Master 2 de Recherche
En Sciences Humaines,
Mention Information et Communication,
Spécialité Gestion de l'Information

Université Paul Valéry Montpellier III

Évaluer l'information sur le web, peut-on arriver à une pertinence sociocognitive satisfaisante ?

Études, analyseset hypothèses à partir du corpus documentaire en Sciences de l'Information et de la Communication.

Mémoire de recherche de
Guillaume-Nicolas Meyer.

Sous la direction du Professeur
Alain Chante.

Mots-clés : évaluer l'information, web, autorité cognitive, autorité informationnelle, pertinence, surcharge cognitive, info-pollution, infobésité, crédibilité, confiance, redocumentarisation.

2011 – 2012

Master 2 de Recherche

En Sciences Humaines,

Mention Information et Communication,

Spécialité Gestion de l'Information

Université Paul Valéry Montpellier III

« Évaluer l'information sur le web, peut-on arriver à une pertinence sociocognitive satisfaisante ? »

Études, analyses et hypothèses à partir du corpus documentaire en Sciences de l'Information et de la Communication.

Mémoire de recherche de Guillaume-Nicolas Meyer.

Sous la direction du professeur Alain Chante.

2011 – 2012

Soutenance le 4 juillet 2012

Mention particulière

L'université n'entend donner aucune approbation ni improbation aux opinions émises dans ce mémoire : ces opinions doivent être considérées comme propres à leur auteur.

Dédicace

Je dédicace ce travail à ma femme et à mes trois enfants. C'est à eux que j'ai emprunté du temps, de la patience, de la disponibilité, de la persévérance… et beaucoup d'amour.

A Aurélie, Luce, Jan, Anaë.

Remerciements

Je remercie,

leProfesseur Alain Chante pour ses conseils avisés,

Madame Céline Paganelli pour ses encouragements,

Madame Coleta Vaisman pour ses pistes inattendues,

Vincent-Étienne pour ses pertinentes réflexions,

tous ceux, spécialistes et non-spécialistes, qui ont enrichis ma vision, mes propos et avec qui j'ai pu confronter mes idées.

Avant-Propos

Nous évaluons en permanence de l'information mais toutes les informations doivent-elles être évaluées ? Notre cerveau nous joue-t-il des tours ?

« La terre est bleue comme uneorange. »

Paul Eluard

Sommaire

Synthèse 12

Introduction 13

1. Partie 1 - Perception et contextualisation 24

1.1 L'environnement socio-cognitif, une typologie de contextes 24

1.1.1 Le contexte 24

1.1.2 La typologie quadripartite d'Armengaud 25

1.1.3 Construction de son échelle d'intentionnalité 26

1.2 Un environnement en permanente mutation 30

1.2.1 Un contexte en pleine mutation 30

1.2.2 Quelle version du web avez-vous ? 33

1.2.3 Web sémantique 34

1.3 Les pollutions informationnelles 40

1.3.1 Un risque écologique de la communication 40

1.3.2 Surcharge cognitive 42

1.3.3 Les invariants 43

1.4 Situer l'information sur le web 45

1.4.1 Quel raisonnement adopter ? 45

1.4.2 Situer dans le temps 47

1.4.3 Evolution des typologies de recherche 53

2. Partie 2 – Langage et désubjectivation 56

2.1 Lire, cette pratique 56

2.1.1 Lisibilité visuelle et lisibilité cognitive 57
2.1.2 Quand les neurones s'en mêlent… mais ils s'en mêlent tout le temps .. 58
2.1.3 Est-ce qu'on lit encore sur le web ? 61
2.1.4 hypertextualité et spatialisation de l'information 62
2.1.5 Naviguer à l'aide d'une carte 63
2.2 Les figures d'autorité 72
2.2.1 Autorité cognitive et autorité informationnelle 72
2.2.2 L'auteur, une figure d'autorité contestable 75
2.3 Faire confiance sur le web 77
2.3.1 Quelle typologie de confiance sur le web ? 77
2.3.2 Les moteurs de confiance 79
3. Partie 3 - Mémoire et redocumentarisation 82
3.1 Des informations pour évaluer d'autres informations 82
3.1.1 Stratégies et mémoires 82
3.1.2 Un stockage protéïforme 82
3.2 Attention et appropriation 84
3.2.1 L'économie de l'attention 84
3.2.2 Le processus d'appropriation 86
3.3 Re-concevoir ses stratégies 87
3.3.1 Répondre à ses besoins 87
3.3.2 Les stratégies de base 89
3.3.3 Les stratégies élaborées 90
3.4 Re-documentariser 93
3.4.1 Les mots-clés, amélioration ou pollution des objets informationnels ? . 94
3.4.2 Partager ses savoirs 97

Conclusion 99

BIBLIOGRAPHIE 102

SITOGRAHIE 109

INDEX 111

TABLE DES ILLUSTRATIONS 112

Synthèse

Evaluer l'information sur le web est un nouveau paradigme cognitif. La profusion des documents, des supports, leurs interconnections, modifient en profondeur nos façons de chercher une information, nos stratégies d'évaluation et d'appropriation et donc nos schèmes cognitifs. La contextualisation et l'identification sont des tâches plus complexes sur le web, de par la nature protéiforme et pluriculturelle de l'information. Les mécanismes de lecture traditionnels ne sont plus opérants dans un espace virtuel sans limites, dé-temporalisé et graphiquement instable. Les mécanismes d'évaluation de l'information, comme la construction de ses autorités informationnelles ou de la pérennité de ses autorités cognitives,prennent maintenant une dimension collective qui passe systématiquement par une redocumentarisation et un brassage des taxonomies personnelles.

114 mots.

Mots-clés : évaluer l'information, web, autorité cognitive, autorité informationnelle, pertinence, surcharge cognitive, objet informationnel, info-pollution, infobésité, crédibilité, confiance, redocumentarisation.

Introduction

« *Le but de l'évaluation est de savoir s'il faut ou non sélectionner, exploiter et s'approprier l'information en question. Il s'agit en fait d'attribuer une valeur, une signification à l'information au moyen d'un jugement*. » (Duplessis, 2007)

« *On croyait que le numérique nous apporterait un accès facilité à la connaissance, il n'en est rien. La littératie se complexifie au contraire et l'illusion de la transparence dissimule délégations techniques et intellectuelles*. » (Le Deuff, 2010)

Axes de recherche

Le jugement de la crédibilité des sources d'information est un thème qui interpelle les chercheurs en sciences de l'information et de la communication depuis toujours (Rouet *et al*, 2008). À l'heure où chacun utilise le web quotidiennement, que ce soit dans la sphère professionnelle ou personnelle, il nous intéresse de comprendre les mécanismes que chacun met en place pour arriver à une pertinence sociocognitive suffisamment satisfaisante pour « classer » un document en terme de confiance et d'utilisabilité : « populaire, mais pas fiable », « crédible, mais pas utile », « pertinent, mais inutile », etc.

Dans notre démarche de recherche, nous nous interrogerons selon trois axes principaux.

Le premier est lié à au media en lui-même et interroge sur nos facultés de perception sur le web. Le web est-il vraiment un espace à part, avec des règles de fonctionnement tellement spécifiques que nos mécanismes cognitifs d'évaluation de l'information en serait perturber ? Est-ce que le web introduit de nouvelles contraintes cognitives ? Est-ce que les mutations quasi-permanentes de cet espace immatériel permettent à l'internaute d'établir des stratégies d'évaluation, de recherche, pérennes ?

Le deuxième axe de recherche tient au concept de pertinence proprement dit. Comment juge-t-on que ce que l'on a trouvé est fiable au point de se dire qu'on va

utiliser cette information, parfois dans un contexte de crise ? Quel sont les mécanismes qui font que l'on investit une information au point de décider de la transmettre à un tiers, de s'en porter garant ? Brigitte Simonnot (2008b) pose la question en ces termes« *Les critères traditionnels d'évaluation de l'information sont-ils toujours opérants, à l'heure où les volumes d'information potentielle ne cessent de croître ?* »

Enfin le troisième axe de recherche porte sur nosfaçons de transformer l'information. Nous partirons de la définition d'Yves Jeanneret (2000) pour qui « *l'information n'existe pas en soi, qu'elle est toujours le produit d'une interaction, d'un regard humain sur un objet, qui est le document* ». Quels outils offre le web aujourd'hui pour co-évaluer une information, pour co-construire un réseau, pour partager des connaissances ? L'évaluation de l'information va-t-il devenir un processus collectif ?

Le web et pas Internet

Un protocole parmi tant d'autres

Nous avons choisi de porter notre recherche sur une partie du maillage global de circulation de l'information que représente Internet : le World Wild Web. Cette partie du réseau des réseaux est un véritable écosystème en soi, constitué, selon une approche systémique, de systèmes ouverts (sites internet) et de systèmes fermés (extranet[1]). Il nous apparaît également que les procédures d'évaluation de l'information ne sont pas les mêmes sur le web que sur les autres protocoles communs utilisés sur Internet. L'accès à « la toile » se fait avec le protocole http (HyperText Transfert Protocol).

Sémiologie du web

On peut noter, sur cet espace que nous allons étudier, sa sémiologie particulière et son héritage Nord-américain. Le protocole utilisé qualifie déjà l'information puisque l'on sait qu'il s'agit d'hyper texte, un texte amélioré,

[1] Sites accessibles depuis une adresse web mais dont tout ou une partie de son contenu est en accès restreint et nécessite une authentification.

enrichi. Le World Wild Web est également présenté comme un espace sauvage, à l'image du « *Wild West* ».

Protocole smtp et pop3

Le courriel est majoritairement évalué avec une composante « confiance » très différente de celle d'un site web, puisque l'on connaît ou identifie facilement l'émetteur du message. Il y a également une dichotomie très forte entre les courriels de la sphère privée et ceux de la sphère professionnelle. Nous verrons que la navigation sur le web échappe à cette dichotomie.

Telnet, IRC et compagnie

Ces protocoles de discussion, les premiers du genre, peu connu du grand public mais encore utilisé, échappent aux usages de masses, ils ont donc, de par leur existence même, une sociologie atypique que l'on ne retrouve pas sur le web. D'autre part, ces protocoles ne présentent que du texte et échappent ainsi aux influences métacommunicationnelles des interfaces graphiques et de des multiples médias qui font la richesse du web.

Les autres protocoles

Bien qu'établissant une communication entre un émetteur et un récepteur selon la théorie de l'Information de Shannon et Weaver (1975), nous estimons que les usages des protocoles tels que p2p (peer-to-peer) ou ftp (File Transfer Protocol) génèrent de trop faibles interactions humaines pour évaluer l'information sur le même paradigme que celle présente sur le web.

Histoire et hypertextualité

A travers les pensées, idées et travaux de trois hommes, nous allons démontrer que « l'esprit » web, que l'on peut qualifier de communautaire, international, interculturel, n'est pas du tout né avec la génération X[2], ne s'est

[2] Personnes nées approximativement entre 1960 et 1980.

pas du tout révélé avec la génération Y[3] mais possède une philosophie, une logique et une conception héritée à la fois du monde de la documentation du début du siècle, de la pensée scientifique de la seconde guerre mondiale et de la contre-culture américaine des années soixante.

La philosophie du web

En 1934, dans son traité de documentation, Paul Otlet imaginait déjà « *des œuvres classiques ou de grande actualité photographiées et mises "en débit constant"dans les annexes des bibliothèques. Chacun pourrait ainsi, à volonté et de loin, obtenir la lecture désirée.* » C'est bien ce que propose les médiathèques quelque soixante-dix ans plus tard.

On retrouve ici une des idées fondamentales du réseau Internet et de son usage principal, le web : rendre accessible des contenus au plus grand nombre et à distance. Pour Paul Otlet « *Le Réseau, de quelque manière que ce soit, doit relier les uns aux autres, les centres producteurs, distributeurs, utilisateurs, de toute spécialisation et de tout lieu*. » Déjà cette vision s'affranchissait des frontières et des cultures.

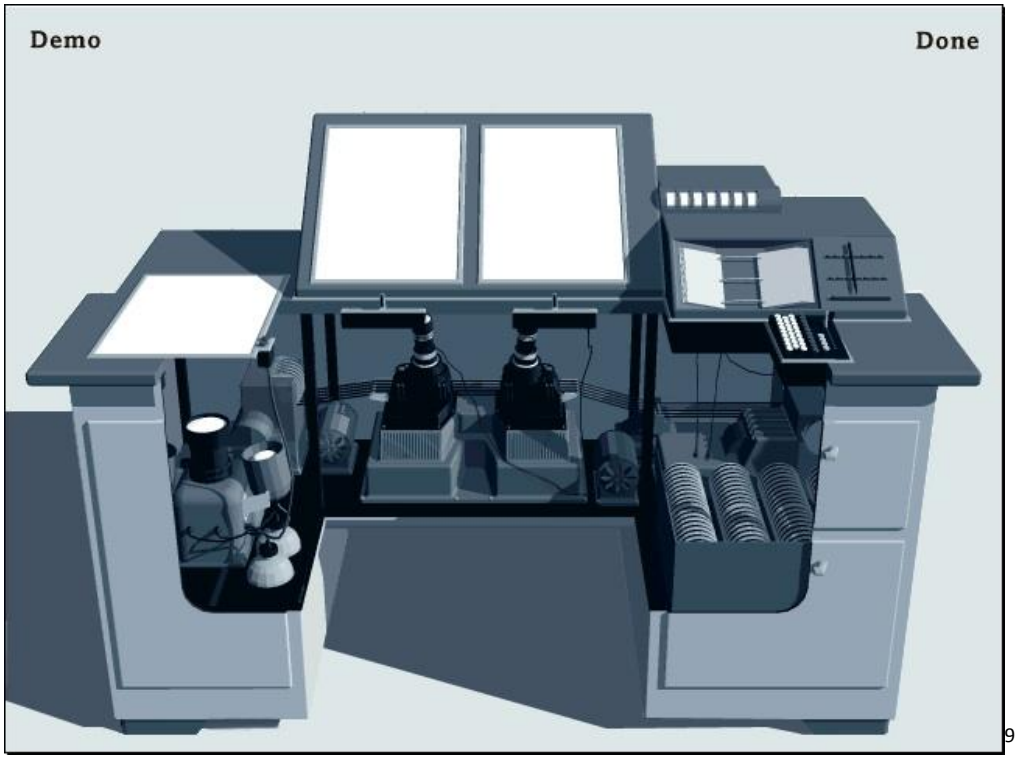
Le réseau de données, l'approche scientifique et rigoureuse

Mathématicien et physicien américain, Vannevar Bush (1890 – 1974) est un visionnaire de la circulation de l'information. Gestionnaire de la recherche scientifique pendant la seconde guerre mondiale, il participe notamment au « Manhattan Project[4] ». Il conçoit le Memex, un appareil qui offre la possibilité à son utilisateur, de créer des liens entre deux documents,d'annoter son parcours de lecture et de conserver la trace de son cheminement. Dans son article publié en juillet 1945 dans le journal « The Atlantic[5] », Vannevar Bush est clairement précurseur de l'ordinateur

[3] Personnes nées approximativement entre 1980 et 2000. Ce sont eux que l'on surnomme les « digitales natives ». Ces jeunes sont « tombés dans la potion numérique dès leur naissance » selon Hervé Sérieyx (2002).
[4] Conception de la bombe atomique
[5] L'article est toujours consultable à l'adresse http://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/3881/

personnel : *« Imaginons un appareil de l'avenir à usage individuel, une sorte de classeur et de bibliothèque personnels et mécaniques. […] Un memex, c'est un appareil dans lequel une personne stocke tous ses livres, ses archives et sa correspondance, et qui est mécanisé de façon à permettre la consultation à une vitesse énorme et avec une grande souplesse. Il s'agit d'un supplément agrandi et intime de sa mémoire. »*

On retrouve aussi dans ces idées ce qui fait l'intérêt et l'attractivité des systèmes de gestion des signets aujourd'hui, tels que Pearltrees[6], Diigo[7], ou anciennement Delicious[8]. Déjà pour Vannevar Bush, il s'agit d'une orthèse cognitive, une extension de sa mémoire.



[9]

Figure 1: Le Memex

[6] www.pearltrees.com
[7] www.diigo.com
[8] www.delicious.com
[9] Infographie trouvée sur le site du groupe de recherche sur les écritures hypertextuelles http://www.hypertexte.org/blog/?p=125

L'hypertextualitélittéraire et utopiste

Chercheur en sciences humaines et en littérature, Ted Nelson[10]invente en 1965 le terme HyperText, pour désigner *« une écriture-lecture non-linéaire donnant à l'utilisateur une liberté de mouvement ».* Il est également à l'origine de l'utopiste projet de la bibliothèque universelle, le projet Xanadu, « *horizon absolu de l'hypertexte*». La vision de Ted Nelson est à l'origine du mythe « d'Internet pour tous », capable de briser les chaînes, d'abolir les distances et de rendre les savoirs accessible à tous. Pour autant, au-delà du mythe, le projet Wikipédia en est un héritage concret et les phénomènes du « Printemps Arabe » de 2011 se sont largement appuyés sur les sciences de l'information et de la communication et les technologies afférentes pour appuyer des mouvements sociaux et politiques de grandes ampleurs.

Une information sur le web

Nous retiendrons la définition que propose la FADBEN sur son site : « *L'information est le produit de la mise en forme, par le recours à un code de transcription particulier (oral, gestuel, textuel, iconique, sonore ou multicodes), d'un élément de connaissance ou de sens, transmis au moyen d'un message, par un canal de transmission. L'information provient d'une source et procède d'une intention de communication.* »[11]

Dans l'exercice périlleux qui consiste à définir ce qu'est une information sur le web, il convient de commencer par la forme. Le web est en effet un espace protéiforme et polysémique. On y trouve des informations de toutes sortes formes et de toutes natures. Le texte côtoie les images, les vidéos et les sons, l'information politique côtoieles produits de tête de gondole, la publicité, les exploits sportifs ainsi que les découvertes scientifiques.

Si le texte est la première forme, la plus importante, elle n'est pas forcément la préférée des internautes. La première hypothèse utilisée par l'équipe de

[10] De son vrai nom Theodor Holmes Nelson.
[11] http://fadben.asso.fr/wikinotions/index.php?n=Accueil.Information

recherche de Gunter*et al* (2009) pour qualifier la génération Google est d'ailleurs « *La génération Google préfère l'information visuelle à l'information textuelle* ».

Les trois dimensions d'un document

Le collectif Roger T. Pédauque propose trois dimensions constitutives d'un document : la forme, le signe et le médium (Pédauque, 2006).

Le document comme forme

« *Sous cette rubrique, nous rangerons les approches qui analysent le document comme un objet, matériel ou immatériel et qui en étudient la structure pour mieux l'analyser, l'utiliser ou le manipuler.* »(Pédauque, 2006).

Alors que les livres ont quasiment tous la même forme, on peut être surpris sur le web, par la variabilité des formes auxquelles qui nous sont proposées.

Le document comme signe

« *Pour ces chercheurs le document est perçu avant tout comme porteur de sens et doté d'une intentionnalité ; ainsi le document est indissociable du sujet en contexte qui le construit ou le reconstruit et lui donne sens ; en même temps, il est pris dans un système documentaire ou un système de connaissances.* »(Pédauque, 2006).

Mais l'intentionnalité d'un message sur Twitter est-il vraiment indissociable du contexte de la plate-forme ?

Le document comme médium

« *Cette dimension enfin pose la question du statut du document dans les relations sociales ; le document est une trace, construite ou retrouvée, d'une communication qui s'est affranchie de l'espace et du temps ; en même temps, il est un élément de systèmes identitaires et un vecteur de pouvoir.* »(Pédauque, 2006).

Quel statut ont les milliards de documents publiés chaque jour par les utilisateurs ? S'agit-il d'une trace à vocation mnésique ou d'une externalisation de la création de son Moi ?

Epistémologie de la pertinence

La pertinence en Sciences de l'Information et de la Communication est souvent abordée via le paradigme du documentaliste : proposer un document pertinent à un usager. De là découle la notion de pertinence des résultats de recherche à une requête donnée. C'est ce paradigme que les usagersont en tête quand ils évaluent les résultats fournis par les moteurs de recherche.Or c'est un tort, le médiateur a changé et avec lui le paradigme.

De la pertinence à la pertinence socio-cognitive

Qu'est-ce qui fait que quelque chose devient une information pertinente pour une ou plusieurs personnes, à un moment donné, dans uncontexte donné ? Et à quelle condition cette information fait-elle sens, se transforme-t-elle en connaissance ? La notion de *pertinence sociocognitive*, proposée par Cosijn et Ingwersen (2000) permet de répondre à ces questions.

Cette notion désigne « *la relation établie entre une situation, une tâche à effectuer ou un problème à résoudre dans un contexte socioculturel donné et les objets informationnels, tels qu'ils sont perçus par un ou plusieurs agents cognitifs.* » (Simonnot, 2008a). Pour Cosijn et Ingwersen la pertinence socio-cognitive est un type de pertinence subjective qui inclut la notion de relation avec l'autre, les autres. Nous le verrons tout au long de notre recherche, mais nous pouvons l'accepter comme postulat de base, le web est une agora, une place sociale. Or les processus d'évaluation de l'information induisent aujourd'hui une dimension collective et les influences qui l'accompagnent, qui n'est pas assez représenté dans les modèles classiques de pertinence. De plus, Cosijn et Ingwersencaractérisent la pertinence socio-cognitive par « *la nouveauté, le degré d'informativité, les préférences et la qualité des objets*

informationnels relatifs au besoin d'information de l'utilisateur à un moment donné »(Simonnot, 2008a), des critères qui, selon nous, caractérisent une activité informationnelle sur le web.

Si la formulation « d'environnement socio-cognitif » est nouvelle, le concept ne l'est pas. A partir du moment où l'on pense à l'Autre, nous sommes dans un contexte d'évaluationsocio-cognitif. A partir du moment où l'on situe l'Autre par rapport à soi, nous sommes dans l'évaluation d'un environnement socio-cognitif. Sur le web, lors d'une recherche, les premières questions qui nous viennent sont « Qui écrit ? Qui est-il par rapport à moi ? Qui suis-je par rapport à lui ? ». Nous verrons que pour évaluer correctement une information dans sa dimension sociale, nous mettons en œuvre plusieurs grandes fonctions cognitives telles que la perception, le langage et la mémoire.

Psycho-anthropologie et néo-darwinisme

Si nous faisons intervenir les sciences cognitives contemporaines dans un mémoire de recherche en sciences de l'information et de la communication, nous tenons néanmoins à en fixer la limite des apports et à en qualifier la teneur philosophique.

De la psycho-anthropologie dite « évolutionnaire », nous retiendrons que la vie psychique et intellectuelle d'un individu est tributaire de son corps physique. « *On n'a en effet aucune raison d'ignorer qu'il existe un corrélat neuro-physiologique pour toutes les expériences et fonctions de l'esprit, du plaisir à la mémoire en passant par l'attachement et le calcul logique*. » (Pharo, 2003)

C'est un fait empirique pour chacun d'entre nous : lorsque nous évaluons une information, notre corps entre en jeu. C'est ainsi que ma perception du monde n'est pas la même que celle d'un enfant de cinq ans, haut de son mètre. Tout comme l'évaluation du film à aller voir avec des amis n'engendre pas la même inquiétude que l'évaluation des thérapies possibles pour une maladie grave dont on est atteint. Nous connaissons maintenant le rôle des neurotransmetteurs dans

les processus cognitifs et, même si une base neurochimique de la mémoire (Laroche, 1998) n'a pas été entièrement validée d'un point de vue expérimental, c'est une piste qui s'inscrit bien dans le modèle néo-darwinien.

Plusieurs études sur les comportements informationnels mettent en évidence les émotions qui accompagnent les usages des navigateurs web, moteurs de recherche et banques de données spécialisées :« *Certaines [nda : études] soulignent les émotions qui les accompagnent, de l'anxiété liée à la difficulté de formuler un sujet à la satisfaction et au renforcement de l'estime de soi lorsque la recherche aboutit* » (Simonnot et Boubée, 2010).C'est pourquoi nous retiendrons, même si nous n'en expliquons pas les tenants et les aboutissants, qu'un individu qui évalue une information est fortement affecté par son corps et par les mécanismes chimiques qui le régule. Hors périmètre de notre recherche, nous retiendrons de cette approche que les mécanismes corporelsont une influence conséquente sur nos capacités à évaluer l'information.

Approche méthodologique

Nous allons chercher à établir des liens entre les différentes étapes identifiées pour évaluer une information sur le web et trois fonctions cognitives universelles : la perception, le langage, et la mémoire. Pour ce faire, nous allons analyser le corpus documentaire existant, aussi bien en sciences de l'information et de la communication que dans les disciplines connexes qui façonnent l'hexagone cognitif d'un internaute. Nous tenterons de garder à l'esprit ces trois fonctions dont nous ferons une courte présentation au début de chaque partie pour répondre à la question majeure : comment un internaute évalue une information sur le web aujourd'hui ?

Nous allons nous intéresser dans une première partie à la contextualisation d'une information. Après avoir présenté une typologie de contextes ainsi que les mutations engendrées par le web en termes de circulation de l'information et, de façon induite, à notre rapport à l'information, nous étudierons les mises en œuvre de la fonction cognitive « perception ». Nous verrons combien il est

de plus en plus difficile d'évaluer correctement une information et que, contrairement aux idées reçues, ceux que l'on qualifie de « *Digitales Natives* » n'ont pas plus de facilités à effectuer ces tâches. En précisant quel est notre environnement socio-cognitif aujourd'hui, notre verrons, à travers des explications simples, comment les sciences cognitives peuvent éclairer les phénomènes de surcharge cognitive et les info-pollution qui y sont associées.

La deuxième partie, basée sur les explications de la fonction cognitive « langage », nous permettra d'étudier les mécanismes de la lecture sur écran et sur le web plus particulièrement, avec, notamment, les effets de l'hypertextualité. Après avoir défini les concepts d'autorité cognitive et d'autorité informationnelle et mis en évidence les critères de construction sur lesquels ils s'appuient sur le web, nous déclinerons les différents modes de confiance que nous mettons en place. Nous verrons que nous attribuons une pertinence à un document, souvent de façon contextuelle et pas empirique, ce qui laisse à penser, que bien que nous construisions des stratégiques d'apprentissage et d'évaluation de l'information sur le web, nous gardons, consciemment ou pas, le libre arbitre de ne pas les utiliser.

Enfin, la troisième partie, basée sur la fonction cognitive « mémoire » sera l'occasion de lister les outils qui nous servent à stocker l'information. Nous étudierons le processus d'appropriation qui passe notamment par la redocumentarisation de l'information trouvée, ainsi que par la mise à jour de sa taxonomie personnelle. Nous nous interrogerons sur l'externalisation croissante de cette fonction cognitive essentielle qui définit le Soi au sens philosophique de l'Être.

1. Partie 1 - Perception et contextualisation

1.1 L'environnement socio-cognitif, une typologie de contextes

1.1.1 Le contexte

« *La notion de contexte est "primaire et fondamentale" dans toute communication.* » Gregory Bateson.

Miermont*(2003)* qualifie la perspective ouverte par Gregory Bateson comme suit *« le contexte est la concaténation d'une série (…) de systèmes de référence, dont certains sont pragmatiques, effectifs ou poïétiques (…) et dont d'autres sont herméneutiques, cognitifs, sémantiques (…)* ».

En naviguant sur le web, en passant de site en site avec une lecture hypertextuelle, on se retrouve très vite dans un contexte différent du contexte initial. Le « bruit » informationnel provoqué par notre propre navigation modifie et redéfinit le contexte initial. A la fois émetteur et récepteur d'information, nos communications avec les différents dispositifs nous soumettent à des flux méta-communicationnels qui perturbent l'évaluation socio-cognitive du contexte, et donc de la pertinence de l'information. Sur le web, « *La prégnance du modèle de la communication de Shannon et Weaver, avec un émetteur, un message et un récepteur [est] très forte et se heurte à l'autre modèle, celui d'un réseau dans lequel il n'y a pas un mais plusieurs émetteurs, dans lequel il n'y a pas un mais plusieurs récepteurs et dans lequel le bruit de fond n'est plus une pénalité, une pénibilité, mais peut-être l'ensemble des interlocuteurs potentiels.[Perriault 01 p.38]»,*cité dans la thèse d'Olivier Ertzscheid (2004).

C'est dans cette acception qu'il nous semble important de s'appuyer sur un outil méthodologique pour contextualiser l'information sur le web.

1.1.2 La typologie quadripartite d'Armengaud

F. Armengaud (1982, 1993) propose une typologie quadripartite des contextes qui devrait nous permettre de contextualiser l'information sur le web. Elle est répartie comme suit : le contexte circonstanciel, le contexte situationnel, le contexte interactionnel et le contexte présuppositionnel. En adéquation avec la théorie de l'Information de Shannon et Weaver (1975), il souligne également que le contexte évolue en même temps que l'information.

Le contexte circonstanciel

Rapporté au web, le contexte circonstanciel correspond à notre perception de l'identité des auteurs, à leur contextualisation physique et temporelle. Elle suffit pour beaucoup à discriminer la pertinence ou non de l'information qu'ils ont produite. Le contexte circonstanciel est extrêmement important sur le web car nous évoluons dans un contexte multiculturel. Ainsi se mélange dans le paysage informationnel français en juin 2011, les primaires socialistes et les révolutions du Printemps Arabe[12]. Nous avons bien là des contextes factuels, des référentiels spatio-temporels différents. De même l'identité de l'émetteur du message propose de fait un contexte circonstanciel différent. Nous n'avons pas la même perception d'un message si l'émetteur est notre chef de l'état que si l'émetteur est le webmestre d'un site commercial grand public.

Le contexte situationnel

Le contexte situationnel, ou paradigmatique, correspond au « bruit » culturel de l'environnement (pays, secteur industriel, etc.) sur le message. Ce contexte établit une adéquation, un permet un delta, entre l'intention culturelle de l'émetteur du message et son interprétation effective par le récepteur. Il est plus difficile sur le web d'établir une ou plusieurs finalité pour un cadre socialement reconnu et présentant un sens immanent. En effet sur le web plus qu'ailleurs, les informations sont (très ?) facilement sorties de leur

[12] Phénomènes politico-sociaux dont l'apogéeà eu lieu au printemps 2011 et qui ont entraînés des changements de régime en Tunisie, en Egypte ainsi que des répercussions au Yémen, en Syrie et en Libye.

contexte initial. Ceci est dû à la nature protéiforme des informations qui circulent, ainsi qu'au phénomène « Tour de Babel » du web puisqu'il s'agit du seul média méta-pluri-culturel.

Le contexte interactionnel

Le contexte interactionnel correspond à la vague produite sur le web par les processus sociaux de génération de contenus, connus notamment sous le terme de web 2.0. Ce sont les réactions provoquées par une information, les commentaires sur un article, les objections, les interrogations et les réponses apportées sur un produit, les rétractations des auteurs lorsqu'ils sont pris en fautes, etc. Le contexte interactionnel du web ne peut se comparer avec les autres médias puisque la notion d'interaction y est en temps réel et est une composante même de ce média. Si pour un journal papier, le « droit de réponse » existe, il est en général usité dans un délai de plusieurs jours. Sur le web ce délai est de l'ordre de l'heure, voire de la minute.

Le contexte présuppositionnel

Le contexte présuppositionnel sur le web met à mal nos capacités d'évaluation de l'information en ce, qu'il est un frein inné, anthropologique. Nous avons tous en effet, des croyances, des a priori, qui influent sur notre perception de l'information. Une des plus dommageables sur le web est de « croire » en la neutralité d'un moteur de recherche. Nous prêtons également des intentions à certains producteurs de contenus et présupposons certaines valeurs : les journalistes qui génèrent du contenu vérifient leurs sources, les informations sur Wikipédia sont moins fiables que qu'une encyclopédie imprimée, Google est altruiste…

1.1.3 Construction de son échelle d'intentionnalité

On évalue de plus en plus d'informations sur le web tout simplement car il y a de plus en plus d'informations disponibles (santé, bricolage, boutique en ligne, etc.). Pour autant, les quatre facteurs les plus importants pour évaluer

l'information sont opérants quel que soit le média : la vigilance, le besoin, l'affectivité liée au besoin, l'environnement immédiat ou la proxémie.

Le pourquoi on accède à l'information est presque aussi important que l'information en elle-même. En effet il permet la contextualisation du contenu et donc son évaluation.

Pyramide de Maslow

La pyramide originale de Maslow définit cinq niveaux de besoins qui constituent la base des motivations des individus. Présents chez chacun de nous de façon inconsciente, ces besoins sont hiérarchisés et régit par une règle simple : un besoin supérieur ne devient conscient que lorsque les besoins inférieurs sont satisfaits. Par exemple, un individu ne peut se sentir en sécurité (niveau 2) si sa première préoccupation est de trouver à boire et à manger (niveau 1).

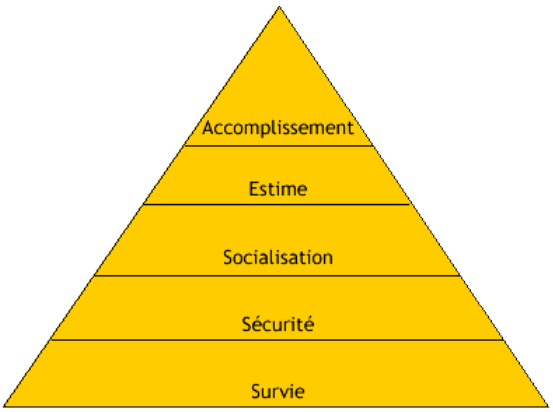


Figure 2 : Pyramide des besoins de Maslow.

Pyramide des besoins informationnels

Nous avons transposé les cinq niveaux de la pyramide afin qu'ils correspondent aux besoins informationnels d'un individu.

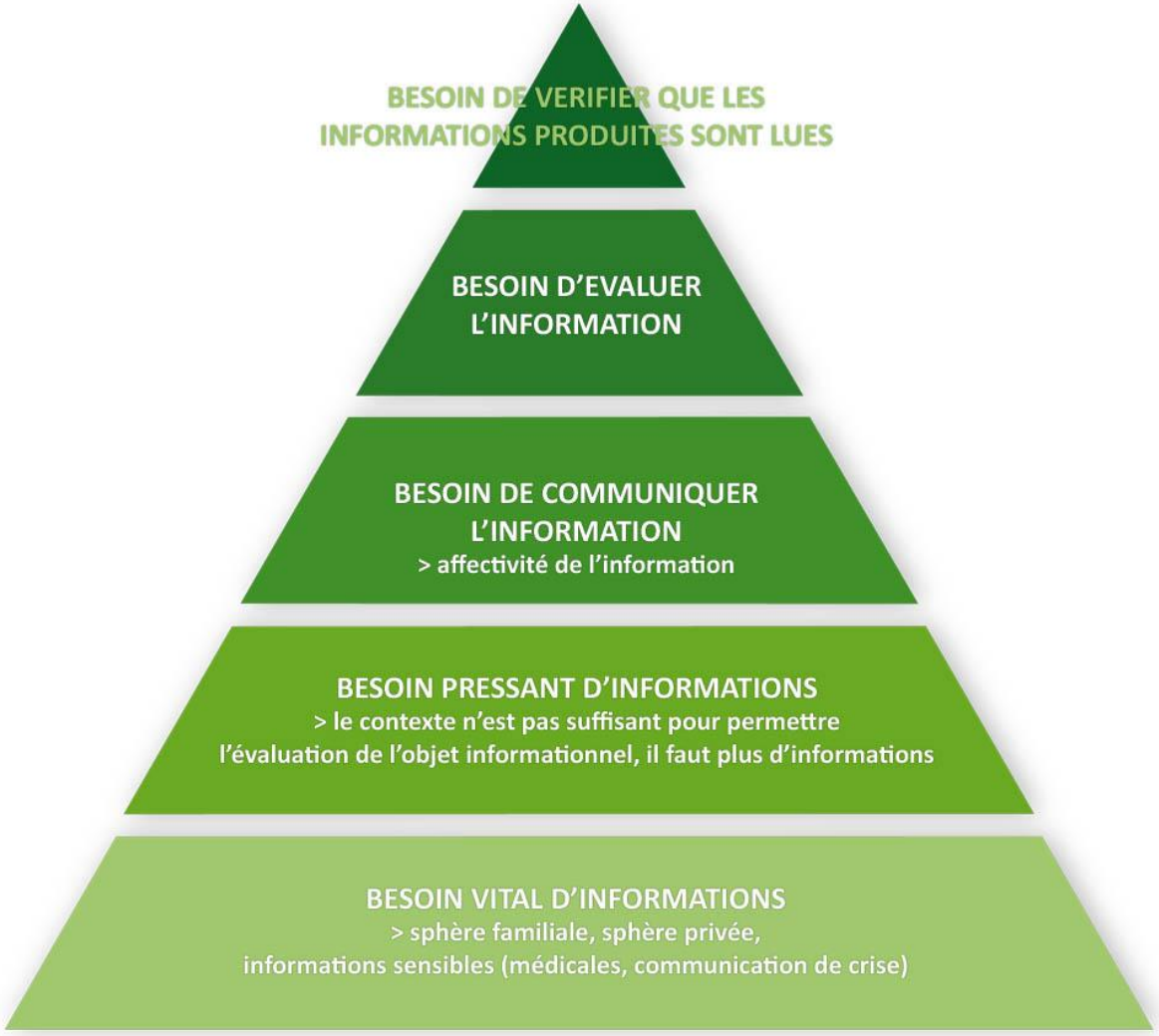


Figure 3 : Pyramide des besoins informationnels.

Nous avons déterminé cinq niveaux de besoin d'informations : le besoin vital, le besoin pressant, le besoin de communiquer une information, le besoin d'évaluer une information, le besoin de vérifier la diffusion d'une information. Faisons le parallèle avec la pyramide originelle de Maslow.

Le besoin vital d'informations correspond à une situation de crise : il faut trouver la bonne information. Il peut s'agir d'un contexte personnel, comme le numéro d'urgence d'un centre antipoison parce que son enfant a avalé un produit dangereux ou bien d'un contexte professionnel comme identifier le

client mécontent dans un groupe hostile afin d'enrayer l'atteinte à la réputation de l'entreprise. Le besoin de trouver est le plus fort des besoins informationnels car il induit obligation de résultat et de pertinence.

Le besoin pressant d'informations n'engage pas la sphère vitale, ce qui ne l'empêche pas d'être important. Dans un contexte professionnel nous pouvons citer le besoin de bien cibler un marché sous peine de voir la concurrence l'emporter, ce qui basculerait peut être les services marketing et commerciaux dans un besoin vital d'informations.

Le besoin de communiquer l'information relève clairement d'un besoin affectif de sociabiliser un ou plusieurs objets informationnels. Il n'y a plus de pression pour identifier une information, il s'agit maintenant de la faire circuler. Des parents qui envoient les photos de leurs enfants ou le plan de communication d'une entreprise relève de ce type de besoin.

Le besoin d'évaluer l'information est celui que nous rencontrons lorsque nous naviguons sur le web. Notons que la notion de navigation est ici comprise comme un acte à part entière et pas une technique d'appropriation de l'information. L'information nous arrive (info-pollutions, surcharge cognitive) et notre seul besoin est d'évaluer cette information.

Enfin le besoin de vérifier la diffusion d'une information fait écho au besoin d'accomplissement d'un individu ou d'une organisation. On mesure sa popularité, on consulte les statistiques de fréquentation d'un site, son Klout[13], son nombre de followers, son nombre d'amis, etc.

La pyramide de nos besoins informationnels nous permet de nous situer dans notre environnement socio-cognitif et informationnel.

[13] www.klout.com

1.2 Un environnement en permanente mutation

1.2.1 Un contexte en pleine mutation

Un changement d'échelle

Le web est un espace dans lequel les services proposés peuvent voir leur consommation littéralement exploser suite à l'adoption massive par les utilisateurs. La société Socialnomics[14] aime à rappeler dans ses vidéos[15] de présentation que pour réunir 50 millions d'utilisateurs, il a fallu 38 ans au média radio, 13 ans au média télévision et 4 ans au média Internet. Quand dans notre histoire avons-nous vu un service passer de 10 à 700 millions d'utilisateurs en moins de 10 ans ? C'est le cas de Facebook[16]. Les principales fonctionnalités plébiscitées par les utilisateurs, sont jeunes : LinkedIn, le réseau social professionnel n'a que 9 ans, Facebook a 8 ans, YouTube, le deuxième moteur de recherche au monde, a 7 ans, Twitter a 6 ans en 140 caractères, et Google+ tout juste un an.

Alors la difficulté réside non seulement dans le changement d'échelle quasi entropique de la production de documents numériques mais également dans la vitesse de mutation des usages. Et dans ce cas, la loi de Moore[17] est dépassée. En 1965, Gordon Moore, co-créateur d'Intel, fait un postulat empirique : les capacités des terminaux, des serveurs et des réseaux doublent tous les 18 mois et corrélativement, cela entraîne une diminution proportionnelle des coûts.

Plusieurs économistes gageaient que le modèle économique des industries liées à Internet serait basé sur cette théorie. Or des sociétés comme Twitter ou Facebook n'ont pas ce genre de modèle économique, tout simplement parce qu'à leur création elles n'avaient pas de modèle économique. Créé par

[14] http://www.socialnomics.net/about-socialnomics/
[15] http://www.socialnomics.net/2011/07/05/social-media-videos-2011/
[16] Nombre d'utilisateurs revendiqués et pas effectifs.
[17] http://www.intel.com/cd/corporate/techtrends/emea/fra/209837.htm

pulsion au lieu de l'analyse d'un marché, les services qu'ils ont proposés ont répondu à un besoin qui n'avait pas été encore anticipé.

Pour preuve de la décorrélation entre la loi de Moore et les économies liées à Internet, il suffit d'observer l'évolution de la courbe du nombre de noms de domaine depuis 1995 fournit par le site Netcraft[18]. On y voit, de façon assez visible, que la progression se fait par à coup brutaux.

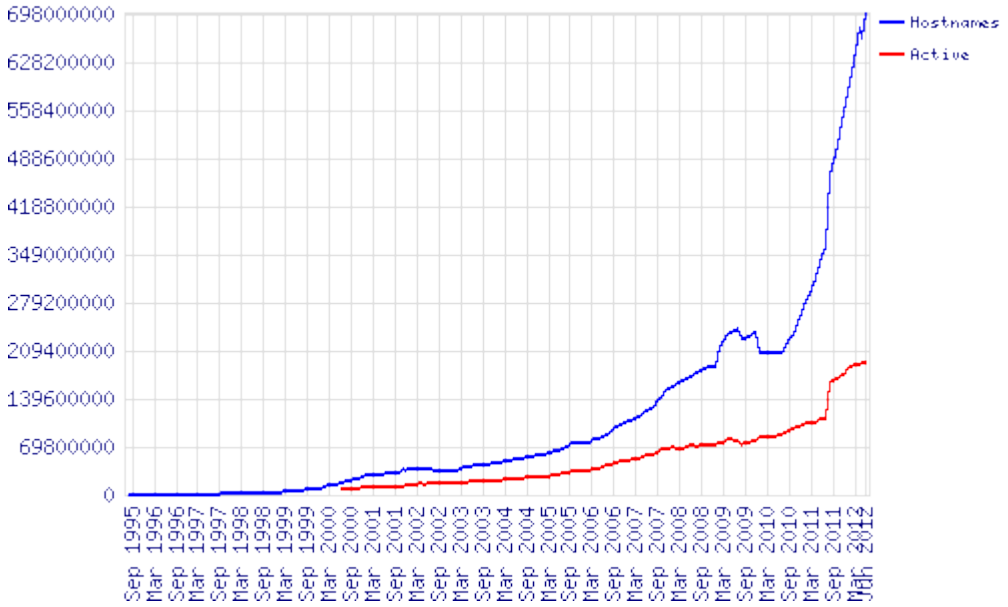


Figure 4 : Evolution du nombre de noms de domaines sur le web.

Un univers en extension

Aucun moteur de recherche n'est en mesure de quantifier précisément le nombre de pages accessibles sur le web. Le site « The Next Web »[19] nous fournit une infographie pour nous rappeler à quel point le web est vaste et à quelle vitesse le nombre de pages progresse.

[18] http://news.netcraft.com/archives/category/web-server-survey/
[19] http://thenextweb.com/shareables/2011/01/11/infographic-how-big-is-the-internet/

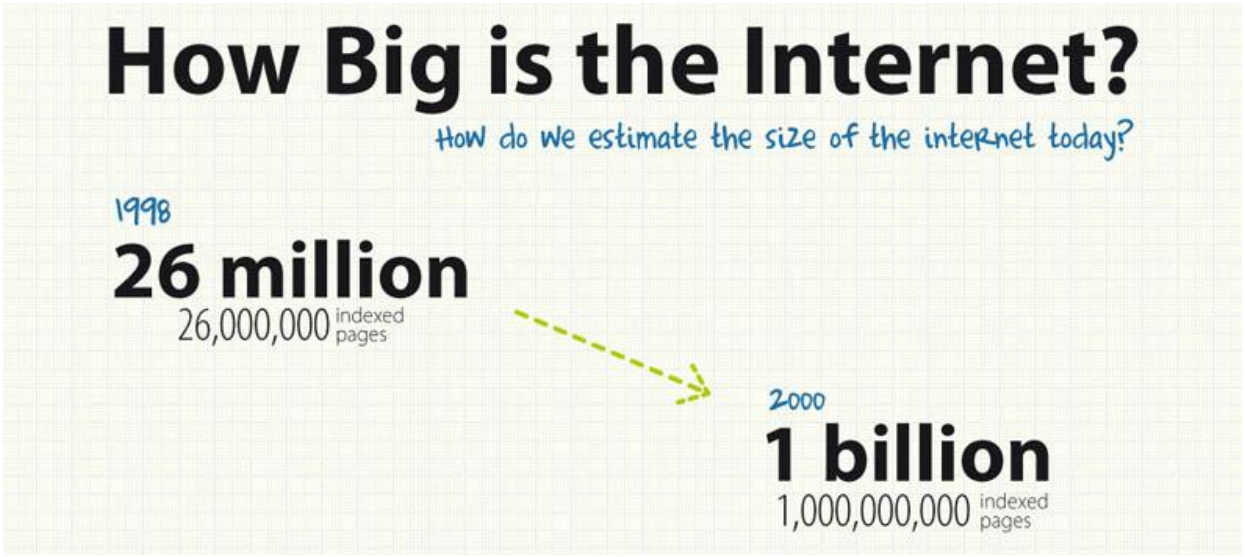


Figure 5 : Peut-on quantifier le nombre de page sur le web ?

Lors de la première indexation officielle de Google, le fameux moteur de recherche recense près de 26 millions de pages, deux ans plus tard, le cap du milliard de pages[20] est franchi. En sachant qu'il peut y avoir plusieurs URL (chemin d'accès) pour une même page (objet informationnel unique), les professionnels du web proposent un ratio de 10 URL pour 1 page tandis que le site TechCrunch[21] propose lui un ratio de 25 URL pour 1 page. Soit, 588 pages uniques par domaine pour le premier et 235 pages uniques par domaine pour le second. Que l'on opte ensuite pour l'un ou pour l'autre, l'évolution du nombre de pages indexées sur la période 2008-2012 est la même, soit 21%.

Si l'on applique le ratio des professionnels du web sur le nombre de domaines évalué en juin 2012 à 697 089 482[22], on arrive à près de 410 milliard de pages[23].

[20] http://googleblog.blogspot.fr/2008/07/we-knew-web-was-big.html
[21] http://techcrunch.com/
[22] http://news.netcraft.com/archives/category/web-server-survey/
[23] 697089482 noms de domaines x 588 pages par nom de domaine = 409 888 615 416 pages.

En juin 2011, soit il y a tout juste un an, Cisco estimait[24] qu'il faudrait 7 millions de DVD pour stocker la somme d'informations correspondant au trafic global d'Internet… pendant 1 heure.

1.2.2 Quelle version du web avez-vous ?

« *L'architecture Internet n'a pas bougé depuis 35 ans...* ». Cette phrase de Louis Pouzin, un des fondateurs des protocoles TCP/IP[25] résume bien l'état d'esprit dans lequel s'est ouverte la conférence sur le futur d'Internet le 28 juin 2010 lors du forum Atena. Et pourtant les usages du web n'ont jamais été aussi différents qu'il y a 35 ans. Les générations X et Y consomment un poids d'information[26] qui n'a jamais été envisagé lorsque le réseau a été conçu.

Nous l'avons dans la sous-partie précédente, cet espace qu'est le web et dans lequel nous évaluons quotidiennement de l'information est en permanente mutation. Les sirènes du marketing le jalonne parfois de grandes étapes comme le fameux web 2.0, mais cela n'a aucun sens.Les premiers à nous en parler, commeFred Cavazza conclut en mai 2010[27] : *« Oui il y a bien eu des grands tournants mais les fondamentaux de l'internet de 2010 étaient déjà présents en 2000*. »

J'expliquais en 2010 dans un article paru sur le site OWNI[28] que « *le web, c'est comme un individu : vous n'êtes pas Raymond 72.0, vous êtes Raymond. Même si Raymond est né (1.0), qu'il a appris à manger (2.0), à marcher (3.0), à parler (4.0), à lire (5.0)…. et à re-porter des couches à un âge avancé (Raymond 71.0)…. c'est toujours Raymond.* » Tim Berners-Lee[29] lui-même

[24] http://techland.time.com/2011/06/01/how-big-is-the-internet-spoiler-not-as-big-as-itll-be-in-2015/

[25] Le signe, qui signifie «Transmission Control Protocol/Internet Protocol» désigne tout simplement les protocoles de communication utilisé sur Internet.

[26] Les photos (3 Mo pour une photo en haute résolution), sons (environ 5 Mo pour un fichier mp3) et vidéos (environ 700 Mo pour un film en qualité divx) sont des médias qui pèsent lourds comparés à une page de texte (quelques octets).

[27] http://www.fredcavazza.net/2010/05/28/bilan-de-10-ans-dinteractivite/

[28] http://owni.fr/2010/07/14/le-web-2-0-est-mort-et-na-jamais-existe/

[29] Il passe pour le fondateur du web et dirige le World Wide Web Consortium (W3C), organisme qui œuvre à la standardisation du web.

dans une entrevue audio, pour developerWorks, le site d'IBM destiné aux programmeurs, expliquait en 2006, que « *L'objet même du Web 1.0 était de faire communiquer les gens entre eux. Le Web 2.0 est avant tout un slogan, personne ne sait ce qu'il signifie. Si pour vous, le Web 2.0 veut dire blogs et wikis, alors il s'agit de faire communiquer les gens entre eux. Ce qui a toujours été l'idée derrière le web.* »

Plutôt que de parler de version du web, on pourrait parler de ses usages. Nous identifions ainsi un « web passif » dans lequel l'internaute ne faisait que consommer des contenus ; un « web participatif » que les utilisateurs ont rendus social, communautaire, collaboratif et interactif etoù chacun produit des contenus ; et enfin le web sémantique, qui profite de la qualification des contenus pour créer des relations entre les contenus.

1.2.3 Web sémantique

Si aujourd'hui vous demandez à un moteur de recherche « quelles sont les plus grandes plages de France ? », il n'est pas en mesure de vous répondre car il ne comprend pas le sens de votre question. A moins qu'une page web contienne l'intitulé exact de la question et la réponse appropriée, les pages de résultats proposeront des résultats basés sur les mots clés « plage » et « France ». Nous allons dans cette partie nous intéresser au fonctionnement du web dit sémantique ou web des objets et étudier de quelle manière il peut nous permettre d'arriver à un niveau de pertinence socio-cognitivesatisfaisant plus rapidement.

Du chaos nait l'ordre

Afin qu'il n'y ait pas d'ambiguïtéTim Berners-Lee nous confirme bien que« *Le Web sémantique n'est pas un Web distinct indépendant du web actuel mais c'est le prolongement du web que l'on connaît et où on attribue à l'information une signification clairement définie, ce qui permet aux ordinateurs et aux humains de travailler en collaboration étroite.* » Nous sommes à un moment privilégié de la vie du web car nous assistons en direct

à la modification de sa structure. De plus en plus d'informations enrichissent les documents :microformats, tag cloud, folksonomie, hashtags, mais le système ne fait pour l'instant, que les visualiser.La philosophie du web sémantique a été présentée dans un article par Tim Berners-Lee, et nous tirons nos citations de la traduction proposée[30] par l'URFIST de Toulouse.*« Le Web sémantique va utiliser la structure pour donner du sens au contenu des pages Web, en créant un environnement où les agents logiciels en parcourant les pages pourront réaliser rapidement des tâches compliquées pour les utilisateurs. »*

Mais pour que le web sémantique fonctionne, les « agents » (algorithmes) qui vont créer les relations entre les objets informationnels doivent pouvoir accéder à des données structurées et à des règles d'inférence. Sans cela, il leur sera impossible de passer d'un espace désordonné à un espace sémantiquement ordonné. Nous allons donc nous intéresser à l'aspect technique de ce qui pourrait représenter une avancée majeure dans l'évaluation de l'information sur le web.

Indexer et structurer pour évaluer

Pour réaliser le web sémantique ou web des objets[31], il faut structurer tous les documents du web grâce au XML[32]. Le XML est métalangage, dans le sens ou il permet à n'importe quel utilisateur de créer lui-même ses propres balises. Le XML est à ce titre l'outil parfait pour structurer des objets informationnels et s'il est utilisé de façon massive, il permettra un vaste espace de documents semi-structurés.

---

[30] http://www.urfist.cict.fr/archive/lettres/lettre28/lettre28-22.html

[31] « *Le terme sémantique prête un peu à confusion car la sémantique s'intéresse au sens du langage pour en déduire des constructions logiques. Du coup, certains ont pensé qu'il s'agissait d'un Web qui permettrait par exemple d'effectuer des recherches sur Internet en posant des questions sous forme de phrases, en langage naturel. Or ce n'est pas son but. En fait, nous aurions dû l'appeler dès le départ « Web de données ». Mais il est trop tard pour changer de nom.* » Tim Berners-Lee dans un article en 2007 :http://www.larecherche.fr/content/recherche/article?id=6566

[32] « Extensible MarkupLanguage » que l'on peut traduire en « langage de balisage extensible »

Il faut ensuite interpréter le contenuavec un formalise à base de connaissance. C'est là qu'intervient RDF[33]qui va permettre de décrire les structures produites en XML. RDF intervient comme un vocabulaire et qualifie les objets informationnels et/ou les relations entre ces objets en leur affectant des métadonnées.

Le dernier composant intervenant dans la structuration des contenus est l'ontologie ou plutôt les ontologies. Il faut effectivement utiliser des ontologies communes et des langages de représentation des connaissances (comme RDF Schema[34] ou OWL[35]) pour annoter et traiter les documents de façon homogène. C'est avec l'ontologie que l'on introduit les règles d'inférences, il est donc normal que ce soit à ce niveau que les difficultés de standardisation soit le plus palpable. Comment s'entendre en effet sur un modèle d'organisation des connaissances dans un domaine donné quand chaque culture, chaque langue, à sa propre façon de l'appréhender ?

Pour illustrer nos propos, nous proposons de regarder le nombre de modèles d'organisation des connaissances en deux infographies, l'une réalisée en 2007 et l'autre en 2009.

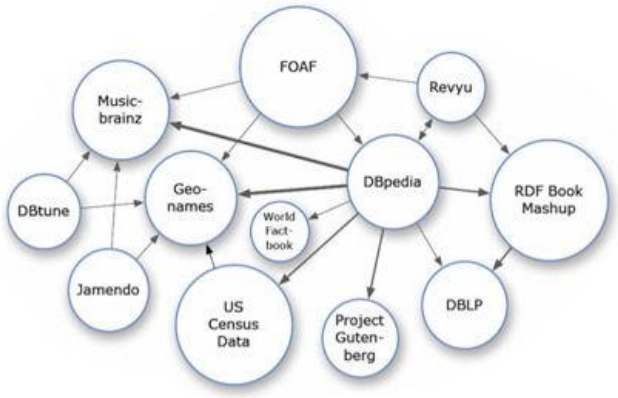


Figure 6 : Nombre de modèles d'organisation des connaissances en 2007.

[33] « Resource Description Framework » que l'on peut traduire en « modèle de description des données ».
[34] http://www.w3.org/TR/rdf-schema/
[35] http://www.w3.org/TR/owl-features/

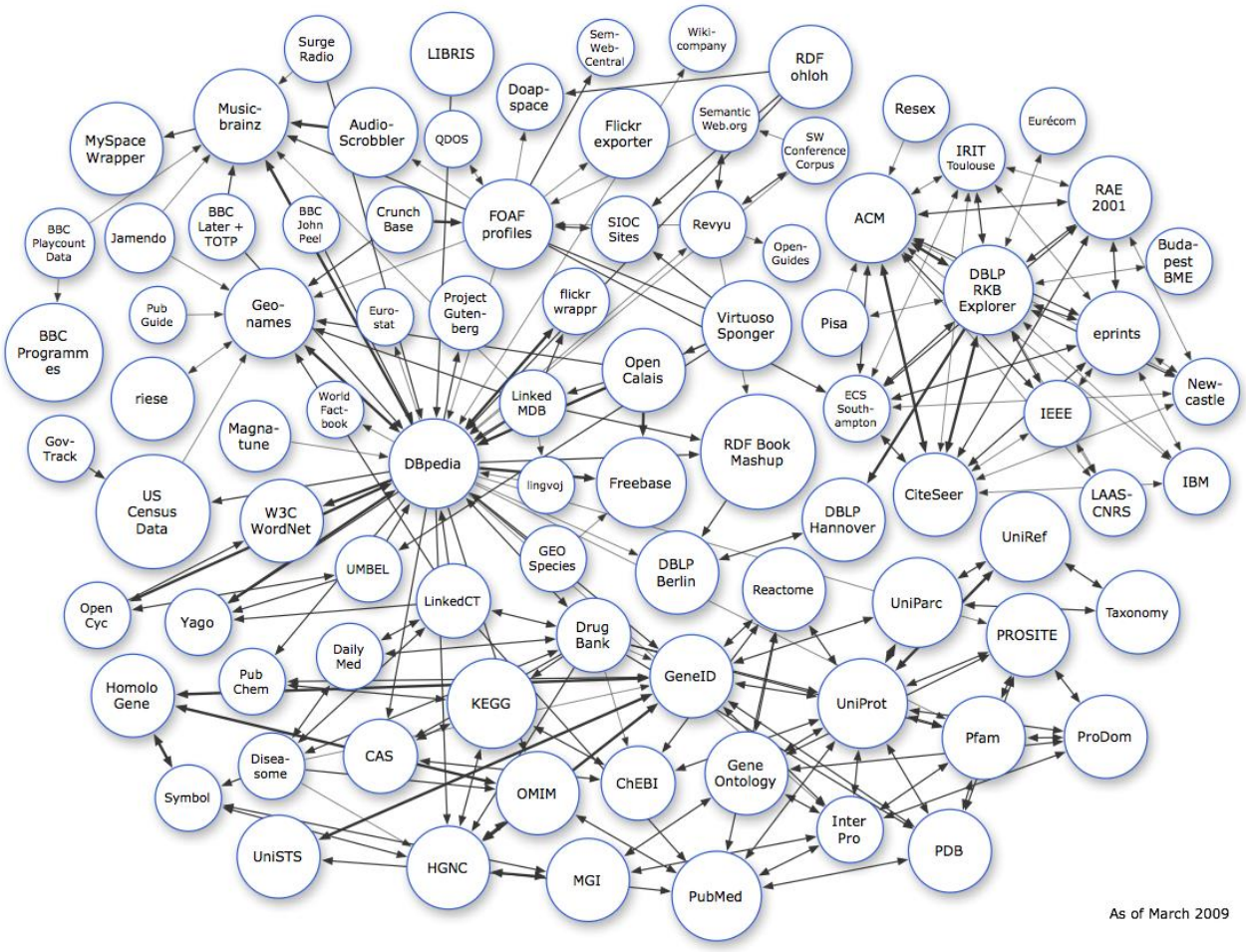


Figure 7 : Nombre de modèles d'organisation des connaissances en 2009.

Finalement, on se retrouve dans le deuxièmedes quatre préceptes que Descartes énonce en 1637 dans son célèbre Discours de la méthode : «*De diviser chacune des difficultés que j'examinerais en autant de parcelles qu'il se pourrait et qu'il serait requis pour les mieux résoudre*. » Ce précepte qui a amené les scientifiques à fragmenter le savoir en autant de domaines qu'il y a de phénomène à observer et à étudier. Et nous nous retrouvons aujourd'hui avec une interopérabilité qui, certes, est possible, mais au prix de quels efforts ?

L'utilisation d'agents logiciels capables de collecter les contenus du Web, de les coordonner entre eux et d'échanger des informations est-elle un mythe ?

Arriver à un graph sémantique

Appréhender aujourd'hui un dispositif d'accès à l'information sur le web, tel qu'un moteur de recherche, avec les conceptions mécaniste ou holiste, revient à le considérer comme un objet technique. Cet objet, grâce à un assemblage de composants pluritechniques interfacés (outils informatiques, indexation et classification, traitement automatique du langage, algorithme évaluant le poids des contenus), réalise des fonctions techniques (indexer un document par exemple) qui permettent des fonctions de services (indiquer à un utilisateur que ce document est pertinent par rapport à sa recherche) qui à leur tour, garantissent la fonction globale recherchée.

Nous sommes bien dans la juxtaposition d'éléments simples issus d'un découpage logique de la réalité (approche mécaniste) : traiter un document, le classer, le mettre à disposition, le conserver, permettre une évaluation de sa pertinence dans un contexte particulier. Certains vont plus loin et n'apprécie la totalité de ces éléments (approche holiste) qu'en ayant une vue générale : « un moteur de recherche permet la pertinence ». Nous estimons que c'est parfois loin d'être le cas.

La firme Google s'intéresse aux relations entre nos connaissances, nos savoirs, et à la sémantique qui lie nos contenus entre eux, il s'agit du « Knowledge Graph ». Dans la présentation vidéo[36] que Google fait de ce nouveau concept on apprend que cet algorithme se déploiera sur plusieurs mois et qu'il est clairement basé sur l'exploitation du web sémantique. Pour l'internaute, c'est la promesse de trouver mieux puisque le sens de ses requêtes en langage naturel sera mieux compris et donc mieux traité.

Preuve par l'image, Google annonce trois avancées majeures pour l'internaute : une meilleure interprétation de la requête, des résultats

[36] http://www.youtube.com/watch?feature=player_embedded&hl=en&v=mmQl6VGvX-c&gl=US

directement sur l'interface du moteur de recherche et des liens vers les objets informationnels jugés connexes.

Très concrètement, pour une meilleure interprétation de la requête soumise, regardons l'exemple proposé par Google[37]. Désirez-vous des informations sur Taj Mahal, le monument, le casino ou le musicien ?

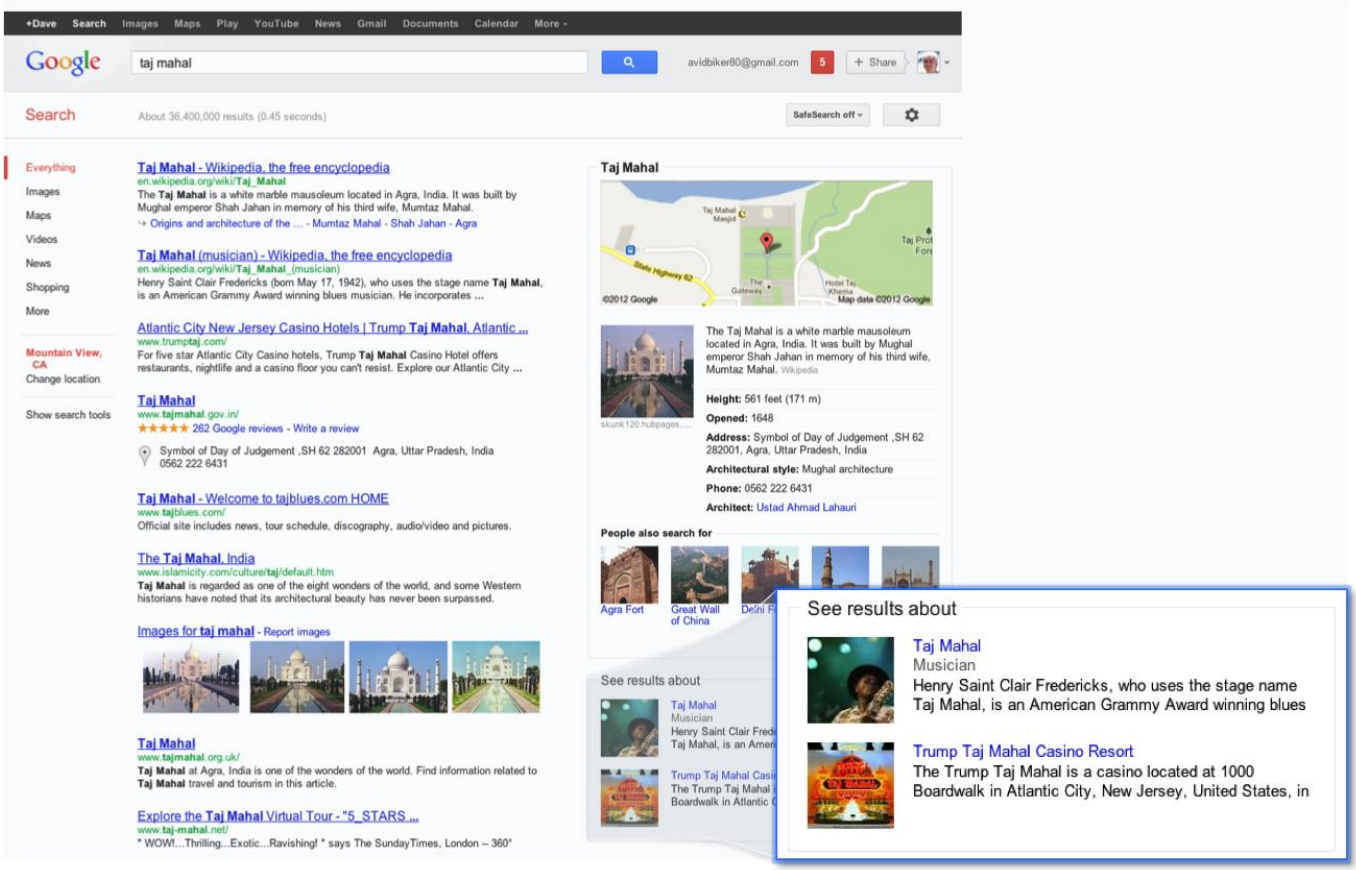


Figure 8 : Utiliser le web sémantique pour améliorer la pertinence des résultats d'une recherche.

[37] http://googleblog.blogspot.fr/2012/05/introducing-knowledge-graph-things-not.html

Pour Yves Simon, auteur de l'étude « Le web sémantique, infrastructure du social média[38] », le web sémantique est l'étape préliminaire à un web plus intelligent.



Figure 9 : Vers un web plus intelligent.

Mais en faisant cela Google nous permet-il vraiment d'accéder plus rapidement à un niveau de pertinence socio-cognitive satisfaisant ou bien s'agit-il d'une nouvelle info-pollution destinée à nous faire jouer le jeu d'un algorithme ?

1.3 Les pollutions informationnelles

1.3.1 Un risque écologique de la communication

Nous sommes tous devenus producteurs d'informations. Pendant longtemps les sites d'information ont été une destination : « *accéder à l'information a été un processus, produire l'information une tâche. Que se passe-t-il quand*

[38] L'étude est téléchargeable sur le site social computing : http://www.social-computing.com/wp-content/uploads/2010/03/2008_2007_web_semantique_web20_web40_ysc.pdf

tout change ? Nous sommes passés des médias de diffusion aux médias en réseaux, ce qui modifie fondamentalement la manière dont s'écoule l'information » (Boyd, 2009). De fait, l'objet informationnel, la relation du lecteur au document ne sont plus les même.

Confusions supplémentaires à cette relation qui n'est plus la même, *« ces difficultés, propres à l'évaluation et sur lesquelles ont toujours buté les professionnels, sont redoublées par les spécificités du numérique et de l'Internet, qui opèrent un quadruple brouillage* » (Serres, 2005) et provoquent une réelle atteinte à la confiance. Alexandre Serres nous rappelle les quatre risques écologiques en matière d'information, tels que définis par EricSutter (1998) : la surabondance, la désinformation, la prolifération d'informations indésirables, et enfin lesabus ou les effets pervers de la publicité.

Si l'on calque ces risques écologiques de la communication sur le modèle computationnel de la Théorie de l'Information de Shannon, on arrive à ce que Joël de Rosnay identifiait, en 1996[39], « *trop d'information crée une sorte de saturation et on ne parvient plus à extraire le signal du bruit*. » Il faut bien sûr accepter comme postulat ce que l'on trouve dans la plupart des traités de linguistique et de sémiotique, à savoir le modèle de la communication ci-dessous[40].

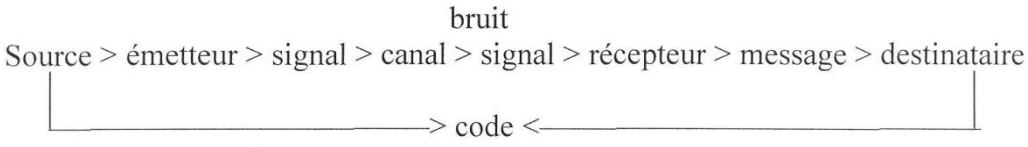


Figure 10 : Théorie de l'information

[39] http://www.cite-sciences.fr/derosnay/articles/Transver.html
[40] Tiré de Eco, U. [1974] Trattato di semioticagenerale, Milan, Bompiani.

1.3.2 Surcharge cognitive

Tout comme Joël de Rosnay, Olivier Ertzscheid (2003) et Gabriel Gallezot (2002) démontrent que la rapidité de transformation de l'information en documents, expliquée ci-dessus, provoque une « *entropie informationnelle* » entraînant une incapacité à traiter l'information, et à plus forte raison, à l'évaluer. Comme le dit Jean-François Rouet (2001) « *la recherche d'informations est une activité stratégique, qui demande un bon niveau de contrôle métacognitif : savoir évaluer ses connaissances et ses besoins d'informations en fonction d'un objectif, contrôler si l'information traitée est pertinente, prendre des décisions utiles en cas de difficultés, etc.* ».En effet, sur le web, les internautes dépensent de plus en plus de ressources cognitives pour identifier une information.

Or, lorsque l'on voit qu'une requête comme « vacances au bord de la mer[41] » propose quelques six millions de résultats, notre postulat est qu'aucun internaute n'est en mesure de traiter cette quantité d'informations. La charge cognitive mobilisée dépasse les ressources disponibles (notre cerveau), il y a alors surcharge cognitive.

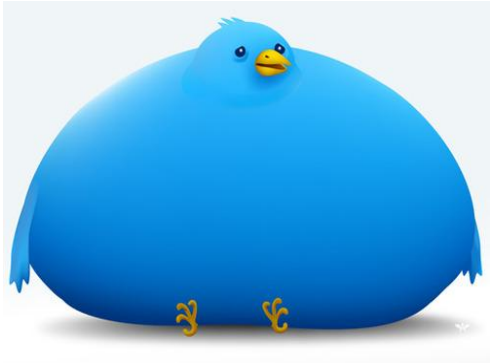


Figure 11 : Une situation d'infobésité ?

Pourtant, selon Anaïs Saint Jude, fondatrice du programme BiblioTech[42] de la bibliothèque de Standford, la surcharge informationnelle ne date pas d'aujourd'hui. Les grecs anciens déjà, comme Sénèque, dénonce l'abondance de livres comme étant une distraction. Au 17ème siècle, c'est Descartes qui s'interroge sur nos capacités de filtre et de tri vis à vis de la masse des documents disponibles en expliquant dans « *Recherche de la vérité par la*

[41] https://www.google.fr/search?q=vacances+au+bord+de+la+mer
[42] http://bibliotech.stanford.edu/

lumière naturelle »[43] que l'on passe plus de temps à choisir les livres qu'à les trouver.

Comme le rappelle la chercheuse, « *la surcharge d'information fait partie de la condition humaine : nous sommes confrontés par trop de possibilité, trop de complexité* » (Saint Jude, 2012[44]). Pour autant le principe de surcharge cognitive n'est pas un phénomène nouveau, alors l'augmentation continue du volume d'information disponible ne peut pas être la seule responsable des phénomènes d'infobésitéet de nos difficultés à identifier une information sur le web.

L'explication tient peut-être qu'entre-temps, si le nombre de documents à bien explosé, le nombre de gens capables d'accéder à ces documents également. Au 17ème siècle, seule une élite intellectuelle restreinte avait accès aux informations, aujourd'hui, on estime que deux milliards de personnes[45] ont accès au web et a ses informations multimédias.

1.3.3 Les invariants

De la position des structuralistes

Tout comme Gödel, les structuralistes tels que Levy-Strauss, Lacan, Piaget ou Bourdieu sont convaincus de la présence d'invariants qui échappent à la conscience du sujet, du système. Car, contrairement au holisme, le structuralisme explique le complexe par l'interaction entre les éléments d'un système.Cette conception particulière se retrouve chez les chercheurs du Mental Research Institute de Palo Alto qui considère exclusivement les systèmes humains comme des systèmes de relations, des systèmes de communication. Pour eux, « *Les liens qui unissent les éléments d'un système sont si étroits qu'une modification de l'un des éléments entraînera une modification de tous les autres, et du système entier* » (Watzlawick, 1972).

[43] http://data.bnf.fr/12417385/rene_descartes_la_recherche_de_la_verite_par_la_lumiere_naturelle/
[44] http://liftconference.com/fr/lift12/program/talk/anais-saint-jude
[45] http://leblog.wcie.fr/2011/11/30/les-petits-cahier-du-nets-le-mobile-en-chiffres/

Cette capacité à laisser la porte ouverte à un phénomène que l'on ne maîtrise pas, voire dont on n'a pas conscience est plus qu'un pied de nez à l'approche scientifique, elle est tout simplement magique. Magie, identifiée par Flawell comme « *La métacognition se rapporte à la connaissance qu'on a de ses propres processus cognitifs, de leurs produits et de tout ce qui y touche, par exemple, les propriétés pertinentes pour l'apprentissage d'information ou de données...* » (Flavell, 1976).

Les influencesméta-cognitives

La métacognition a deux aspects complémentaires : la gestion de ses processus mentaux d'une part et les connaissances métacognitives d'autre part. En psychologie cognitive, l´étude du système de traitement de l´information chez l´être humain a fait apparaître que celui-ci est et devient un acte de connaissance. Dès lors que nous évaluons une information sur le web, nous transformons l'objet informationnel évalué en connaissance. Tous les processus cognitifs que nous mobilisons lorsque nous utilisons un dispositif d'accès à l'information de type moteur de recherche est donc acte de connaissance.

Mais comment faire acte de connaissance lorsque nous-sommes en permanence sollicités visuellement sur les dispositifs d'accès à l'information ? En posant la question à Google, celui-ci n'a pas manqué de nous répondre… de la meilleure façon. Lorsque nous avons soumis la requête « pollution visuelle sur le web », voici ce qui apparaissait sur la première page de résultat : une publicité pour la régie publicitaire de Google, Google Adwords.



Figure 12 : La réponse de Google à la pollution visuelle sur le web.

C'est cette régie qui affiche près de 45 milliards de publicités en ligne par jour(chiffres septembre 2010[46]).

1.4 Situer l'information sur le web

1.4.1 Quel raisonnement adopter ?

Gregory Bateson aimait à reprendre ce syllogisme de E. vonDomarus « *l'herbe est mortelle ; les hommes sont mortels ; donc les hommes sont de l'herbe* ».

Une forte mobilisation cognitive

Internet en général, mais plus particulièrement le web « *est un nouveau média interactif, dont la particularité par rapport à l'écrit, la radio et la télévision, est de comprimer dans un temps plus restreint, et dans un espace compacté, un grand nombre de canaux de communication humaine : l'image, l'écrit, le son, la vidéo, etc.* » (De Rosnay, 1998). Avec la multiplicité des producteurs d'informations et des finalités de publication, arriver à remettre une information dans son contexte sur le web nécessite une forte mobilisation cognitive, si tenté qu'on ait envie de le faire. Or le contexte est primordial. Il contient la stratégie organisationnelle sous-jacente à l'environnement de l'utilisateur et influence de facto les stratégies d'évaluation de l'information pour la qualification d'une pertinence socio-cognitive d'un document.

[46] http://searchengineland.com/schmidt-great-stage-search-is-autonomous-personal-50014

Si nous cherchons à identifier le meilleur produit à acheter parmi une liste présentée par un site de e-commerce, notre stratégie organisationnelle va s'appuyer sur des paramètres (identification et réputation du vendeur, nombre d'étoiles, commentaires positifs, prix, services associés, etc.) basée sur une stratégie d'évaluation commerciale. Nous ne pourrons pas utiliser cette stratégie pour évaluer la pertinence socio-cognitive d'un message sur un forum d'informations médicales. Pour Cosijn et Ingwersen (2000) chaque contexte contient sa propre stratégie.

Types de raisonnement

Face à une situation nouvelle, l'être humain raisonne de trois manières différentes : par déduction, par induction ou par abduction.

Fonctionner par déduction signifie que la situation vécue, comme évaluer une information sur le web dans un contexte donné, est déjàcomplètement incluse dans une classe de situations semblables. L'internaute n'est soumis à aucune angoisse car il maîtrise, ou croit maîtriser la situation.

Fonctionner par induction signifie que l'on assimile la situation vécue à une bibliothèque de situations présentant les mêmes caractéristiques. « J'ai déjà évalué une information sur le web, même si le contexte est différent, je sais le faire. » L'internaute pré-suppose que la nouvelle situation présente suffisamment de propriétés identiques avec une situation déjà vécue et remet peu en cause l'influence du nouveau contexte.

Enfin, fonctionner par abduction signifie que l'on estime ne pas pouvoir faire de lien entre la situation vécue et l'historique de nos situations antérieures. C'est en général source d'angoisse car nous ne possédons pas de schéma (pattern) présentant les mêmes propriétés et nous sommes donc dans l'incapacité de prédire l'issue de la situation. Nous devons de plus mobiliser plus de ressources cognitives pour construire une nouvelle stratégie de raisonnement.

Lorsque nous sommes confrontés à un nouveau contexte, nous optons pour la stratégie qui nous semble la plus appropriée. Pour Miermont (2003), « *L'abduction se caractérise précisément par la tentative de compréhension de la singularité d'une situation, dont les indices peuvent être diversement interprétés.* » Si l'on considère que les internautes qui naviguent et recherchent sur le web sont confrontés à des situations nouvelles (un postulat difficile à contester), la stratégie d'abduction devrait donc être prédominante. Notre hypothèse est qu'au contraire, la majorité des utilisateurs procède par induction, voire même par déduction car le contexte proposé par le dispositif majeur (le moteur de recherche de Google) est identique quel que soit la recherche et quelles que soient les réponses.

De ce qui devrait être un processus a subjectivation palpable, on assiste plutôt à une désubjectivation, ce qui d'un point de cognitif ne favorise évidemment pas l'évaluation de l'information.

1.4.2 Situer dans le temps

Un changement de temporalité

« *Il y a quelque paradoxe dans le constat que l'étude de la dimension temporelle de la communication interpersonnelle et interculturelle est encore fort peu active, alors que le temps est un des investissements majeurs de la société occidentale contemporaine.* » Quand Yves Winkin, écrit ceci en 1993, dans le dictionnaire critique de la communication sous la direction de Lucien Sfez, il regrette que cette perspective soit peu utilisée dans le domaine des Sciences de l'Information et de la Communication.

Pourtant, Michel Foucault montre bien dans sa réflexion sur les systèmes disciplinaires développée dans *Surveiller et Punir* (1975) que le pouvoir « *s'articule directement sur le temps ; il en assure le contrôle et en garantit l'usage* ». Avant lui, tous les stratèges, de Sun Tzu à Clausewitz, avaient compris que le facteur temps était un élément décisionnel majeur dans l'évaluation de l'information et la prise de décision.

Socialement, nous sommes passés en un siècle d'une temporalité basée sur un cycle annuel, saisonnier, agraire, à une temporalité de l'immédiateté, de l'urgence systématique, dans laquelle la généralisation de l'usage des technologies de communication semble être en rapport avec cette « tyrannie de l'urgence » (Laidi, 1999). Cela a bien évidemment des incidences sur nos capacités et nos méthodes pour évaluer une information.

Quand on lit à propos d'un public que c'est *« la simultanéité de leur conviction ou de leur passion, la conscience possédée par chacun d'eux que cette idée ou cette volonté est partagée au même moment par un grand nombre d'hommes* » (Tarde, 1901), on pourrait se dire qu'il s'agit de l'analyse d'un groupe Facebook ou LinkedIn. Mais quand Gabriel Tarde écrit ceci au début du siècle dernier, on ne peut que « *souligner l'ancienneté de la mise en évidence d'un rapport étroit entre pratiques communicationnelles et pratiques du temps* » (Carayol, 2005). De fait, on retrouve ce rapport dans la publication sur le web aujourd'hui avec les dates liées à un document numérique.

Historicité du document numérique

Pour situer une information dans son contexte, il faut préalablement la situer dans le temps. Déjà hier, quand les journalistes devaient boucler leurs articles en fonction de l'heure de lancement des rotatives, nous avions un délai entre la production de l'information et sa publication qui pouvait être important. Les rédacteurs en chef décalaient un sujet pour le prochain hors-série car il fallait valoriser autre chose. Aujourd'hui les systèmes de publication de type CMS[47] ou plate-forme de blogs permettent de programmer, en deux clics, la publication d'un article des mois à l'avance. On peut également programmer la « dépublication » d'un article, c'est-à-dire choisir sa visibilité temporelle. L'auteur-publicateur possède lui aussi le pouvoir de situer dans le temps ou plutôt de re-situer. Qu'est-ce qui empêche en effet un auteur de publier un

[47] « Content Management System » ou Système de Gestion de Contenu.

article datant d'il-y-a dix ans ? Personne ! Je l'ai fait sur mon blog[48] et cette page a été indexée comme les autres.

En tordant l'espace-temps de publication comme le faisait le professeur Samuel Beckett dans la série télévisé Code Quantum[49], est-ce que je n'interdis pas aux autres de situer correctement l'information dans le temps ?

Caractéristiques temporelles d'un article

Prenons le cas d'un article de blog et regardons si les caractéristiques temporelles proposées par Ballard &Seibold (2004) se retrouvent dans ses métadonnées. Ballard &Seibold distingue la flexibilité, le tempo, la protection, la linéarité, la programmation, la ponctualité et l'ajustabilité. Du côté des métadonnées, on trouve couramment : date d'écriture, date de publication, date de modification, et date de commentaire s'ils sont autorisés. Ces dates permettent de situer temporellement l'article et sont très regardées par les internautes.

On retrouve dans la date d'écriture la caractéristique de ponctualité, qui représente la précision qui est fait de l'usage du temps. Un article écrit aujourd'hui sur la guerre du Golfe sera jugé moins pertinent qu'un article rédigé à l'époque, tout simplement parce qu'il n'est pas « ponctuel » avec l'évènement en rapport.

La date de publication, elle, regroupe les caractéristiques temporelles de linéarité, d'ajustabilité et de programmation. Etant donné qu'on peut la déterminer de façon arbitraire, la date de publication nous interdit de juger de la capacité d'un objet informationnel de servir à une chose à la fois. L'ajustabilité du document est de fait, puisqu'il s'agit bien d'une variable ajustable. Quant à la formalisation dans la programmation des activités, c'est

[48] http://www.guillaume-nicolas-meyer.fr/actualite-personnelle/actualite-professionnelle/ohe-ohe-matelot
[49] http://www.allocine.fr/series/ficheserie_gen_cserie=281.html

bien la date de publication qui détermine l'espace-temps dans lequel l'objet informationnel s'insère, et pas la date d'écriture.

La date de modification d'un document, parfois mentionnée dans les pages d'un blog ou d'un site permet d'évaluer la caractéristique temporelle de flexibilité. C'est un indicateur de rigidité ou au contraire d'adaptation. Un objet informationnel qui possède une date de modification visible informe également sur sa pérennité. « Je suis flexible, j'évolue, je serais toujours pertinent dans dix, cinquante ans » nous disent les pages de l'encyclopédie en ligne Wikipédia, car leur contenu n'est jamais figé.

Enfin, la date de publication des commentaires, qui ne correspond pas forcément à la date d'écriture du commentaire puisque les outils de publication en ligne permettent de valider à posteriori les commentaires, correspond aux caractéristiques temporelles de tempo et de protection. Le rythme de parution des commentaires, décorélé du nombre, permet de juger de la viralité de l'objet informationnel. Enfin nous jugeons que le degré de protection des séquences temporelles contre les intrusions correspond aux dates d'éventuelles réactions d'autres objets informationnels, qui sont alors, non plus un article sur une autre plate-forme, mais bien un commentaire de l'article initial.

Nous voyons bien, à travers ces explications que la notion de temporalité est extrêmement importante dans les processus de qualification de l'information. Et s'il est un outil de publication qui a engendré des décalages puissants dans l'espace-temps informationnel sur le web, c'est bien Twitter.

Twitter et la nouvelle fracture temporelle



Ce service de MicroBlogging qui permet d'envoyer des messages (tweets, gazouillis) de 140 caractères à ses correspondants (followers) en mode privé ou public a en effet bousculé les règles en vigueur dans la circulation de l'information entre

médias classique et média web.

Comme la plupart des autres plates-formes de réseau social à succès, Twitter ne connaît pas une progression linéaire, mais exponentielle.

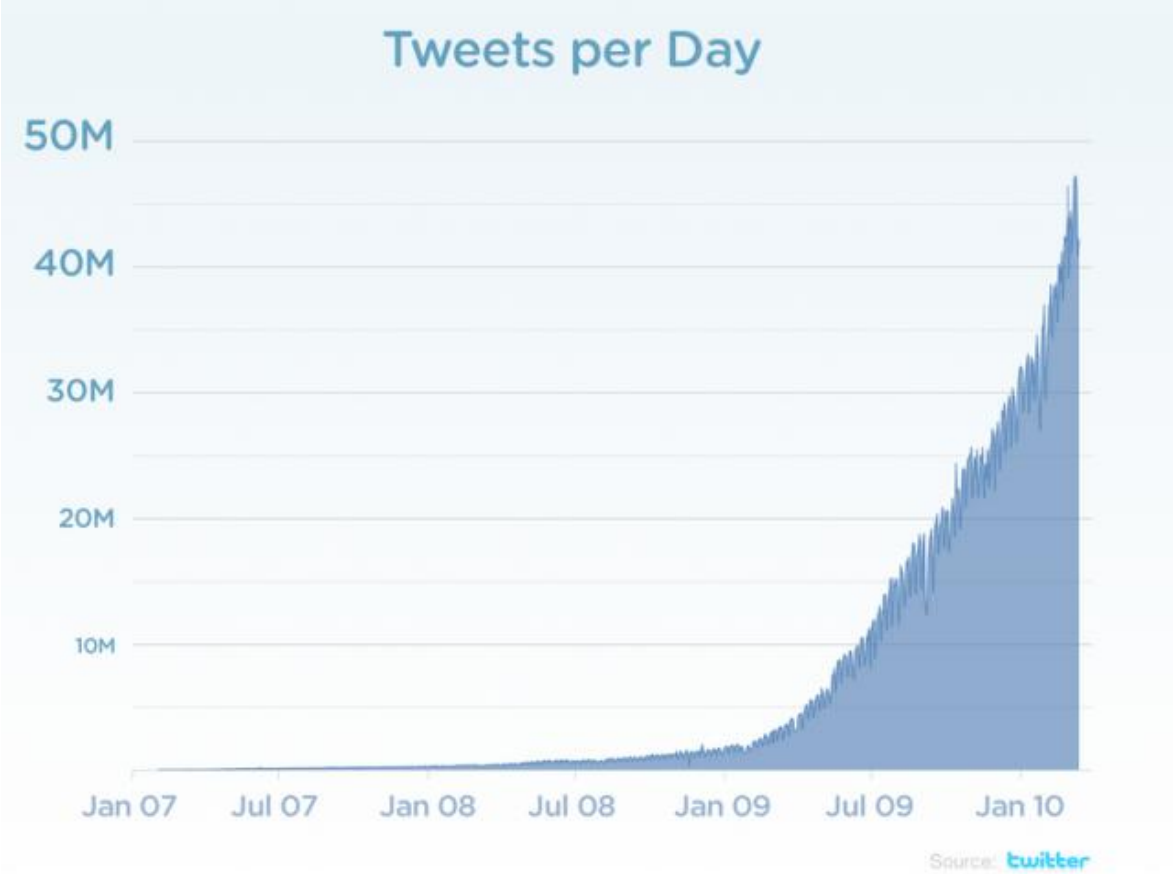


Figure 13 : Évolution du nombre de tweets par jour.

Tout commence avec un fait divers en janvier 2009, ce qui correspond au début de l'augmentation du nombre de message sur Twitter. Le 16 janvier 2009, le pilote du vol 1549 de US Airways décide de poser son Airbus A-320 dans le fleuve Hudson. L'amerrissage se passe très bien et un ferry vient porter assistance aux passagers. A son bord un homme, Janis Krums, prend une photo avec son téléphone portable et la publie sur son compte Twitter[50]. C'est le premier grand buzz de Twitter. Après, plus rien ne sera pareil.

[50] https://twitter.com/jkrums/status/1121915133

Twitterdevient l'endroit où l'information est publiée en premier. On peut voir sur l'infographie ci-dessous qu'il-y-a maintenant un décalage de près de huit heures entre la publication d'une information sur Twitter et dans la presse écrite.

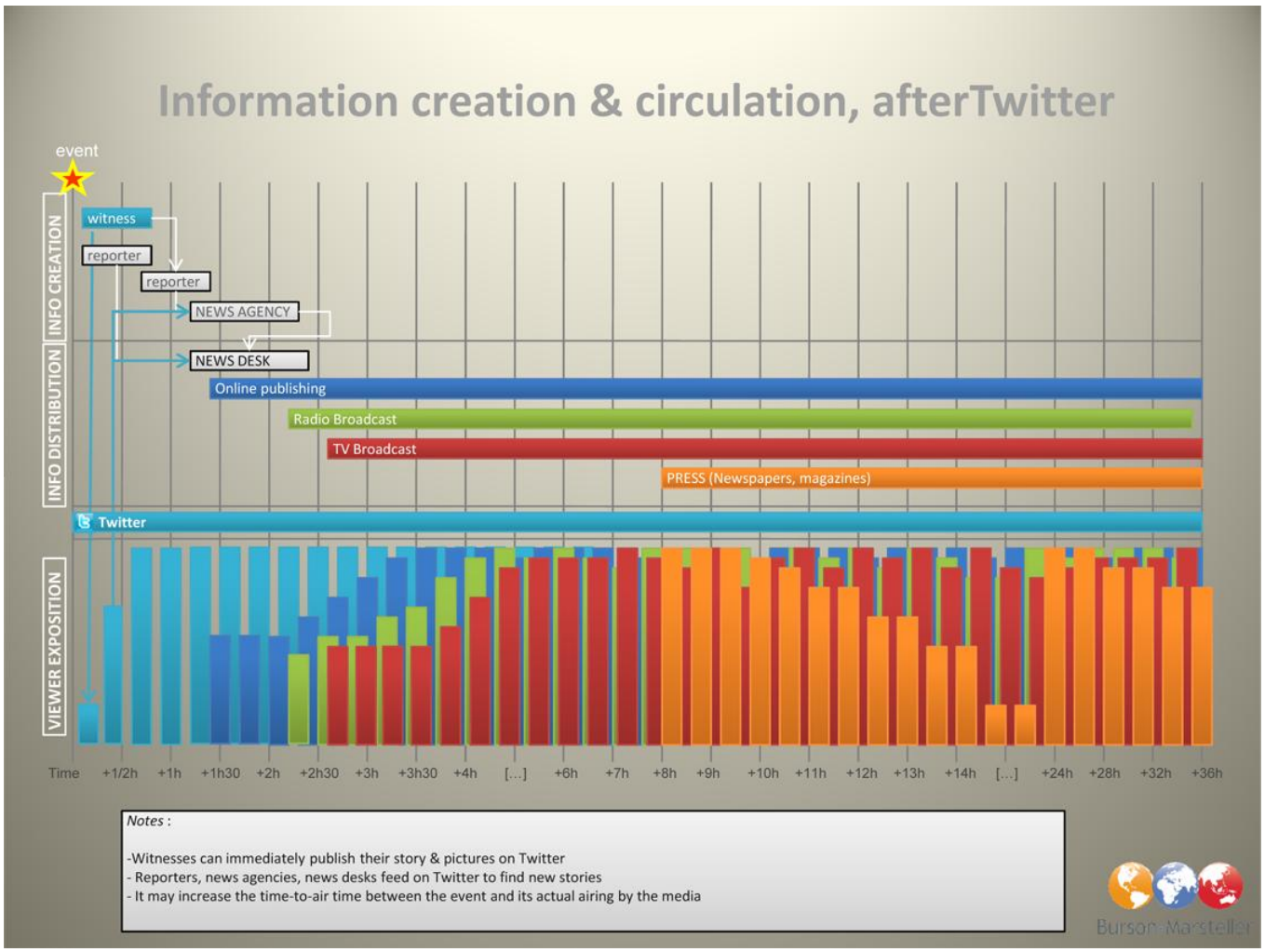


Figure 14 : Circulation de l'information après le phénomène Twitter.

Lorsque j'ai commencé à réaliser des formations sur Twitter, le logiciel GigaTweet[51] indiquait qu'il existait près de 2,5 milliards de tweets. L'année suivante, le compteur indiquait un peu plus de 10 milliards de tweets. Enfin, le 11 juin 2010, l'interface a cessé de fonctionner en indiquant plus de 29 milliards de tweets. Et pour cause, en 2011, 1 milliard de tweets a été

[51] http://gigatweeter.com/counter

enregistrés tous les… cinq jours. Le record de tweets par secondes 2011 a été atteint le 9 décembre, au Japon, lors de la diffusion du film « Le château dans le ciel ». A ce moment la plate-forme sociale a enregistré près 25 088 tweets par seconde. Avec de tels chiffres, il n'est plus besoin de rappeler le concept d'info-pollution que nous présentions dans la sous-partie précédente.

Son cousin chinois : www.weibo.com a vu son nombre d'utilisateurs augmenter de 185% en 6 mois, soit 195 millions d'utilisateurs supplémentaires de juin 2011 à novembre 2011. Ce qui nous amène à contextualiser une information dans sa culture.

1.4.3 Evolution des typologies de recherche

Situer dans une culture

Sur le web une information est protéiforme mais également multiculturelle. Alors qu'auparavant nos lectures étaient essentiellement celles d'auteurs de notre pays d'origine, le web permet en quelques clics de trouver de la littéracie internationale sur un sujet donné. Alors qu'auparavant nos lectures étaient essentiellement celles d'auteurs de notre pays d'origine, le web permet en quelques clics de trouver de la littéracie internationale sur un sujet donné. Comme le dit le Docteur Nawal (2009) « *Dans le cadre de la société mondiale de l'information, les pratiques de construction de connaissances renvoient désormais à de nouvelles dimensions cognitives et culturelles.* »

Comment contextualiser une information produite par une culture pour laquelle on n'est pas capable d'établir une pertinence socio-cognitive ? C'est bien parce que c'est difficile que Google se décline par pays. Bien sûr il y a les lois et les systèmes politiques, qui ne sont pas tous démocratiques et auxquels il faut s'adapter, mais il y a surtout le fossé culturel qui fait qu'une information dans un pays n'a pas le même poids, la même pertinence, la même portée que dans un autre. Or, à moins d'être biculturel, la plupart des utilisateurs des systèmes d'informations ne sont pas à même d'évaluer correctement une information issue d'une autre culture. Ce qui pose un véritable problème

puisque nous pouvons tout de même, lors d'une recherche, nous voir proposer de l'information d'origine étrangère et non re-contextualisée.

L'accès à d'autres cultures est important puisque cela nous permet de mieux évaluer l'information. En effet, la diversité de culture induit la diversité de points de vue, essentielle pour regarder un objet informationnel d'une autre manière. Sutter le dit d'ailleurs « *trop d'informations similaires ou mono-culturelles (consanguines ?) est nuisible* » (Sutter, 1998).

De nouveaux filtres

Nous avons vu combien il est difficile sur le web de situer un objet informationnel dans son temps, dans sa culture, de lui rendre son contexte. Dans le même temps, les dispositifs qui nous permettent d'accéder aux informations nous offrent de plus en plus de paramètres pour affiner les résultats d'une recherche. Est-ce qu'il s'agit de véritables outils nous permettant d'accéder plus rapidement à un niveau de pertinence socio-cognitif satisfaisant ou bien agissent-ils comme un trompe l'œil cognitif ?

Ce qui est sûr c'est que l'on est passé de « Madame, il me faudrait un livre d'histoire récent pour une photo du président des USA. » à « Je cherche une photo du visage de Barack Obama, de grande taille, au format jpeg, et dont la licence autorise des modifications. » (lien vers la requête[52]). Ainsi, la question de la distance sur le web est relancée. Nous proposons de voir la distance comme la probabilité qu'un internaute parvienne à une information qu'il jugera suffisamment pertinente pour s'en servir.

Ces distances sont variables pour plusieurs raisons. Tout d'abord rien ne garantit que la requête effectuée un jour renvoie les mêmes résultats le lendemain car l'indexation d'un objet informationnel peut changer, sa temporalité aussi. Ensuite, les moteurs de recherche, qui agissent, nous en

[52] https://www.google.fr/search?as_q=barack+obama&hl=fr&btnG=Recherche+Google&as_epq=&as_oq=&as_eq=&as_sitesearch=&safe=images&as_st=y&tbs=isch:1,isz:l,itp:face,ic:color,ift:jpg,iur:fm&biw=1600&bih=779&sei=1_PrT7vYCOSk0QWPkJH6DA&tbm=isch

sommes convaincus, comme des médiateurs, sont également soumis à une contextualisation de l'information basées sur les publicités en cours. Si l'évaluation de la rentabilité d'un mot-clé ou d'une requête change, alors les informations méta-communicationnelles qui l'accompagnent changeront également, ainsi que les chemins pour y accéder.

Sans lister ici la liste des outils que Googlemet à notre disposition sur l'interface de son moteur de recherche, nous pouvons tout de même noter son adaptation face à l'abondance d'objets informationnels sur le web. Alors qu'elle ne distinguait auparavant que le format des objets (textes, images, vidéos), l'interface propose maintenant de distinguer les actualités des discussions, des blogs et nous offre la possibilité de chercher directement des adresses, des emplacements dans une carte, etc.

A noter également que Google s'est adapté à la nouvelle temporalité du web que nous avons vu dans la sous-partie précédente. Nous le voyons dans la figure 16 ci-contre, son interface permet dorénavant de filtrer de l'information datant de moins d'une heure. Ce genre de filtre n'était pas disponible avant l'avènement de Twitter.



Figure 15 : Google s'adapte à la nouvelle temporalité du web.

Notre nouvel environnement socio-cognitif, nos nouvelles façons de s'en servir remettent-ils en cause la façon dont nous nous approprions l'information ? Nous allons étudier dans la deuxième partie nos façons de lire, de transformer l'information et les mécanismes intervenant dans l'évaluation.

2. Partie 2 –Langage etdésubjectivation

L'activité informationnelle s'appuie en général prioritairementsur la compétence cognitive lecture qui, elle, fait appel aux trois fonctions cognitives de perception, du langage et de la mémoire.Nous allons nous intéresser dans une première sous-partie au fonctionnement cognitif de cette compétence si spéciale. Dans une deuxième sous-partie nous étudierons les paramètres tangibles sur lesquels nous nous fondons, consciemment ou non, pour accorder une crédibilité sur un objet informationnel. Enfin dans une troisième partie, nous verrons quels critères de confiance nous retenons pour construire nos autorités informationnelles et, par extension, nos autorités cognitives.

2.1 Lire, cette pratique[53]

L'écriture n'a été inventée qu'il y a environ 5 400 ans et son accès était, jusqu'à très récemment, réservé à une élite intellectuelle. C'est pourquoi le Professeur Stanislas Dehaene affirme que « *le cerveau humain ne peut donc en aucun cas avoir fait l'objet d'une pression sélective, au cours de son évolution, pour en faciliter l'apprentissage.* » (Dehaene, 2008)

Lire est donc une « nouvelle » compétence cognitive et n'a pas toujours eu la signification qu'on lui prête aujourd'hui. Saint Augustin, à la fin du IV^{e} siècle, trouve étrange que son maître, Saint Ambroise, pratique la lecture à voix basse, la lecture étant à cette époque, essentiellement pratiquée à voix haute. Pour les humanistes de la Renaissance, la lecture est un entretien avec des grands hommes et non pas un processus d'amélioration de ses connaissances. Descartes disait à ce sujet : « *La lecture de tous les bons livres est comme une conversation avec les plus honnêtes gens des siècles passés qui en ont été les auteurs*. »

[53] Mallarmé.

Mais lire sur le web, c'est avant tout lire un hypermédia et notre « nouvelle » compétence cognitive n'a peut-être pas encore fait le lien entre nos différentes ressources cognitives et ce nouvel enjeu.

2.1.1 Lisibilité visuelle et lisibilité cognitive

Pour définir un critère de lisibilité d'une page sur le web, il convient de tenir compte de deux dimensions : d'une part l'aspect matériel, visuel, d'une page et d'autre part la lisibilité cognitive.

Lisibilité visuelle

Nous sommes habitués à lire sur papier et nous avons acquis des réflexes de lecture qui sont difficilement transposables lors de la lecture sur écran. De plus, nos caractéristiques physiologiques ne nous permettent pas de lire à l'écran comme sur papier. Ainsi la lecture à l'écran est en général plus lente de 25% par rapport à la lecture papier. La lecture à l'écran entraîne également une fatigue visuelle importante due à la luminosité du support.

Le Professeur Dehaene nous explique que les traits universels des écritures présentent tous dans la fovéa de la rétine, une haute densité de traits hautement contrastés (typiquement noir sur blanc). D'où le fait que les sites web proposant une altération du contraste soient déclarés moins lisibles par leurs utilisateurs. Cela perturbe clairement notre lecture, mobilise plus de ressource cognitive et amoindrit le confort visuel.

Lisibilité cognitive

La lisibilité cognitive représente l'effort de mobilisation de ressources cognitives pour la lecture d'une page web.

Tout d'abord, il convient de rappeler qu'on ne lit pas sur le web comme dans d'autre espace de lecture. Selon l'étude de Jakob Nielsen[54], 79% des lecteurs sur le web utilisent une lecture de type balayage. Le lecteur survole le texte à la recherche d'informations, lit par groupe de mots et passe d'un groupe à un

[54] http://www.useit.com/alertbox/reading_pattern.html

autre, afin d'essayer d'assimiler le plus d'information possible tout en mobilisant le moins de ressources cognitives possibles. À peine 16% des lecteurs sur le web auraient une lecture mot à mot.

Nielsen démontre également dans son étude que la lecture sur le web est modulaire et correspond à un schéma de lecture en forme de F. Notre œil balaye l'ensemble de la page, puis il est attiré par les titres, les images, les encadrés, menus, et publicité bien sûr. Finalement, est-ce que notre « nouvelle » compétence cognitive qu'est la lecture ne s'est pas déjà adaptée à ce média ?



Figure 16 : La lecture en F de Nielsen.

2.1.2 Quand les neurones s'en mêlent… mais ils s'en mêlent tout le temps

Psychologie cognitive expérimentale

Les mécanismes d'évaluation de l'information se jouent dans la plus fantastique des « machines » traitant de l'information en permanence, le cerveau. Il nous apparaît important de tirer profit des avancées des disciplines transverses comme les sciences cognitives pour poser le décor physique, matériel, des processus que nous allons chercher à identifier.

Les actes de notre quotidien sur le web comme s'installer confortablement dans son fauteuil et positionner son clavier, saisir une requête dans un moteur de recherche sur une destination idyllique, apprécier les photos que propose l'interface, sont fait le plus souvent sans y penser. Ils sont devenus des actes banals, au même titre que feuilleter un livre. Pourtant, ces actes que l'on peut qualifier d'élémentaires, « *impliquent en réalité au niveau neurophysiologique une cascade d'opérations allant des plus élémentaires à des mécanismes opérant à grande échelle (du point de vue du système nerveux), au travers de multiples relais.* » (Lalanne, 2005)

Même si nous sommes encore loin d'avoir compris tous les mécanismes fonctionnels opérants dans le cerveau, les trente dernières années ont permis des avancées considérables, que ce soit avec les neurosciences ou la psychologie cognitive expérimentale.

Évaluer l'information sur le web modifie notre cerveau

Le docteur Gary Small et son équipe de l'Université de Californie à Los Angeles ont publiés en 2008, dans l'*American Journal of GeriatricPsychiatry*[55] les résultats de leur étude portant sur la mesure des fonctions cérébrales lors de recherches effectuées sur le web. Le résultat de ces études nous permet d'affirmer d'un point de vue clinique que l'évaluation d'une information sur le web est bien une activité cognitive très élaborée. Nous mobilisons pour ce faireles centres clés du cerveau qui contrôlent le processus de décision et de raisonnement complexe.

[55] http://journals.lww.com/ajgponline/Abstract/2009/02000/Your_Brain_on_Google__Patterns_of_Cerebral.4.aspx

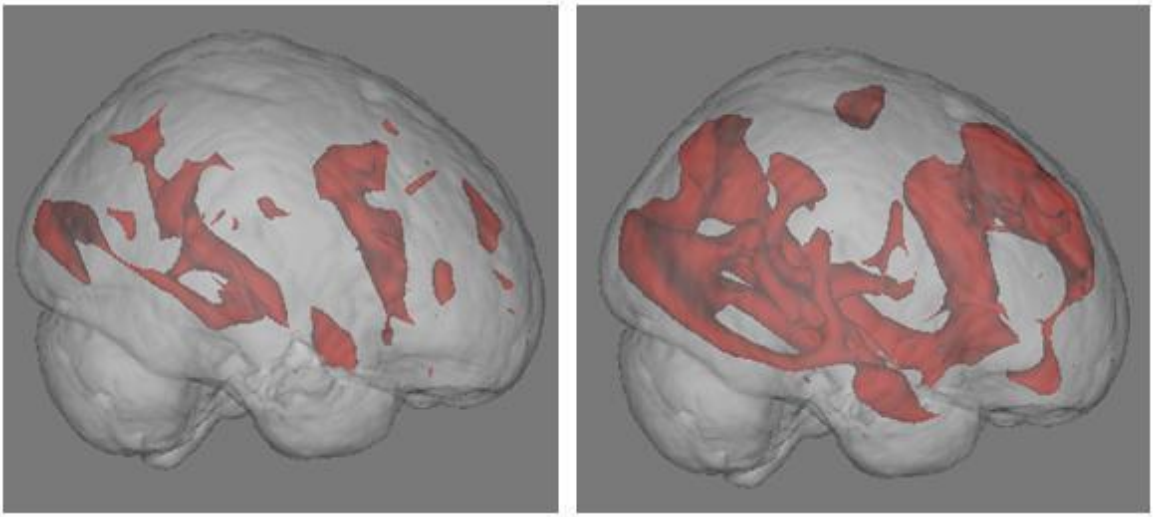


Figure 17 : A gauche, l'activité cérébrale de la lecture d'un livre, à droite l'activité cérébrale lors de l'évaluation d'une information sur le web.

« *Notre découverte la plus frappante a été que les sujets faisant des recherches sur Internet ont paru mobiliser davantage de circuits neuronaux qui ne sont pas stimulés par la lecture, mais seulement chez ceux ayant une expérience de recherche sur Internet*. » — Dc Gary Small, UCLA, 2008[56].

Pour autant cette étude à une limite. Il existe en effet deux types de lecture : la lecture savante et la lecture privée. Si on oppose une recherche sur le web à une lecture, encore faudrait-il qu'il s'agisse d'une lecture savante, pendant laquelle le lecteur est beaucoup plus actif que lorsqu'il lit un roman. Pour être tout à fait objectif, il conviendrait de compléter ces observations par d'autres dans laquelle on comparerait un épisode delecture plaisir et un épisode de lecture savantesur les deux types de support. On pourrait surement alors constater non seulement des écarts entre les supports mais également entre les types de lecture. Ce n'est surement pas pour rien que déjà

[56] http://www.uclahealth.org/body.cfm?id=403&action=detail&ref=1100

Cicéronopposait ceux qui aiment lirepour le *voluptas*[57] de la lecture, et ceux qui lisent pour son *utilitas*[58] (Cicéron, De Fin., V, 2).

Des prédictions qui nous dérangent

Quelles sont donc les différences fondamentales qui existent entre une lecture papier et une lecture à l'écran ? Tout d'abord on ne fait pas que lire quand on est devant un écran. Même en sélectionnant un article sur la version numérique d'un journal, rien ne nous prévient que quelques lignes plus loin, l'article « papier » ne va pas proposer une vidéo, une galerie d'images ou autre support multimédia. D'une étape « choisie » de lecture nous pouvons/allons passer en un instant à une ou plusieurs étapes, plus ou moins imposées, d'écoute et/ou de visionnage d'images. Ce changement de médias n'est tout simplement pas possible avec un journal papier, la radio ou la télévision. Nous verrons plus loin l'impact de l'hypertextualité et des suggestions méta-communicationnelles.

Ainsi, lorsque l'on aborde une lecture papier, notre cerveau prédit une lecture linéaire et plus sereine que lorsque l'on aborde une lecture à l'écran. Dans le cas d'une lecture sur le web, notre cerveau prédit une lecture compliquée. Les principaux facteurs qui perturbent la lecture à l'écran, et plus spécifiquement sur le web peuvent être de façon non exhaustive : l'ergonomie, le choix des couleurs, la typographie, la mobilité de l'espace de lecture, les info-pollutions (publicités notamment) au sein du contenu à lire, les perturbations auditives et visuelles (encarts vidéos et sonores). Si bien qu'on peut se poser la question : est-ce qu'on lit encore sur le web ?

2.1.3 Est-ce qu'on lit encore sur le web ?

D'un acte individuel vers un acte collectif

Nous sommes tous devenus des producteurs de contenus et cela se mesure. Sur le web, plus de 70% des nouveaux contenus sont produits par les

[57] plaisir
[58] utilité

utilisateurs, on parle d'UGC (User Generated Content). Paradoxalement l'essentiel des internautes n'ont pas le sentiment d'être actif sur le web. Normal, la frontière entre lecture et écriture s'est diluée dans les usages. Lire sur le web n'est plus un acte passif, car lire, c'est déjà être remarqué « 203 personnes ont lu ça » et donc avoir un impact sur d'autres lecteurs potentiels. Auriez-vous lu l'article s'il était indiqué « 0 personne a lu ça » ? On ne fait plus « que » lire, on « aime ça » (bouton like de Facebook), on gazouille (twitt, re-twitt et hastag sur Twitter), on « +1 » (Google +), on « trackbacks », on « backlinks », etc.

Comme le dit le sociologue Dominique Cardon (2010) : « *Il faut être attentif au fait que les internautes produisent, par leurs clics, leurs liens, leur retweet, la visibilité des informations sur internet.* » Cette visibilité, délégitimé par la décontextualisation qu'elle engendre,a un réel impact sur le degré de pertinence socio-cognitive (Cosign et Ingwersen, 2000) des internautes.

2.1.4 hypertextualité et spatialisation de l'information

Les mutations provoquées par la lecture de document numérique, notamment la transversalité due à l'hypertexte, induisent de nouveaux comportements et ont un réel impact sur la confiance. Quand je lis une page, est-ce que je ne suis pas influencé par l'autorité du contenu de la page d'où j'arrive ?

D'un espace euclidien à un espace sémantique

Michel Serres nous le rappelle (Serres, 2005), nous sommes passé d'un espace fini, métrique, et donc par essence maîtrisable, à un espace infini et sémiotiquement in-maîtrisable (nous rappelons ici la notion de l'état « sauvage » du Word Wild Web).

Le web est donc un espace qui n'a pas de limites[59] et cela n'est pas sans poser de problèmes aux internautes que nous sommes car nous sommes habitués à dresser nos repères dans un espaceeuclidien. En effet, « *… l'espace de navigation n'est pas un espace physique dans lequel on s'oriente, mais un espace sémantique dans lequel on traite des contenus* » (Tricot *et al*. 1998).Cette particularité du web comme espace à la fois documentaire, médium et méta-plate-forme de dispositifs d'accès à l'information engendre une confusion cognitive.

Si nous recherchons le sens d'un mot dans un dictionnaire, nous pouvons être tentés de regarder les autres mots de la page mais, étant dans un espace fini, les lectures « suggérées » ne sont pas nombreuses. En comparaison, sur une « page » web, dont l'espace de lecture peut varier du simple (deux cents pixels de haut sur six cents de large) au triple, nous pouvons « rebondir » à l'infini grâce aux liens hypertextes.

Cette transversalité peut facilement nous faire « sortir » de notre contexte initial de recherche.

2.1.5 Naviguer à l'aide d'une carte

Les nœuds d'autorité

Face à ce déluge informationnel et pour remédier au problème de spatialisation de l'information nous avons réintroduit un outil puissant s'il en est, la carte. L'Homme utilise la carte depuis des millénaires et lorsqu'il est perdu, il fait une carte. En inventant le web, nous avons inventé un nouvel espace dans lequel nous perdre et, logiquement, nous avons améliorénotre appréhension de cet univers numérique avec des techniques de visualisation de données.

[59] Par contre, on ne peut pas véritablement dire qu'il n'y a pas de frontières au vu des différences de normalisation d'indexation et du caractère propriétaire de certains formats de données, qui sont clairement, à mon sens, des frontières techniques, sémantiques, parfois même éthique.

Au lieu de se repérer dans une carte grâce à des repères spatiaux, ce qui ne ferait aucun sens puisque nombre de site européens sont hébergés sur des serveurs américains, les cartes positionnent les sites ou les gens les uns par rapport aux autres en se basant sur leur influence, leur notoriété, leur popularité. De repères objectifs, mesurables, quantifiables, nous avons basculés sur des repères flous, variants et discutables. Pour autant, cette nouvelle mécanique cognitive qu'est la cartographie d'information nous permets d'identifier les nœuds des réseaux qui nous environnent. Être un « nœud » névralgique d'un réseau n'est pas nouveau. Les gardiens des portes au Moyen-Âge en étaient : sans eux, impossible de pénétrer dans le réseau social de la ville.

Sur le web, la transposition fonctionne. Impossible en effet de bien référencer son site sur un sujet donné sans qu'il y ait de lien (de préférence entrant) avec un des nœuds (site, blog, forum) référent sur le sujet. On peut illustrer très concrètement les nœuds d'autorité informationnelle en analysant ces cartographies de la blogosphère italienne réalisée en 2011. La première infographie(figure 19) présente l'intégralité des nœuds de la blogosphère italienne. Sans catégorisation, on peut déjà identifier le nœud le plus important du réseau (point jaune légèrement excentré sur la gauche de l'infographie).

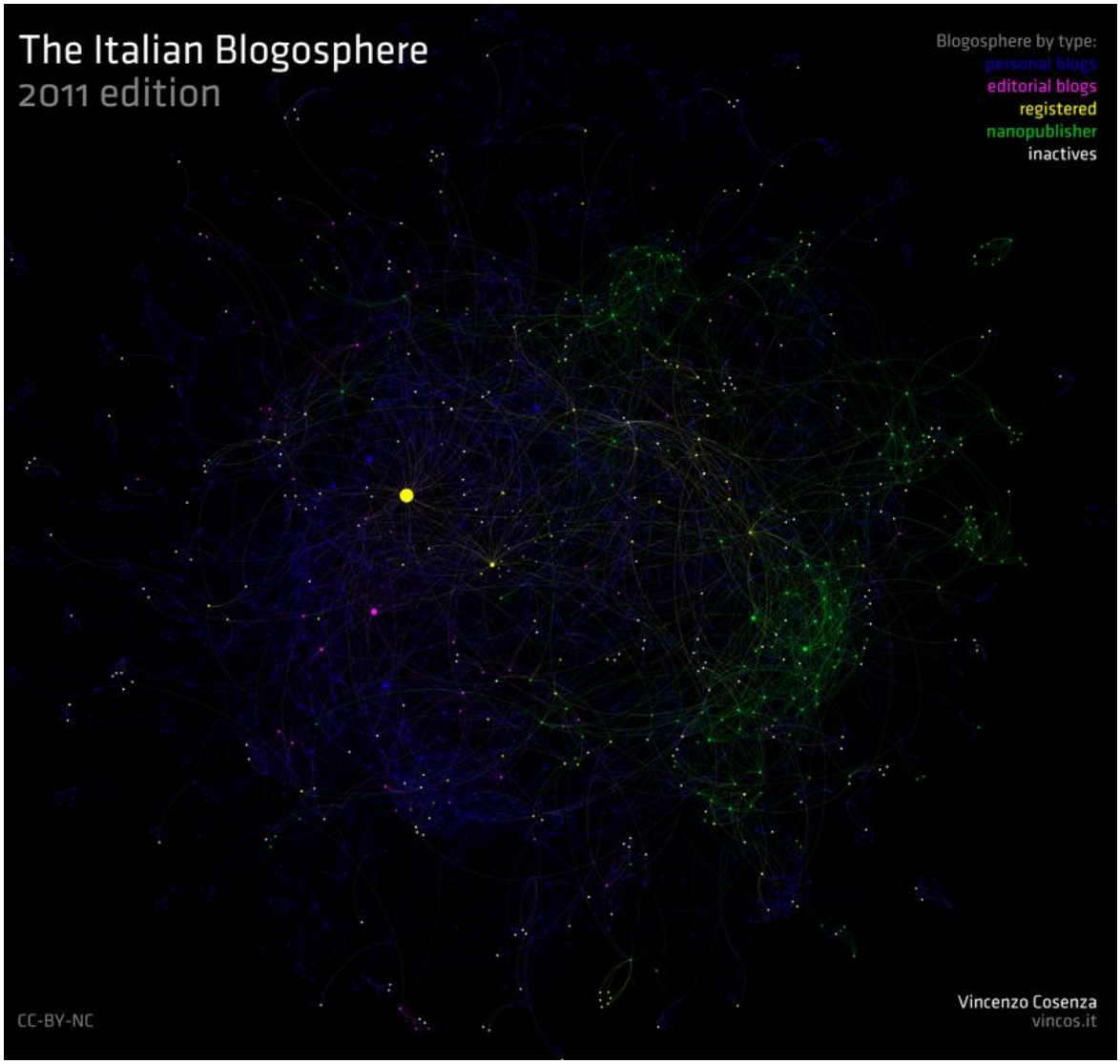


Figure 18 : Cartographie de la blogosphère italienne en 2011.

Un traitement algorithmique permet d'identifier les principaux nœuds du réseau. L'infographie de la figure 20ci-après,présente les 139 nœuds les plus importants de la blogosphère italienne.

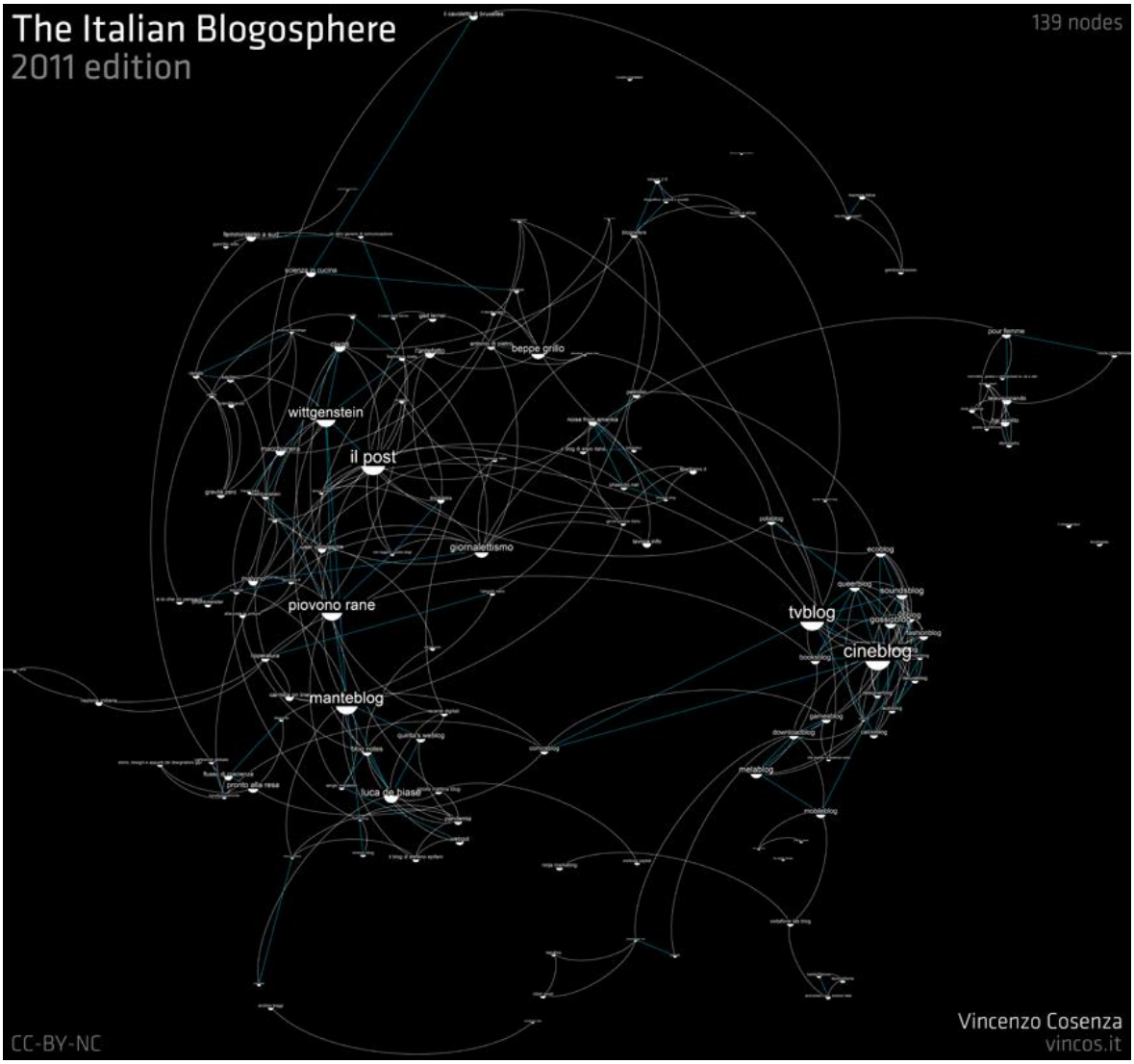


Figure 19 : Nœuds importants de la blogosphère italienne.

Cartographier des réseaux de personnes

Pour Dana Boyd, la transposition avec des réseaux de personnes est évidente. *« **Nous donnons du pouvoir aux gens** quand nous leur accordons notre attention et les gens gagnent du pouvoir quand ils font le pont entre des mondes différents et déterminent quelles informations seront reversées dans les réseaux »* dont ils sont l'un des nœuds (Boyd, 2009). Joël de Rosnay en son temps (1996) avait clairement expliqué ce décalage, cette mutation du pouvoir informationnel[60] « … *chacun d'entre nous à titre individuel, comme*

[60] http://www.cite-sciences.fr/derosnay/articles/Transver.html

nœud de réseau, peut maintenant agir, ce qui était auparavant l'apanage des seuls puissants, des riches et des politiques. »

Et les outils permettant à un simple individu d'identifier les nœuds importants de son réseau se multiplient. Dans les Labs[61] de LinkedIn, le réseau social professionnel qui revendique 135 millions de membres à travers le monde, se trouve l'outil LinkedInmaps[62]. Il suffit de synchroniser votre compte etl'application se charge de cartographier l'ensemble de vos contacts en les évaluant, et en les répartissant dans un graphe. Le « poids » de chaque nœud correspond au nombre de contact de l'individu, ce qui n'est pas satisfaisant d'un point de vue pertinence pour qualifier les nœuds, mais c'est un point de départ intéressant.

Autre point, l'application crée des sphères de différentes couleurs dans lesquelles elles agrègent les individus en fonction de thématiques communes, de leur profil professionnel. On est complètement dans les « écritures de soi » dont parle Olivier Le Deuff dans sa formation à l'Urfist de Strasbourg[63].

[61] Espace de laboratoire dans lequel les applications sont testées avant d'être proposées à grande échelle.
[62] http://inmaps.linkedinlabs.com/network
[63] Formation suivie en décembre 2011, la présentation est accessible sur le blog « le guide des égarés » http://www.guidedesegares.info/2011/03/08/controle-de-ses-donnees-et-ecritures-de-soi/

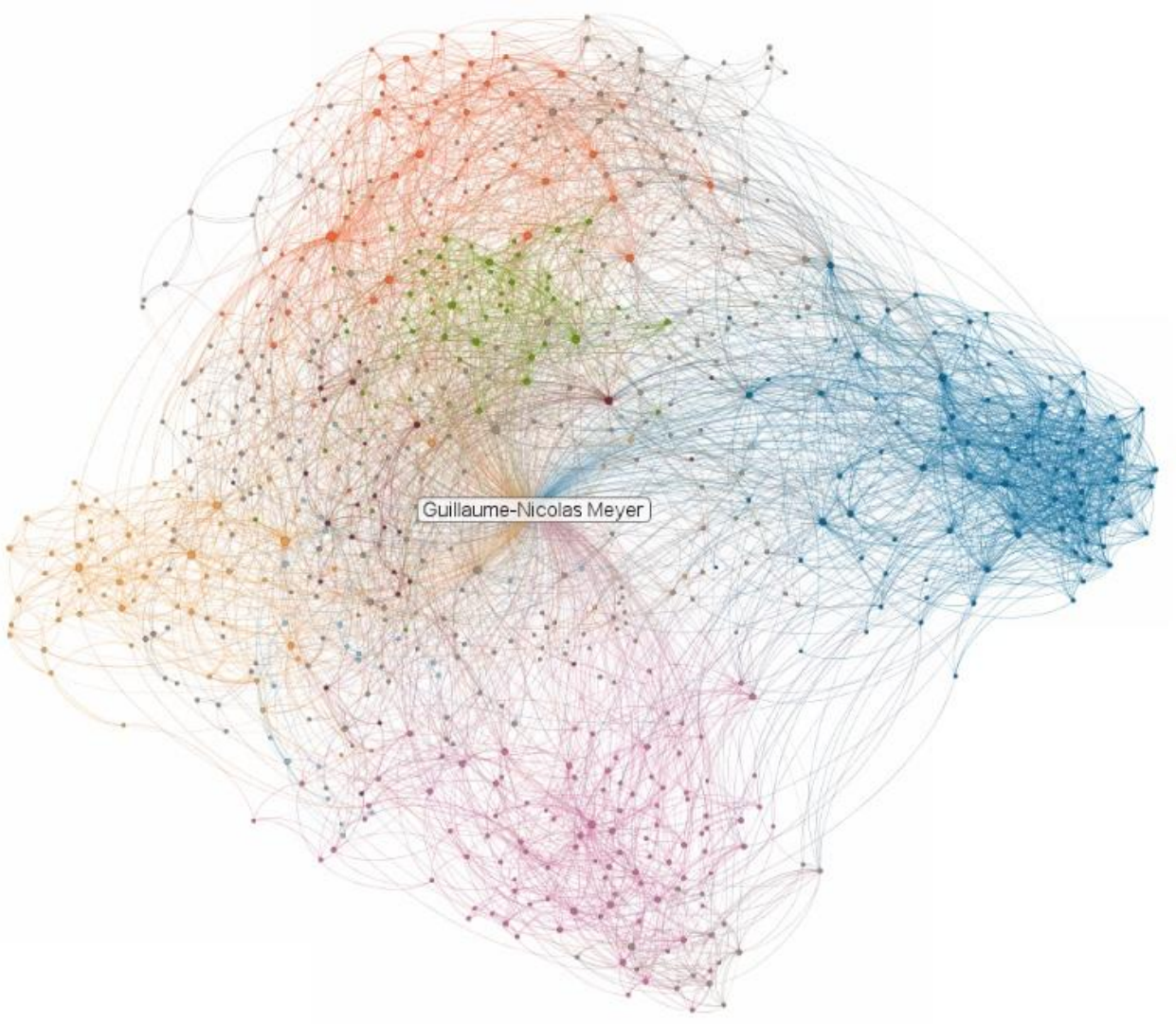


Figure 20 : Cartographie des contacts LinkedIn (681 contacts sur l'infographie).

Un zoom dans la carte permet de se rendre compte de la densité de ses contacts pour une thématique donnée.

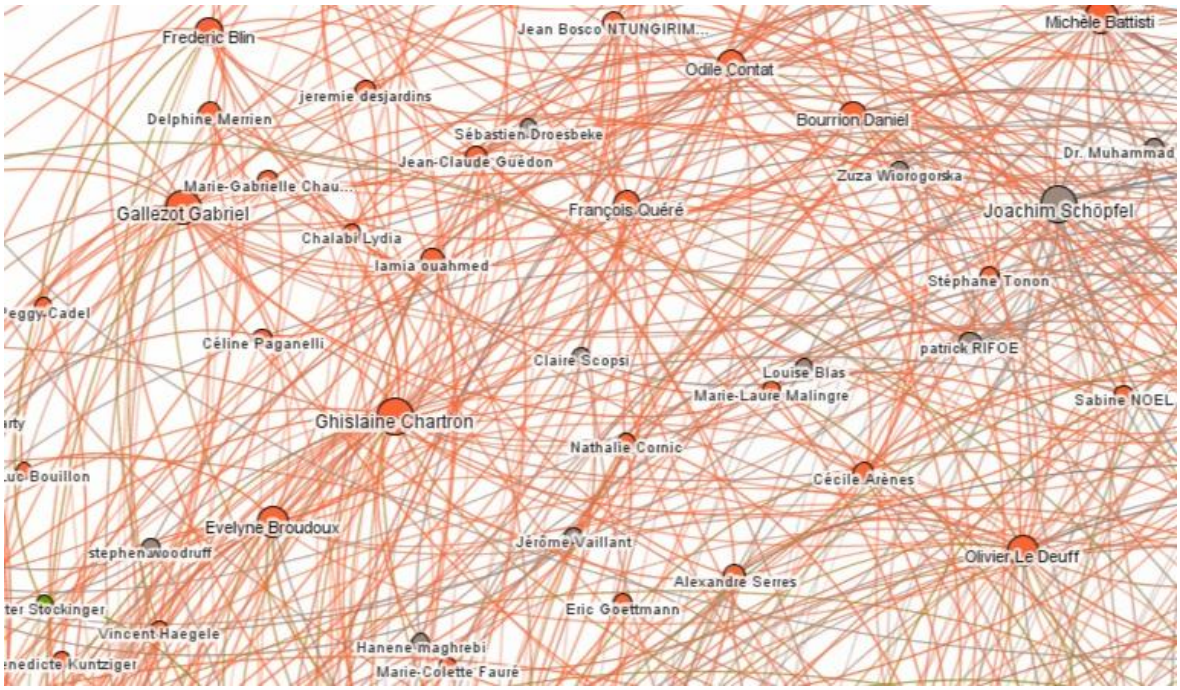


Figure 21 : Mesurer la densité de ses contacts LinkedIn.

Pour identifier un nœud du réseau, il suffit de positionner le curseur sur le nœud et une fenêtre d'information présente l'individu à partir des informations disponibles sur son profil.

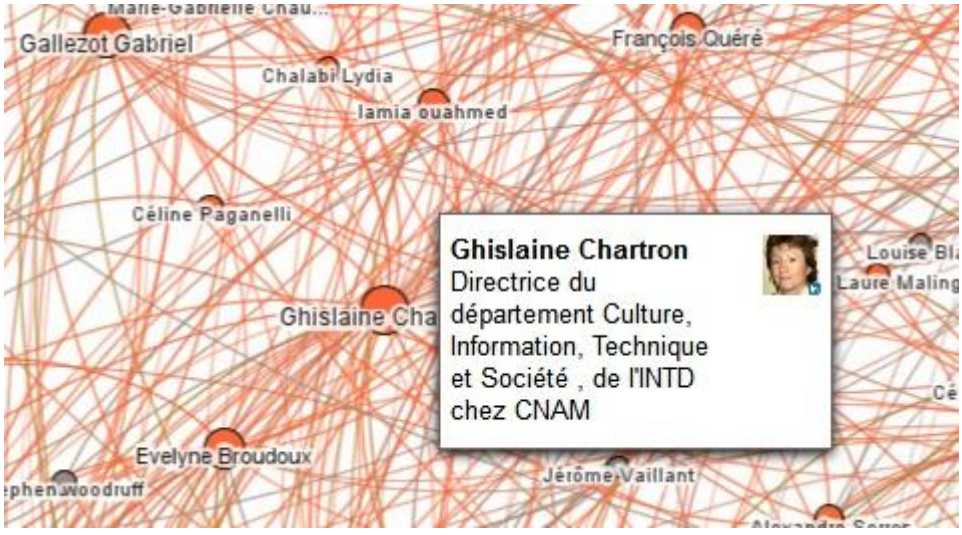


Figure 22 : LinkedIn, identifier un nœud de son réseau.

Enfin, un clic sur un nœud de son réseau permet d'identifier les nœuds les plus importants, ceux qui sont le plus interconnectés.

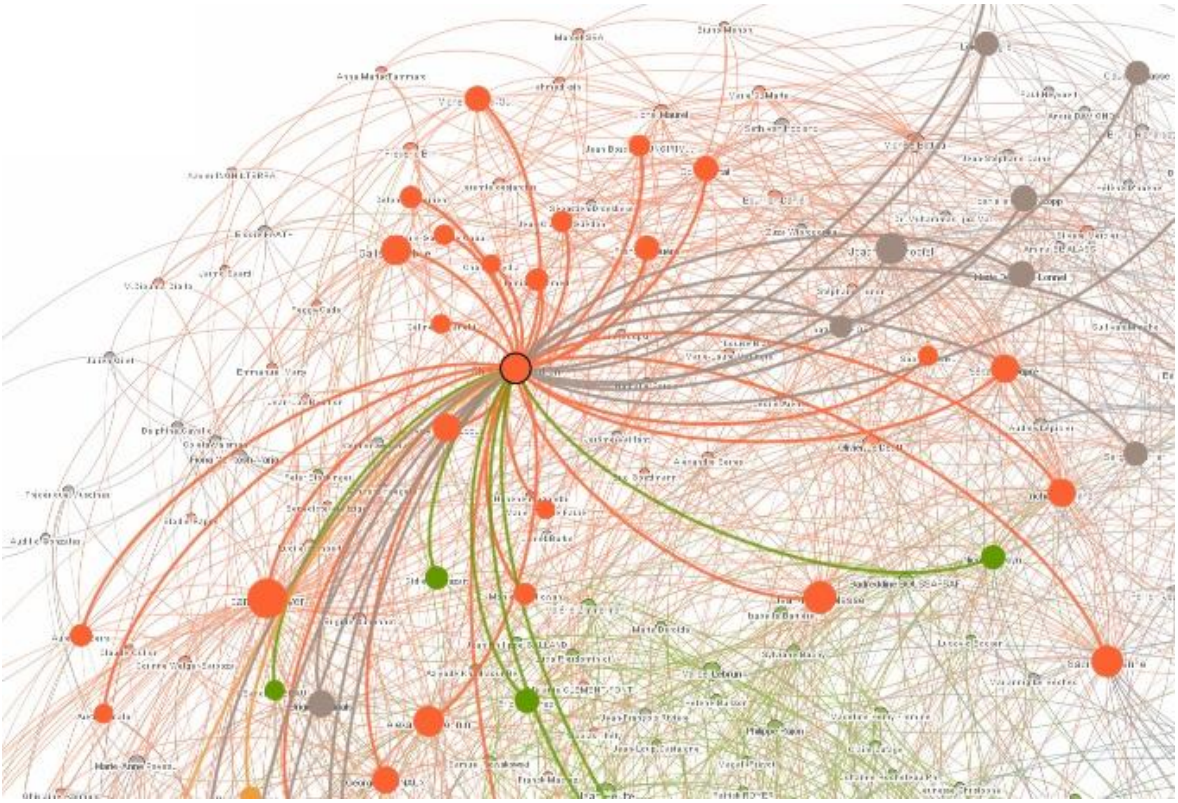


Figure 23 : Trouver les nœuds importants de son réseau professionnel.

En quelques clics, on parvient à identifier les nœuds stratégiques, ceux qui permettent d'accéder à de nouvelles sphères, à de nouveaux nœuds. Ce qui nous amène à penser que nous inventons sans cesse de nouvelles grammairesinteractives et méta-narratives. Celles-ci se répandent dans nos contenus et permettent, avec un nouveau niveau d'acuité, intelligence économique et datajournalisme.

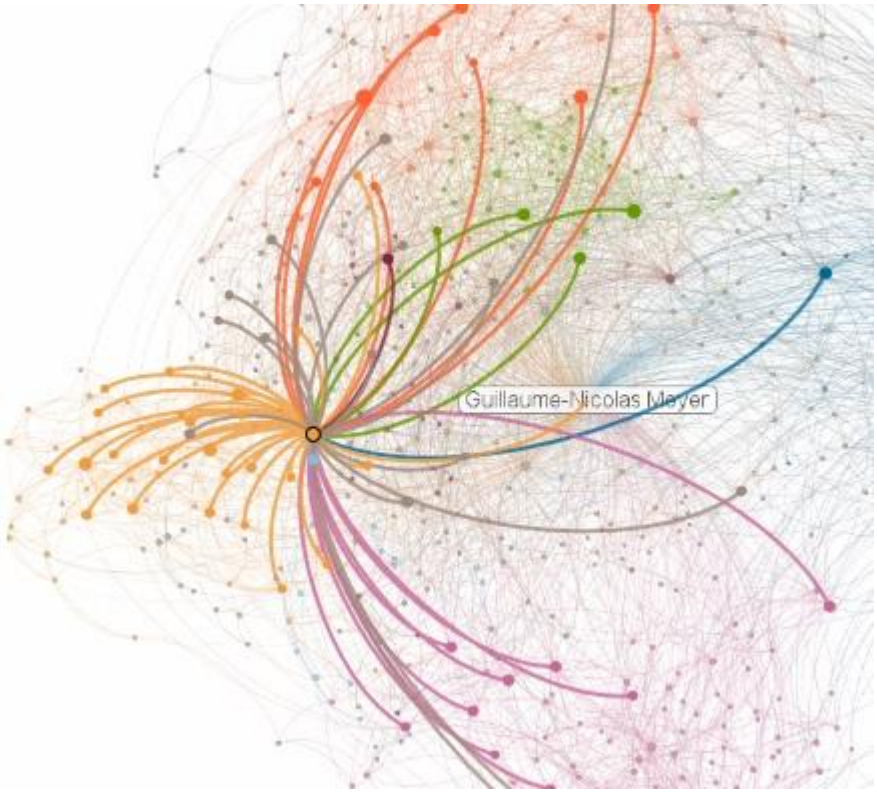


Figure 24 : Identifier un nœud stratégique dans une cartographie.

Nous venons de démontrer, par l'utilisation d'une simple cartographie basée sur l'exploitation des données de son réseau social professionnel que « l'homme est un document comme les autres » (Ertzscheid, 2009). Mais attention, si on en croit l'auteur, « *l'intérêt marchand du processus est d'enrichir la granularité et la rentabilité des documents et des identités via les usages et les recompositions qui seront opérées hors de la sphère de contrôle et de diffusion initiale*. » On se rapproche alors indubitablement du graphe social tant désiré par Mark Zuckerberg, le fondateur de Facebook. Il est ici cité dans un article de Francis Pisani[64] (2007) alors qu'il répond à une question de Michael Arrington, le patron de TechCrunch[65], « *C'est l'ensemble des relations de toutes les personnes dans le monde. Il y en a un seul et il comprend tout le monde. Personne ne le possède. Ce que nous essayons de faire c'est de le modéliser* ».

[64] http://pisani.blog.lemonde.fr/2007/09/18/facebook-cest-pour-mieux-te-connaitre-mon-enfant/
[65] http://www.techcrunch.com/

Pour s'y retrouver dans ce graph social dont nous faisons partie, bon gré, mal gré, il faut identifier les nœuds du réseau, en quoi la carte peut nous aider, puis les qualifier en reconnaissant ou non leur autorité sur un sujet donné.

2.2 Les figures d'autorité

Reconnaitre l'autorité d'une entité, c'est avant tout reconnaître son expertise et accepter son influence. Rappelons que 70% des contenus publiés sur le web le sont par les utilisateurs, qui étaient auparavant simples lecteurs. Il faut donc identifier nos autorités.

2.2.1 Autorité cognitive et autorité informationnelle

Autorité cognitive

C'est Patrick Wilson qui a défini le principe d'autorité cognitive comme « *une relation d'influence de pensée, impliquant au minimum deux personnes, l'une accordant à l'autre sa confiance parce qu'elle maîtrise un domaine spécifique de compétences* » (Broudoux, 2007). En effet nous dépendons pour une très grande part, d'idées et d'informations provenant de l'autre, des autres. Fondamentalement l'information vient toujours d'un autre agent cognitif que soi-même.

Nous avons tous des autorités cognitives et l'influence qu'elles exercent varient dans le temps et de façon contextuelle. Comme le disent Patrick Wilson et Soo Young Rieh, les autorités cognitives que nous nous choisissons ne le sont que sur un sujet donné et à un instant précis. Il y a fort à parier que notre autorité cognitive en coloriage à l'âge de cinq ans (notre maman vraisemblablement) ne sera pas la même que nos autorités cognitives en menuiserie, comptabilité ou médecine, quelques trente ans plus tard.

Nous investissons les gens d'une certaine influence, et pas seulement ceux que nous côtoyons. Nous prenons en effet pour autorités cognitives des personnes que nous n'avons jamais vues. Les messages du Dalaï Lama me paraissent toujours plus sages et altruistes que ceux de ma voisine, peut-

êtresimplement parce que j'ai attribué au Dalaï Lama une influence dans ces domaines.

Évelyne Broudoux et Alexandre Serres semble voir le concept d'autorité cognitive comme un modèle propre au web :

- « *Cette définition de l'autorité vue en tant qu'influence choisie est plus à même de s'accorder avec les pratiques de pointage des sites et des blogs entre eux (blogroll).*» (Broudoux, 2007)
- « *le nouveau régime d'autorité cognitive, en émergence sur le web, inverse le rapport et c'est la notoriété qui semble conférer désormais l'autorité, la compétence venant ensuite.* » (Serres, 2010)

Pourtant le concept d'influence choisi a toujours existé et il n'y a pas de raison de supposer qu'il n'existera plus à moins d'envisager un monde dans lequel nos échanges auraient lieu au travers d'avatars. Ainsi j'ai moi-même choisi d'investir de l'influence dans leurs écrits respectifs, alors même qu'ils n'ont pas une notoriété énorme sur le web. Dans mon cas, il s'agit bien d'une autorité investie sur la base de leurs compétences et non de la notoriété. Mais leur point de vue est peut être une illustration de l'amalgame qui est fait par nombre d'entre nous entre réputation et e-reputation, terme auquel Alexandre Serres préfère la notion de « présence numérique ».

Dans sa présentation sur la construction de l'autorité dans les médias sociaux[66], Martin Lessardpose une question pertinente: « *comment fait-on pour choisir une autorité cognitive en l'absence de toute compétence soi-même sur le sujet ?* » Et bien dans ce cas, nous essayons de déterminer l'autorité informationnelle de l'objet que nous sommes en train d'évaluer.

Autorité informationnelle

Évelyne Broudoux distingue l'autorité informationnelle de l'autorité cognitive car « *Contrairement à l'autorité cognitive, l'autorité informationnelle,*

[66] http://zeroseconde.blogspot.fr/2010/09/ma-presentation-credibilites.html

susceptible d'être portée par un individu ou un groupe, un objet ou un outil cognitif ou encore un média, n'a pas pour fonction principale l'influence mais celle d'in-former (donner une forme) » (Broudoux, 2007).

La chercheuse arrive ainsi à une représentation de l'autorité informationnelle que nous avons modélisée dans une carte heuristique.



Figure 25 : L'autorité informationnelle selon Broudoux.

Alexandre Serres, lui propose une modélisation légèrement différente. Nous avons conservé les mêmes couleurs pour pouvoir comparer les thématiques similaires.



Figure 26 : L'autorité informationnelle selon Serres.

Il nous apparaît étrange qu'aucun des deux modèles envisagés ne contiennent l'autorité cognitive. Car après tout qu'il s'agisse de l'autorité énonciative ou de la fiabilité de la source, c'est bien l'autorité cognitive qui est évaluée. Et même si nous ne connaissons pas l'auteur ou ne sommes pas en capacité de l'identifier, cela revient à définir la variable auteur de l'objet informationnel à « nulle »[67].

Enfin, si « *le passage au numérique exige une redéfinition du document* » (Pedauque, 2003), il nous apparaît essentiel de redéfinir également les deux figures d'autorité fondamentales que sont l'auteur et l'éditeur.

2.2.2 L'auteur, une figure d'autorité contestable

Nous allons considérer ici l'auteur du point de vue de Foucault, c'est-à-dire comme « *l'unité minimale d'information conditionnant l'existence du système juridique et institutionnel qui enserre, détermine et articule l'univers des discours* » (Foucault, 1994).

Être auteur sur le web

Selon Évelyne Broudoux, le modèle dominant du web participatif est « *l'amateur diffuseur* ». Si l'on en croit l'avis de Robert Samuelson, du Washington Post, qui dit que le web est « *la plus grande manifestation d'exhibitionnisme populaire de l'histoire* », on est tenté de le croire.

Nous croyons en effet que la figure d'autorité ancestrale de l'auteur « il fait autorité en la matière », s'est profondément diluée sur le web, mais qu'elle n'a pas disparue. Le fait de pouvoir écrire est une chose, le fait de pouvoir publier et simuler des commentaires de façon « anonyme » sur sa propre œuvre pour en faire la promotion, en est une autre. A tel point qu'« *Aujourd'hui, la figure de l'auteur apparaît affaiblie face à la profusion des productions dont il devient de plus en plus difficile de déterminer les sources et*

[67] Dans le sens du « null » d'une variable informatique ou d'une base de données.

les origines ; la publication par tous entraînant ce que l'on pourrait appeler un plagiat généralisé » (Broudoux, 2007).

De nos jours donc, la variable « auteur » d'un objet informationnel sur le web correspond à un indice de fiabilité. On peut observer que les propos d'un expert ont un indice de fiabilité supérieur à celui d'un amateur, mais ils n'ont plus le poids, le caractère sacro-saint d'autrefois.

Je sais comment tu me lis, les mutations de l'auteur

La dernière mutation de l'auteur que nous analyserons est sans doute la plus puissante. Quand vous lisez un roman dans votre espace personnel (fauteuil, canapé, lit), l'auteur n'a aucune idée de qui vous êtes ni de ce que vous pensez de son œuvre, il ignore jusqu'à votre existence. Il ne pourra vous quantifier que lorsque, quelques mois plus tard, il percevra ses droits d'auteur et, avec eux, le décompte d'acheteurs, par défaut le nombre de lecteurs.

Aujourd'hui, avec les statistiques géolocalisées proposés par des services gratuits tels que Google Analytics, n'importe quel auteur-publicateur peut quantifier ses visiteurs et analyser leur comportement. Il peut ainsi tricher, s'adapter et surtout, il est en mesure de vous *« imposer un parcours hypertexte et cognitif* » (Heinrich2009). Les moyens dont disposent les entreprises spécialisées sont énormes : eyetracking, ciblage sociodémographique des visiteurs, mesure de fidélité des lecteurs, étude des chemins de navigation, etc.

On peut alors se demander si c'est toujours l'auteur qui écrit l'œuvre ou bien si c'est l'œuvre, esquissé par les contraintes du marketing qui cherche un auteur pour se faire écrire.

L'éditeur, le gardien à qui on a volé les clefs

La fonction même de l'éditeur n'existe plus pour un amateur-diffuseur, puisqu'il s'autopublie. Ainsi les deux fonctions majeures de l'éditeur qui était la sélection et la publication ont été transférées à l'auteur sur le web. Nous

perdons ainsi une des variables identifiées dans l'élaboration de nos autorités informationnelles.

2.3 Faire confiance sur le web

Quel terme a une acception populaire aussi évidente que la confiance ? « Je lui fais confiance. J'ai confiance en son jugement. Il a trahi ma confiance. J'avais confiance en lui. » C'est un mot de tous les jours, un de ces mots sur lesquels on ne s'interroge pas tant il nous paraît évident. Mais l'est-il vraiment ?

2.3.1 Quelle typologie de confiance sur le web ?

Comme un serpent qui se mord la queue, l'entropie informationnelle change notre rapport à l'information, nos façons de la rechercher, de la percevoir, de l'évaluer, de la stocker, de l'annoter. Les impacts de ces modifications de structure et les brouillages cognitifs associés nous permettent-il encore de faire confiance et si oui, de quelle manière ? Nous avons tous en tête les jugements récents d'agences de cotations capables de décréter que tel ou tel pays n'est pas solvable, les précipitants un peu plus dans une crise financière au risque d'en entraîner d'autre. A la base de ces indices de réputation économique, il-y-a une évaluation d'information. Les différents épisodes de cette crise ont permis de mettre en évidence une réalité, une relation entre un ou plusieurs agents cognitifs : le pouvoir que nos systèmes financiers interconnectés ont attribué à ceux capable de modéliser la confiance, à savoir les agences de notations. Sur le web, l'algorithme PageRank de Google est une modélisation de la confiance.

Multidimensionnel, récursif et téléologique

La confiance et la défiance sont des attributs que nous utilisons tous et qui ont une grande influence sur notre perception et notre traitement de

l'information. Lors du colloque[68]« du mépris à la Confiance, de nouveaux comportements pour faire face à la complexité », en 1991, J.L. Lemoigne y a avancé que ce concept est à la fois multidimensionnel (il n'est pas réductible à une composition), récursif (il désigne à la fois l'action en elle-même et son résultat) et téléologique (il s'entend dans le temps, se transforme et est irréversible).

Pour Mellinger (1956), la confiance a indéniablement une dimension interpersonnelle en tant que facteur de communication. Ce qui permet à Loomis (1995) d'hypothétiser un rapport entre la communication,le développement de la confiance et le comportement coopératif. Car, de fait, les communautés sont maintenant les principaux relais pour créer du lien. « *Cela coûte si peu cher d'avoir des liens, qu'on peut les démultiplier facilement* » nous assure Azzem Azhar sur la scène de l'édition 2001 de la conférence Lift[69]. Et Hubert Guillaud de conclure dans son article[70] que *« la contrepartie est que la confiance générale finit par diminuer* ».

Les quatre formes de la confiance :

Adam Franck Tyar propose dans son ouvrage « Les aléas de la confiance » (Tyar, 1997) une typologie de la confiance ou plutôt des confiances. Il identifie les quatre formes suivantes de confiance : originelle, idéologique, méthodique et éthique. Nous nous proposons de transposer ces formes de confiance sur le web et nous allons pour cela en modifier l'ordre d'apparition : d'abord la confiance originelle, celle qui nous a fait naviguer la première fois, puis la confiance méthodique, plus construite et empirique, suivi de la confiance

[68]Colloque organisé par l'Université de Technologie de Compiègne (TSH/IDTH) ; Actes publiés en 1991 par l'UTC sous la direction de G. Le Cardinal et de J.F. Guyonnet

[69]http://liftconference.com/fr/news/new-video-azeem-azhar-online-communities-and-reputation-management

[70]http://www.internetactu.net/2011/02/17/reputation-influence-et-monetisation/

éthique avec les connaissances et les prises de conscience qui l'accompagne, et enfin la confiance magico-idéologique, celle que malheureusement le plus grand nombre investit aujourd'hui, notamment dans les moteurs de recherche.

2.3.2 Les moteurs de confiance

Une confiance aveugle ?

Le point de départ de notre recherche est que« *La plupart desinternautes n'identifient pas toujours avec précision ce qu'ils veulent (ou doivent) évaluerdans une ressource. Est-ce la source, l'auteur, l'information elle-même, l'organisation del'information, sa facilité d'usage, sa présentation… ?* » (Serres, 2010). Faisant cela, l'internaute passe à côté du véritable objet à évaluer : le dispositif d'accès à l'information, qui est, le plus souvent, un moteur de recherche et, encore plus souvent, Google.

Selon une étude sur la vie privée menée par l'agence média GroupM[71], une large majorité d'internautes français dit se sentir en sécurité sur Internet. Partant de ce constat, il n'est pas étonnant que les internautes aient une confiance magico-idéologique envers Google. Pourtant, « *Dans les milliers d'échanges quotidiens qui tissent ensemble le fond socioculturel du sujet postmoderne, il n'existe aucune volonté de promouvoir et de constituer un univers cognitif et moral* » (Lianos, 2001).

Une confiance aveugle !

Selon une enquête menée en 2004 aux États-Unis (Fallows, 2005), 62 % des personnes interrogées affirment ne pas distinguer les résultats naturels des liens commerciaux. Pourtant du côté des moteurs de recherche, la perception est différente. Google explique par exemple dans les 10 principes de sa philosophie[72] « *Sur Google, les publicités sont toujours clairement identifiées*

[71] http://www.lemondeinformatique.fr/actualites/lire-les-francais-se-sentent-en-securite-sur-internet-a-tort-31743.html
[72] http://www.google.fr/intl/fr/about/corporate/company/tenthings.html

comme "liens commerciaux". (..) Nos utilisateurs font confiance à l'objectivité de Google et rien ne pourrait justifier la remise en cause de cette confiance. » Alors qui croire, des utilisateurs ou une firme à but lucratif ?

Ce qui est sûr c'est que les utilisateurs font confiance à « l'objectivité » de Google car les nouvelles typologies de recherche ont créé un lien de dépendance très fort des utilisateurs vis-à-vis des moteurs de recherche. Comme le souligne Évelyne Broudoux (2007), « *Une écrasante majorité d'utilisateurs considère que les résultats donnés par les moteurs de recherche sont fiables et questionne peu la pertinence des résultats*. » Nous nous retrouvons donc dans une situation gênante dans laquelle la confiance est accordée sans l'exercice d'un sens critique, mais de fait.

Une évolution inquiétante des IHM[73]

Avec « Google Suggest » qui prédit les requêtes de recherche en temps réel et « Google Instant Search », qui affiche automatiquement les résultats associés à une requête populaire commençant par ces lettres on peut légitimement se dire que « *suggérer n'est pas prédire, mais peut fortement y ressembler* » (Ertzscheid, 2010).

« *L'un des outils méthodologiques les plus surveillés par l'équipe IHM ("Usabilitylabs") de chez Google est la fonction "refinesearch" qui (...) est un remarquable indicateur permettant d'établir des corrélations entre ce que les gens tapent et ce qu'ils cherchent "en fait" parmi la liste proposée. Entre les deux se construit une échelle d'intentionnalité qui semble observée de très près par Dan Russel*. » (Ertzscheid, 2010) L'échelle d'intentionnalité que nous avons bâtie dans la partie 1.3.3 sur le modèle de la pyramide de Maslow s'avère en fait être le nerf de la guerre pour le moteur de recherche majoritaire. Au plus il se rapproche de l'échelle d'intentionnalité réelle d'un

[73] Interface Homme Machine, domaine portant aussi bien sur l'ergonomie d'une interface que sur la mobilisation des ressources cognitives pour se servir de l'interface.

internaute, au plus il peut proposer à ses annonceurs un ciblage précis pour répondre à leur besoin de diffusion en publicité.

Dès lors, on peut se poser deux questions :

- « *Les moteurs de recherche ne seraient-ils que des dispositifs visant à transformer les personnes à la recherche d'information en consommateurs ?* » (Giffard, 2007)
- « *la capacité prédictive de nos systèmes sociotechniques va-t-elle tuer notre libre arbitre ?* » (Schirrmacher, 2009)

En conclusion, et comme Alexandre Serres (2005), nous pensons que « *l'évaluation de l'information reste surtout un questionnement personnel, incertain et permanent* ». A ce titre, les grilles d'évaluation proposées par de plus en plus d'université, ne sont pas à même de déterminer la relation entre une situation dans un contexte socioculturel donné et les objets informationnels, tels qu'ils sont perçus par un ou plusieurs agents cognitifs. De fait, il apparaît que ces outils ne permettent pas d'établir la pertinence socio-cognitive d'un objet informationnel à un moment donné. Ils sont, par contre, extrêmement utiles pour construire un process clair d'identification de la ou des sources, première étape dans l'évaluation d'une information. Nous estimons qu'à ce titre, les grilles d'évaluation peuvent faire partie d'une stratégie cognitive d'évaluation de l'information sur le web.

3. Partie 3 - Mémoire et redocumentarisation

3.1 Des informations pour évaluer d'autres informations

3.1.1 Stratégies et mémoires

Pour évaluer une information, nous nous basons, de façon consciente et inconsciente, sur nos savoirs, sur la masse d'informations que nous avons déjà évaluée. Même lorsque nous sommes complètement incompétents sur un sujet, nous avons en général suffisamment d'éléments pour construire une stratégie de recherche et d'être en mesure d'évaluer les informations trouvées et d'améliorer notre savoir.

Chacun, consciemment ou inconsciemment, nous élaborons des stratégies pour évaluer une information. Cela peut être en attribuant une confiance élevée sur un sujet donné à l'une de nos autorités cognitives, « *s'il me l'a dit, c'est que c'est vrai !* », ou bien en conférant à une autorité informationnelle un poids élevé « *je l'ai lu ce matin dans tel journal, alors ça doit être vrai.* », ou bien encore parce que nous jugeons notre cheminement, et ses itérations pour accéder à l'information, suffisamment pertinent pour la retenir et l'exploiter. Nos différentes stratégies se construisent dans un même but, prévenir, et s'appuient toutes sur la fonction cognitive mémoire. « *La mémoire du passé n'est pas faite pour se souvenir du passé, elle est faite pour prévenir le futur. La mémoire est un instrument de prédiction* » (Alain Berthoz, 2003[74]).

3.1.2 Un stockage protéiforme

Nous ne stockons pas toutes les informations qui nous parviennent sous forme sémantique, nous ré-encodons une information sous une autre forme ou sous de multiples formes.

[74] http://www.automatesintelligents.com/interviews/2003/octobre/berthoz.html

Un stockage qui n'est pas que sémantique

Quand vous pensez à un éléphant, ce n'est pas le mot en lui-même qui vous vient à l'esprit. Il s'agit en général de plusieurs informations comme l'image de l'animal, le son du barrissement, le mouvement de la trompe, une scène multimédia tirée d'un film, l'évocation de l'Afrique ou de l'Inde, des cultures…Notre cerveau reconstruit à chaque fois les relations entre forme, couleur, odeur, son, etc. d'un objet informationnel. Dans le cas d'une lecture isolée du mot éléphant, et en fonction du contexte, votre champ sémantique, votre taxonomie personnelle, mais pas que, va tout de suite rentrer en action, et c'est ainsi que vous allez penser, par association, aux lions, girafes, zèbres, gazelles car ces informations, ont des points communs avec l'information que vous venez d'évaluer. Tout d'abord ces informations sont classés dans la même catégorie sémantique, il s'agit d'animaux, ensuite ils sont reliés entre eux par une méta-information, il s'agit d'animaux de la savane, puis une autre, il s'agit d'animaux d'Afrique, etc. L'information qu'est l'image d'un éléphant a bien été ré-encodée et enrichie de marqueurs et de liens en fonction de nos apprentissages éducatifs, familiaux, socio-culturels, spirituels.

De même, si vous pensez à votre parfum de glace préféré, ce n'est pas le mot qui va vous venir à l'esprit mais le goût, la forme d'un cornet ou d'un pot, une évocation de la météo (en général on mange des glaces quand il fait beau et chaud) et parfois même un contexte de dégustation (en terrasse avec des amis, en soirée pyjamas entre filles, la première glace de votre enfant, etc.). On a donc bien un stockage de l'information qui est protéiforme. C'est pourquoi, dans notre système mémoriel, une information isolée est plus difficile d'accès tandis que les informations protéiformes, avec des marqueurs mnésiques variés, sont plus facilement accessibles.

Une diversité qui augmente la difficulté

Or cette variété de forme de stockage d'un objet informationnel et cette abondance de liens sémantique que nos neurones créent sont autant de critères qu'il va falloir discriminer, de façon consciente ou inconsciente pour

bien évaluer une information. C'est sans doute pour nous simplifier les tâches de traitement de l'information que nous avons optés pour des plates-formes spécialisées dans une forme. Ainsi, alors que l'on nous prédisait un web universel, interopérable à l'accès, il n'en est rien, nous optons pour Flickr pour stocker des photos, pour YouTube quand il s'agit de vidéos, pour des plates-formes comme Blogger ou Wordpress pour publier des contenus textuels et Deezer pour mettre de la musique en ligne. Nous avons reproduit sur le web des systèmes très différents du fonctionnement de notre cerveau.

3.2 Attention et appropriation

3.2.1 L'économie de l'attention

Sur le web, comme dans la vie de tous les jours, nous avons une attention très sélective. Par exemple, lors d'un dîner avec de nombreux convives, et alors que l'on discute avec quelqu'un, il nous arrive de capter la prononciation de notre prénom à l'autre bout de la table, au travers le brouhaha ambiant : les autres conversations, la musique et les bruits de fond. Ce phénomène, a été étudié pour la première fois par Cherry, en 1953 qui l'a désigné par « effet cocktail party ».

L'attention, cette ouverture

Nous retiendrons que « *L'attention est une ouverture vers l'extérieur. Il s'agit d'une rencontre qui s'effectue dans la relation : l'attention est construite conjointement par le sujet, en corrélation avec le cerveau, le corps et la myriade de connexions dans lesquelles nous sommes pris*. » Cette intervention de Tiziana Terranova(qui enseigne les media studies à l'Université de East London)a eu lieu dans le cadre de la conférence « Paying Attention[75] » organisée par l'European Science Foundationen Suède en septembre 2010. On retrouve ici les principales composantesde la pertinence socio-cognitive telle que présentée par Cosijn et Ingwersen (2000).

[75] http://www.esf.org/index.php?id=6527

Grâce aux outils de publication en ligne, qui s'affranchissent des barrières technologiques et financières, notre capacité à produire et à reproduire l'objet informationnel numérique est à l'origine de la surcharge cognitive que nous avons identifiée dans la première partie. Selon l'économiste et sociologue Herbert Simon (1971) : *« Ce que l'information consomme est assez évident : l'information consomme l'attention de ceux qui la reçoivent.* » Il ajoute : « *Du même coup, une grande quantité d'information créée une pauvreté de l'attention et le besoin de répartir efficacement cette attention entre des sources très nombreuses au milieu desquelles elle pourrait se dissoudre.* »

Du temps de cerveau disponible

C'est avec ses travaux qu'est apparu ce que l'on nomme aujourd'hui « l'économie de l'attention », dramatiquement illustré par les propos du PDG de TF1 en 2004[76] « *Ce que nous vendons à Coca-Cola, c'est du temps de cerveau humain disponible.* » Mais cette citation bien connue est encore pire lorsqu'on cite l'ensemble du propos : « *Or pour qu'un message publicitaire soit perçu, il faut que le cerveau du téléspectateur soit disponible. Nos émissions ont pour vocation de le rendre disponible : c'est-à-dire de le divertir, de le détendre pour le préparer entre deux messages. Ce que nous vendons à Coca-Cola, c'est du temps de cerveau humain disponible (...).* » (Lelay, 2004).Cette attitude est confirmée par Dana Boyd, sociologue de la société Microsoft, qui expliquaità l'occasion de la conférence SXSW[77] qui se tenait en mars 2012 à Austin : « *Dans l'économie d'attention, la volonté des gens pour distribuer leur attention à des stimuli d'information divers crée de la valeur pour lesdits stimuli. L'importance économique de la publicité repose sur l'idée qu'amener les gens à prêter attention à quelque chose a une valeur.* »

Maintenant que le média vidéo est omniprésent sur le média web, nous pouvons légitimement nous poser la question : est-ce qu'il en est de même

[76] Extraits de l'ouvrage *Les dirigeants face au changement* (Ed. Huitième jour)
[77] http://sxsw.com/interactive

sur le web ? C'est David Eun, responsable des partenariats de contenus chez Google à New York (en janvier 2008), qui nous apporte la réponse « *Les internautes vous paient avec le temps qu'ils passent sur vos contenus, ils vous paient avec leur attention. C'est cette attention, que les annonceurs veulent* ». Il n'y a plus aucun doute, les dispositifs de recherche d'informations, d'accès à l'information et de diffusion de l'information n'ont bien qu'un seul but : rendre notre cerveau « disponible », perméable aux annonces qui les font vivre. Quelle que soit notre intention en nous connectant, nous sommes l'objet d'influences méta-communicationnelles puissantes qui cherche à nous transformer en consommateur.

On comprend alors mieux les propos d'Olivier Ertzscheid (2009) : « *Dans le contexte d'une économie de l'accès et de l'attention totalement préemptée par les moteurs de recherche, tout est mis en œuvre pour accroître les possibilités de recouper systématiquement les données ainsi collectées jusqu'à constituer une « base de donnée des intentions » (Battelle, 2003) couvrant l'ensemble des données, informations et connaissances indexables.* »

3.2.2 Le processus d'appropriation

L'apprentissage est la principale activité du cerveau. C'est un processus qui permet de stocker des connaissances (une information évaluée et contextualisée devient connaissance), des émotions (un état affectif dont on a ou pas pris conscience) et des impressions (une perception évaluée comme permettant d'influencer notre comportement). Pour Jean-Philippe Lachaux, neurobiologiste au centre de recherche en neurosciences de Lyon la condition principale d'un bon apprentissage reste l'attention. « *L'expérience montre ainsi à quel point l'attention ne peut être focalisée que sur un seul objet, et qu'elle inhibe ainsi la prise en compte de tout autre stimulus*. »

3.3 Re-concevoirses stratégies

3.3.1 Répondre à ses besoins

Stratégies inconscientes

On peut retenir que, « *savoir s'informer, c'est accepter de remettre en cause ses connaissances antérieures et être prêt à intégrer ce qui n'irait pas dans le sens de ce qu'on sait déjà (ou de ce que l'on croit savoir). Être prêt à une telle remise en cause est inconfortable, mais nécessaire* » (Simonnot, 2008b).

Quand nous cherchons à évaluer une information, nous passons souvent par les mêmes étapes, surtout s'il s'agit de la même sphère de contenu. Par exemple si je cherche une information en équitation, je regarderai dans un des livres spécialisés de ma fille (jeune écuyère). Si je ne parviens pas à obtenir l'information que je recherche, je demanderai prioritairement à ma petite sœur, qui est une de mes autorités cognitives avérées en équitation et enfin j'effectuerai une recherche ciblée sur le web si je n'ai pas trouvé tout ce qu'il me fallait.Cela s'appelle une stratégie de recherche de l'information. La plupart sont inconscientes car peu d'entre nous s'interroge sur stratégies méta-cognitives.

Nos stratégies d'évaluation sont souvent identiques à nos stratégies de recherches. Nous avons classifiés deux types de stratégies permettant d'évaluer une information sur le web : les stratégies de bases et les stratégies élaborées.

Pyramide des besoins informationnels

Grâce à la pyramide de Maslowde nos besoins informationnels construite dans la sous-partie 1.1.3, nous pouvons dégager une frontière pour le choix d'une stratégie de base ou d'une stratégie avancée.

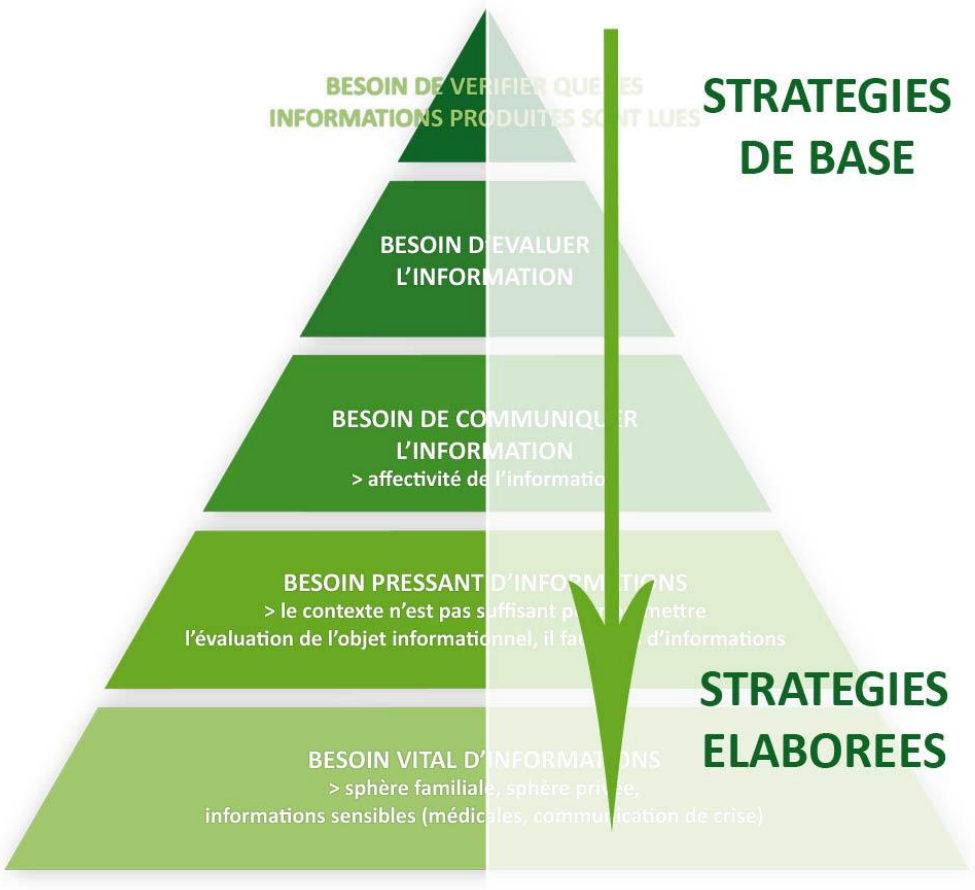


Figure 27 : Pyramide des besoins informationnels et des stratégies utilisées.

Nous estimons que les stratégies avancées sont à la portée de tous et qu'il n'y a pas de critères socio-professionnels discriminant pour l'usage de ce type de stratégie d'évaluation de l'information (cela reste à prouver). Par exemple, pour évaluerune information comme le choix d'un film à regarder, une stratégie de base suffira tandis que pour évaluer une information comme identifier une thérapie pour une maladie grave dont un de ses proches est atteint nécessitera une stratégie élaborée. Nous voyons bien, à travers de simples exemples choisi dans la sphère personnelle, que nous n'opterons pas pour la même stratégie d'évaluation en fonction des enjeux sociaux-cognitifs auxquels nous sommes soumis.

3.3.2 Les stratégies de base

Favoriser le hasard – J'ai de la chance

L'usage massif, voire unique, de l'interface épurée du moteur de recherche de Google et autre « Google Like »[78] aura permis, de façon graduelle, de passer d'habitudes de recherche très précises à très floues. Les interfaces des logiciels des bibliothèques obligent en effet l'internaute à structurer sa recherche, à formaliser sa pensée, à faire un travail précis d'identification de son besoin et des moyens d'y accéder. Notre cerveau nous prédit, nous avertit, qu'il existe même une probabilité élevée de ne pas trouver l'information, le livre, qui nous intéresse. A l'inverse, les moteurs de recherche sur le web offrent un modèle cognitif flou. Non seulement notre cerveau nous prédit que nous allons trouver un résultat, mais pire, il tolère le concept de chance, d'approximatif.



Car si l'on clique sur le bouton « *J'ai de la chance* », c'est bien que l'on octroie aux résultats de recherche un indice de confiance élevé. En toute logique, le flou cognitif lors de la recherche engendre un flou au niveau des résultats. En effet, comment évaluer impartialement, correctement, des résultats dont on se dit à l'avance qu'ils vont être pertinents ? « *Du mode recherche flou, on*

[78] Interfaces ressemblant à la page d'accueil par défaut du moteur de recherche qui ne présente qu'un seul champ de saisie et un bouton.

passe alors au mode résultats floutés. Ce constat est évidemment d'autant plus accablant qu'il s'accompagne de l'autosatisfaction aveuglante liée à une découverte par soi-même effectuée. » (Gallezot*et al*, 2008).

Tous les étudiants en Sciences de l'Information et de la Communication connaisse la citation de Roy Tennant « *Seuls les bibliothécaires aiment chercher, tous les autres aiment trouver.* », c'est peut être une des raisons pour lesquels s'est répandu l'usage massif de l'interface qui « trouve » toujours quelque chose.

Parmi les autres stratégies de base identifiées par Simonnot et Boubée (2010), nous pouvons citer la navigation qui consiste à se faire diriger par la transversalité des documents et la cueillette ou *berry-picking*, quiconsiste àprélever çà et là des bribes d'information qu'il faudra relier pour leur donner du sens.

3.3.3 Les stratégies élaborées

Dans les stratégies élaborées, nous trouvons la sérendipité, les équations de recherche, la rétroaction de pertinence, le chaînage et la veille spécialisée. Nous allons nous intéresser à la sérendipité car, c'est une stratégie, qui bien qu'existant avant l'avènement du web, prend tout son sens sur ce média.

Histoire de la sérendipité

Le mot sérendipité vient d'un conte persan du 16ème siècle. Il s'agit du nom d'un royaume (Serendib) situé dans l'île de Ceylan[79]. Quelques siècles plus tard, c'est l'écrivain Horace Walpole qui a décliné le mot en serendipity. Nos amis québécois ont choisis comme traduction : fortuité et nous sérendipité. Henri Kaufman a consacré un livre « *Carnets de Sérendipité* » à ce sujet mais nous laissons le soin à Olivier Ertzscheid et Gabriel Gallezot d'en illustrer la portée : « *Nombre d'autres découvertes tout aussi essentielles pour l'humanité ont partie liées avec la sérendipité. En voici une liste non-exhaustive : le*

[79]Le Sri Lanka actuel.

principe de champagnisation (Dom Pérignon), la pasteurisation (L. Pasteur), la pénicilline (A. Fleming), les rayons X (W. Röntgen), la vulcanisation du caoutchouc (Ch. Goodyear). Citons également d'autres découvertes moins "essentielles" comme le "post-it" (où comment répondre à la question : "que faire d'une colle qui ne colle pas ?"), le Caprice des dieux, le Coca-Cola, le Zyban et le Viagra… » (Ertzscheid&Gallezot, 2003).

Hasard ou chance ?

Peut-être les deux, mais ce qui est sûr, c'est que la sérendipité n'arrive qu'aux esprits bien préparés, c'est pourquoi nous la plaçons dans les stratégies élaborées. Si « *L'effet "serendip" (…) consiste à trouver par hasard et avec agilité une chose que l'on ne cherche pas. On est alors conduit à pratiquer l'inférence abductive, à construire un cadre théorique qui englobe grâce à un "bricolage" approprié des informations jusqu'alors disparates.*» (Perriault, 2000).

Nous préférons la version d'Olivier Ertzscheid et Gabriel Gallezot qui définissent la sérendipité « *comme la propagation d'un style cognitif stable (mis en place au début de la session de navigation) dans un environnement différent mais contenant de l'information pertinente pour l'usager dans le contexte initial de sa navigation et au vu de la tâche qu'il s'était assigné*. »(Ertzscheid&Gallezot, 2003)

Cela correspond plus à l'idée que nous nous faisons de ce qui est identifié comme une stratégie de recherche… parallèle à mon sens. C'est-à-dire qu'on choisit, consciemment ou inconsciemment, une stratégie de recherche, et on active ou non son mode sérendipité. Cela nous permettra de trouver quelque part, quelque chose que l'on ne cherchait pas. Est-ce encore du hasard ? Nous avons dans la partie ci-dessous, étudié comment cela se passe ou peut se passer dans notre cerveau.

Comment notre cerveau génère de la chance

Notre cerveau reçoit tous les signaux qui l'environne, mais, en fantastique machine à traiter de l'information qu'il est, il est capable de mettre en place une série de filtres sémantiques très élaborés.De nombreuses expériences en psychologie induisent une représentation mentale sémantique. Le sens des mots joue un rôle essentiel dans leur mémorisation. En 1975 Collins et Loftus suggéraient qu'il pouvait s'agir d'un réseau neuronal sémantique. En 1996, Damasio, Grabowski, Hichwa&Damasio conclu à l'existence d'un lexique mental à trois niveaux : un niveau conceptuel contenant les informations sémantiques, un niveau lexical contenant les mots eux-mêmes et un niveau phonologique correspondant au son des mots.

Ce qui est sûr c'est que les processus attentionnels sont multiples et qu'ils s'appuient sur un réseau neuronal étendu, impliquant de nombreuses régions cérébrales, permettant ainsi à l'attention d'être sélective, soutenue, orientée, partagée, exogène ou endogène, consciente et inconsciente.

Grâce au modèle de Posner et Raichle établi en 1994, on a une information intéressante sur le phénomène « d'inhibition de retour ». Pour illustrer ce phénomène, disons qu'après avoir permis une attention vers une cible, nos neurones ont besoin de décompresser**.** Pour y arriver, ils octroient des « vacances » (un peu plus de 250 ms) à nos neuromédiateurs, afin de pouvoir les mobiliser sur un nouveau projet. Ces temps de repos sont nécessaires pour conserver une certaine santé mentale. En effet, le cerveau étant assailli toute la journée par de nombreux stimuli, s'il ne pouvait pas limiter son attention dans le temps sur un sujet donné, nous aurions du mal à dormir et bien sûrnous ne pourrions être réceptifs à d'autres stimuli.

La capacité de notre cerveau à bien gérer nos processus attentionnels (je me focalise, je passe à autre chose), entre clairement en jeu pour comprendre l'effet « serendip ».Le cerveau, une fois qu'il a déjà traité un mot, a du mal, après le phénomène d'inhibition de retour, à se focaliser de nouveau vers ce

mot, peut-être parce qu'il a déjà effectué les phases d'analyse conceptuelles, lexicales et phonologiques. C'est pourquoi il cherche d'autres points d'attention. Une page web classique ou une page de résultats d'une requête sur un moteur de recherche offre de multiples nouveaux points d'attention. Est-ce pour cela que notre cerveau nous suggère d'y porter notre attention ? Est-ce parce que notre cerveau a détecté à l'autre bout de la page un mot, une image, qui fait écho dans notre lexique mental que nous y dirigeons notre attention ? La sérendipité peut-elle être réduite à un processus neuronal d'alerte attentionnelle ?

En conclusion nous citerons la thèse d'Olivier Ertzscheid (2002) pour qui « *Le facteur déterminant dans les stratégies de navigation qui seront alors mises en place par l'internaute est celui qui lui permettra d'éviter le bruit et non plus le silence.* »

3.4 Re-documentariser

Documentariser un document implique de le classer, de le cataloguer. L'objectif de cette étape par laquelle passe tous les professionnels de la documentation est bien de qualifier le document, et d'en permettre un accès plus rapide. La documentarisation crée, renforce, la pertinence d'un document par le simple fait qu'elle l'indexe, le résume, le présente.

Dès qu'ils sont numérisés, et encore plus spécifiquement sur le web, les documents sont redocumentarisés par les auteurs, les producteurs, éditeurs, voire simples lecteurs. En effet dès lors que nous stockons l'URL d'une page dans notre gestionnaire de signets, de marques-pages, nous le redocumentarisons car nous le classons, l'indexons dans nos catégories, avec notre vocabulaire, notre taxonomie. Selon Manuel Zacklad (2007), « *Redocumentariser, c'est documentariser à nouveau un document ou une collection en permettant à un bénéficiaire de réarticuler les contenus sémiotiques selon son interprétation et ses usagesà la fois selon la dimension interne (…) ou externe (…)* ».

3.4.1 Les mots-clés, amélioration ou pollution des objets informationnels ?

Folksonomies

Le terme folksonomy a été créé par l'architecte de l'information Thomas Vander Wal, le 23 juillet 2004 pour être précis[80]. Le terme est construit du mot « taxinomie » (les règles de classification, taxonomy en anglais) et du mot « usagers » (folk). Comme le dit Olivier Le Deuff (2006) « *Le concept est lié à l'accroissement et à l'accélération de la production d'informations.* » A l'inverse des thesaurus et ontologies, les folksonomies se caractérisent par un vocabulaire non-normalisé et non-structuré, dont les termes n'ont pas de relations sémantiques ou hiérarchiques entre eux.

L'ensemble des tags, des mots-clés, des catégories, des thématiques que nous créons, chacun à l'image de notre façon de penser, participe à augmenter le flou cognitif ambiant. Nous créons donc de plus en plus de bruit autour du message. Ainsi, le processus de redocumentarisation du web, point nodal du web sémantique, qui aurait pu permettre son avènement, va, avec l'explosion des folksonomies, tout simplement lui interdire d'atteindre un niveau de pertinence sociocognitive satisfaisant.Adam Mathes, dans un article paru en décembre 2004, résume le problème en une phrase : « *Une folksonomie représente en même temps ce qu'il y a de meilleur et de pire dans l'organisation de l'information*. »

Ontologie formelle et web sémantique

Sans aller jusqu'au mythe de l'ontologie formelle de Husserl (Drummond, 2009), notre incapacité à s'entendre sur des ontologies communes et leur déploiement n'est rien comparé aux freins générés par les utilisateurs. Quel webmaster va se plier aux contraintes documentaires d'un web sémantique, et dans quel but ? C'est sans doute à cause de cette inaptitude typiquement humaine que Tim Berners-Lee parle maintenant de web des objets. Là, pas de

[80] On July 24, 2004 I responded just after that with, "*So the user-created bottom-up categorical structure development with an emergent thesaurus would become a Folksonomy ?*". Citation présentée sur le site de l'auteur : http://vanderwal.net/folksonomy.html

frein humain, on assigne une ou plusieurs propriétés sémantiquesdans la variable d'un objet informationnel et on laisse travailler les algorithmes de traitement automatique du langage.

Le succès des folksonomies réside dans plusieurs facteurs : tout d'abord, il n'est plus besoin de connaître l'ensemble d'un thesaurus pour savoir comment indexer un document, ensuite, l'investissement cognitif est minime et permet ainsi un ancrage dans un contexte socio-cognitif immédiat. Je vois une image de cheval, j'ajoute simplement le mot-clé « cheval ». Je n'ai pas à me soucier de la pertinence de mon objet informationnel pour savoir comment il va répondre à une requête sur « équitation » ou « équidés ». Si bien qu'aujourd'hui les folksonomies sont apposées partout, sur les textes, les images, les vidéos.



Figure 28 : Folksonomies pour un article publié sur le site lemonde.fr.



Figure 29 : Folksonomies pour une image postée sur Flickr



Figure 30 : Folksonomies pour une vidéo postée sur YouTube.

Or si ce web des objets permet un traitement automatisé de l'information, nous savons bien, chercheurs et professionnels de l'information, que les meilleurs algorithmes n'atteindront jamais le niveau de pertinence d'un expert qualifié. Il y a plusieurs raisons à cela et l'une d'elle est qu'une information peut s'avérer pertinente pour le domaine ABC, même si son corpus sémantique la rattache à l'ontologie du domaine XYZ. Franck Bulingue (2004) nous le rappelle, *« l'exploitation de l'information est un processus complexe qui, de ce fait, n'est pas infaillible.* » C'est bien pour cela qu'ont perduré les métiers destinés à qualifier l'information : bibliothécaire, documentaliste, journaliste. Mais sur le web, les professionnels côtoient l'amateur et l'autorité du médiateur s'est diluée.

De la modification interactive de l'indexation

Tous ces métiers utilisent un même processus pour identifier, qualifier et retrouver un objet informationnel, le mot-clé. Dans sa thèse, Hervé Le Crosnier présente la modification interactive de l'indexation comme un atout majeur permettant d'améliorer la description d'un objet informationnel dans un processus continu. « *Une autre démarche d'apprentissage à long terme concerne la capacité à modifier la description d'un document en fonction des*

utilisations qui en sont faites durant les différentes recherches consécutives qui ont jugé ce document. » (Le Crosnier, 1990)

Toujours dans sa thèse, Hervé Le Crosnier fait référence à une expérience menée par Jean Tague[81] sur 503 documents, 64 recherches et engageant 8 utilisateurs. Le chercheur estime que les résultats ont été améliorés d'environ 25 % grâce à la modification interactive de l'indexation. Dans le même registre, mais cette fois sur le web, l'étude de Golder&Huberman (2005),établit que plus une ressource est partagée, plus le nombre de tags la décrivant se stabilise ainsi que leur nature.

C'est tout à fait dans cet esprit que ce sont créés des services comme Delicious[82], Diigo[83] ou Pearltrees[84].

3.4.2 Partager ses savoirs

Lorsque l'on évalue une information sur le web, c'est souvent pour la transmettre. Évelyne Broudoux nous le rappelle, « *Les enjeux de l'autoritativité résident dans la constitution d'une mémoire de ses pratiques qui sortirait de l'espace technologique des ordinateurs pour alimenter de ses expériences la mémoire collective.* »(Broudoux, 2003)

Or la mémoire collective sur le web se construit autour des plates-formes de publication. C'est pourquoi nous allons nous intéresser à uneplate-forme de curation,*Paper.li*[85], qui propose de présenter de façon périodique des éléments qu'elle a agencé elle-même en fonction de source d'informations qu'on lui a passé.

[81]Tague Jean M., 1981, « User-responsive subject control in bibliographicretrievalsystems. », Information Processing& Management, N° 17, pp. 149-159.

[82]http://delicious.com/
[83]http://www.diigo.com/
[84]http://www.pearltrees.com/
[85]www.paper.li

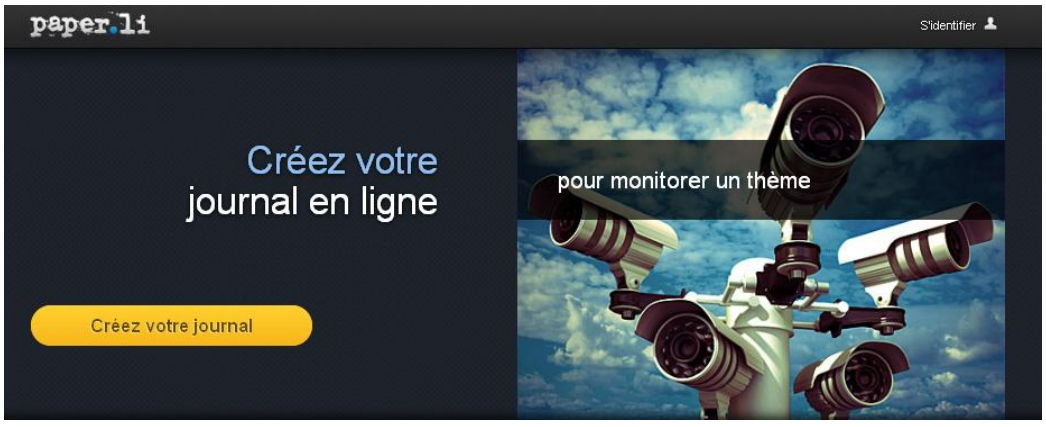
Dans notre échelle des besoins informationnels, nous sommes ici dans un besoin de communiquer l'information.

Paper.li

Paper.li est une plate-forme de curation intéressante car elle propose de s'occuper elle-même de la curation. Une fois que vous avez sélectionné vos sources

Votre travail de curation, votre valeur ajoutée, réside donc votre capacité à identifier et à sélectionner les différents émetteurs pertinents potentiels. On retrouve ici l'idée d'autorité énonciative annoncée par Evelyne Broudoux (2007)dans son concept d'autorité informationnelle. Le créateur d'un « journal » Paper.li attribue une pertinence suffisante à plusieurs autorités énonciatives pour faire reposer la sélection des contenus sur une règle mathématique.

La plate-forme de curation sélectionne alors, dans le flux des tweets, les éléments qu'elle juge les plus pertinents. Sans détail sur le fonctionnement de l'algorithme, nous ne pouvons que supposer qu'il tient compte du nombre de re-tweets[86], du nombre de followers[87] et sans doute de l'environnement social. Même si toutes les autorités énonciatives choisies émettaient le même nombre de messages, nous estimons que Paper.li re-définit l'autorité énonciative en attribuant un poids en fonction du nombre d'abonnés.



[86] Message publié par quelqu'un que l'on suit et que l'on décide de transférer à ses abonnés.
[87] Personnes abonnées à un compte Twitter.

Conclusion

L'évaluation de l'information sur le web est un enjeu sociétal majeur

Nous l'avons largement démontré, cette information que nous avions le temps de digérer parce qu'elle était rare, parce qu'elle était extrêmement qualifiée, par des sources d'autorités reconnues, nous ne sommes plus en mesure de l'évaluer correctement.

Nous avons produit en deux ans plus de contenu que dans toute l'histoire de l'humanité. 93 % de ce contenu est électronique et 70% des données sont créées par des utilisateurs (UGC).

Jay Rosen illustre bien la perte de contrôle du message de l'auteur face au lecteur : « *le journalisme est désormais une relation mutuelle avec des gens qui étaient considérés auparavant comme un public*. » (Rosen, 2010)

Plus que le contrôle du message, nous perdons la mémoire de la communication puisque le support (papyrus, livre, web) est là pour se souvenir de tout. Nous passons d'une mémoire interne à des mémoires externes, d'une mémoire subjective à des mémoires objectives. On peut donc se demander quels impacts l'externalisation des facultés cognitives ont ou vont avoir sur notre société ?

Pour autant, nous estimons que nous avons tout à fait les ressources cognitives pour accéder à un niveau de pertinence socio-cognitive satisfaisant.Peut-être que la situation dans laquelle nous nous trouvons ne nous permet pas de prendre du recul et de dégager une théorie de la pertinence sur le web.

Permettre la citoyenneté

Notre hypothèse est qu'il y a bien un nouveau paradigme d'évaluation de l'information, spécifique au web. Or le web est un espace multimodal, inter-civilisationnel et pluriculturel. De fait,« *L'importance qu'a la littéracie de l'information pour fournir les compétences et capacités essentielles à une citoyenneté active et aux apprentissages de toute une vie est au cœur d'un discours sociétal d'envergure.* » (IFLA, 2005)

Les débats publics sont complexes, on en mesure bien la portée lorsqu'on essaye de comprendre le rôle des TIC dans les révolutions du « Printemps Arabe » ou l'usage des techniques de multimédiatisation de la dernière campagne présidentielle en France. Pour Olivier Ertzscheid (2003), « *Appréhender la complexité d'un phénomène nécessite le repérage et la gestion d'un corpus de documents pertinent et volumineux pour extraire des informations et ensuite les transformer en nouvelles connaissances.* » Or ces conditions ne peuvent plus être réunies ni par le citoyen lambda, ni même par ses représentants élus. Nous vivons une ère de surcharge cognitive et cette *entropie informationnelle*(Gallezot, 2002) change notre rapport aux débats publics, nos façons de rechercher, de comprendre et d'évaluer les informations qu'ils contiennent. Les modifications des structures médiatiques et les *info-pollutions* (Serres, 2005) associées ne nous permettent plus d'évaluer l'information correctement, et donc d'appréhender la complexité du débat public, et donc de s'en saisir et de faire preuve de réflexion ….

Une compétence attendue

Selon l'Information LiteracyCommittee (1995) « *La 'culture' ou la 'maîtrise' de l'information (information literacy) pourrait être définie comme étant un ensemble d'habiletés permettant d'identifier quelle information est nécessaire, ainsi que de localiser, d'évaluer et d'utiliser l'information trouvée dans une démarche de résolution de problème aboutissant à une communication de l'information retenue et traitée. Cet ensemble peut aussi se présenter comme*

une série de compétences qui permettront à l'individu de survivre et d'avoir du succès dans la société de l'information. »

Arriver à une pertinence sociocognitive satisfaisante est un processus itératif complexe que nous mettons en œuvre à toutes les étapes de notre vie. De l'enfance à l'analyse stratégique, nous devons régulièrement évaluer la pertinence sociocognitive d'une information et donc accepter de, parfois, nous tromper. Mais ne nous y trompons pas, « *La maîtrise de l'information […] est l'une des cinq habiletés essentielles pour pouvoir intégrer le marché de travail dans l'avenir*. » (Information LiteracyCommittee,1995)

Zone d'autonomie perceptive

Nous hypothétisons qu'il existe sur le web une zone d'autonomie perceptive, basé sur les trois grandes fonctions cognitives que sont la perception, le langage et la mémoire. Cette zone, fortement soumises aux influences méta-communicationnelles du média web, se construit au fil des expériences de navigation et de la prise en compte du fonctionnement de ses stratégies de recherche et d'identification de l'information à un niveau métacognitif.

Il s'agit d'un processus individuel, chaque internaute a sa propre zone d'autonomie perceptive, qui fluctue en fonction de tous les paramètres que l'on a vu (contexte, sujet, affectivité en jeu, etc.), pour autant, nous en sommes convaincu, cette zone peut se définir par des critères valables pour tout un chacun, incluant les variabilités culturelles. Notre zone d'autonomie perceptive, une fois identifiée, servira de socle pourconstruire des autorités informationnelles stables, nous permettant ainsi d'atteindre un niveau de pertinence socio-cognitive satisfaisant en mobilisant moins de ressources cognitives.

BIBLIOGRAPHIE

ALA (American Library Association), 1989, « PresidentialCommittee on Information Literacy: Final Report », accessible [en ligne], http://www.ala.org/ala/mgrps/divs/acrl/publications/whitepapers/presidential.cfm, consulté le 5 novembre 2010

L'Atelier BNP Paribas « Les émotions collectives se repèrent aussi sur les réseaux sociaux » [en ligne] http://www.atelier.fr/reseaux/10/30062010/emotion-collective-partage-pic-marketing-communication-crise-produit-reseaux-sociaux-39965-.html, consulté le 16/10/10.

BALLARD, D. I. et SEIBOLD, D. R., 2004, « Organizationalmembrer's communication and temporal experience. Scaledevelopment and validation », Communication research, vol. 31, n° 2, April 2004, pp. 135-172.

BERNERS-LEE Tim, HENDLER James, LASSILAOra, 2001, « The Semantic Web », [en ligne], http://www.urfist.cict.fr/archive/lettres/lettre28/lettre28-22.html, consulté le 02 mars 2008.

BORNARELFrédéric, 2007, « La confiance comme instrument d'analyse de l'organisation », Revue française de gestion, n° 175, pp. 95-109.

BULINGE Franck, 2004, « Rôle et faiblesse de l'analyse dans la culture française de l'information. », Regards sur l'Intelligence Économique, N°5, Septembre / Octobre 2004.

BOYDDanah, 2009, « Ce qu'implique de vivre dans un monde de flux. », Web 2.0 Expo 2009, tiré de l'article de Hubert Guillaud du 06/01/10 paru sur InternetActu, [en ligne], http://www.internetactu.net/2010/01/06/danah-boyd-ce-quimplique-de-vivre-dans-un-monde-de-flux/, consulté le 13 novembre 2010.

BROUDOUX Évelyne, 2003, « Autoritativité, support informatique, mémoire », Journées " hypertextes, mémoire, fiction ", Montréal, novembre 2003.

BROUDOUXÉvelyne, GRESILLAUDSylvie ,LE CROSNIER Hervé , LUX-POGODALLAVéronika, 2005, Construction de l'auteur autour de ses modes d'écriture et de publication, [en ligne], http://archivesic.ccsd.cnrs.fr/sic_00001552, consulté le 17 novembre 2011.

BROUDOUX Évelyne, 2007, « Construction de l'autorité informationnelle sur le web. », accessible [en ligne], http://hal.archives-ouvertes.fr/sic_00120710/, consulté le 16 octobre 2010.

CARAYOL Valérie, 2005, « Principe de contrôle, communication et temporalités organisationnelles », Études de communication, 28, [en ligne], http://edc.revues.org/index276.html, mis en ligne le 03 novembre 2011, consulté le 20 mars 2012.

CARDON Dominique, 2010, « Pourquoi l'internet n'a-t-il pas changé la politique ? », propos recueillis par Hubert Guillaud, le 24 septembre 2010 pour InternetActu, [en ligne], http://www.internetactu.net/2010/10/07/dominique-cardon-pourquoi-linternet-na-t-il-pas-change-la-politique/, consulté le 26 septembre 2010.

COSIJNErica, INGWERSEN Peter, 2000, « Dimension of relevance. », Information Processing and Management, vol. 36 (4), p. 533-550.

DEHAENE Stanislas, 2008, cours sur "les mécanismes cérébraux de la lecture", Chaire de psychologie cognitive expérimentale du Collège de France, [en ligne], http://www.college-de-france.fr/site/stanislas-dehaene/les_mecanismes_cerebraux.htm, consulté le 06 septembre 2011.

DRUMMONDJohn J., 2009, « La limitation de l'ontologie par la logique », Methodos N°9, [En ligne], mis en ligne le 20 février 2009, http://methodos.revues.org/2131, consulté le 24 janvier 2012.

ERTZSCHEID Olivier, GALLEZOT Gabriel, 2003, « Chercher faux et trouver juste, Serendipité et recherche d'information. », CIFSIC, Bucarest 2003 – Atelier D2 - «Communication et complexité ».

ERTZSCHEID Olivier, 2004, « Les enjeux cognitifs et stylistiques de l'organisation hypertextuelle. », Thèse en Sciences de l'information et de la communication, Université de

Toulouse 2, sous la dir. de FC Gaudard& J. Link-Pezet, 450 pages, [en ligne] http://www.ertzscheid.net, consulté le 14 octobre 2010.

ERTZSCHEID Olivier, 2008, « Indexation sociale et folksonomies : le monde comme catalogue », présentation pour les« Journées ABES », Montpellier, 20 et 21 Mai 2008.

ERTZSCHEID Olivier, 2009, « L'homme est un document comme les autres : du World Wide Web au World Life Web », Hermès N°53, accessible [en ligne], http://archivesic.ccsd.cnrs.fr/sic_00377457/en/, consulté le 15 octobre 2011.

ERTZSCHEID Olivier, 2010, « La réponse avant la question. Ou le complexe du scribe. », blogaffordance.info, [en ligne], http://affordance.typepad.com/mon_weblog/2010/09/la-reponse-avant-la-question.html, consulté le 30 septembre 2010.

FALLOWS Deborah, 2005, « Search Engine Users: Internet searchers are confident, satisfied and trusting – but they are also unaware and naïve », Pew Internet & American Life Project, [en ligne], http://www.pewinternet.org/~/media/Files/Reports/2005/PIP_Searchengine_users.pdf.pdf , consulté le 11 novembre 2010.

FLAVELL J.H., 1976, « Metacognitive aspects of problem-solving », In Resnick and all : The nature of intelligence : Lawrence Erlbaum Associates.

GALLEZOT Gabriel, 2002, « La recherche in silico » In : Chartron G. (dir.) *« Les chercheurs et la documentation électronique : nouveaux services, nouveaux usages* », Edition du cercle de la Librairie, Coll. Bibliothèque.

GALLEZOT Gabriel, ROLAND Michel, ARASZKIEWIEZ Jacques, 2008, « La recherche floue» In : « Traitements et pratiques documentaires : vers un changement de paradigme ? » – Document numérique et Société, Paris, CNAM, France.

GIFFARD Alain, 2007, «*Enchaîner les esthètes, gouverner les Français*.», cycle de conférence Esthétique et société, Marseille, mai 2007.

HEIDERICH Didier, 2009, « Influence sur Internet. Perceptions et mécanismes d'influence sur Internet dans la société de l'urgence. », Publié par l'Observatoire International des Crises, accessible [en ligne], http://www.communication-sensible.com/download/influence-sur-internet-didier-heiderich.pdf, consulté en janvier 2010.

Information LiteracyCommittee, 1995, « *Information Literacy* », Ocotillo Report '94, Arizona, USA, Maricopa Center for Learning and Instruction (MCLI), disponible [en ligne] à l'adresse : http://hakatai.mcli.dist.maricopa.edu/ocotillo/report94/rep7.html, traduction disponible [en ligne] sur http://www.ebsi.umontreal.ca/formanet/culture.html, consulté le 12 novembre 2010.

JEANNERET Yves, 2000, « *Y a-t-il (vraiment) des technologies de l'information ?* », Villeneuve d'Ascq : Presses Universitaires du Septentrion.

LAROCHE Y., 1998, « Les mécanismes de la mémoire », in « *Pour la science*. *L'intelligence* », n° 254, décembre 1998, pp. 94-101.

LE CROSNIERHervé, 1990, Thèse « Systèmes d'accès à des ressources documentaires – vers des anteserveurs intelligents ».

LE DEUFF Olivier, 2006, « Autorité et pertinence vs popularité et influence : réseauxsociaux sur Internet et mutations, institutionnelles. » 7 p., [en ligne], http://archivesic.ccsd.cnrs.fr/sic_00122603, consulté le 26 octobre 2010.

LE DEUFF Olivier, 2010, « Nous avons de plus en plus besoin de lecteurs de crânes de licorne », [en ligne], http://www.guidedesegares.info/2010/12/31/nous-avons-de-plus-en-plus-besoin-de-lecteurs-de-cranes-de-licorne/, consulté le 2 janvier 2011.

LORENZENMichaël, 2001, « The land of confusion? High school students and their use of the World Wide Web for research », Research Strategies, vol.18, n°2, p151-163, accessible en partie [en ligne],http://www.libraryinstruction.com/confusion.html, consulté le 11 novembre 2010.

LOOMIS James L., 1995, « Communication, the developpement of trust and coopérative behavior », Human relation, 12, p 305-315.

MELLLINGER G. D., 1956, « Interpersonal trust as a factor in communication », Journal of Abnormal Psychology 52, p 304-309.

MIERMONT Jacques, 2003, « Contextualisations, communication et cognition. », Atelier N°2 - « Le complexe cognition-communication » du Programme européen MCX « Modélisation de la CompleXité ».

NAWALBoudechiche, 2009, « Traitement de l'écrit et construction de connaissances en contexte plurilingue : défis et opportunités. », in Synergies Algérie n° 6- 2009 pp. 29-40, Centre Universitaire d'El Tarf, accessible [en ligne], http://ressources-cla.univ-fcomte.fr/gerflint/Algerie6/nawal.pdf, consulté le 1er novembre 2010.

PAQUET Sébastien, mis à jour en 2009, « La confiance et la crédibilité à l'ère du réseau. » in « Le web social et la circulation de l'information. », cours en ligne de l'université du Québec à Montréal, [en ligne], http://benhur.teluq.uquebec.ca/SPIP/inf6107/spip.php?article=67&rubrique=8, consulté le 18 octobre 2010.

PEDAUQUE R.T., 2006, « *Le Document à la lumière du numérique* », Caen, C&F Editions.

PERRIAULT Jacques, 2001, « Historique de la notion de réseau. », in Actes du colloque L'éducation et sesréseaux, sous la dir. de Ferrero M..

PERRIAULT Jacques, « *Effet diligence, effet serendip et autres défis pour les sciences de l'information.* », [en ligne], http://archives.limsi.fr/WkG/PCD2000/textes/perriault.html, consulté le 15 octobre 2010.

PHARO Patrick, 2003, « *Naturalité et civilité du monde social* », Revue européenne des sciences sociales, XLI-127, [En ligne], mis en ligne le 30 novembre 2009, http://ress.revues.org/504, consulté le 08 avril 2012.

RASTIERFrançois, 2007, « Communication, interprétation, transmission », *Semen* [En ligne], 23 | 2007, mis en ligne le 22 août 2007, http://semen.revues.org/5341, consulté le 24 juin 2012.

RIEH, Soo Young, 2002, « *Judgment of Information Quality and Cognitive Authority in the Web.* », Journal of the American Society for Information Science and Technology, 53(2), 145-161.

ROUET Jean-François, MACEDO-ROUETMônica, ZAMPAVirgiie, 2008, « *L'information Internet : le jugement de crédibilité des usagers*. », Communications du 16ème Congrès de la SFSIC (Société française des sciences de l'information et de la communication) [en ligne], http://www.sfsic.org/congres_2008/spip.php?article132, consulté le 30 octobre 2010.

ROSEN Jay, 2010, Tiré de la leçon inaugurale de l'école de journalisme de Sciences-Po Paris le 02 septembre 2010, [en ligne], http://www.journalisme.sciences-po.fr/index.php?option=com_content&task=view&id=653&Itemid=1, consulté le 06 septembre 2010.

SAINT-JUDE Anaïs, 2012, « De Gutenberg à Zuckerberg », présentation Lift 2012, [en ligne], http://liftconference.com/fr/lift12/program/talk/anais-saint-jude, consulté le 19 mars 2012.

SALANSKIS Jean-Michel, 2003, « Herméneutique et cognition. », Villeneuve-d'Ascq : Presses universitaires du septentrion.

SCHIRRMACHER Franck, 2009, « La capacité prédictive de nos systèmes socio-techniques va-t-elle tuer notre libre arbitre ? », article d'Hubert Guillaud, le 18 novembre 2009 pour InternetActu, [en ligne], http://www.internetactu.net/2009/11/18/la-capacite-predictive-de-nos-systemes-socio-techniques-va-t-elle-tuer-notre-libre-arbitre/, consulté le 07 novembre 2010.

SERIEYX Hervé, 2002, « *Jeunes et entreprises, des noces ambiguës* », Paris, Eyrolles.

SERRES Alexandre, 2005, « *Évaluation de l'information sur Internet, le défi de la formation*. », Bulettin des Bibliothèque de France - Paris, t. 50, n° 6, dossier : Formation des usagers, [en ligne], http://bbf.enssib.fr/consulter/bbf-2005-06-0038-006, consulté le 10 octobre 2010.

SERRES Alexandre, 2008, « *La culture informationnelle*. », ERELLIF - Equipe de Recherches sur la diversité Linguistique et Littéraire du Monde Francophone. UHB, accessible [en ligne], http://archivesic.ccsd.cnrs.fr/sic_00267115/fr/, consulté le 27 octobre 2010.

SERRES Michel, 2005, « *Les nouvelles technologies, que nous apportent-elles ?* », conférence enregistrée à l'École Polytechnique le 1er décembre 2005, cycle Culture Web dans le cadre des Thématiques INRIA, [en ligne], http://interstices.info/jcms/c_15918/les-nouvelles-technologies-que-nous-apportent-elles, consulté le 17 octobre 2010.

SIMONNOT Brigitte, 2008a, « *La pertinence en sciences de l'information : des modèles, une théorie ?* », pp. 161-182 ch 6 in Papy F. (dir.), Problématiques émergentes dans les Sciences de l'Information, Paris, HermèsLavoisier.

SIMONNOT Brigitte, 2008b, « *Être usager de l'information en ligne nécessite-t-il de nouvelles compétences ?* », Ch1, p 21-39 in Dinet J. (dir, « Usages, usages et compétences informationnelles au 21ème siècle », Paris, Hermès Lavoisier.

SIMONNOT Brigitte, BOUBEE Nicole, 2010, « Les activités informationnelles et documentaires :des pratiques primordiales au coeur des apprentissages », in « Repères pour la mise en œuvre du Parcours de formation à laculture de l'information », Ministère de l'Éducation Nationale.

SUTTEREric, 1998, « Pour une écologie de l'information », Documentaliste - Sciences de l'information, vol. 35, n° 2.

TILMAN Francis, 2005, « *Les concepts de l'évaluation* », Atelier de pédagogie sociale Le Grain ASBL, [en ligne], http://www.legrainasbl.org/Les-concepts-de-l-evaluation, consulté le 04 novembre 2010.

TYARAdam Franck, 1997, « *Les aléas de la confiance* », chapitre XIII, Paris,L'Harmattan.

WANGPeiling, SOERGELDagobert, 1998, « A Cognitive Model of Document Use during a Research Project. Study I. », Document Selection, Journal of the American Society for Information Science and Technology, 49(2), 115-133.

WANGPeiling, DOMAS WHITE Marilyn, 1999, « A Cognitive Model of Document Use during a Research Project. Study II. », Decisions at the Reading and Citing Stages, Journal of the American Society for Information Science and Technology, 50(2), pp. 98-114.

WATZLAWICK Paul, HELMICKBEAVIN Janet, JACKSON Don D., 1972, *« Une logique de la communication* », Paris, Seuil.

WINKIN Yves, 1993, « Chronémique », in : Sfez L. (1993), « *Dictionnaire critique de la communication* », Paris, P.U.F., tome 1, p.434.

WILSON Patrick, 1983, « *Second Hand Knowledge : An Inquiry into Cognitive Authority*. », Westport, Connecticut : Greenport Press.

ZACKLADManuel, 2007,« Réseaux et communautés d'imaginaire documédiatisées », in « Skare », R., Lund, W. L., Varheim, A., A Document (Re)turn, Peter Lang, Frankfurt am Main, pp. 279-297.

SITOGRAHIE

Affordance.info

http://affordance.typepad.com

Blog d'Olivier Ertzscheid, maître de conférences en sciences de l'information.

Archive Ouverte en Sciences de l'Information et de la Communication

http://hal.archives-ouvertes.fr

@rchiveSIC a vu le jour en Mai 2002. Il s'inscrit dans le mouvement mondial des Archives ouvertes pluridisciplinaire HAL, et est destinée au dépôt et à la diffusion d'articles scientifiques de niveau recherche, publiés ou non, et de thèses, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Association des professionnels de l'information et de la documentation

http://www.adbs.fr

Créée en 1963, l'ADBS, forte de ses 5 000 adhérents professionnels, est la première association professionnelle de l'information et de la documentation en Europe.

Huyghe.fr

http://www.huyghe.fr

Le blog de François-Bernard Huyghe, Docteur d'État en Sciences Politiques Habilité à diriger des recherches en Sciences de l'Information et Communication.

Internet Actu

www.internetactu.net

Internet Actu, l'un des principaux sites francophones de veille sur l'innovation dans le domaine des technologies de l'information et de la communication, est soutenu par 15 acteurs de l'innovation et de la recherche pour garantir son indépendance, assurer son développement et enrichir son contenu.

Le Web social

http://benhur.teluq.uquebec.ca/SPIP/inf6107/index.php

Cours en ligne de l'université du Québec à Montréal.

Société française des sciences de l'information et de la communication

http://www.sfsic.org

La Société Française des Sciences de l'Information et de la Communication, créée en 1974, est une Association régie par la loi de 1901.

Zéro seconde

http://zeroseconde.blogspot.com

Blog de Martin Lessard, conseiller, conférencier, spécialiste en stratégie web et médias sociaux.

INDEX

Autorité cognitive 71
Autorité informationnelle 63, 72, 73
cartographies 63
cerveau 57
confiance 40, 76, 78, 88
contexte 30, 44, 46
contextualisation 24
curation 97
échelle d'intentionnalité 26, 79
externalisation des facultés cognitives 98
folksonomies 94
Google 32, 38, 54
Google Instant Search 79
Google Suggest 79
graph sémantique 38
hypertexte 61
hypertextualité 60
influence 15
influences méta-cognitives 43
influences méta-communicationnelles 100
infobésité 42
L'économie de l'attention 83, 84
métacognition 43
neurones 57
nœuds d'autorité 62
objet informationnel 95
ontologie formelle 93
pertinence socio-cognitive 45, 61, 83, 100
Protocole smtp 15
Pyramide de Maslow 27
Pyramide des besoins informationnels 27, 86
redocumentarisation 81
Redocumentariser 92
sérendipité 90
stratégies 81
surcharge cognitive 42
temporalité 46
Théorie de l'Information 15, 40
Twitter 49
User Generated Content 60
Web sémantique 34, 35

TABLE DES ILLUSTRATIONS

Figure 1: Le Memex 17
Figure 2 : Pyramide des besoins de Maslow. 27
Figure 3 : Pyramide des besoins informationnels. 28
Figure 4 : Evolution du nombre de noms de domaines sur le web. 31
Figure 5 : Peut-on quantifier le nombre de page sur le web ? 32
Figure 6 : Nombre de modèles d'organisation des connaissances en 2007. 36
Figure 7 : Nombre de modèles d'organisation des connaissances en 2009. 37
Figure 8 : Utiliser le web sémantique pour améliorer la pertinence des résultats d'une recherche. 39
Figure 9 : Vers un web plus intelligent. 40
Figure 10 : Théorie de l'information 41
Figure 11 : Une situation d'infobésité ? 42
Figure 12 : La réponse de Google à la pollution visuelle sur le web. 45
Figure 13 : Évolution du nombre de tweets par jour. 51
Figure 14 : Circulation de l'information après le phénomène Twitter. 52
Figure 15 : Google s'adapte à la nouvelle temporalité du web. 55
Figure 16 : La lecture en F de Nielsen. 58
Figure 17 : A gauche, l'activité cérébrale de la lecture d'un livre, à droite l'activité cérébrale lors de l'évaluation d'une information sur le web. 60
Figure 18 : Cartographie de la blogosphère italienne en 2011. 65
Figure 19 : Nœuds importants de la blogosphère italienne. 66
Figure 20 : Cartographie des contacts LinkedIn (681 contacts sur l'infographie). 68
Figure 21 : Mesurer la densité de ses contacts LinkedIn. 69
Figure 22 : LinkedIn, identifier un nœud de son réseau. 69
Figure 23 : Trouver les nœuds importants de son réseau professionnel. 70
Figure 24 : Identifier un nœud stratégique dans une cartographie. 71
Figure 25 : L'autorité informationnelle selon Broudoux. 74

Figure 26 : L'autorité informationnelle selon Serres. 74
Figure 27 : Pyramide des besoins informationnels et des stratégies utilisées. 88
Figure 28 : Folksonomies pour un article publié sur le site lemonde.fr. 95
Figure 29 : Folksonomies pour une image postée sur Flickr 96
Figure 30 : Folksonomies pour une vidéo postée sur YouTube. 96