniers millénaires, l'homme a élaboré pour échapper à cette contrainte [l'événement] un anti-événement qui est le **document** [...] le pas décisif a été franchi lorsque l'homme a institué le **document**, cumulation de traces fixes et permanentes [...] où les réponses données en feed-back, à travers le temps, aux expériences antérieures, restent disponibles pour une lecture, c'est-à-dire une

exploration libre de toute contrainte événementielle ou chronologique, en fonction du projet et de la stratégie destinée à le réaliser (Escarpit, 1991, 203).

Bien que marquée par la terminologie héritée de la théorie mathématique de l'**information**-communication, cette citation de Robert Escarpit rejoint la distinction fameuse faite par Jean Meyriat entre **document** par intention (l'objet produit en vue d'informer) et **document** par attribution (l'objet élaboré dans la recherche d'**information**) (Meyriat, 2001a, 144).

Selon cette distinction, l'objet documentaire dépend bien d'une « stratégie destinée à le réaliser ». Cette stratégie évoquée, qui est double et asymétrique, situe l'objet documentaire à deux moments différents : le moment de son intentionnalité et le moment de son utilisation. À partir de ces deux visées communicationnelles distinctes, nous avons insisté sur la notion de virtualité comme fondatrice de l'objet documentaire, le moment de l'interprétation du **document** primant sur celui de sa production ; tout objet se voulant **document** ne le sera qu'à partir du moment où il sera **(re)construit en tant que tel**, l'absence de **(re)construction du sens** d'un objet documentaire invalidant son statut même de **document** (Courbières, 2008). Un objet documentaire n'existe donc qu'à partir du moment où il **signifie pour quelqu'un**. Pointer l'**action institutionnalisante du récepteur** dans la définition de l'objet documentaire impose en même temps de l'appréhender à partir de son actualisation en tant qu'**objet signifiant**. L'on doit ainsi concevoir une sorte de pré-statut documentaire potentiel d'un objet pouvant être soumis à la description ou l'analyse, puisque, à l'image du signe, le **document** dépend de « ce moment instituant, celui du **regard prêt à interpréter** » que pose Yves Jeanneret à propos du support documentaire (Jeanneret, 2007, 100).

C'est suivant cette virtualisation/actualisation de l'objet documentaire que nous avons avancé le principe de l'horizon documentaire (Courbières, 2004 ; Courbières et Couzinet, 2006), en reprenant les concepts d'horizon d'attente développé en esthétique de la réception : l'horizon d'attente tel que le définit Hans Robert Jauss embrasse les **phénomènes de réception** pendant la production de l'objet artistique et après sa création ; il concerne à la fois le producteur de l'œuvre et ses récepteurs. L'objet implique un « horizon d'attente littéraire », c'est-à-dire « tout un ensemble d'attente et de règles du jeu avec lesquelles les textes antérieurs l'ont familiarisé et qui, au fil de la lecture peuvent être modulées, corrigées, modifiées ou tout simplement reproduites » (Jauss, 1978, 51), tout en impliquant un « horizon d'attente social », à savoir, « la **disposition d'esprit** ou le code esthétique des lecteurs, qui conditionne la réception » (Jauss, 1978, 259). L'attente en question présuppose l'activité du sujet récepteur et son appréhension de l'œuvre dans la médiation d'une intertextualité. L'horizon documentaire, à l'image de l'horizon d'attente, n'est pas figé, il se modifie et projette la **com-préhension de l'objet documentaire** dans un ensemble d'objets que **le sujet**

récepteur (ré)active. Or le caractère variable de l'horizon documentaire remet en question la stabilité informationnelle de l'objet documentaire, non pas en raison de la forme ou du support de ce dernier, mais parce que le document se construit à partir de son **contexte de réception**. C'est sur cette (non) définition du document (pris comme un artefact) et sur l'horizon documentaire qui la subsume que la notion de représentation documentaire prend appui.

La problématique de la représentation documentaire

Cette problématique peut prendre place au centre du débat classique pour les théoriciens du langage entre la fonction représentationnelle et la fonction constructive (du langage). Nous avons eu l'occasion d'examiner la question de la **représentation** à travers l'analyse des langages documentaires. Ces outils professionnels, qui peuvent également être appréhendés en tant qu'objets documentaires constituent en effet un espace privilégié par leur double statut, à la fois outils de description des objets documentaires et discours construisant leurs propres représentations ; on a ainsi pu montrer comment ces langages, censés représenter des objets référentiels (de la connaissance), **construisent leur propre contenu informationnel** à travers lequel ils deviennent **objets d'interprétation**. La représentation documentaire réside alors dans cette tension entre un dit documentaire et un dire documentaire (Courbières, 2011). En d'autres termes, et à partir de la distinction posée en 1892 par Frege entre le **sens** (der Sinn) et la **dénotation** (die Bedeutung), on peut convenir que, d'une part, la saisie des langages documentaires en tant qu'outils vise à déterminer la dénotation des propositions documentaires, c'est-à-dire leur fonction de désignation des connaissances à représenter. D'autre part, ces langages, en tant que systèmes de signes, autorisent une seconde saisie qui s'intéresse à leur **sens**, résultant de la configuration **signifiante** constituée par des unités documentaires définies les unes par rapport aux autres. Si nous faisons référence aux travaux du philosophe et logicien Gottlob Frege en liaison avec la tension observée dans le fonctionnement même des langages documentaires, c'est parce qu'ils nous semblent trouver un écho contemporain dans la tension évoquée par Bernd Frohmann entre la matérialité du document et sa fonction représentationnelle (Frohmann, 2011). Néanmoins, il faut noter que la représentation documentaire ne renvoie pas à ce que Gottlob Frege désigne sous le terme de représentation – qui reste pour lui essentiellement individuelle (Frege, 1971, 105-106) –, mais bien au **sens** et à la **dénotation** tout à la fois.

Plus précisément, et en prolongeant cette approche philosophique, on peut affirmer que la représentation à l'œuvre dans les systèmes documentaires combine les différences de fonction qu'assument le langage et le langage technique : tandis que le langage va **de la signification à la désignation**, le langage technique opère le trajet inverse, **de la désignation à la signification** (Coseriu, 2001). L'objet documentaire lui-même se place à l'intersection de ces deux fonctions : un document répond à la fonction de désignation par les données qui constituent son expression, et **c'est par l'actualisation de son contenu qu'il acquiert un**

sens, devenant un signe à part entière. Si l'on consulte la définition technique de l'objet documentaire, on constate que la notion de signe utilisée par les pères de la documentation se trouve pulvérisée à travers celles d'**information** (documentaire), d'**informations** (au sens de données), de la **signification attribuée**, et de la visée communicationnelle plurielle de l'objet : un **document** est « l'ensemble d'un support d'**information**, quel qu'il soit, des données enregistrées sur ce support et de leur **signification**, servant à la consultation, l'étude, la preuve ou la trace, etc. : livre, échantillon de parfum, tissus, film, etc. Le tout constitue une unité autonome » (Boulogne, INTD, 2004). On retrouve dans cette définition l'aspect indissociable de l'objet documentaire et de son **pouvoir-signifier**, à la manière dont Ferdinand de Saussure reliait le signe linguistique à la notion de **signification** : « Ce sont deux formes du même concept de **l'esprit**, vu que la **signification** n'existerait pas sans un signe, et qu'elle n'est que l'expérience à rebours du signe, comme on ne peut pas découper une feuille de papier sans entamer l'envers et l'endroit de ce papier, du même coup de ciseaux » (Saussure, 2002, 96). Si tout objet peut prétendre à un statut documentaire – à l'image de l'antilope prise comme exemple par Suzanne Briet ou de pièces vestimentaires présentées dans des expositions muséales –, c'est en vertu de la fonction informative qu'**un récepteur est susceptible de lui attribuer**.

La virtualité de l'objet documentaire et sa charge sémantique intrinsèque l'inscrivent explicitement au centre du mécanisme dynamique de la communication. Le linguiste Bernard Pottier a modélisé ce double processus par une représentation des parcours onomasiologique – qui rend compte des **phénomènes de désignation (conceptualisation et sémiotisation)** – et sémasiologique – qui rend compte des **phénomènes d'interprétation (identification et compréhension)**.

Figure 1 : le papillon sémantique de Bernard Pottier

À partir de l'articulation de ces deux parcours, Bernard Pottier distingue quatre lieux d'analyse possible d'**appréhension du sens** : le référentiel (externe) comme « monde de départ ou d'arrivée, facultatif », le conceptuel (ou référentiel interne) comme lieu de la représentation mentale, la langue comme savoir, et le discours, « dans sa double fonction de résultat observable [...] et de base de départ pour **l'interprétant** » (Pottier, 1992a, 18). Cette mise à plat des possibilités de **production et d'interprétation du sens** permet de dresser l'étendue d'une théorie générale qui peut aussi bien viser à comprendre le **sens** d'une expression sur la base de la **structure sémantique** – objet de la sémantique formelle d'inspiration logique dite vériconditionnelle –, que s'occuper des relations entre les niveaux référentiels et le système de la langue – tâche relevant de la **sémantique référentielle** –, ou encore s'attacher de manière privilégiée à l'analyse des formes **signifiantes** résultant de la mise en signe opérée.

Ces deux derniers angles d'attaque doivent ainsi être choisis pour déterminer la **double**

nature des relations sémantiques qui structurent tout langage documentaire : « relations between concepts, senses or meanings should not be confused with relations between the terms, words, expressions or signs that are used to express the concepts. It is, however, common to mix both of these kinds of relations under the heading ‘semantic relations’ » (Hjørland, 2007). Concernant l’objet documentaire, c’est bien à sa **matérialité signifiante** que l’on doit s’attacher en priorité, à la fois comme manifestation d’une représentation et comme actualisation de connaissances, **joignant dès lors le phénomène de la connotation au fait de dénotation** : « la dénotation renvoie essentiellement à la **conceptualisation du référent** ; la connotation est dépendante du savoir (**cognitif**), du vouloir (**intentionnel**), de l’énonciation, et, éventuellement du contexte et de la situation » (Pottier, 1992b, 74). Le problème de la désignation abordé précédemment se trouve alors rejeté de l’analyse première puisque c’est la signification de l’objet lui-même qui prime, indépendamment de ce à quoi il est susceptible de renvoyer aux niveaux conceptuel ou référentiel : « un signe n’est pas pourvu de **sens** parce qu’il désigne un référent » (Rastier, 2009, 21). Cette problématique, qui s’appuie sur les apports de la sémantique linguistique et textuelle, nous invite à réfléchir plus largement aux soubassements sémiologiques qui sous-tendent le cadre général d’une théorie de la documentation.

La documentation comme discipline : la question du sens au cœur des Sciences de l’information et de la communication

Apparu à la fin du XIX^e siècle pour signifier l’action de réunir des **documents**, le terme documentation est adopté au début du XX^e siècle par l’Institut international de bibliographie fondé par Paul Otlet et qui deviendra en 1931 l’Institut international de documentation. Rapidement imposé, documentation devient dans les années 1930 « le mot-centre d’une activité de recherche et de traitement des **documents**, organisée à cette époque » (Rey, 1998, 1114). Ce terme exprime bien à la fois l’objet en question (le **document** – terme dérivé, faut-il le rappeler, du verbe latin docere) et la nécessaire action qu’il implique (explicite à travers la dérivation nominale produite par le suffixe -tion) – : la documentation, en tant que discipline, ne se résume pas à l’application de techniques documentaires, elle suppose une certaine posture qui prend appui sur l’élaboration même de son objet (l’objet documentaire). Si peu d’études hexagonales portent précisément sur cette « discipline scientifique qui fournit les connaissances fondamentales et appliquées sur lesquelles s’appuie la pratique » (Meyriat, 1993c, 151), la documentation doit pourtant affirmer sa place au sein des Sciences de l’**information** et de la communication, dont la spécificité même a amené ses spécialistes à lister, classer, organiser ses diverses composantes (Couzinet, 2008).

Une cartographie des Sciences de l’**information** et de la communication reste extrêmement mouvante ; c’est ainsi qu’à l’origine de ce champ, les recherches sur les médias ou en journalisme s’inscrivaient sur le volet **information**, tandis qu’elles peuvent apparaître

aujourd'hui, pour certaines, comme le fer de lance du volet communication. Exception culturelle aidant, la situation est différente à l'étranger où les termes **information** et communication ne sont pas couplés. De fait, la revendication possible d'une science de l'**information** s'explique par l'importation du modèle américain qui, à partir de la seconde moitié du XXe siècle, impose l'Information Science (Fondin, 2005). L'étude menée récemment sur la représentation des Sciences de l'**information** et de la communication à travers le langage documentaire RAMEAU témoigne du statut instable de ces Sciences de l'**information** dans l'ensemble de ce champ et constate la prédominance de « termes associés au domaine des sciences de l'**information** en général et aux métiers de la documentation en particulier » (Boutin et al., 2008). Si l'on ne saurait taire l'asservissement de la documentation aux techniques documentaires dans la représentation commune de cette discipline, celle-ci ne peut être arrachée au substrat théorique qui oriente sa pratique : la documentation, en tant que discipline scientifique, doit bénéficier des « acquis d'une science pour laquelle **le sens est la valeur décisive** » (Meyriat, 2001b : 261). L'étude de **documents**, de leur traitement informationnel et de leur circulation implique en effet d'étudier la **construction du sens des systèmes et des processus documentaires**. L'exigence sémiologique, dont la visée est de décrire les **conditions de production et de saisie du sens** nous semble devoir guider, de fait, les études actuelles sur l'organisation des connaissances.

Selon Jean Meyriat, la sémiologie devait légitimement être considérée comme une des sciences de la communication, puisqu'elle « étudie les moyens symboliques par lesquels se réalise la communication » (Meyriat, 1993d, 334). Projetée comme « une science qui étudie la vie des signes au sein de la vie sociale » (Saussure, 1972, 33), la sémiologie possède plusieurs histoires, et se dit aussi sémiotique selon l'approche que l'on souhaite mettre en valeur : outre l'institutionnalisation du terme sémiotique lors de la fondation en 1969 à Paris de l'Association Internationale de Sémiotique (International Association for Semiotic Studies), sous l'égide notamment de Roman Jakobson, André Martinet, Roland Barthes, Algirdas Julien Greimas et Umberto Eco, il existe une vision oppositionnelle qui consiste à renvoyer la sémiologie du côté des systèmes de signes, la sémiotique s'occupant des **processus de signification**. Par ailleurs, la sémiotique peut encore principalement se référer à la semiotics du philosophe et logicien américain Charles Sanders Peirce, ou renvoyer aux travaux du Cercle de Paris fondé en 1967 par A. J. Greimas dans le prolongement de la linguistique structurale. La sémiologie (ou sémiotique) européenne tire son origine du geste fondateur du linguiste suisse Ferdinand de Saussure qui consista à déconnecter le signe linguistique de l'objet extra-linguistique auquel il réfère. La notion saussurienne de **signifié** – liée à celle de **signifiant** et distincte de celle de concept –, permet d'analyser la **construction du sens** à partir du seul **contenu linguistique**, le signe n'ayant de valeur que par les rapports oppositionnels qu'il entretient avec les autres signes du système. Figure emblématique de ce courant européen, Roland Barthes conçut la sémiologie comme une branche de la linguistique, inversant par là même la proposition saussurienne : « la langue est le domaine des articulations, et le **sens** est avant tout découpage » (Barthes, 1964, 114). Ce programme se

retrouve de fait si l'on veut penser la documentation dont « la question essentielle [est] celle des possibilités et modalités de catégorisation de l'**information** et de sa représentation » (Meyriat, 1993c, 151).

Dans un article sur la science de l'**information**, Jean Meyriat évoquait ainsi la linguistique et la sémiologie comme sciences pourvoyeuses d'un « équipement méthodologique » (Meyriat, 2001b, 256), un peu à l'image de l'outillage sémiotique qui sert les études marketing (Boutaud, 2007). De son côté, et mettant en perspective l'importance des semiotic studies par rapport au nombre peu élevé des départements de sémiotique de par le monde, le danois Niels Winfeld Lund souligne également l'apport de la théorie sémiologique à nombre de disciplines scientifiques (Lund, 2010). Certains travaux de linguistique et sémiotique peuvent d'ailleurs relever de la compétence des Sciences de l'**information** et de la communication « si l'objet de l'étude porte sur le fonctionnement d'un processus de communication, ou sur l'extraction et la représentation des connaissances en vue de la conception de systèmes d'**information** » (Miège, Odin, 1993, 8). La généalogie même des Sciences de l'**information** et de la communication rappelle enfin les liens privilégiés que celles-ci entretiennent avec le courant sémiologique ; en effet,

c'est à la rencontre [...] de personnalités représentant des courants scientifiques différents (A. G. Greimas et R. Barthes, pour la sémiologie et la sémiotique, Jean Meyriat pour les Sciences de l'**information** et de la documentation, Robert Escarpit, pour les sciences de la communication...), que tient l'idée d'un projet de constitution d'interdiscipline regroupant ces courants (Ollivier, 2001, 340-341).

Si la documentation en tant que discipline se situe précisément au cœur de cette interdiscipline, l'action documentaire, à la charnière d'un monde à représenter et d'un monde **signifié**, doit être saisie dans cette **articulation du sens** puisque son rôle est de permettre et faciliter l'accès aux connaissances à partir de l'organisation et de la représentation d'objets informationnels.

Conclusion

La documentation, discipline trop souvent déracinée de ses soubassements théoriques et rivée à ses aspects techniques, ne peut faire l'économie des principes de l'**étude du sens et des significations**. L'**appréhension** de l'**information** en tant que signe disponible pour la constitution d'une mémoire documentaire rejoint la théorisation du **document** dont la virtualité consubstantielle achève de proclamer son impossible définition. Si la voie sémiologique continue d'être empruntée comme cadre général d'une théorie de la documentation (Mai, 2001 ; Raber et Budd, 2003 ; Lund, 2010), c'est bien parce que les objets, systèmes ou processus documentaires ne peuvent être envisagés hors de cette « nébuleuse du **sens** » qu'évoquait Saussure.

Bibliographie

- **Barthes R.**, (1964), « Éléments de sémiologie », *Communications*, 4, pp. 91-135.
- **Boulogne A. et INTD-ER**, (2004), *Vocabulaire de la documentation*, Patis, ADBS Éditions, <http://www.adbs.fr/vocabulaire-de-la-documentation-41820>.
- **Boutaud J.-J.**, (2007), « Du sens, des sens : Sémiotique, marketing et communication en terrain sensible », *Semen*, <http://semen.revues.org> : (consulté le 24/02/2012).
- **Boutin E., Liu P., Gorla S., Dumas P. et Amos D.**, (2008), « Les SIC à partir du thésaurus Rameau : représentation ou interprétation ? », in *Les sciences de l'information et de la communication : affirmation et pluralité*, 16e Congrès de la SFSIC, 11-13 juin 2008, Compiègne, http://www.sfsic.org/congres_2008/spip.php?article125.
- **Briet S.**, (1951), *Qu'est-ce que la documentation ?*, Paris, EDIT.
- **Coseriu E.**, (2001), « Dix thèses à propos de l'essence du langage et du signifié », *Texte !*, vol. VI, n° 2, http://www.revue-texto.net/Inedits/Coseriu_Theses.html.
- **Courbières C.**, (2011), « Stéréotypage et artifice documentaires », *Communication & langages*, n° 170, pp. 97-106.
- **Courbières C.**, (2008), « La question documentaire à l'épreuve du numérique : le recours aux fondamentaux », *Sciences de la Société*, n° 75, pp. 40-51.
- **Courbières C.**, (2004), « Documents, signes et savoirs : retour sur l'analyse documentaire », in Metzger J.-P. (dir.), *Médiation et représentation des savoirs*, Paris, L'Harmattan, pp. 159-170.
- **Courbières C. et Couzinet V.**, (2006), « Du bleu horizon à l'horizon documentaire : représentation des connaissances à l'aube de la construction européenne », in Timimi I. et Kovacs S. (dir.), *Indice, index, indexation : Actes du colloque international organisé les 3 et 4 novembre 2005 à l'Université Lille 3*, Paris, ADBS, pp. 81-91.
- **Couzinet V.**, (2008), « Représenter, répertorier, transmettre : formes d'institutionnalisation d'une discipline », in Marteleto R. M. et Thiesen I. (dirs.), *Médiations et usages des savoirs et de l'information : un dialogue France-Brésil*, actes du 1er colloque du réseau MUSSI, Rio de Janeiro, 4-7 novembre 2008, pp. 63-81.
- **Escarpit R.**, (1991), *L'information et la communication : théorie générale*, Paris, Hachette.
- **Fondin H.**, (2005), « La science de l'information ou le poids de l'histoire », *Les enjeux de la communication*, http://w3.u-grenoble3.fr/les_enjeux/2005/Fondin/home.html.
- **Frege F. L. G.**, (1971), *Écrits logiques et philosophiques*, Paris, Points Seuil.
- **Frohmann B.**, (2011), « Reference, representation, and the materiality of documents », in Couzinet V. et Courbières C. (dirs.) *Médiations et hybridations : construction sociale des savoirs et de l'information*, actes du 2e colloque scientifique international du réseau MUSSI, 15-17 juin 2011, Toulouse, Manifestations scientifiques du Réseau MUSSI, pp. 55-67.

- **Jauss H. R.**, (1978), *Pour une esthétique de la réception*, Paris, Gallimard.
- Hjørland B., (2007), « Semantics and Knowledge Organization », *Annual Review of Information Science and Technology (ARIST)*, vol. 41, pp. 367-405, <http://dlist.sir.arizona.edu/2312>.
- **Jeanneret Y.**, (2007), *Y-a-t-il vraiment des technologies de l'information?*, 2e édition, Villeneuve d'Ascq, Presses Universitaires du Septentrion.
- **Jeanneret Y.**, (2008), « La relation entre médiation et usage dans les recherches en information-communication », in Marteleto R. M. et Thiesen I. (dirs.), *Médiations et usages des savoirs et de l'information : un dialogue France-Brésil*, actes du 1er colloque du réseau MUSSI, 4-7 novembre 2008, Rio de Janeiro, pp. 36-59.
- Lund N. W., (2010), « Document, text and medium : concepts, theories and disciplines », *Journal of Documentation*, vol. 66, n° 5, pp. 734-749.
- Mai J. E., (2001), « Semiotics and indexing : an analysis of the subject indexing process », *Journal of Documentation*, vol. 57, n° 5, pp. 591-622.
- **Meyriat J.**, (1993a), « Langages documentaires », in Estivals R. (dir.), *Les sciences de l'écrit : Encyclopédie internationale de bibliologie*, Paris, Retz, pp. 351-352.
- **Meyriat J.**, (1993b), « Documents photographiques et électroniques », in Estivals R. (dir.), *Les sciences de l'écrit : Encyclopédie internationale de bibliologie*, pp. 152-154.
- **Meyriat J.**, (1993c), « Documentation », in Estivals R. (dir.), *Les sciences de l'écrit : Encyclopédie internationale de bibliologie*, Paris, Retz, pp. 150-151.
- **Meyriat J.**, (1993d), « La bibliologie parmi les sciences de l'information et de la communication », in Estivals R. (dir.), *Les sciences de l'écrit : Encyclopédie internationale de bibliologie*, Paris, Retz, pp. 326-335.
- **Meyriat J.**, (2001a), « Document, documentation, documentologie », in Couzinet V. (dir.), *Jean Meyriat, théoricien et praticien de l'Information Documentation*, Paris, ADBS Éditions, pp. 143-159.
- **Meyriat J.**, (2001b), « La science de l'information », in Couzinet V. (dir.), *Jean Meyriat, théoricien et praticien de l'Information Documentation*, Paris, ADBS Éditions, pp. 249-265.
- **Miège B. et Odin R.**, (1993), « Les domaines de compétence de la 71e section », *La Lettre d'Inforcom*, n° 42, pp. 7-8.
- Ollivier B., (2001), « Enjeux de l'interdiscipline », *L'année sociologique*, vol. 51, pp. 337-354.
- **Otlet P.**, (1934), *Traité de documentation : le livre sur le livre*, Brussels, Éditions Mundaneum, Centre de Lecture Publique de la Communauté Française.
- **Pottier B.**, (1992a), *Sémantique générale*, Paris, PUF.
- **Pottier B.**, (1992b), *Théorie et analyse en linguistique*, 2e éd., Paris, Hachette.
- **Raber D. et Budd J. M.**, (2003), « Information as sign : semiotics and information sciences », *Journal of Documentation*, vol. 59, n° 5, pp. 507-522.

- **Rastier F.**, (2009), *Sémantique interprétative*, 3e éd., Paris, PUF.
- **Rey A.** (dir.), (1998), *Dictionnaire historique de la langue française*, Paris, Dictionnaire Le Robert.
- **Saussure F. de**, (1972), *Cours de linguistique générale*, Paris, Payot.
- **Saussure F. de**, (2002), *Écrits de linguistique générale*, établis et édités par Simon Bouquet et Rudolph Engler avec la collaboration d'Antoinette Weil, Paris, Gallimard.

Les concepts information et document à l'épreuve du fait numérique approche épistémologique

Auteur : Isabelle Couturier

Directrice de mémoire : Cécile Gardiès

Année : 2015

Nombre de pages : 147

Résumé :

Approche épistémologique des controverses scientifiques autour des concepts information et document au sein des Sciences de l'information et de la communication. Une vision « sociale », issue des théoriciens fondateurs des sciences de l'information en France, définit l'information comme le contenu cognitif de la communication. Cette conception s'est constituée d'emblée en opposition à l'autre vision « mécaniste » de l'information, élaborée par les ingénieurs de la *théorie mathématique de l'information* au milieu du XX^e siècle, la considérant comme une mesure des phénomènes techniques de communication. L'influence de cette acception transparaît à travers les débats actuels au sein des SIC. Les pratiques et usages informationnels nouveaux induits par la prégnance de la technologie numérique dans notre société donnent ainsi une opportunité réelle aux discours techno-centrés – avec une portée médiatique certaine –, notamment au sujet du document numérique. Néanmoins, les paradigmes originels des sciences de l'information, en mettant l'accent sur la production de sens dans le processus communicationnel, indépendamment de toute pression économique, technique ou idéologique, transcendent le fait numérique.

Mots clés : information – document – épistémologie – sciences de l'information et de la communication – numérique – théorie mathématique de l'information

Summary :

My work is an epistemological approach to the scientific debate on the concepts of information and document in French Informational science. The pioneers of Informational science defined information as the intellectual content of the communicational process. This conception is the opposite of the definition of information developed, in the mid-twentieth century, by the engineers in the mathematical theory of communication, as a measure of the technical process of communication. This theory influences of the current debates in SIC. New uses of information, induced by the importance of digital technology in our society, open the door for media-friendly techno-centric discussions, particularly concerning digital document. Nevertheless, Information Science theories buck this digital trend by emphasising the pursuit of reason in human communication, independent of economic, technical and ideological pressure.

Key words : information – document – epistemology – Informational science – digital – mathematical theory of communication