



introduction à l'ingénierie documentaire et aux sciences de l'information

Brigitte Guyot

► **To cite this version:**

Brigitte Guyot. introduction à l'ingénierie documentaire et aux sciences de l'information. polycopié introductif pour la formation chef de projet en ingénierie documentaire, Cnam. 2011. <sic_00665267>

HAL Id: sic_00665267

https://archivesic.ccsd.cnrs.fr/sic_00665267

Submitted on 1 Feb 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Introduction à l'ingénierie documentaire et aux sciences de l'information

Brigitte Guyot

Maître de conférences, INTD / CNAM

Octobre 2011

Brigitte.guyot@cnam.fr

Tome 1

Sommaire

AVANT-PROPOS	4
QU'EST-CE QUE L'INGENIERIE DOCUMENTAIRE ?	5
CHAPITRE I – ELLE CONVOQUE PLUSIEURS DISCIPLINES	8
1. POSITIONNEMENT GENERAL : UN SECTEUR EN EVOLUTION PERMANENTE	8
2. ARTICULER PLUSIEURS APPROCHES DISCIPLINAIRES	10
CHAPITRE 2 – LA SCIENCE DU DOCUMENT : LES OBJETS DOCUMENTAIRE	
A. TYPOLOGIE DES DOCUMENTS	20
B. LES ELEMENTS CONSTITUTIFS D'UN DOCUMENT	21
CHAPITRE 3 - LES PROCESSUS DOCUMENTAIRES	27
CHAPITRE 4 - ANALYSE DE L'INFORMATION ET ANALYSE DOCUMENTAIRE	32
40. POSITIONNEMENTS :	ERREUR ! SIGNET NON DEFINI.
41. L'ANALYSE DOCUMENTAIRE : IDENTIFIER, DECRIRE, CARACTERISER, TRANSCRIRE	35
4.2 QUELQUES MODELES DE FORMALISATION / REPRESENTATION	ERREUR ! SIGNET NON DEFINI.
POUR CONCLURE LES TECHNOLOGIES INTELLECTUELLES	44
GRILLE D'ANALYSE D'UN DOCUMENT	42
GRILLE D'ANALYSE D'UNE PERSONNE	43
CHAPITRE 5 - LES ACTEURS DU SECTEUR DE L'INFORMATION	45
CHAPITRE 6 – SOCIOLOGIE DE LA PRODUCTION SCIENTIFIQUE	52
I - LE CADRE : FONCTIONNEMENT DE LA RECHERCHE SCIENTIFIQUE EN FRANCE	52
II - LES CONDITIONS DE PRODUCTION SCIENTIFIQUE : FONCTIONNEMENT D'UN MILIEU	54
III - OBSERVER LA PRODUCTION ET LES COOPERATIONS DANS LE SECTEUR SCIENTIFIQUE	55
IV - LES CHERCHEURS UTILISATEURS D'INFORMATION	57
CHAPITRE 7 – UN PEU D'HISTOIRE...	59
II – EVOLUTION DU RAPPORT AUX CONNAISSANCES	63
III - PETITE HISTOIRE DES OUTILS DE MEMOIRE DOCUMENTAIRE :	65
CHAPITRE 8 - POLITIQUES PUBLIQUES EN MATIERE D'IST	68
CHAPITRE 9 - CONCEPTS ET RECHERCHES EN SCIENCES DE L'INFORMATION	70
I - QUELQUES CONCEPTS ET THEORIES QUI ONT ALIMENTE LES SCIENCES DE L'INFORMATION	70
I.1. LIGNEE DES TRANSPORTS (COMMUNICATION DES MESSAGES)	70
I.2- LA LIGNEE ARTS DE L'INGENIEUR, SCIENCE DES SYSTEMES :	71
I.3 - LA LIGNEE SEMANTIQUE	73
I.4 - LA LIGNEE DES SCIENCES POLITIQUES :	73
I.5 - LA LIGNEE BIBLIOTHECONOMIQUE	73
II – LA RECHERCHE EN SCIENCES DE L'INFORMATION AUJOURD'HUI	74
CONCLUSION	76

ANNEXES :QUELQUES LECTURES

Buckland What is a document ?	76
Bruno Latour Ces réseaux que la raison ignore	83
Glossaire	87

Avant-propos

Devenir un spécialiste de l'information demande de s'intéresser en permanence à de nombreux aspects ; certains forment le « cœur de métier », alors que d'autres sont issus d'autres disciplines. Mais dans tous les cas, une vigilance est essentielle pour suivre l'évolution d'un contexte extrêmement rapide dans lequel ces fonctions s'exercent.

Citons en quelques unes :

- évolution des modes d'accès et de transmission avec les technologies (passage d'une médiation humaine à une médiation technique, du mode en ligne) ;
- évolution des objets documentaires : tous ne sont plus uniquement liés à la connaissance, mais de plus en plus à l'action (documents d'entreprises, connaissances, bonnes pratiques...) ; non plus seulement écrits mais aussi multimédia) ;
- évolution des unités d'information et des modes de traitement (documents structurés) ;
- évolution des supports numériques, des formats.
- Evolution des dispositifs d'information : KM, RM, IE s'implantent dans les grandes organisations, sont intégrés dans l'Intranet de la structure, au côté de site documentaire, collaboratif...

La liste de ces transformations est loin d'être exhaustive.

Une autre évolution d'importance concerne les méthodes de travail : travail à distance, travail collaboratif, partage des dossiers... Toute personne au travail devient productrice et gestionnaire d'information, la sienne, celle des autres, qu'il lui faut retrouver, rendre accessible, donner à lire ou à modifier, qu'il lui faut aussi classer..

Les compétences exigées de ceux qu'on nomme encore « utilisateurs » se multiplient : il faut maintenant jongler avec des agendas collectifs, écrire sur un wiki, mettre en partage des documents, valider l'information postée sur un Intranet, rédiger des documents pour des personnes qui n'en sont pas la cible première....

Tous ces éléments repositionnent le statut de prestataire de service qu'est le spécialiste de l'information. En effet, certains utilisateurs se révèlent encore plus connaisseurs que lui en matière de technologie, alors que d'autres vont s'appuyer sur lui pour les guider dans les méandres d'Internet, la prolifération d'informations de nature et de statut divers..

C'est en partant de sa connaissance de ces transformations qu'il pourra caler son action, se positionner pour s'adapter à l'évolution des modes de travail et d'échanges, et affirmer son rôle et le décliner en fonction de la culture de l'entreprise à laquelle il appartient...

Les éléments qui figurent dans ce tome 1 constituent des points de repère qui donnent une cohérence aux nombreux cours que vous aurez durant l'année.

En savoir plus :

Guyot B, Les dynamiques informationnelles dans les organisations (2006), Hermès-Lavoisier

Introduction : qu'est-ce que l'ingénierie documentaire ?

Un **ingénieur** assure l'étude globale d'un projet industriel, depuis sa conception jusqu'à sa mise en œuvre. Il rassemble les moyens techniques, humains et financiers pour parvenir au résultat. Il mobilise donc de multiples savoir-faire en plus de ses connaissances techniques. En outre, puisqu'il rassemble l'ensemble des moyens, il y a une forte composante relationnelle, notamment pour animer une équipe. D'où la pédagogie de la formation orientée projet et imposant le travail d'équipe.

Son rôle est pluriel : organisationnel + méthodologique + technique
--

L'ingénierie documentaire consiste alors à :

- **mettre en place** ou à gérer une chaîne **de production documentaire** : alimentation d'un site documentaire sur un Intranet, produits documentaires (bulletins, listes de nouveautés, revue de presse, synthèses, bulletin de veille, etc). Cette gestion documentaire courante nécessite d'évaluer chaque opération, de suivre le processus de production, de le contrôler et de rendre compte du travail réalisé ; de budgéter une action, de choisir les outils adaptés ou de faire appel à la sous-traitance, d'argumenter ses choix, de mobiliser une équipe, d'assurer l'adéquation des prestations aux besoins de sa cible ou de faire connaître ses activités... et la liste n'est pas exhaustive.

- **mettre en place des dispositifs d'information** ou des services documentaires. Ce sont de gros projets basés sur l'évaluation préalable de la situation, l'identification des contraintes, la recherche de solution (résolution de problème) ou de propositions pour justifier une nouvelle implantation ou l'ouverture d'un nouveau système d'information. Les savoir-faire sont ici ceux de responsable de projet ou de participant à un groupe projet : bonne connaissance de l'entreprise et des autres dispositifs d'information existants, insertion dans son fonctionnement, positionnement stratégique, capacité à faire travailler les gens ensemble ou à s'insérer dans un projet, identification des problématiques d'information sous-jacentes au projet, analyse des forces en présence, argumentation de son point de vue, prise en compte des cibles...(cf tome II janvier 2010).

Dans une telle optique, les seuls savoir-faire techniques sont insuffisants : il faut articuler plusieurs points de vue, organisationnel, utilisateur et technique. Même si la formation insiste sur le cœur de métier que constituent les méthodes documentaires, il ne faudra jamais perdre de vue qu'elles s'insèrent dans un fonctionnement d'ensemble (un secteur d'activité, une entreprise, un service), et qu'elles doivent se rendre transparentes pour les utilisateurs...

Ce que recouvre le qualificatif de « documentaire » :

L'activité documentaire s'exerce sur les documents produits en interne ou collectés à l'extérieur, ainsi que sur l'information qu'ils contiennent. D'où l'importance de toujours distinguer les différents niveaux, celui de l'information et celui du document-support de cette information

Selon les auteurs [L'information](#) recouvre plusieurs réalités. C'est :

1. Une **matière première** pour les professionnels, sous la forme de flux (information courante) à organiser ou de documents / supports à gérer. Elle se collecte, se traite, se transforme (on la munit de clés d'accès). D'où les notions de *fonds documentaire* lorsqu'on la mémorise et de *processus*, ou enchaînement d'*opérations*, pour la rechercher, la traiter et la diffuser

Ces opérations donnent lieu à un dispositif : *ensemble de moyens rassemblés pour assurer cette gestion*. Un dispositif documentaire (cf tome II) prendra la forme d'une banque de données, d'un service documentation ou de veille...

2. une **valeur d'usage** : pour un « utilisateur », l'information est **ce qui lui manque pour agir et résoudre un problème**, c'est-à-dire qu'elle est un facteur de réduction d'incertitude. Elle est donc indissociable de l'ACTION, (coordination, réflexion, argument...)
3. une **valeur économique**, qui s'échange et se monnaie.
4. une **valeur ajoutée**, enfin, dans une problématique de connaissance, qui s'ajoute aux connaissances et représentations d'une personne, s'y articule, les transforme. Elle devient donc un atout pour la personne qui la possède et qui s'en sert pour assurer sa position (cela rejoint le concept d'*asymétrie de l'information* des économistes).

L'ingénierie documentaire tient donc compte de tous ces aspects : méthodologique (les techniques documentaires), cognitif, stratégique (ce qu'on fait de l'information dans son activité professionnelle).

INFORMATION (Informare = mettre en forme)

Pour les économistes :

C'est un bien immatériel car :

Ne disparaît pas lorsqu'on s'en sert

Ne se possède pas (comme un bien matériel) : en location

Avec une valeur :

Sociale

D'usage (action, pouvoir)

D'échange (économique, atout dans une négociation)

Marchande (dans le cas du marché de l'information)

A aussi une dimension matérielle (il faut la gérer, notamment sous ses différents supports d'inscription)

Selon les cognitivistes :

Un signal qui modifie ses représentations

Selon les sociologues :

Donne lieu à une activité (échanges entre des personnes)

Document (*Voir typologie des documents au ch2*)

Une intention qui s'exprime (énoncé), se donne à voir

Un sens

Un ensemble de signes qui peuvent être traités

Une forme (support), surface d'inscription d'énoncés

Une expression de relation sociales (producteur, médiateur, lecteur)

Cf Suzanne Briet

Références

Briet Suzanne 1951 qu'est ce qu'un document, disponible sur <http://martinetl.free.fr/suzannebriet/questcequeladocumentation/> consulté le 10 sept 2009

Foray dominique 2001, La gestion des connaissances, Dominique, la découverte

Prendre la mesure du positionnement de l'ingénierie documentaire

1. Un secteur en évolution permanente

L'information fait partie de l'économie de l'immatériel, aux côtés de la formation, du conseil, et plus généralement des **prestations intellectuelles**, qui ont trait au **rapport au savoir**, à l'apprentissage, aux activités de consultance, à la formation. On parle aussi d'ingénierie des connaissances. On travaille à la fois sur le sens et sur le support qui le véhicule (le document).

« Documentation » sera ici entendu de façon large comme étant l'organisation de l'information en vue **de la rendre accessible et exploitable**. On peut se positionner soit en amont, du côté de gestionnaires et des dispositifs qui permettront un usage de cette information soit en aval, du côté de l'utilisateur.

Ce secteur émerge également au secteur de **l'économie des services** (qui analyse les relations qui s'établissent entre un prestataire, un produit et un client).

Il importe alors de rester vigilant sur plusieurs types d'évolution

1.1 – évolution sociétale : TIC et société de l'information

Les discours actuels sur la **société de l'information** recouvrent des réalités très différentes : les TIC (technologies d'information et de communication) seraient un facteur de développement économique censées assurer une démocratisation, recouvrent souvent des stratégies industrielles ou politiques, dont il importe de vérifier les incidences sur le terrain concret des pratiques professionnelles.

1.2 - L'évolution technique la plus spectaculaire concerne l'interconnexion des réseaux de télécommunication, dont Internet est la manifestation la plus récente. Mondialisation des ressources, accès universel et permanent en temps réel sans intermédiaires. Le web 2 est basé sur la participation de tous pour produire de l'information et de connaissances (wikipédia), la visibilité, chacun devient éditeur (blog, netvibes), partage ses signets, annote un site. Les logiciels libres se développent.

Les médiateurs documentaires voient alors leur rôle se transformer : les utilisateurs accèdent directement aux informations, ils développent des compétences, ce qui laisse supposer qu'ils ne feront plus appel à eux ; mais, en même temps, leurs attentes pour des prestations à valeur ajoutée augmentent. Le direct fait oublier qu'un site web ou une banque de données s'alimentent en permanence, selon certaines règles qui garantissent leur performance... repositionnement vers des fonctions de production, de conseil...

1.3 – Evolution économique : le marché de l'information

Celui-ci se développe avec une « marchandisation de l'information » qui fait évoluer les modèles économiques : abonnement, forfait, paiement à la demande, contractualisation, location, sont des modes de rémunération des éditeurs, auteurs ou des ayants droit. On peut constater des tensions entre une information gratuite et payante, ouverte à tous ou réservée à une cible particulière ; des tensions autour de la propriété intellectuelle qui oblige à des réajustements juridiques ; développement de nouveaux critères économiques pour évaluer la productivité des services d'information qui doivent dorénavant trouver leurs propres moyens de financement ou justifier de leur rentabilité.

1.4 – Evolution des entreprises,

Une place grandissante est portée à l'information : que celle-ci provienne de l'extérieur, avec l'intelligence économique ou la veille documentaire, stratégique ; ou qu'elle soit produite en interne, les termes de management des connaissances, de capitalisation des savoirs, de travail collectif autour de groupware, de partage des connaissances sont au cœur des discours. S'y montre un mouvement de « rationalisation » des circuits et dispositifs d'information. (cf tome II). D'où l'importance de développer un regard d'ensemble sur les modalités qui s'offrent à un organisme pour gérer son information (systémique)

Ces espaces de vigilance socio-économique sont aussi importants que le suivi des évolutions techniques parce qu'ils conditionnent la mise en œuvre des technologies ou des méthodes documentaires, du fait qu'ils concernent les utilisateurs et les partenaires. Ceux-ci appartiennent en effet à trois environnements qui s'emboîtent réciproquement :

- **l'entreprise**, qui constitue le point d'insertion immédiate (sur le plan institutionnel interne), qu'elle soit entreprise industrielle, société de service ou administration, pour laquelle travaille le centre de documentation ;
- le **secteur d'activités** dans lequel l'entreprise s'insère et agit ; clients, concurrents, cœur de métier...
- le **marché de l'information**, le secteur d'activités de l'information, qui constitue le cadre dans lequel évolue tout spécialiste documentaire, selon les moments client, observateur ou producteur ou partenaire. L'information donne lieu à une activité professionnelle spécifique, aux côtés des journalistes, hommes de communication, informaticiens dont elle emprunte certaines de leurs méthodes.

1.5 - la dimension cognitive est centrale, c'est-à-dire la façon dont fonctionne une personne dans ses relations avec la réalité et avec les autres (aspect psychologique, symbolique et communicationnel). L'information fait partie des **processus cognitifs**, ces mécanismes d'appropriation et de formation des connaissances dont est doté chaque individu. De ce point de vue, **se regarder fonctionner soi-même** et observer comment fonctionne l'autre aide à comprendre celui qu'on dénomme utilisateur. Il y a un travail d'auto-analyse et d'observation indispensable.

1.6 - Les méthodes, et des outillages documentaires font partie des **technologies intellectuelles** c'est-à-dire les facultés de classer, répertorier, calculer propres à l'esprit humain : l'invention de l'écriture et des supports d'inscription, depuis le papier jusqu'à l'informatique ou Internet, ont donné à l'homme des moyens d'étendre sa mémoire dans le temps et dans l'espace, comme autant de possibilités pour mémoriser, comparer, réfléchir, c'est-à-dire sortir de l'immédiateté, du réel, pour se le représenter et agir sur lui. L'extension

de ces espaces-temps ouvre des possibilités qui n'existaient pas dans une communication en face-à-face (Goody, Chartier).

Ces six points ont des impacts forts sur le « cadre d'exercice » de ce métier, comme le résume le tableau ci-dessous

Une activité de service positionnée en relation avec

- l'interne (se traduit par des fonctions, des intitulés)

- l'externe (car tantôt client, tantôt fournisseur ou encore interlocuteur auprès des acteurs du marché de l'information et du secteur d'activité de l'organisme)

2. Mobiliser plusieurs approches disciplinaires

A. Les sciences de l'information communication

Elles s'appuient sur plusieurs théories fondatrices :

a) le **schéma de communication** énoncé par Shannon (théorie mathématique de l'information, 1948) : Emetteur, Message, Récepteur. Cette vision issue d'ingénieurs des télécommunications, s'attache au système général qui permet de faire circuler l'information. L'accent sur le *système de transmission*, et l'information y apparaît comme un *message* à transmettre le plus fidèlement possible, en cherchant à limiter la déperdition, la distorsion. L'information est définie comme un contenu qui circule, sans qu'on s'intéresse au contenu c'est-à-dire au sens. Les informaticiens, par exemple, considèrent l'information comme un ensemble de 0 et de 1.

Lasswell, en 1936, commence à se préoccuper des *effets* du message sur le récepteur, percevant qu'une information modifie les représentations du récepteur (cf plus loin l'activité cognitive)

De ce point de vue, l'activité documentaire devient médiatrice dans un processus de circulation, comme une interface facilitant l'accès depuis la production du message vers l'utilisateur, mais aussi comme **interface** qui assure l'accès au producteur/auteur). Cela engage à « tenir les deux bouts de la chaîne » lorsqu'on réfléchit à l'information

b) dans la lignée du **modèle cybernétique** centré sur le pilotage des machines (Wiener 1948), l'information est un *signal* qui modifie un *état* antérieur et qui renseigne sur l'état du système technique (notamment avec le feed-back/réaction). Tout système possède de ce fait une forte composante d'information, laquelle apparaît comme un facteur de régulation et de maintien de l'équilibre du système.

L'approche systémique définit un *système* comme un ensemble d'éléments en interaction, regroupés et agencés pour parvenir à un but.

c) un **modèle industriel**, basé sur les trois pôles **entrée-sortie-traitement**, est issu de cette vision globale : des matières premières entrent dans un système, y sont transformées à l'aide de moyens matériels et humains, et en sortent sous forme de produits. La notion de *processus* est centrale,

processus : séquençement d'opérations pour parvenir au but fixé, réalisées selon des procédures.

processus documentaire : *ressource entrante (input) + transformation / traitement + production de produit ou de prestations en sortie (= output).*

Ce n'est pas sans rappeler l'expression de « chaîne documentaire » (cf chapitre 2)

B. Les sciences cognitives

L'approche cognitive définit elle aussi l'information comme un *signal* en tant qu'il contribue à mettre en forme ou modifier un système de pensée et de représentations. Cela pose directement la question de la construction du sens (et des connaissances), ce qui tend à faire de l'homme un système de traitement de l'information. Toute activité mentale est déclenchée par un signal (l'information) qui (ré)organise, construit des cartes cognitives, élabore de nouvelles connaissances. Cette gestion de symboles passe par un traitement de l'information en les reliant aux représentations mentales. (voir le schéma *système cognitif*)

L'intelligence artificielle, les cognisciences et la psychologie cognitive ont modélisé le fonctionnement cognitif en distinguant plusieurs phases : perception et captage du signal (expérience cognitive) ; reconnaissance et interprétation (traitement) ; mémorisation (sous différentes formes de stockage) et de rappel.

Parmi les multiples activités cognitives, citons celles qui permettent d'identifier, de reconnaître, d'établir des discriminations (sélection), de nommer, de raisonner, d'interpréter, de décoder, de faire des liens, des regroupements. Elles renvoient à des aspects d'organisation de la mémoire (et des clés d'accès mises en place pour stocker et retrouver l'information (organisation chronologique, par classifications, etc)

Dans cette approche, l'information est considérée comme un signal mais surtout un stimulus qui conduit à une réponse chez une personne. L'information modifie celui qui la reçoit, notamment sa carte cognitive, c'est-à-dire ses propres représentations du monde (ou d'une situation).

Cela montre que: l'information n'est jamais un en-soi désincarné. Le sens se « construit » par des opérations d'interprétation (en fonction des connaissances antérieures, du type de cheminement habituel, de la personne).

Plusieurs types cognitifs ont été repérés en ce qui concerne l'information :

- Profil visuel : on n'intègre une information qu'en la lisant
- Profil auditif : on intègre une information en l'écoutant seulement

Auxquels s'ajoute

- intuitif (vision, extrapolation),
- sensitif (sensible au présent, aux données, aux faits),
- rationnel (vrai ou faux),
- affective (bon ou mauvais).

Cela jouera évidemment un grand rôle dans le choix des informations.

L'information est également liée à l'action

Connaître, s'informer, se représenter sont des opérations cognitives singulières liées à l'action. Le schéma *individu en situation* rend compte de la complexité de ce qui entre en jeu quand on parle d'information. Chaque pôle identifie certaines caractéristiques, tout en étant en relation avec les deux autres pôles. Les axes qui relient deux pôles permettent de dresser quelques principes de vigilance. Elle permet de construire une relation à la CONNAISSANCE, en tenant compte de la situation de l'individu et l'action qu'il compte mener et du contexte propre à chacun d'eux.

Le pôle INFORMATION permet d'insister sur le fait que toute information est produite dans un certain contexte (condition de production), circule selon des canaux particuliers, et que tout ceci vient la connoter (cf chapitre 4)

Le pôle HOMME : insiste sur le système cognitif qui organise son rapport au monde. Les représentations mentales individuelles en dépendent mais aussi de son histoire, de son expérience, des connaissances antérieures, et constituent le « cadre de référence (*référentiel personnel*) »

L'axe Homme- information correspond à l'approche cognitive, c'est-à-dire à la relation qu'il entretient à l'information (cf les profils), les attitudes qu'il a par rapport à elle

Le pôle ACTION signifie que tout individu est engagé dans le « faire » qui englobe aussi *l'intention* de faire, prenant des formes aussi diverses que décider, vérifier, connaître, (ce qui fait apparaître la notion de « fonction de l'information »). L'action est conditionnée par une finalité (produire), par un contexte (économique, juridique), par des techniques et des méthodes. L'individu mobilise autour de lui les moyens d'effectuer son action, et l'information fait partie de ces ressources. L'action à mener oriente sa relation à l'information : connaître le cours d'une action avant une opération financière, la santé d'un concurrent avant de prendre contact avec lui.

Tout spécialiste de l'information est donc tenu de tenir compte, de façon conjointe :

- **le contexte individuel** : la place de la personne et son histoire conditionne sa relation et son attitude face à l'information- contexte collectif : prend en compte à la fois l'appartenance à l'entité de travail et l'organisation de dispositifs d'information
- **le contexte collectif** : dans un même milieu professionnel, si chaque individu a son propre référentiel individuel, selon son horizon antérieur (formation, type d'activité, appartenance institutionnelle), il appartient aussi et en même temps à un collectif, défini par des objectifs et par un métier, qui a ses habitudes, sa culture.

C'est ainsi que tout individu cherchant de l'information est porteur d'un double référentiel, personnel et collectif, ces deux niveaux venant forger une certaine relation, ou attitude face à

l'information ; on voit immédiatement la nécessité, pour un documentaliste, d'être vigilant sur ces deux niveaux, personnel et collectif, d'une personne qu'on appelle utilisateur. L'activité d'information est inséparable de la SITUATION, tant dans son usage que dans sa production ou que dans ses modes de traitement (activité des spécialistes de l'information-documentation).

Ce schéma exprime donc plusieurs niveaux d'attention pour le spécialiste de l'information : l'interprétation d'une information se bâtit sur la relation personnelle avec un contenu la situation, c'est-à-dire le contexte (personnel et organisationnel) variable selon l'usage et l'exploitation attendus

D'où l'importance, pour un spécialiste de l'information, de s'attacher à caractériser au mieux les informations dont il dispose pour donner à l'utilisateur le maximum de clarté sur l'information susceptible de lui convenir dans le contexte d'action qui est le sien. L'étroite interdépendance entre l'individu, son action et l'information est constitutive du *rapport à l'information*. L'individu a une certaine relation à l'information, et une intention par rapport à une action. Il veut pouvoir se cultiver, discuter, négocier...).

Tout système documentaire organise la rencontre entre une personne dotée d'une intention (d'action) et une information qui a, elle aussi, été produite dans un certain contexte (spatio-temporel)

C - La sociologie

La sociologie des organisations est mobilisée pour comprendre les jeux de pouvoir et les stratégies d'acteurs, l'évolution des organisations au fil du temps ; la sociologie du travail pour analyser les composantes de l'activité, la coopération, l'action collective, les collectifs de travail (cf tome II management de l'information, janvier 2012).

Etudier un système d'information suppose d'étudier la stratégie des différents acteurs impliqués, qui ont des statuts ou des places différentes dans l'organisation, avec des rôles particuliers (par exemple de producteur d'information, de gestionnaire, d'utilisateur ou d'usager, de contributeur, de valideur, d'éditeur).

D. - Les sciences de gestion

Enfin, toute réflexion sur l'information adopte nécessairement un *point de vue gestionnaire et managérial* soucieux d'organiser au mieux la gestion de l'information indispensable aux activités de l'entreprise. L'information y est considérée comme un « intrant » (diraient les économistes) qui contribue, au même titre que d'autres moyens, à produire des biens ou des services. (Pensons aux métaphores du genre « l'information est une ressource » comparable à une matière première si l'on adopte une approche basée sur le modèle industriel.

Une **organisation** est un système social (le terme d'entreprise recouvrant aussi bien une Administration, une entreprise commerciale, industrielle ou une association) structurée autour d'objectifs (par exemple vendre un service, un produit, assurer une mission de service public, etc). Elle œuvre dans un certain contexte / environnement spatio-temporel.

Etant en relation avec ses environnements (interne et externe) elle a besoin d'information pour s'alimenter. Elle met donc en place des processus de collecte, de traitement et de stockage dans des **dispositifs** ou systèmes d'information, appuyés aujourd'hui par le développement des technologies. Le management de l'information intègre nécessairement un management des technologies (informatiques et (télé) communicationnelles). Nous reviendrons longuement sur la question des dispositifs et des systèmes d'information (tome II).

De nombreux systèmes d'information

Un **système d'information** se construit autour d'acteurs, dans un contexte d'action, avec une organisation (des règles de fonctionnement, des règles économiques), éventuellement des technologies ; il a des relations avec d'autres environnements ou d'autres systèmes (de complémentarité, de concurrence, de spécificité...).

- des systèmes individuels :

Toute personne, en effet, est engagée dans son travail, avec des objectifs et tâches à remplir. Elle travaille selon des directives, règles, échéances, tout en ayant une marge de manoeuvre personnelle. Elle a, notamment, des besoins d'informations et met alors en place une stratégie pour les obtenir. Son activité génère en effet des relations plus ou moins formalisées, et il utilise des circuits d'information plus ou moins formels pour s'informer. Ce système de relations se donne à voir lorsqu'on interroge quelqu'un sur son activité et sur ses activités d'information (avec des questions comme : avec qui travaille-t-il, avec qui échange-t-il, auprès de qui demande-t-il de l'information, etc). Chacun est donc amené à chercher et à gérer des informations pour ses besoins propres.

Nous parlons alors de « système d'information personnel » (SIP) pour désigner l'ensemble de ces démarches et moyens mis en place individuellement pour rechercher, traiter de l'information, en produire et l'échanger. Il est totalement adapté à l'activité ou à l'action à mener. Cela s'ajoute à l'utilisation éventuelle de systèmes d'information collectifs (SIC) organisés par l'organisme pour assurer l'accès à des ressources d'information (*Les dynamiques informationnelles, Brigitte Guyot, 2006, Hermès-Lavoisier*).

- des systèmes collectifs (SIC):

Ces systèmes collectifs sont mis en place pour servir un certain public, et sont centrés sur la fonction de médiation et de prestation de service. Les objectifs de la structure, ses contraintes, orienteront leurs activités et leurs priorités.

Plusieurs cas de figure peuvent se rencontrer en ce qui concerne les unités documentaires. Ce peut être

- un service unique, qui centralise l'ensemble des ressources documentaires (assurant les acquisitions, le traitement ou le pointage des lieux où se trouvent ces ressources) ;

Systèmes d'information collectifs (SIC) dans une organisation : typologie

1) Ceux qui ont comme finalité d'organiser le travail : ce sont des outils de travail

Workflow : informatise un processus de travail

ERP (système de gestion intégré)

Serveur commun, avec ou non des espaces privés (GED)

2) Ceux qui gèrent des ressources d'information = outils d'information

Internes (mémoire / stockage, preuve) :

Archives (conservation à des fins administratives, historiques, culturelles)

records management (gestion de la documentation produite en interne)

gestion des connaissances ou des compétences (expériences, savoir-faire)

Externes (filtrage, exploitation) :

Bibliothèque, médiathèque, photothèque (préservation du patrimoine)

Services de documentation,

Services de veille

3) Mixtes (concentre des ressources internes et externes et des outils de travail)

Groupware / travail coopératif, dispositifs de travail à distance, Intranet

Une unité documentaire est en général dévolue à un service et ne travaille que pour lui. Historiquement, les premières unités documentaires furent implantées auprès de la fonction Recherche (ou R&D) ; aujourd'hui, elles tendent à se créer au plus près des utilisateurs (au sein d'un service marketing, service juridique, direction des ressources humaines, stratégie, etc). D'autres entités gèrent le patrimoine informationnel de l'entreprise (mémoire de l'entreprise) complétant le service classique des archives, ou encore les services brevets, normalisation ou qualité. C'est ainsi que dans les grosses entreprises plusieurs services de documentation, travaillent de façon autonome ou en réseau, selon la politique de la maison. Ils sont ou non associés au service de presse, ou celui de la communication externe ou interne, plus spécialement chargés de gérer l'image de l'entreprise en direction de l'extérieur, ou le climat social (journal interne). Aujourd'hui, la gestion des flux de données s'accompagne de pratiques collectives pour organiser et faire circuler les connaissances.

L'évolution récente montre un accroissement du nombre d'acteurs chargés de gérer, rechercher ou capitaliser l'information, qui constituent autant de partenaires pour les spécialistes de l'information documentaire :

- journalistes ou responsables de communication (qui écrivent le journal interne ou externe, conçoivent le web de la société) ;
- informaticiens, qui mettent en place le système informatique ou l'intranet ;
- ingénieurs responsables d'intelligence économique, qui animent des groupes de veille ou organisent le management des connaissances ;
- qualityciens, qui organisent la qualité et produisent des procédures et documents qualité ; documentalistes, qui s'occupent de propriété industrielle, de veille documentaire ;
- archivistes ou records managers, qui s'occupent de rationaliser la production de documents de travail et gèrent leurs cycles de vie
- contributeurs, responsables information d'un site intranet...

L'espace s'accroît, se diversifie, et il importe de prendre la mesure des positionnements de chacun pour s'y insérer, ou pour travailler en collaboration avec ces "corps de métiers" qui, tout en ayant tous l'information comme objet, ont parfois des représentations et pré-supposés différents.

Premières conclusions : être prestataire de service : positionnement et contraintes

Un système d'information documentaire occupe une position de prestataire de service, puisqu'il travaille au profit d'autres personnes ou métiers. A ce titre, il est soumis à plusieurs logiques.

- L'une qui adopte le point de vue des usagers (pôle utilisateur / service)
- une autre qui se concentre sur les ressources (pôle gestionnaire)
- une troisième qui correspond à la médiation elle-même.

Chacune obéit à des objectifs particuliers et possède son propre mode de fonctionnement. Tout gestionnaire d'information adopte alternativement le point de vue de chaque pôle :

- *Le pôle usager*, utilisateur, client ou demandeur : il est « muni » de besoins en informations, exprime des demandes plus ou moins explicites, a des habitudes et un comportement en matière d'information. Ce pôle est prescripteur, c'est-à-dire qu'il oriente et fait travailler les deux autres (qui ne se justifient pas sans lui). C'est ici que se mesure la satisfaction face aux services rendus. D'où la nécessité de connaître les différents comportements individuels face à l'information, d'être à l'écoute de leurs besoins, pour préparer un service qui leur soit adapté.

- *Le pôle ressources*, celui de l'offre, est constitué de l'ensemble des sources, fournisseurs, producteurs d'information rassemblés et rendus accessibles. Il met en œuvre des techniques et des méthodes documentaires pour organiser l'accès à ces ressources (cf ch 2).

- *Le pôle médiation*, enfin, assure l'interface entre le demandeur et les ressources pouvant l'intéresser : elle peut être directe, sans intermédiaire, car les technologies d'information

assurent maintenant un accès en ligne à des banques de données distantes ; elle peut aussi être médiatisée par une personne ou un centre de documentation qui assurera la recherche à la place de la personne et lui apportera les réponses. Cette activité de service est fortement liée au pôle de l'offre et à celui de l'usage.

Chaque pôle ayant sa propre logique de fonctionnement et ses contraintes spécifiques, il peut y avoir parfois des contradictions entre eux. Ainsi l'utilisateur veut une réponse rapide alors que le documentaliste cherche à maîtriser son budget, ce qui viendra jouer sur la qualité du service rendu (par exemple moins de temps à y consacrer).

Toute unité documentaire est donc tenue de prendre en compte à la fois les pôles ressources et usager. La gestion des ressources et leur traitement et diffusion constitue ce que le marketing nomme le *back-office* ou base-arrière (là où s'effectue la préparation du service) (1) ; le *front-office* étant l'espace où s'organise la relation directe avec le client (la vente, le conseil et le suivi de la prestation de service) (2). Le *back-office* obéit à une logique de gestion alors que le *front-office* privilégie la communication, la réponse, la diffusion.

Le versant technique (entendu comme l'ensemble des méthodes et des outils) s'organise autour des opérations documentaires (1) ; il s'agit d'acquérir de l'information, de la transformer et de l'organiser pour la rendre exploitable, en créant ou mettant en oeuvre des outils (fichiers, banques de données) ou des produits (bulletins documentaire, cédéroms, etc.). Ces opérations sont chiffrables économiquement : coûts d'acquisition, d'interrogation, de matériel, de gestion, par exemple. Tout service d'information se dote d'indicateurs quantitatifs pour évaluer son activité, calculer ses investissements, les coûts de traitement, de stockage, de production...

Par contre, le pôle d'*intermédiation* (2), orienté vers la communication et la relation directe avec le client, est confronté à une opération plus difficilement chiffrable : prendre en charge une demande (la comprendre, l'exprimer, la traduire) et assurer la réponse la mieux adaptée, c'est créer une relation de service, dans laquelle le client participe, crée lui-même une partie du produit ou du service (par exemple en donnant les éléments qui vont constituer la requête, en indiquant le niveau et la forme des résultats). La difficulté vient du fait que le spécialiste de l'information doit faire face à des demandes aléatoires, fluctuantes, imprévisibles. Il ne s'agit pas seulement d'assurer un accès technique, mais aussi de comprendre les processus cognitifs de ceux qui viennent (mais aussi de ceux qui ne viennent pas). Pourtant, c'est de ce pôle que proviennent la SATISFACTION et l'évaluation du service rendu qui, seule, justifie l'existence de la documentation.

On le voit, les intérêts, et l'organisation du service fonctionnent sur des modes différents. Toute la question est d'assurer la cohérence entre ces deux pôles du *back-office*, essentiel, mais qui n'a de sens qu'en fonction du *front-office*, lieu de la transaction où se mesure la satisfaction du client. C'est à partir des besoins ou demandes de ce dernier qu'il organise la collecte et le traitement, et non l'inverse. Peu importe au client la lourdeur des tâches documentaires et leurs contraintes de fonctionnement, alors qu'elles sont encore trop souvent "mises en avant" par les professionnels. Ce qu'il attend, c'est une qualité de réponse du système, se désintéressant du système lui-même qu'il veut sans contrainte pour lui-même.

L'évolution actuelle joue principalement sur cette relation client-médiateur. Les technologies de communication (TIC), en permettant d'accéder directement aux informations contenues dans des banques de données distantes, pourraient laisser entendre que le médiateur humain n'aurait plus sa place dans cette nouvelle organisation. La machine remplacerait

l'intermédiaire. Ils peuvent donc intervenir en amont, sur la production (1) ou en aval, dans la relation de service (2)

En effet ces produits ou outils d'information ont été conçus et structurés par une personne, qui y a mis sa logique d'interrogation (langages de commande, façon de formuler la question, etc). Certains documentalistes passent ainsi du côté d'une médiation qui d'humaine est devenue technique, oeuvrant dans le back-office en concevant des banques de données, des cédéroms ou un intranet, travaillant l'interfaçage pour garder une certaine relation directe.

Parallèlement, la proximité qu'entretiennent les documentalistes avec leurs clients leur procure une bonne connaissance des cultures professionnelles en présence : cela les rend précieux pour dialoguer avec les informaticiens, pour développer un nouvel outil, dialoguer avec les utilisateurs (un ingénieur n'aborde pas l'information de la même manière qu'un commercial). Cette situation privilégiée permet d'aider à l'appropriation des nouveaux outils techniques par l'utilisateur : sens pédagogique, accompagnement, souci de formation, deviennent des qualités indispensables.

→ Face à la diversification des ressources, des moyens d'accès, des acteurs, des technologies, le rôle du médiateur se transforme

→ Il sera amené à toujours **situer le cadre** dans lequel il se place, et à spécifier la focale, le regard qu'il va utiliser pour analyser ou rendre compte d'une situation d'information :

- sociologique (exploitant alors les apports de la sociologie des organisations, de la sociologie des sciences ou de la sociologie du travail),
- économiste (économie de l'information) ou
- systémique (qui considère une entité dans ses relations avec ses environnements) ;
- organisationnelle (sciences de gestion et management de l'information et des systèmes d'information)
- cognitive, travaillant alors sur les mécanismes d'appropriation des connaissances ;
- technique (l'instrumentalisation de la communication et du traitement de l'information).

Même si ce dernier aspect semble le plus développé dans la formation INTD, il ne faut jamais oublier d'entrecroiser ces approches : l'information est à la fois la matière première qu'il s'agit de collecter et de mettre à disposition. Elle donne lieu à une activité spécifique devant impérativement prendre en compte le fait que l'information n'est jamais un contenu EN-SOI, mais qu'il est connoté par le contexte dans lequel il a été produit et sera utilisé (que ce soit au niveau micro-social, intermédiaire ou macro-économique). L'ingénierie documentaire exploite toutes ces dimensions.

Quelques lectures :

pour évaluer le chemin parcouru :

Archier, Serieyx 1984 L'entreprise du troisième type, Seuil

Levy Pierre 1990, Les technologies de l'intelligence, Ed La découverte (

Nora Simon, Minc Alain 1978, Rapport sur l'informatisation de la société, La documentation française,

Paché Gilles, Paraponaris Claude, 1993. L'entreprise en réseau, Paris, PUF, (Que sais-je n°2704)

Serieyx Hervé 1992, *Le zéro mépris*, Seuil,

Sur le management, la gestion, l'environnement organisationnel :

Bernoux Philippe, 1995, *La sociologie des entreprises*, Seuil

Darbelet, Izard, Scaramuzza 1995 *Notions fondamentales de gestion d'entreprise : organisation, fonctions et stratégie*, Paris, Foucher,

Ghertman. Michel *Management stratégique de l'entreprise/* Paris, PUF (Que sais-je n°2463)

Guyot Brigitte 2006, *Dynamiques informationnelles dans les organisations /*, Hermès-sciences – Lavoisier

Lesca Humbert, Lesca Elisabeth, 1995, *La gestion de l'information : qualité de l'information et performances de l'entreprise /* Paris, éditions Litec, 210 p.-

l'immatériel : une nouvelle approche de la gestion (2000), revue française de gestion n° 130

Michel Jean 1992, *Pratique du management de l'information*, Paris, ADBS,

Segrestin Denis 1996 *Sociologie de l'entreprise*, Paris, Armand Colin (Coll. U),

Thiétart Raymond-Alain 1992 *Le management /*. Paris, PUF (5ème édition corrigée) (Que sais-je n°1860)

Dans un premier temps, consulter des revues généralistes comme :

Les objets documentaires - La science du document

Comme son nom l'indique, la **démarche documentaire** s'applique à des documents. Ces objets documentaires se sont multipliés au fil du temps, tant dans leur forme / support que leur nature. Ce terme d'objet oblige à caractériser ce qu'on observe : un livre n'est ni une brochure ni une revue. La forme numérique vient encore « brouiller » les repères classiques.

Après les livres (en général concernant un thème et écrits par un auteur unique (d'où le terme de monographie), vinrent les périodiques (plusieurs thèmes, plusieurs auteurs) qui se développent de façon exponentielle au XIX^{ème} siècle. A l'origine, ils étaient produits à des fins de connaissance (philosophique d'abord, puis scientifique au XVIII^{ème}, ensuite technique au XIX^{ème}). De nombreux documents sont dorénavant liés à l'action (pour la réflexion, décision, stratégie, la surveillance technologique ou celle des marchés). Leurs formes se diversifient elles aussi, notamment avec l'informatique.

Aujourd'hui, leurs intitulés et leurs formes sont extrêmement hétérogènes et fluctuantes : études de marché, bilans comptables, documents statistiques, annuaires, Référentiels métiers, profils de poste.... Tous demandent à être classés pour être retrouvés : classement physique (même sur un bureau électronique !) et classement intellectuel (dans quel fichier, sous quel thème / catégorie, décrits avec quels mots, retrouvables avec quelles clés d'accès...).

Document : Support de données enregistrées à des fins de preuve ou de renseignement. Il possède deux niveaux constitutifs : l'information et le support d'enregistrement.

a. Typologie des documents

Etablir une **typologie** consiste à regrouper des objets selon un même critère. Dans le cas d'un document, ceux-ci sont multiples, et peuvent se décliner selon :

- la forme des données enregistrées : texte écrit, image, son, données...
- la nature des données : statistiques, références bibliographiques,
- le support d'enregistrement : papier, audio-visuel (magnétique, argentique), numérique
- la forme éditoriale, liée au mode de production du document : livre, périodique, rapport, thèse, brevet, carte, annuaire, répertoire ...
- le contenu documentaire : document primaire (document originel/source), secondaire (références bibliographiques, métadonnées), tertiaire (synthèses, bilans réalisés à partir de plusieurs documents ou données...)
- le champ disciplinaire : juridique, scientifique, ce qui indique les usages/utilisateurs possibles
- la source : document interne (compte-rendu de réunion, rapport d'activités) administratif (courrier, méil), document externe (littérature éditée, périodiques.)...
- le statut : confidentiel, (accessible en lecture seule ou aussi en écriture) ou du moment de son cycle de vie (en relecture, validé...)

Ces typologies sont très souples et dépendent de l'usage souhaité. Le repérage documentaire tiendra compte de ces critères qui peuvent se combiner. Ainsi un document pris sur un site est un document numérique (avec un format, par ex .pdf), peut être un rapport officiel (forme) réalisé par un organisme international (source externe), avoir un statut (public), mais c'est aussi un document statistique avec des tableaux, des schémas....

La prise en compte de tous ces éléments donne lieu à des [métadonnées](#), c'est-à-dire à des éléments descriptifs (de la forme, de la structure, du type d'information, de son origine, etc) qui sont en général rassemblés dans une « notice bibliographique » : associée au document elle permet à un lecteur de sélectionner celui-ci sans avoir à en lire tout le contenu...

b. Les éléments constitutifs d'un document

il y a ainsi plusieurs strates. Un document, c'est

1. Une expression, un énoncé porteur de sens (un « contenu ») : l'œuvre d'un auteur
2. Écrit selon un code d'écriture, un langage plus ou moins formalisé : formules mathématiques, écriture musicale ; alphabétique, visuelle...
3. Auxquels s'ajoutent des règles techniques : logiciel (.pdf, .doc), formatage informatique (ASCII), écran.
4. Enregistré sur un support (de mémoire) : papier, électronique, objet (échantillon, œuvre muséale)...
5. Mis en forme selon des règles éditoriales conventionnelles : organisation interne (texte et paratexte (cf plus loin)
6. muni de marquages documentaires apposés pour permettre de le retrouver (accès) : identifiants, liens, classements selon le sens, selon les intentions présumées..
7. Doté de règles socio-organisationnelles : il a un statut, il est entouré d'un cadre d'usage, de règles de circulation
8. Contextualisé (réalisé et circulant dans certaines conditions (de production, de communication)

Un document conjugué :

un énoncé, écrit avec un code, + une structure intellectuelle + une structure formelle-logique (le faisant appartenir à un genre) + un système de navigation (paratexte / marquages documentaires) + mis sur support

Un support est réalisé selon des règles techniques (donnant lieu à un dispositif technique : formatage en livre, dimension de l'écran, fenêtrage, etc). Les documents numériques demandent un appareil pour être lus (ordinateur, tablette), un logiciel formate le texte, la taille des caractères, ainsi que les fonctionnalités, c'est-à-dire les actions proposées (touches fonction, icônes, type de navigation (par allers-retours, etc)...

Des règles de mise en forme (= éditoriales) s'appliquent à la PRESENTATION, à la mise en scène de texte, d'images, d'un multimédia, de graphiques... Un texte imprimé (par exemple) est en général structuré en chapitres, en paragraphes avec un titre. Il possède une organisation interne, linéaire ou structurée : un roman se déroule séquentiellement, chapitre par chapitre ; une thèse commence par l'énoncé de la problématique, puis la méthodologie, les résultats, les conclusions... Cette organisation se repère par les paragraphes, par les sous-titres, par la ponctuation. On voit déjà que forme et fond se répondent.

Cet énoncé peut être entouré de *paratexte*, c'est-à-dire d'un ensemble de données complémentaires qui viennent l'enrichir : appareil de notes, liens avec d'autres parties du document (internes) ou liens vers d'autres documents (externes), bibliographie. Elles complètent ce texte, comme une bibliographie, proposent des ouvertures ; un index rend possible une navigation d'un chapitre ou d'un paragraphe à un autre grâce à des liens/renvois) ou encore offre le moyen de retrouver une information (c'est la principale fonction d'un index).

Le mode d'organisation de l'information et des textes est conventionnel, et donne lieu à la notion de *genre* (un article scientifique, un dictionnaire) auxquels sont associés des codes d'écriture et de lecture. Le genre de document découle pour une part de la relation entre forme textuelle et paratexte: un roman suit une logique linéaire (peu de titres de chapitres) et possède peu de paratexte. A l'inverse, une encyclopédie est très riche en éléments paratextuels (mots-clés dans un index, renvois, bibliographie, définitions), qui permettent de retrouver une information à l'extérieur du document (sur un site web) ou dans une autre partie du document.

La logique de l'hypertexte n'est pas récente : les travaux de Paul Otlet au début du XXème siècle ou de Vannevar Bush dans les années 1940 insistaient déjà sur la nécessité de faire des liens entre documents. Ce type de navigation intertextuelle permet de sauter d'un document, d'un mot ou d'un chapitre à l'autre en utilisant les associations d'idées traduites en termes ou "boutons" qui renvoient vers d'autres textes ou qui utilisent le même mot pour le retrouver dans un texte (indexicalité). Les collections (rassemblement de documents autour d'un thème commun) ainsi constituées peuvent être virtuelles, créées le temps de l'interrogation par le lecteur.

type de document	caractéristiques	organisation	mode d'accès recherche navigation
texte littéraire	1 texte 1 auteur	séquentielle	linéaire
document collectif	X textes, X auteurs	texte + index + t matières	3 modes d'accès
		1 éditeur organise	
encyclopédie	texte, images, définitions, légendes	idem + liens organisés	clés d'accès multiples 1 éditeur et X auteurs
multimédia	idem + son	idem	idem
site internet	coll de documents	idem, liens internes / externes	

Quelques catégories de documents

Aujourd'hui, des balises de description sont incluses dans le texte pour indiquer le type de texte et son statut : titre, introduction, résumé, chapitre, paragraphe, bibliographie etc. Les normes SGML, HTML et XML servent à marquer les éléments significatifs des textes, tant sur le plan formel que sémantique.

Ces marques identitaires, énonciatives (le discours) et éditoriales (forme, structure) viennent jouer sur la « réception », en l'occurrence la lecture. Ces marques suivent en général des conventions d'écriture et de structure qui demandent à être repérées par le lecteur. Cela induit un certain nombre de « compétences » appropriées qu'il ne possède pas toujours. Le propre d'un texte électronique est de disposer toutes ces marques sur la même surface, brouillant leur repérage par le lecteur. Un site Internet offre une surface de « publication » de textes et d'informations de toute nature, selon un ou plusieurs modes de structuration.

Ainsi, la production d'un document suppose plusieurs dispositifs, entendus comme un assemblage / organisation de moyens et d'opérations :

Dispositif d'écriture, selon un code sémiotique (mathématique, musical..) Dispositif d'enregistrement, pour coder le contenu sur un support Dispositif de restitution, pour construire la forme sous laquelle le document pourra être lu Dispositif de lecture (le sens est véhiculé par la forme)

Qu'est-ce alors qu'un document structuré ?

C'est un document qui contient (ou auquel on a ajouté) un certain nombre d'informations à propos de sa structure :

- physique (blocs paragraphe)
- logique (liens),
- sémantique.

Ainsi les « marques » qu'on peut faire apparaître dans un logiciel de traitement de texte (structure physique) ; un CV est structuré, tout comme une grille d'analyse, une recette de cuisine, une fiche technique, parfois un compte-rendu, car le texte s'inscrit dans des champs identifiés et nommés (exemple : formation, expérience professionnelle, intérêts...)

Dans le cas d'une banque de données, l'information est découpée et répartie dans des *champs* munis de balises qui permettent de la retrouver (dans un paragraphe, dans un champ auteur, géographique, titre etc). Ainsi, on pourra distinguer les ouvrages écrits par Piaget auteur (on interrogera le champ auteur), de ceux qui parlent de Piaget (on interrogera le champ matière), thèses dirigées par Piaget (on interrogera le champ directeur de recherche), jury auxquels il a participé (jury..) etc

Faire une recherche dans un **document non structuré**, consiste simplement à trouver le terme qui figure dans le texte lui-même (recherche en texte intégral). Dans un document structuré, on peut par exemple distinguer une citation (car le texte est précédé et suivi de « »), d'une source (le nom d'un auteur sera mis entre crochets []), une référence bibliographique (à la fin d'un texte, du titre (balise titre), auteur (balise auteur), etc

HTML et XML sont des langages qui marquent un document selon une certaine grammaire

(*ML = mark up language*)

```
TEI.2 => teiHeader text
teiHeader => fileDesc
fileDesc => titleStmt
titleStmt => author title
author => <suite de caractères>
title => <suite de caractères>
text => group | front body
group => text+
front => head dateline
head => <suite de caractères>
dateline => <suite de caractères>
body => lg
lg => lg+ | l+
l => <suite de caractères>
```

exemple de balises indiquant le type d'information

Ce type de marquage (cf ch 4) permet de n'exploiter qu'une partie du document puisque celle-ci est aisément identifiable.

Ainsi, par exemple, on peut n'envoyer qu'une partie d'un compte rendu de réunion : le relevé de décision concernant une personne. C'est pourquoi on parle de « déstructuration » (un découpage en entités plus fines et significatives, qu'on appelle *unité d'information*) qui peuvent être ré-organisées selon l'objectif du lecteur. (cf Bachimont)

Ainsi on peut alors retenir que [l'information](#) se caractérise

- par la *nature des données enregistrées* : texte, son, image, chacune jouant un rôle différent sur l'esprit (comme l'enseigne la sémiotique).
- Par ses supports ; par ses fonctions (c'est-à-dire l'usage attendu ou effectif, par exemple se cultiver, décider) ;
- par la portée de ses interventions (plus ou moins étendue dans l'espace et dans le temps).
- L'information donne lieu à une activité particulière personnelle (celle que chacun fait pour son propre usage) ou une activité de service (réalisée pour d'autres, comme le font toutes les unités documentaires).

➔ la démarche documentaire

- est appuyée sur le contenu, sur le sens (étude du message en lui-même), utilisant alors les apports de la linguistique et la sémantique.

- décrit ce contenu « représenté », à l'aide de termes, de symboles et selon des formats conventionnels qui le rendent communicable et retrouvable.

Un document est un support qui inscrit une intention selon une certaine forme avant d'être mis en circulation.

Références

Terminologie documentaire, définitions

- <http://www.ebsi.umontreal.ca/etud/guides.html> , Ecole de bibliothéconomie de Montréal

- Vocabulaire de la documentation, Adbs, 2004

Buckland « what is a document ?

<http://www.sims.berkeley.edu/%7Ebuckland>

Buckland : what is a digital document ? 1997

<http://people.ischool.berkeley.edu/~buckland/digdoc.html>

Cacaly (ed. 2004) Dictionnaire encyclopedique de l'information et de la documentation, Nathan,

Sur les transformations documentaires :

Consulter le bloc-notes de Jean-Michel Salaün, Repérage de données sur l'économie des documents dans un environnement numérique

<http://blogues.ebsi.umontreal.ca/jms/index.php/>

Bachimont Bruno, Crozat Stéphane 2004 Instrumentation numérique des documents : pour une séparation fonds/forme,

http://archivesic.ccsd.cnrs.fr/docs/00/06/23/99/PDF/sic_00001017.pdf

Bertrand-Gastaldy Suzanne, 2002 Des lectures sur papier aux lectures numériques : quelles mutations ? (disponible http://archivesic.ccsd.cnrs.fr/docs/00/06/21/05/PDF/sic_00000256.pdf

Bruno Emmanuel, Calabretto Sylvie, Muriasco Elisabeth 2007, Documents textuels multistructurés : un état de l'art , Information - Interaction – Intelligence Vol 7 n° 1,

http://www.revue-i3.org/volume07/numero01/revue_i3_07_01_07.pdf

Chartier Roger « Du codex à l'écran », revue solaris, 1994

<http://biblio-fr.info.unicaen.fr/bnum/jelec/Solaris/d01/1chartier.html>

Chartier, Roger «De l'écrit sur l'écran.». *Imageson.org*, mai 2005, disponible sur

<http://www.imageson.org/document591.html>

Cotte D 2002, l'approche néophyte de la page web /. In les cahiers du numérique, n° 3, p 17-32

Goody Jacques 1979 La raison graphique, éd de minuit

Lainé-Cruzet Sylvie 2004, Documents, ressources, données : les avatars du numérique, Revue I3 4(1). 05 juillet. http://www.revue-i3.org/volume04/numero01/revue_i3_04_01_07.pdf

Lévy Pierre, l'intelligence collective, la découverte, 1994

Pédauque Roger T. 2006 Le document à la lumière du numérique, cféditions.

Salaün JM : Document et protodocument, in : Repérage de données sur l'économie des documents dans un environnement numérique, 19 sept 2011

<http://blogues.ebsi.umontreal.ca/jms/index.php/post/2011/09/19/Document-et-protodocument>

Souchier, Jeanneret, Le Marec (éd) Lire, écrire, récrire,. Paris, édi BPI, 2003

Chapitre 3

Les processus documentaires

Ce sont les **opérations** réalisées en back-office pour organiser l'accès (en amont) c'est-à-dire de faciliter la **recherche** d'information ou de documents (en aval). Chacune d'elles met en œuvre un ensemble de techniques et de méthodes ainsi que des outils spécifiques : cela constitue le savoir-faire de base spécifique du métier. Rappelons qu'un **métier** est un *ensemble de savoir-faire (acquis par une formation ou par l'expérience) partagés par les personnes qui se réclament de ce métier, avec une déontologie, et qui en général se constituent en réseau.*

Chaque opération se décline

en processus (enchaînement de tâches)

en procédures (comment faire),

met en oeuvre

une ou plusieurs méthodes,

des outils (intellectuels ou techniques)

mobilise un ou plusieurs types d'acteurs.

Cela suppose au préalable d'avoir identifié les objectifs (ceux de l'entreprise, ceux du service client), de connaître le contexte dans lequel on travaille (pour qui on travaille), de réaliser une évaluation permanente de travail, de développer des critères de qualité s'appliquant aux services ou produits. Cela permettra d'orienter son action, et d'effectuer une évaluation permanente des besoins des utilisateurs

Le processus chronologique est traditionnellement appelé « chaîne documentaire » et se découpe en opération et en phases. Cette représentation simplifiée ne correspond cependant pas tout à fait à la réalité rencontrée, puisque plusieurs opérations peuvent être effectuées en même temps (décrire, indexer, saisir les données dans une base de données)...

1) Recherche et collecte de l'information :

Elle s'appuie sur la connaissance des différents supports, de leurs conditions de production (qui, quand, où) et de leurs circuits de diffusion. Cela demande un suivi permanent des canaux et modes de production et de distribution de l'information, pour repérer, trouver, et acquérir de [l'information primaire](#), c'est-à-dire le document-source dans lequel se trouve l'information. L'information est dite **secondaire** quand il s'agit d'*information sur l'information* (ex bulletins bibliographiques, références, métadonnées). Il faudra ensuite mettre en œuvre une stratégie

d'acquisition ou d'accès au document primaire. L'information primaire (ou plus exactement le document primaire) est la source dans laquelle figure l'information.

Cela suppose d'en connaître un minimum pour pouvoir sélectionner l'information ou le document qui semble pertinent, en identifiant, pour commencer :

Conditions de production CP

Qui ? quand ? où (contexte), avec quelle intention ?

Conditions de communication CCOM

par quel canal ce document est il accessible (facilité d'accès, comment, où)

conditions d'usage : le cadre pour l'interroger (selon le support, le lieu...)

Il existe un certain nombre d'outils, certains

- identifient l'information (via des outils de recherche) et en indiquent les références.
- Indiquent comment accéder à un document (outils de localisation) : certains, comme les catalogues collectifs, permettent de localiser le document primaire, (titres de revues conservés et leur localisation dans une bibliothèque).

Certaines références donnent directement accès à l'information, alors que dans d'autres cas (banques de données bibliographiques par exemple), il faut se procurer le document primaire (via le prêt entre bibliothèques, ou encore le cat-INIST).

Puis interviennent des procédures d'acquisition (achat /commande, prêt, don, échange).

2) Traitement documentaire : préparer l'accès / exploitation

Le terme général d'analyse définit à la fois l'opération intellectuelle qui consiste à prendre connaissance d'un objet, et le résultat matériel de cette opération (par exemple un condensé, résumé, une analyse documentaire)

Analyse : opération qui consiste à réduire un objet en ses éléments constitutifs pour ensuite recomposer le schéma d'ensemble ou en démontrer la logique. Cela induit de disséquer, résumer, étudier, examiner. Opération intellectuelle qui suppose la notion d'observation, de recueil d'information sur un objet, de déduction suivie de conclusions. Il faut donc pouvoir identifier, décrire et nommer et, éventuellement, évaluer.*

Le traitement s'appuie sur cette démarche et se décompose en plusieurs opérations

1. L'identification / description : identifier les éléments qui vont rendre compte, de l'origine, de la forme, des conditions d'accès du document
2. L'analyse consiste à repérer les éléments constitutifs de contenu,
3. La représentation consiste à traduire dans des formats le résultats de l'analyse : par exemple la notice bibliographique, l'indexation et les résumés
4. La catégorisation : consiste à insérer le document dans un ensemble pour le rendre diffusable ou disponible (insertion dans une collection).

Cela suppose, au préalable, d'avoir défini ses critères de sélection (vérifiés régulièrement en fonction de l'évolution des axes de travail des utilisateurs).

Cette activité se réalise à plusieurs niveaux.

21) Décrire la forme du document, son origine, ses éléments caractéristiques : le support et ses conditions de production et de communication. Cette *description bibliographique* donne lieu à la rédaction d'une **notice bibliographique**, à l'aide d'outils que sont les normes internationales (leur importance pour les échanges de fichiers, l'homogénéité des données) (cf cours de description bibliographique)

22) Prise de connaissance du contenu :

Elle comprend deux niveaux : l'indexation (choix des thèmes et concepts traduits en termes significatifs) et la condensation (résumé du contenu)

22a. L'indexation consiste à choisir, en amont, les clés d'accès qui permettront de retrouver ultérieurement l'information. La principale difficulté est de faire coïncider les termes retenus à l'entrée (termes d'indexation choisis par le documentaliste) avec les termes qui seront utilisés par l'utilisateur lors de la recherche (« retrieval », ou output). Il faut donc prévoir toutes les façons d'interroger de l'usager selon sa démarche :

- il veut par exemple accéder directement à l'information recherchée (accès par mot-clé ou en langage naturel)
- il veut cheminer pas à pas (par exemple en choisissant un thème dans une rubrique)
- il souhaite naviguer, vagabonder (pour se donner des idées, faire le tour d'une question : c'est le principe de la navigation hypertextuelle, par allers et retours)

Cette opération d'indexation se fait à l'aide d'outils d'indexation ou outils linguistiques mis au point par les documentalistes avec le concours des utilisateurs (pour coller au vocabulaire de ceux-ci) : lexiques, listes contrôlées, thésaurus, classification qu'on regroupe sous le terme de langages documentaires. On utilisera les mêmes termes pour décrire les concepts présents dans un document (à l'entrée) et pour retrouver ce document ou l'information qu'il contient (en aval, lors de la formulation de la question). Ces *clés d'accès* sont donc des outils de communication, des passerelles entre l'utilisateur et le document puisqu'ils sont utilisés lors de la saisie dans le fonds et à la sortie (recherche).

22b. Analyse du contenu : elle consiste à condenser le contenu du document, en rédigeant un résumé documentaire qui comporte, au minimum, les principaux thèmes abordés.

La **condensation** indique en général le point de vue (par exemple, essai sociologique), les thèmes abordés dans le document. Elle complète l'indexation pour que l'utilisateur puisse juger de la pertinence du document par rapport à ce qu'il cherche.

Les résumés sont réalisés pour figurer dans les produits documentaires (banques de données bibliographiques, bulletins bibliographiques), qui sont alors dits « analytiques ».

Il existe plusieurs formats de résumés dans des variations de taille mais surtout de type d'information) :

- le résumé indicatif liste les principaux thèmes abordés dans le document, en indiquant la thèse principale de l'auteur, en les liant dans une ou plusieurs phrases

- le résumé informatif fournit tous les informations qui figurent dans le document, ce qui évite de recourir au document source : les documents statistiques, les enquêtes se prêtent bien à ce genre de résumé ; ou encore principaux faits ou éléments d'un article insérés dans une revue de presse..

- le résumé critique, enfin, indique la présence de l'auteur du résumé qui prend parti par rapport au contenu du document. Il exprime son point de vue.

23 – Catégorisation : l'identification des concepts sur un texte rend déjà possible son rapprochement avec d'autres documents puisque tous les documents indexés avec le même mot-clé « sortiront » ensemble lors de l'interrogation. L'opération de classification / catégorisation va plus loin : elle consiste à lui donner une place à côté d'autres documents. L'insérer dans une catégorie, dans une classe d'un plan de classement, dans un dossier documentaire, dans une rubrique de la revue de presse, ou encore dans une rubrique de site web ou de l'intranet documentaire permet de l'associer explicitement à d'autres documents. Cette opération demande de construire au préalable une structure classificatoire. (cf ch 4 organisation/ structuration documentaire)

3) *Stockage*

Il est de deux ordres :

31- mise en mémoire l'information secondaire résultant de l'opération d'analyse : dans une banque de données (fichiers structurés), le fait qu'elle soit organisée en champs (par exemple description, résumé, mots-clés) permettra de porter l'interrogation sur chacun d'entre eux (date, auteur, titre, etc).

32- mise en mémoire du document primaire : il sera classé, rangé en fonction d'une structure établie au préalable

- Soit dans une structure physique (rayonnages, dossiers)
- Soit dans un système de gestion électronique (GED, serveur de fichiers) sous un format choisi

4) *Diffusion*

(également appelée sortie, ou retrieval /exploitation du système) : les formes en sont multiples :

41 – simple mise à disposition (mise en place de procédures d'accès), reproduction, duplication de documents.

42 - réalisation de produits documentaires

1. d'*information courante* (pour se tenir régulièrement au courant), telle une revue de presse, un bulletin de sommaires
2. d'*information ponctuelle* (bibliographie sur un thème) ou *pérenne* (alimentation de dossiers documentaires). Cela suppose, en amont, un mode d'organisation de l'information (plan de classement) et la sélection de termes descriptifs ou clés d'accès (par ex pour faire un index)

3. de produits sur-mesure à forte valeur ajoutée : synthèses, état de l'art qui sont le résultat d'une recherche d'information sur le sujet (retraitement des informations, réorganisation et reformatage..)
4. de produits "rétrospectifs" : banques de données : mémoire du fonds documentaire qu'on exploite à la demande

43 – la diffusion peut être générale, hiérarchique, sur profil (selon les intérêts déclarés de chacun), en étoile : selon la culture de l'entreprise, les habitudes de travail de chacun.

44 - services sur mesure : ces prestations documentaires sont rendue plus ou moins régulièrement à un public ciblé dont on a sélectionné les besoins.

- sous forme de recherches/veilles ponctuelles réalisées dans différents outils, dans le fonds du centre, à l'extérieur (banques de données, via des réseaux documentaires, en faisant appel à des sources spécifiques, des personnes, des organismes)

- sous forme de conseil personnalisé : aiguiller vers la source pertinente, aider une personne dans l'organisation de sa documentation personnelle, former à l'utilisation (rendre autonome...)

Toutes ces opérations, réalisées par le prestataire, nécessitent d'utiliser en permanence des méthodes de marketing pour positionner son activité, évaluer les sources, les résultats (pertinence), la satisfaction des usagers, par des mesures de satisfaction, des méthodes de conception de nouveaux produits ou services.

Rappelons que pour l'utilisateur, tout ce back-office est "transparent" ; les contraintes internes ne doivent pas peser sur lui. Il est souvent fort étonné lorsqu'on lui présente tout le travail qui précède la mise à disposition d'un document... les journées portes ouvertes peuvent avoir du bon !!

Références bibliographiques :

L'information scientifique et technique / F Jakobiak. Paris, PUF que sais-je n° 3015, 1995
 La recherche documentaire / M Darobers, N Le Pottier. Nathan
 Les techniques documentaires /Chaumier. PUF que sais-je, 2005
 Guide pratique des techniques documentaires / C Guinchat et Y Skouri. Edicef, 1996, nouvelle éd
 Vocabulaire de la documentation, coord A. Boulogne, 2004, éd adbs
 Mesguich V., Thomas A, 2009, Net recherche : le guide pratique pour mieux trouver l'information utile, éd adbs

Analyse de l'information et analyse documentaire

Elle constitue l'un des centres du travail documentaire qui consiste à rendre accessibles et à mettre en relation des ressources avec des usages multiples plus ou moins prévisibles. Face au développement de systèmes d'information comme l'Internet, il devient primordial de pouvoir *qualifier* l'information (c'est-à-dire de fournir une information sur elle (information secondaire), pour aider le lecteur à savoir si elle lui sera pertinente ou non.

En philosophie l'analyse vise à comprendre un objet en le décomposant en ses constituants. Elle commence par établir les critères permettant d'identifier ces composants.

Tout objet peut ainsi être mis sous un regard analytique : personnes, objets matériels, photos, documents, information, données, sites, etc.

Pourquoi développer une attitude analytique ?

Le développement de multiples sources d'information, ainsi que des liens entre une information et une autre (hypertexte), rend cette opération d'autant plus nécessaire si l'on veut connaître l'origine de l'information elle-même, sa provenance, son contexte de production. Qui n'a jamais, sur un site Internet, été dans l'incapacité de savoir qui avait produit tel texte, pour quel propos, pour décider si l'on peut lui faire confiance ? c'est la validité et la fiabilité de l'information qui est ici en jeu.

Trois faits sont à garder en mémoire :

1. l'information est liée à l'action humaine. Ce n'est pas une donnée en soi, qui existerait indépendamment de celui qui l'a produite, ni indépendamment de celui qui va en faire usage. C'est bien la grande difficulté pour les spécialistes de l'information que de devoir souvent juger à priori (en fonction de ce qu'ils connaissent de leurs clients) de la valeur de l'information, de son utilité potentielle. Or les besoins de chaque individu changent dans le temps (même s'ils travaillent sur le même sujet), il ne les exprime pas toujours de la même façon (il utilisera des termes différents pour indiquer ce qu'il cherche). Autant de difficultés pour constituer un système stable susceptible cependant de s'adapter à chaque cas.

Cela signifie que l'information prend sens en fonction de l'usage qui en sera fait (intention de l'utilisateur), tout en étant définie par l'intention de celui qui l'a produite (intention du locuteur), dans un certain contexte (temps, espace) et avec une certaine intentionnalité (convaincre, expliquer, critiquer...). L'information est donc « encapsulée » dans le système de pensée de l'auteur et dans le mode de construction de sens de l'utilisateur (interprétation).

L'une des premières conséquences, pour le médiateur - documentaliste, est de devoir déceler l'intention et le mode d'expression de l'auteur tout en ayant à l'esprit les différents usages possibles de cette information (alors qu'ils ne se sont pas encore manifestés). Il est à la jonction de deux univers, d'où le terme de **médiation**.

2. Une information n'est jamais isolée. Elle se construit à l'intérieur d'un système de connaissances, elle se rattache à d'autres éléments d'information. Le rôle du documentaliste est donc de repérer à quel domaine elle appartient, et de l'associer à d'autres documents, ou à une collection* existante (par une opération de classification). Il construit ainsi des ensembles d'informations pour une exploitation ultérieure. Etablir ces liens, ces relations entre unités d'information dispersées est bien l'une des fonctions d'un documentaliste. Il y a donc construction d'un espace intermédiaire de relations entre des documents ou informations qui rassemble à posteriori les documents indexés avec le même terme. Cet assemblage se réalise également lors de la recherche, lorsque l'utilisateur retrouve tous les documents qui ont été indexés avec le même terme (collection a posteriori)

3. la difficulté de constituer ces ensembles d'informations (constitution d'un fonds documentaire, d'une banque de données) vient du fait qu'on présuppose des modes d'exploitation possibles. On fait donc un pari, d'où l'importance d'évaluer en permanence l'adéquation entre cet espace documentaire construit à priori et le mode d'exploitation réel de ses utilisateurs (c'est-à-dire leurs pratiques).

Ce travail concerne aussi bien les documentalistes que les utilisateurs

Ainsi, ce rôle central d'analyse de l'information – identification, description, caractériser - permet d'aider l'utilisateur à sélectionner celle qui lui convient.

L'opération d'analyse comme repérage des caractéristiques d'un objet est au carrefour des démarches de deux acteurs (schéma):

1) le professionnel, qui caractérise l'information en explicitant les critères descriptifs : qui a dit cela, à qui, avec quelle intention, par quel biais et sous quelle forme ; explicitation des conditions de production (CP) de l'information, de son accessibilité (CCOM), de sa forme et de son contenu

2) le demandeur, qui, lui, sait plus ou moins ce qu'il recherche, pour quel type d'action (recherche, action, vérification...), qu'il formule avec des phrases du type : je veux ..., pour faire Il a donc lui aussi un travail à faire pour clarifier ses propres critères d'évaluation de la pertinence de l'information. Il peut chercher à savoir tout ce qui a été écrit sur une question, vouloir actualiser ou valider une information, comparer des situations, apprendre sur un sujet mal connu, décider de quelque chose... Les situations sont variées et chacune est spécifique. C'est en mettant en relation les résultats de ces deux démarches que pourra se réaliser l'adéquation entre la source et la demande.

L'utilisateur :

a) repère les éléments caractéristiques d'un objet

Le paratexte qui entoure un objet constitue autant de caractéristiques en facilitant l'accès et le retrouver. Ces *métadonnées* prennent la forme

- de mots clés, issues d'un vocabulaire contrôlé
- d'une insertion dans une catégorie ou une classification.

C'est évoquer les relations qui existent entre des objets :

- l'intertextualité : relations entre plusieurs textes, en donnant le même mot-clé à plusieurs documents qui traitent du même sujet. Cela les fera « sortir » ensemble à l'interrogation
- l'intra-textualité : à l'intérieur du document lui-même, (par exemple avec un index)

- l'inter modalité : insertion de codes liés à des genres de texte.

Le lecteur navigue donc dans une organisation qui a été structurée pour l'aider. Il lui faut les repérer pour s'en servir (c'est une aide et non une contrainte).

b) repérer la structure d'un document (et la place de l'information)

Le fait que l'information ait été insérée dans une structure permet de s'y retrouver à l'intérieur d'un site, d'un document, d'une banque de données, etc. Il existe deux grandes familles de **structuration** :

- La première ne résulte d'aucune analyse conceptuelle du contenu, elle ne fait qu'organiser la PRESENTATION des informations, (forme). par exemple :

a) un **sommaire** reprend les titres des chapitres et pointe vers un endroit du texte : moyen d'accès qui agit comme un « menu ». C'est un facteur de **localisation**.

C'est également le cas de nombreux liens hyper textuels (ainsi, en cliquant sur un titre dans une liste on accède directement au texte de ce document). La liste constitue une étape intermédiaire (niveau 1 d'une référence bibliographique) avant de parvenir à l'information finale (niveau 2, celui du texte intégral)

b) Un **répertoire** organise une masse d'informations *selon un critère* défini à l'avance, en général thématique ou géographique, ce qui suppose que le lecteur sache où rechercher l'information (par ville, par thème). En général on y ajoute un second mode d'accès, par exemple un index alphabétique (par auteur, par concept, nom, etc) qui permet de pointer directement vers l'endroit du répertoire où ce terme est mentionné. Tous les annuaires sont bâtis sur ce modèle

- Le second type de structuration résulte d'un travail conceptuel

Un cadre d'ensemble a été constitué au préalable, cette organisation intellectuelle apportant du sens. Par exemple :

- arborescence d'un site web, classification spécialisée, plan de classement de dossiers ou de bulletins, ils procèdent d'une mise en catégories préalable qui permettra d'insérer ensuite l'information sous la rubrique la plus adaptée (cela suppose un repérage du contenu et de sa signification). La recherche d'information se fait en avançant pas à pas dans ces catégories, dans une démarche qui part du plus général pour aller vers le plus particulier. (sur le web de tels outils se nomment annuaires ou répertoires). Il s'agit d'une logique classificatoire.

- Index à la fin d'un ouvrage : il se construit en sélectionnant les concepts ou les thèmes abordés dans les pages de l'ouvrage. (cf la revue « the indexer »)

- organisation en **banque de données** dite aussi **information structurée** : les informations, données ou documents ont été préalablement identifiés, sélectionnés puis découpés puis organisés :

- selon des critères descriptifs de forme (éléments de données de description bibliographique : titre, auteur, éditeur, date)

- selon des concepts traduits par des TERMES (opération d'indexation)
Ensuite, l'information est insérée dans des CHAMPS (titre, auteur, résumé, mot-clé) qui serviront de critères à l'interrogation. (on interrogera le champ qui convient le mieux)

Des *formulaires*, ou masques (de saisie ; de recherche ; de visualisation) servent à l'interaction homme-machine. Le format de visualisation présentant le « résultat de la recherche » laisse voir cette structuration faite en amont : on y voit le titre, le nom de l'auteur, les mots clés, par exemple. Les moteurs de recherche fonctionnent sur une indexation effectuée sur chaque document puis mise en mémoire dans l'index d'une banque de données.

Prendre conscience de ces deux types de structuration permet de remarquer qu'un **lien** peut avoir plusieurs fonctions :

- soit il renvoie à un objet et n'est qu'un simple pointeur technique (un mot donne accès au même mot à plusieurs endroits d'un texte),
- soit il résulte d'une analyse sémantique préalable qui tiendra compte des synonymies ou des formes grammaticales qui ont le même sens).

En effet, aux modes d'accès / d'usage correspondent deux modes de structuration

- **interrogation par menu** ou par arborescence : on se rapproche progressivement de l'information recherchée, en suivant la classification, le plan de classement, par des choix successifs. C'est le cas de l'interrogation par menu ou par catégorie. Utile quand on ne sait pas exactement ce qu'on cherche ou qu'on veut se donner des idées.
 - En amont, le spécialiste organise donc un classement
- **interrogation par mots clés** (accès direct) qui donne immédiatement un résultat. Utile quand on effectue couramment une recherche et qu'on sait ce qu'on cherche. En général on saisit une requête dans une boîte de dialogue (formulaire)
 - en amont, le spécialiste sélectionne les termes qui permettront l'accès

La navigation est devenue de plus en plus complexe, entre textes électroniques et entre genres différents. L'utilisateur doit connaître la « langue des textes », établir des critères de choix entre outils, entre services ; ce sont autant de compétences à développer sur les fonctionnalités qu'ils offrent (par exemple connaître le langage de commande d'un serveur).

Ces préalables expliquent que l'opération de caractérisation de l'information soit fondamentale, puisqu'elle permet d'évaluer l'utilité, la pertinence d'un contenu, et sert aussi à choisir l'exploitation qui peut en être faite (diffusion, stockage, élimination...). Elle permet de sélectionner et de choisir le mode de traitement le plus approprié (pour un texte général, pas de résumé, mais des termes d'indexation ; pour un document portant sur un sujet nouveau, résumé et indexation ; décider de l'information à faire circuler avant de jeter le document...)

Le travail réalisé par le spécialiste

En amont, le travail consiste à repérer et à décrire ces différents plans, formel et conceptuel, et à faire des liens avec d'autres documents (inter-textualité) et éventuellement insérer l'objet dans une structure.

Ce travail s'effectue en général dès l'arrivée d'un document, ou dès le repérage d'une nouvelle source d'information : il s'agit d'une démarche « a priori ». C'est bien parce que les documentalistes ont repéré et évalué des sources d'information possibles, que les utilisateurs pourront les activer au bon moment et rapidement. Les spécialistes de l'information évaluent ainsi aussi bien des documents que des sources potentielles.

Ces opérations peuvent se voir comme des traductions successives centrées sur le référentiel métier le plus probable de l'utilisateur potentiel.

Concrètement, la démarche procède de l'extérieur vers l'interne (c'est-à-dire du contenant vers le contenu).

Démarche générale

Choisir son objet (un support, un contenu, une personne, une situation, un dispositif, un outil....)

1. Observer, identifier
2. Décrire / nommer selon des critères d'analyse (mise en place des mots pour le dire)
3. Transcrire / Représenter selon différents formats (schématisation, condensation)
4. Organiser et structurer : ajouter des liens, insertion dans des familles, des classes

Avec l'aide d'outils intellectuels, langages documentaires, vocabulaires contrôlés voir le cours sur les langages, les classifications et plans de classement)

1. identifier l'objet à étudier : s'agit-il

- D'un document (primaire : une revue, un article, ou secondaire comme un bulletin bibliographique, un outil comme une banque de données), image, multimédia ;
- D'une personne (pour dresser un fichier d'expert) ;
- D'un organisme (source d'information).

2. Décrire les différents plans à l'aide de plusieurs familles de critères :

a) critères descriptifs externes : ces éléments "extérieurs" éclairent le document avant même d'en lire le contenu ; ils donnent un statut à l'information, en la positionnant dans un ensemble.

- **conditions de production (CP)**, comme par exemple, revue généraliste ou spécialisée, contexte et cadre de production (à l'occasion d'un congrès), appartenance de l'auteur à un organisme, une discipline. Ensuite, en identifiant les « marques identitaires » et des intentions, en repérant la façon dont les concepteurs de site (mais aussi de cédéroms) se rendent présents sur le site : producteurs (non seulement le webmestre, mais aussi les auteurs), par l'anonymat ou la signature. Repérage de la façon dont ils s'adressent au « lecteur », leur neutralité, ou l'exhortation directe (« vous pouvez », « faites ») et à quelle fin (pour l'aider, l'orienter).

- **conditions de communication (CCOM)** : la façon dont cette information ou ce document est rendu accessible (mode et circuits de diffusion, facilité d'accès, coût...)

Cela suppose, ici, de repérer les 4 principaux circuits de communication : contacts interpersonnels informels, colloques, littérature souterraine, édition. Chacun suppose un type de validation, un type de démarche pour obtenir l'information, un coût, autant de critères qu'il importe d'identifier parce qu'ils pèsent sur l'identité du document. (schéma circuits / canaux éditoriaux)

- **caractéristiques formelles** : ces critères se basent sur l'observation du document (du contenant, et pas encore de son contenu). Par delà le support, il s'agit de repérer le type de document (thèse, article, exposé). Ces éléments de données sont trouvés en manipulant le document (sur la pochette du cédérom, page de titre, comité de rédaction, quatrième de couverture, etc)

b) Critères descriptifs internes du contenu : repérage de l'organisation du texte, description du contenu, en séparant les faits de l'interprétation. Lorsqu'ils travaillent sur le contenu, les documentalistes s'appuient sur le texte lui-même, lorsqu'ils en font par exemple un résumé, mais s'appuient surtout sur les éléments de paratexte qui connectent ce document et son contenu à d'autres unités d'information ou d'autres documents. C'est la raison d'être de l'indexation, qui consiste à identifier les principaux concepts d'un texte et à choisir le ou les termes qui vont l'insérer dans un ensemble constitué d'autres documents décrits eux aussi avec ce même terme. L'indexation peut être entrevue comme des termes « passeurs » qui assurent l'accès à la classe ainsi identifiée par ce mot.

. **Transcrire ces données relatives au contenu**

Nb ne sont pas évoquées ici les données de codage concernant la structure d'un document (XML)

La transcription se réalise à l'aide de :

a) **normes de description et de représentation** d'un texte ou d'un document. Citons-en les principaux qui font l'objet de normes internationales et qui permettent les échanges de données :

- normes ISBD et notice bibliographique
- Langages balisés pour indiquer qu'il s'agit d'un titre, d'un résumé...
- champs structurés d'une banque de données : ces éléments sont intégrés dans des champs sur lesquels portera la recherche
- grilles d'analyse, sur lesquelles sont regroupées toutes les caractéristiques externes et internes utiles sur le même objet
- réalisation de produits documentaires (synthèses, fiches techniques, cartes)

b) outils intellectuels :

- ou langages documentaires : ce sont des outils de description, de traduction en une forme condensée signifiante. Qu'il s'agisse d'une simple liste / lexique de mots clé (liste d'autorité), d'un thésaurus ou d'une ontologie* tous visent à fournir un cadre commun tant pour l'entrée que pour la sortie / recherche. Ils réduisent l'hétérogénéité des façons de faire personnelles en les rapportant à des types c'est-à-dire à des rapprochements.

Ces « informations sur l'information » (ou information secondaire) vont permettre une gestion facilitée de l'information ou d'un document en réduisant sa taille en le représentant dans un format réduit, et traitable par une machine.

. Représenter

Outre cette analyse documentaire, le spécialiste de l'information réalise un travail permanent de repérage et d'évaluation des objets qui l'entourent :

- ses propres outils de travail,
- ses processus ou opération
- le secteur d'activité dans lequel il évolue,
- les personnes / experts / organismes auxquelles il peut faire appel.

Un certain nombre d'outils l'aident à représenter de façon synthétique concepts, idées, domaine....

a) représenter de façon graphique : schéma et carte

La construction d'un schéma ou d'une carte utilise la capacité à se ré-présenter une situation ou un lieu (topographie). L'opération consiste à représenter (positionner) plusieurs informations / objets sur un espace plan avec des effets de proximité afin de pouvoir structurer ses connaissances, les mémoriser. Cette démarche est précieuse dans un contexte d'aide à la recherche d'information, d'analyse des objectifs, d'aide à la décision. Contexte où l'information attendue doit être synthétique, directement et rapidement échangeable et exploitable, savoir réaliser schémas ou cartes est un impératif pour les spécialistes de l'information.

Ce travail de construction offre :

- une vision synthétique

Parce qu'elle présente un maximum d'informations sous un format réduit, visible, d'une façon structurée sur un espace réduit, elle convient bien aux intelligences intuitives/graphiques, moins aux intelligences analytiques/textuelles. Incontestablement la représentation cartographique suscite la créativité et augmente le champ des "possibles". La vision globale est propre au cerveau droit, le gauche étant consacré à une approche analytique

- vision contextualisante

En rapprochant visuellement des thèmes dont la proximité de sens ne transparaît pas dans la masse des données brutes, la cartographie permet de faire émerger des familles contextuelles.

Elle **met en évidence les relations** entre les points. Elle répond au vieil adage selon lequel "un bon dessin vaut mieux qu'un long discours".

L'effort initial de structuration n'est pas négligeable. Mais vouloir rendre l'information accessible à d'autres exige de structurer sa propre pensée pour la diffuser efficacement. Cartographier c'est faire émerger le sens d'une masse d'information

Trois phases de construction :

- Identifier (reconnaître et sélectionner les éléments les plus importants)
- Les décrire à l'aide de typologies (mises en catégories) selon des critères que l'on a définis
- *construire* : schématiser (réduire aux éléments signifiants) pour faire ressortir les grandes lignes significatives (c'est à dire qui font sens) = modéliser = traduire une réalité et la représenter à l'aide d'un schéma (qui servira ensuite de modèle explicatif dans d'autres situations).
- Une carte se présente donc sous forme de « patates », dans lesquelles sont regroupés les acteurs. Des traits ou flèches matérialisent les relations (une légende est donc nécessaire pour identifier ces relations, par des couleurs ou l'épaisseur des traits).

Différents formats de représentation graphique :

1. le tableau : ses doubles entrées permettent déjà une première classification. En particulier les matrices vides et pleines du tableau (selon que les intersections lignes/colonnes sont remplies ou non) procurent une visualisation immédiate
2. Le sociogramme
3. La carte conceptuelle (graphes en bibliométrie, schémas fléchés des thésaurus, cartes de connaissances...) : indique une position d'un concept par rapport aux autres et par rapport à l'ensemble
4. Un ensemble de concepts (les nœuds) sont représentés par un certain nombre de formes géométriques, et sont associés par des relations sémantiques entre concepts, formant une structure hiérarchique.
5. La carte heuristique : représentation des idées (pour préparer un plan d'intervention, un texte...

La mise en relief des relations permet une vision à la fois globale et particulière. Cela permet de saisir les similarités et les différences.

Les logiciels de représentation par cartes permettent de gérer la complexité, rendent visibles des données grâce aux connexions visibles sur un même espace. exemple : préparation d'une réunion, de son déroulement et de son compte rendu. support visuel qui sert de référentiel pour exécuter l'ensemble de ces activités.

Ils peuvent faciliter l'expression et le partage des idées et des données entre les collaborateurs de l'entreprise. Il permet aux équipes de se synchroniser en mettant en phase leurs plans d'actions, en les rendant plus clairs et partageables. En outre la présentation visuelle facilite la compréhension des stratégies et leur mémorisation.

Un certain nombre de ces outils sont disponibles dans les salles informatiques

Le classement des résultats d'une recherche, pour des besoins de veille (cf Exalead)
Applications : besoins de veille, d'identification automatique des thèmes et des "signaux faibles". Les thèmes sont classés par ordre d'importance décroissante.
sous forme de carte représentant des thèmes clés

. Organiser

C'est-à-dire positionner l'objet dans un cadre pré-établi, comme une classification, un plan de classement, ontologie (outil conceptuel) qui permettra de naviguer dans les catégories et rubriques.

Plusieurs conclusions :

De l'utilité de se doter de grilles d'analyse

L'une des compétences du spécialiste de l'information, est de suivre une méthode d'analyse formalisée. Il se dote généralement de grilles, organisées selon des critères pertinents qui diffèrent en fonction de l'objet analysé et surtout du contexte d'utilisation.

Ces grilles sont précieuses car elles permettent de décider rapidement d'un achat ou d'un abonnement, de choisir l'expert à interroger : elles constituent, certes, un investissement, qui fera ensuite de gagner du temps, lorsqu'on n'a ni le document, la personne ni le site sous les yeux. Ce sont des « images de l'objet ».

Elles servent par exemple à comparer des objets semblables afin d'évaluer leur utilité ou leur pertinence les uns par rapport aux autres.

Toutes mettent en œuvre des capacités d'observation, pour dégager rapidement des critères distinctifs. Elles seront orientées différemment selon qu'elles constituent un outil personnel (rôle d'aide-mémoire ou d'aide à la décision) ou collectif (guide pour d'autres, telle une fiche de lecture ou une fiche sur un site). Dans le travail quotidien, il s'agira d'une grille mentale, sorte de check-list qui permet de repérer très rapidement les éléments significatifs de l'objet.

Typologie

- grilles de sources d'information : personnes, organismes producteurs d'informations. Un fichier d'experts, de contacts, est organisé selon ces caractéristiques
- grilles d'outils (logiciels, répertoires, outils de recherche, etc) permettent de s'en servir facilement ou de choisir celui qui est le plus adapté au besoin du moment
- grilles de manifestations, ou de congrès, compte-rendus, qui mémorisent les points essentiels et faire circuler l'information, etc.
- grilles de lecture

Composition

Toute grille comporte 4 parties :

- description externe ;
- description interne (le contenu) ;
- évaluation
- exploitation possible.

Selon l'objet, les critères diffèrent.

Éléments constitutifs d'une grille d'analyse générique

(à moduler selon les besoins et le type d'objet sur laquelle elle porte)

I – identification / description de l'objet

- conditions de production = CP
- conditions de communication / accessibilité / diffusion = CCOM
- FORME : description physique (caractéristiques du contenant)

II - Contenu

- description des éléments caractéristiques (points essentiels) ; faits, idées
- mots-clés de repérage, permettant l'accès, la mémorisation / la recherche
- condensation (résumé) : rappel des principaux éléments

III - Evaluation

en fonction du référentiel et du contexte d'action envisagé : valeur de l'information (qualité, coût), éléments critiques.

IV - Exploitation en fonction des usages possibles

a – utilité

pour qui, quand (durée de vie)

pour quoi faire

b - usabilité = facilité d'usage

c – exploitation (point de vue gestionnaire) : à intégrer dans un ensemble d'informations du même ordre, à stocker, garder trace, diffuser puis jeter ...

Grille d'analyse d'un DOCUMENT

I - Description

CP : conditions de production
auteur (statut, origine, secteur d'activités)
cadre de production (à quelle occasion)
prévu pour quel public

C COM : canaux de distribution ou d'accès à ce document
origine :

- directe (contacts personnels) ==> littérature souterraine
- réunions => compte-rendu
- via des salons..
- dans des revues spécialisées
- édition

facilité d'accès, coût d'acquisition

II -analyse formelle + support

cf typologie des documents = forme matérielle
structure et organisation : titre, sommaire, glossaire, bibliographie
partie d'un élément

→ résultat : notice ou référence bibliographique

III – analyse du contenu

domaine, thèmes traités
niveau de traitement : information brute, digérée (info primaire, secondaire, tertiaire)
typologie des données (faits, chiffres, rumeurs, idées, tendances == > déjà validées ou à valider
= permet de mesurer le niveau de confiance à accorder au document

IV- Evaluation

niveau : technique, vulgarisation..
type de production observations, critiques
information courante, information de fond
utilité : quel usage ? utile pour faire quoi (recherche, se tenir au courant)
fiabilité de l'information (confiance) => validation
durée de vie / décalage entre les dates, mise à jour, etc
usabilité (conditions pour exploiter correctement : exemple maîtrise du langage d'interrogation pour une banque de données)

IV - Exploitation

suite à donner : réflexion, intuitions à vérifier,
garder ou jeter
traiter : indexer, résumer, insérer dans le fonds
à diffuser (sous quelle forme)

Grille d'analyse d'une PERSONNE

Cas d'un fichier d'experts

I – Identification / description

- identité
- coordonnées de l'organisme, secteur d'activités

CP

- statut :
- titre : (ex : directeur de laboratoire)
- fonction : (ex : chercheur, fournisseur, client, partenaire, contact...)
- lieu d'exercice (service)
- typologie (observateur, acteur, expert...)

origine du contact (par une personne, une revue, un colloque....)

C COM

- accessibilité (disponibilité)
- langue
- aire d'intervention (géographique, réseau personnel)

FORME

statut de l'information produite (confidentielle / publique, contrôlée ou à contrôler)

II – Contenu

- de quelles informations cette personne est elle porteuse ?
- ses compétences, expertise (personnelle, en fonction de son secteur d'activités)
- suivi de la relation (son historique)
- réunions (objet, résultats)
- son réseau (ce qui permet d'élargir ses propres sources)

III - Évaluation

= valeur d'usage

notoriété, POIDS

degré de confiance à accorder

IV - usage

utile pour quoi, dans quel type de situation

A intégrer dans un fichier de contacts, à suivre régulièrement, ponctuellement, à rencontrer pour avis, conseil, etc

Se rappeler que les technologies intellectuelles sont constitutives du travail documentaire

Les méthodes qui viennent d'être évoquées appartiennent aux « technologies intellectuelles » constitutives de l'esprit humain, c'est-à-dire aux opérations cognitives : penser, discriminer, calculer, faire des liens, classer (cf Goody, Lévy)

Elles sont technologiques au sens où ce sont des outils pour l'homme. Leur instrumentation par les technologies (informatiques) amplifie leur portée : traitements statistiques, recherche rétrospective, s'inscrivent effectivement dans cet élargissement des possibilités humaines. L'écriture par exemple, a permis une mise en mémoire, et une circulation dans le temps et l'espace.

C'est pourquoi il importe d'en revenir à ces méthodes intellectuelles (dont il sera question en début de formation) comme des « fondamentaux » qui sont au cœur de tout dispositif documentaire.

La lecture engage à la fois une gestuelle (attitudes physiques, gestes pour manipuler, prendre des notes, éventuellement se déplacer), une activité cognitive de reconnaissance d'un texte avant lecture, son repérage,

L'activité de lecture est une activité active, d'interaction avec l'auteur, une construction de la part du lecteur. Une interaction entre une surface d'inscription et un lecteur

Concernant la lecture elle-même, plusieurs stratégies sont possibles : le balayage engage un repérage rapide pour y revenir ultérieurement : survol des contenus, repérage de l'organisation générale (place du texte dans un ensemble, ou d'unités d'information dans un texte). il y a donc une attention à l'organisation et à la démarche mise en œuvre. Deuxième type de lecture, la focalisation, qui engage une attention finalisée au contenu recherché (l'organisation des textes, les clés d'accès, ne sont qu'un moyen d'y parvenir).

*Types de lecture : survol, découverte, repérage, vérification, lecture linéaire ou sélective...
Lecture critique, rapide...*

Références bibliographiques :

L'analyse documentaire / Suzanne Waller. Adbs, 1999

Sur la représentation graphique :

Pour en savoir plus (et visualiser des exemples de ces concepts)

<http://www.mmdfrance.fr/mindmanager/>

Déchargement d'une version de démo valable 21 jours

Lectures complémentaires :

Pour en savoir plus <http://www.grimmersoft.com>

<http://www.knowledgeboard.com/download/3114/4-Cartographie-conceptuelle-ABD.pdf>

2 cours de l'université de Laval : La cartographie conceptuelle

« Aspects théoriques et pratiques »

<http://www.cours.fse.ulaval.ca/did-62344/module2/demarche/mod2b/doc/ref12b.html>

<http://www.cours.fse.ulaval.ca/did-62344/module2/ref12b.html>

Les acteurs du secteur de l'information

Ils sont fort nombreux, qu'ils produisent des matériels, des logiciels, assurent de la maintenance, des prestations diverses utiles à un spécialiste de l'information ou encore soient des observateurs ou des donneurs d'ordre (subventions, appels d'offres, politique publique).

Marché de l'information

On parle de marché de l'information devant l'ampleur du phénomène d'industrialisation, qui possède les caractéristiques suivantes :

- une offre conséquente
- demande solvable ;
- des usages avérés, repérés ;
- des règles de fonctionnement
- des règles économiques ;
- des acteurs identifiés que l'on peut suivre, qui ont des stratégies.

En effet, il y a bien un marché de l'information, car l'information, bien immatériel et non matériel, est devenue un bien marchand comme un autre, qui se vend, s'achète, s'évalue (on parle aussi de marchandisation de l'information). Produite en grande quantité, reproduite, dupliquée sur de nombreux supports, etc.

Ce secteur d'activités se compose de plusieurs types d'acteurs :

- Les producteurs qui conçoivent et proposent des produits ;
- Les politiques qui orientent le secteur et ceux qui observent (observatoires)
- Les intermédiaires, (dont les documentalistes) prestataires qui assurent des prestations d'information (les professionnels de l'information)
- L'utilisateur final : les clients / utilisateurs / usagers, qui sont soumis à des règles économiques (achat, contrat, abonnement, etc)

Avec un cycle d'innovation / production...

On parle de cycle d'innovation / production pour signaler les phase allant de la conception à l'usage, n a été activé. Avant d'être mis sur le marché (c'est-à-dire en vente),

- le produit est conçu (c'est la phase recherche-développement),
- puis testé et expérimenté,
- mis en vente ;

Lorsqu'il a beaucoup de clients, le produit tend à se banaliser. Il est donc important de repérer dans quelle phase se situe le produit étudié (logiciel par exemple) : phase de croissance (nouveau et recherche de clients encore peu nombreux) ; phase de maturité - stabilisation ; phase de déclin (moins de clients, perte de profit) en le comparant à des produits proches ou concurrents.

Les économistes désignent par « courbe en S » ce cycle - montée en puissance, apogée et décroissance - d'un produit ou d'une société commerciale. Elle permet d'expliquer pourquoi

les producteurs tendent à réduire les délais de mise sur le marché (par exemple des versions de logiciels) pour que leur courbe de croissance reste haute. Ils sortent un nouveau produit alors même que l'ancien est encore en maturité, sans attendre sa décroissance, pour continuer à faire du profit ou pour rester au même niveau que la concurrence.

Très concrètement, suivre le marché de l'information demande plusieurs opérations :

- identifier les acteurs en leur associant leurs produits (leurs modes de communication, de commercialisation, etc),
- les regrouper par type d'acteur, par exemple les éditeurs (c'est l'opération de catégorisation) en y associant leurs produits et surtout les relations qui lient chaque acteur avec d'autres.
- identifier leurs relations (concurrence, partenariat économique, institutionnel..)
- construire une représentation graphique (opération de cartographie (cf ch 4).

Cette carte se construit progressivement en fonction des informations rencontrées, ce qui permet d'avoir une vue d'ensemble du secteur. Elle s'enrichit lors de lecture de la presse spécialisée, de rencontre, en ajoutant un acteur quand un nouvel entrant se manifeste, et en établissant des liens entre acteurs (par exemple quand il y a rachat) ; il s'agit donc d'avoir une vision synthétique et dynamique du secteur.

Etablir la carte d'un secteur d'activité met en oeuvre trois notions essentielles :

Acteur : il agit (a une certaine activité) depuis une PLACE (par rapport à d'autres) ; il joue un certain RÔLE (d'observateur, de producteur, de décideur politique) ; il est soumis (ou soumet les autres) à des pressions (économiques, politiques) : il est donc dépendant, ou au contraire contrôle.

Catégorisation* : (mise en classe / groupe) : établir une typologie c'est regrouper, organiser autour de spécificités. C'est la notion d'ensemble fermé qui possède des frontières (étanches ou perméables) ; cela inclut une définition des éléments communs (caractéristiques) qui peuvent être des comportements, des produits, des rôles (plus que des statuts) ; en effet, une catégorie regroupe des objets possédant les mêmes caractéristiques.

Il importe également d'utiliser la typologie de l'analyse stratégique () :

L'analyse de Porter invite à prendre en compte les acteurs suivants

Client, fournisseur, concurrents, nouveaux entrants

C'est ainsi qu'il y a des acteurs traditionnels, institutionnels présents depuis longtemps, et de « nouveaux entrants » qui peuvent venir modifier l'équilibre du moment : concurrents, ou innovations (techniques ou sociales) viennent peser sur la scène.

Relations : les liens qui existent entre un acteur et un autre. Elles découlent de questions du genre : qui travaille avec qui (collaboration) ; concurrence, complémentarité. Ces relations peuvent être économiques, hiérarchiques ... Il convient aussi de préciser ce qui s'échange (des documents, un flux économique), le volume de ces relations (leur importance).

I – Acteurs du marché :

La typologie suivante est succincte mais elle permet à chacun de poser le cadre général et de construire une carte précise qui sera alimentée progressivement par ajout du nom des acteurs rencontrés ou mentionnés, ainsi que le produit proposé par chacun.

a) Acteurs liés à l'informatique :

Ce sont les fabricants informatiques et bureautiques. (eux-mêmes liés aux fabricants d'électronique (processeurs) qui produisent :

- des matériels (hard) : ordinateurs de calcul (supercalculateurs, ordinateurs neuroniques, après le développement de la micro-informatique depuis 1973), périphériques (scanners (numériseurs), imprimantes, lecteurs externes (de CD, disques durs), ou encore stations de travail (de GED, de Conception CAO pour le dessin industriel), etc ou encore de cartes de communication rendant un ordinateur communicant.
- les logiciels (soft) : logiciels intégrés, logiciels documentaires, hypertexte, de communication, logiciels d'analyse de données, de téléchargement, de reformatage, de PAO.
- les sociétés d'ingénierie informatique (SSII) et installateurs de systèmes
- les distributeurs, diffuseurs ou assembleurs de ces matériels et logiciels (commerciaux et techniciens)

Evolution : un phénomène d'intégration progressive des différentes fonctions (machines multitâches ou multifonctions), comme par exemple la GED où l'on traite toute la chaîne de commande, d'enregistrement, de numérisation, de stockage en machine, de duplication et de diffusion.

b) les opérateurs de réseaux

Ils installent l'infrastructure physique (câblage de fils de cuivre, de fibres optiques)

Il convient de distinguer les réseaux informatiques des réseaux de télécommunication, en distinguant le support et les procédures de transmission (cf chapitre Réseaux)

- diversification des opérateurs de réseaux nationaux (l'opérateur historique France telecom en charge des infrastructures, produits Transpac, Numéris, ADSL)
- les fabricants de réseaux locaux LAN ou RLE (ethernet,) qui sont souvent des constructeurs informatiques (IBM par exemple) ou de réseaux étendus (WAN)
- les industriels des satellites, de la télédistribution et du câble,
- les instances qui édictent les protocoles de communication (techniques et normatifs), au niveau national et international, comme le CCITT
- les sociétés qui donnent accès aux réseaux (fournisseurs d'accès comme, wanadoo, cybercable, etc)

c) les industriels du contenu :

ils conçoivent le produit d'information : ils l'organisent, la présentent, conçoivent les logiciels d'accès. Ils travaillent souvent de façon artisanale mais la tendance est à la concentration et à la fabrication industrielle.

Ne pas oublier ceux qui fabriquent matériellement le produit (sociétés de pressage de cédéroms par exemple)

Plusieurs secteurs peuvent être distingués :

- Édition (livre et information électronique) : diversification progressive pour décliner leurs produits de base sur de nouveaux supports (les premiers cédéroms datent de 1985). Souvent les acteurs traditionnels se sont diversifiés vers le multimédia (encyclopédies Hachette, dictionnaires). Les « métiers du livre » regroupent les éditeurs, les libraires. Il y a actuellement cohabitation du marché traditionnel avec des nouveaux venus qui ne sont pas éditeurs de métier.

- Information en ligne : historiquement, plusieurs phases à garder à l'esprit (cf chapitre « repères historiques »)

1965 -> 1979 : banques de données internes à une entreprise

1979 à l'ouverture du réseau national Transpac, mise en ligne de ces banques de données (en distinguant les banques bibliographiques, factuelles, en texte intégral)

1985 : ouverture du kiosque vidéotex : paiement par palier

mouvement vers le cédérom (620 titres français en 1995) pendant longtemps appelé « hors-ligne »)

Ce secteur s'est intitulé Télématique (néologisme créé en 1978 par Nora et Minc, traduisant le mariage entre l'informatique et les télécommunications),

Il faut distinguer différents types d'acteurs : les producteurs, les serveurs, les opérateurs de réseaux, les intermédiaires commerciaux (brokers, courtiers).

Le terme de Serveur reste ambigu : il désigne soit une machine qui abrite la banque de données, soit les sociétés intermédiaires qui assurent sa commercialisation (par exemple Questel). Le terme de producteur se réfère à ceux qui définissent le contenu, les outils de structuration et de navigation, ou encore les logiciels de traitement (analyse statistique des données)

Des associations regroupent ces types d'acteurs : par exemple l'association des producteurs, (GFII : groupement de fournisseurs d'information en ligne), des consommateurs (clubs d'utilisateurs), l'AFTEL (association française pour la télématique).

m

- Les observateurs de ce secteur sont principalement l'observatoire IMO (information market observatory) de l'Union européenne. Il lance régulièrement des études statistiques sur ce secteur (avec le concours, en France, du GFII et du Ministère de l'Industrie)

d) Prestataires de service :

- Sociétés d'audit, de conseil en organisation ou aussi prestataires de services documentaires sous-traitants pour un service de documentation (réalisation de produits documentaires, ou d'outils comme les thésaurus)

- Intermédiaires commerciaux (courtiers, brokers), de prestataires de services (lié phénomène sous-traitance des entreprises) SVP, innovation 128...

e) Fournisseurs de documents primaires

centrales d'achat, abonnements, agrégateurs..

f) Métiers de la mise en forme

(graphisme, imprimerie), sociétés de routage

g) organismes de formation

privés et publics, diplômants ou non,

h) recherche (publique, privée) (cf chapitre Production scientifique)

Tous ces acteurs sont des partenaires pour le documentaliste ; d'où la nécessité de suivre les évolutions (lors des congrès, colloques, organisés par des associations, sociétés savantes, universités, industriels, etc)

Cette complexité grandissante des relations entre ces acteurs rend nécessaire de développer des critères de comparaison, d'évaluation, pour choisir : grille d'analyse comparative (dans les revues spécialisées et leurs tests)

II - Les professionnels de l'information

Leur typologie diffère selon le point de vue adopté : elle se fait selon la fonction assurée dans le processus de transfert de l'information.

a) Leurs appellations sont très nombreuses.

- en amont : du côté de la production et du traitement :

production de services ou de produits d'information (B d D, Cédérom), webmestre, gestionnaire de site..

- du côté du stockage archivage : conservation, mémoire de l'entreprise, patrimoine : archivistes, records managers

- du côté de l'accès (usager)

services de renseignement ou d'aiguillage

- mise à disposition : Bibliothèques Municipales, B Univ, BCP (centrale de prêt)

- prestataire de services (ex d'études bibliométriques, de reprise de fonds, de produits documentaires ciblés, de revues de presse électroniques...)

- du côté du retraitement (valeur ajoutée) : groupe d'experts, veilleurs

- assurant un rôle d'organisation et de management : directeur des ressources documentaires, Management de l'information, consultants internes ou externes, knowledge managers, Records managers.

Il existe désormais une certification des personnes organisée par l'ADBS, qui a mis au point un "référentiel des métiers et des emplois-type", et a fortement contribué à formaliser un "euro-référentiel". 4 niveaux de compétences ont été retenus : assistant, technicien, ingénieur (budget, équipe), expert. Un référentiel des métiers et des emplois-type permet à chacun de se situer.

b) typologie des unités documentaires

- bibliothèques :

BNF : bibliothèque nationale

BU et BIU (bibliothèque inter-universitaires au niveau universitaire (rattachement enseignement supérieur)

BM bibliothèques municipales et BDP (bibliothèques départementales de prêt, gérant les bibliobus et mettant régulièrement en dépôt des ouvrages dans des bibliothèques municipales de petites villes)

bibliothèques privées, de musée..

niveau de diplômes : conservateurs, bibliothécaires (organismes de formation (ENSSIB et IFB))

- Archives : nationales, départementales (Ecole des chartes), archives d'entreprise (Tourcoing)

- Centres de documentation
peuvent être :

- pluridisciplinaires, nationaux (INIST) ou internationaux
- spécialisés (sectoriels), nationaux ou internationaux
- internes à une institution (administration, association, entreprise publique ou privée, gros instituts de recherche)
- appartenir à des réseaux documentaires

Répertoire des centres de documentation (ADBS)

c) Les associations professionnelles :

elles facilitent les contacts, les discussions sur le métier, organisent Congrès, journées d'études et de formation, éditent souvent des bulletins ou des études..

- pour les bibliothécaires : ABF (bulletin de l'ABF) ; bulletin des bibliothèques de France, congrès annuel
- pour les Documentalistes : ADBS (Documentaliste, congrès IDT), APRONET (documentalistes des collectivités territoriales <http://www.apronet.info/>), FADBEN (documentalistes de l'éducation nationale)
- pour les spécialistes de la veille : SCIP : association des professionnels de l'intelligence économique
- pour les spécialistes du Knowledge management,
- pour les prestataires et consultants en information : SYNAPI : syndicat des www.synapi.com/

Au plan international,

- IFLA (international federation of library associations www.ifla.org), ASLIB www.aslib.com FID (fédération internationale de documentation). Au plan européen, Eusidic (European Association of Information Services)
- USA : ASIS (association for scientific information science), ALA (american library association)

A noter que dans le monde anglo-saxon, le terme de documentaliste n'existe pas, on parle de bibliothécaire spécialisé (special librarian) (cf chapitre « politiques publiques en matière d'information »)

Références bibliographiques :

les techniques documentaires au fil de l'histoire 1950-2000, J Chaumier, éd adbs, 2003

rapports réalisés par le GFII sur l'industrie de l'information électronique (le dernier date de 2004)

www.gfii.asso.fr/

apoged.fr pour les professionnels de la GED

Consulter les revues : Archimag, Bases, O1 informatique, Internet professionnel, bulletin de l'INIST et des organisations internationales (UNISIST) et leurs sites web (cf web INTD, ressources documentaires)

L'euro-référentiel, ECIA, 2004 www.aslib.co.uk/ecia/ [European Council of Information Associations](#)

les sites de formation

<http://formist.enssib.fr>

les urfist (unités régionales de formation à l'information scientifique et technique)

www.urfist.info/

voir le site INTD pour les sources professionnelles

consulter le netvibes et le Delicious pour suivre l'activité du secteur documentaire

Sociologie de la production scientifique

La constitution des savoirs scientifiques est, pour une grande part, liée au secteur de la recherche scientifique. Le fonctionnement de celui-ci se caractérise, notamment, par une organisation en communautés et par une production évaluée par les pairs à l'intérieur de chaque discipline.

C'est le domaine de la sociologie de la science et sociologie de l'innovation, mais aussi de la scientométrie (mesure de la science et de la production scientifique)

I - Le cadre : fonctionnement de la recherche scientifique en France

rappel historique :

1660 Colbert crée l'Académie des sciences

1665 Création du journal des savants

1945 Création du CNRS

1958 DGRST et comité interministériel pour la recherche scientifique

a) **Recherche fondamentale** est assurée en grande partie par la puissance publique

a1 - Recherche spécialisée au sein d'organismes publics d'Etat :

les grands instituts de recherche spécialisés, sous tutelle des différents ministères, comme INRA, INSERM, INRIA, CNET, CEA, ont été fondés, pour la plupart, en 1945.

a2 - Recherche académique au sein des universités :

les enseignants-chercheurs (qui assurent des missions d'enseignement et de recherche) sont regroupés au sein d'unités d'enseignement et de recherche, sous la tutelle du Ministère Education nationale, dont l'organisation se fait par sections disciplinaires. La nomination d'un enseignant est faite par les pairs (via le CNU - Comité national des universités - qui habilite et qualifie les enseignants-chercheurs, Maîtres de conférence et Professeurs des Universités)

a3 – Laboratoires CNRS :

organisés selon des sections disciplinaires, ayant des statuts multiples : laboratoires, laboratoires associés, regroupement d'équipes (GDR). Recrutement sur concours spécifiques.

Le financement de la recherche publique se fait par réponse à des appels d'offres, notamment ceux de l'ANR, agence nationale de la recherche, qui regroupe les fonds destinés à la financer. Dans le domaine de l'information, le CNRS via l'ISCC (institut de la communication) publie chaque année des appels d'offres.

Les laboratoires de recherche sont évalués tous les 4 ans par l'AERES (agence d'évaluation) qui note leurs résultats et leur degré d'audience nationale et internationale. Rappelons aussi les critères de Shanghai qui établissent eux aussi une notation internationale fondée sur le volume

de publications dans des revues d'audience internationale reconnues (cf scientométrie et revues de citations comme l'ISI –international science index)

b - Recherche privée :

b1 - au sein des entreprises : R&D (utilisation de la recherche fondamentale pour atteindre le stade de l'exploitation industrielle). Cycle de l'innovation

Souvent en liaison avec les universitaires, par le biais de bourses CIFRE (étudiants en thèse insérés dans une entreprise pour 3 ans), de conventions.

b2 - les collectivités professionnelles (qui peuvent recouvrer des cotisations volontaires) : centres de recherche par secteur d'activité (ex bâtiment, textile, papier, lait etc), centres techniques industriels (CTI) comme le CETIM, le CTBA). Ils assurent des missions de test, de certification, pour les PME.

Les principales tendances sont de développer des collaborations entre le secteur privé et l'université, par le biais de conventions sur des projets précis (financements privés) pour la recherche fondamentale ou appliquée, par l'utilisation des équipements des laboratoires publics ou d'une mise à disposition de personnels.

c - Rôle de l'Etat :

Il se matérialise par des subventions de fonctionnement, au travers d'appels d'offres ministériels (Ministère de l'Industrie, de la recherche par exemple), ou des programmes nationaux et internationaux (comme IMPACT, ESPRIT, EUREKA pour l'Europe)

- Par la mise en place d'un dispositif d'organismes publics d'aide à la recherche : tels OSEO (orientée vers les PME), l'ANRT, les CRITT centres de recherche et de technologies (cf cours politiques publiques en matière d'IST)

-

d - Sociétés savantes

en général de statut associatif, elles regroupent les chercheurs d'une discipline. Elles organisent congrès et regroupements, éditent parfois des lettres d'information ou revues.

- SFSIC, (société française en sciences de l'information et de la communication www.sfsic.org/)
- ADEST, l'association « schéma et schématisation) (cf chapitre recherches en sciences de l'information) www.upmf-grenoble.fr/adest/
- CENTRE DES HAUTES ÉTUDES INTERNATIONALES D'INFORMATIQUE DOCUMENTAIRE (CID) www.riao.org/fr/
- ISKO (knowledge organisation)
- FULBI (bibliothèques)

Regarder les publications des groupes de recherche comme l'enssib www.enssib.fr/

La revue du gresec www.u-grenoble3.fr/gresec/

e- Brevets

protègent l'innovation technologique ; procédure de dépôt via :

INPI Institut national de la propriété industrielle, OEB office européen des brevets délivrent les brevets (innovation scientifique et technique)

II - Les conditions de production scientifique : fonctionnement d'un milieu

Ce milieu possède des règles de fonctionnement spécifiques de la communauté scientifique : il se caractérise notamment par l'importance du contrôle par les pairs, par des relations directes entre ses membres. Tout enseignant-chercheur, recruté après sa thèse, est tenu de publier régulièrement et de participer à la vie de sa discipline (via les sociétés savantes, notamment)

II.1- Des circuits et canaux de production et d'échanges

- thèses
 - universitaires (déposées dans la BU du lieu de soutenance). Le fichier central des thèses à Nanterre signale les thèses en cours et les thèses soutenues. téléthèse
 - Thèses d'ingénieurs (signalées par exemple sur le site du CNRS science)
 - La communication des résultats de recherche et des recherches en cours se fait via :
 - **Les congrès** : un comité scientifique sélectionne puis valide les propositions de communication parvenues suite à une « appel à communiquer ». les communications seront édités dans les actes de congrès qu'on peut se procurer auprès des organisateurs et qui sont remis aux participants. Avec Internet, ils sont souvent disponibles avant le congrès
 - **Les Working papers**, work in progress qui font le point sur une recherche en cours. Signalés sur les sites des Laboratoires, et maintenant sur les Archives ouvertes (OAI)
 - Les « collègues invisibles », réseaux interpersonnels entre les chercheurs
 - littérature grise : ce sont des rapports intermédiaires (état d'avancement des travaux), rapports de fin de contrat (rapport de recherche), notes ponctuelles, compte-rendus de séminaires, points sur une question)
- Ces informations ont un statut de document de travail avec un degré de confidentialité possible. Parfois la confidentialité est demandée par le commanditaire (qui se réserve le droit d'éditer ou non ce rapport ultérieurement).

On se tient au courant de cette actualité en consultant, dans les Bulletins d'institutions (laboratoires et groupes de recherche), les rubriques : qui fait quoi, recherches en cours, bulletins plus ou moins réguliers, parfois leurs catalogues et les archives ouvertes. Le moteur de recherche Google scholar donne accès à ces travaux.

II.2 – L'édition scientifique

elle est organisée en collections, et les auteurs sont sollicités.

- importance des bibliographies et des citations dans la littérature scientifique (notoriété, insertion dans la discipline)
- Les périodiques scientifiques ont un comité de rédaction, avec ou sans comité de lecture ; la notoriété des titres est importante dans la stratégie de publication) (voir par exemple le rôle de l'ISI (institute for scientific information et ses « citation index »)

- Pre-prints : articles soumis à l'accord d'un comité de lecture, qui seront édités (ou non) dans un délais plus ou moins long, sont maintenant disponibles sur certains sites.
- Littérature sous-terrain : Système pour l'Information en Littérature Grise en Europe: base de données SIGLE <http://opensigle.inist.fr/>

L'Etat accorde une aide aux périodiques scientifiques, par le biais du comité national des Lettres, du CNRS par exemple

Evolution récente :

le développement des revues électroniques : montre plusieurs modifications :

L'avenir de l'édition scientifique classique (Elsevier par exemple) : certains éditeurs conservent leurs titres sur support papier et développent un support électronique. D'autres n'ont plus que des titres électroniques.. On peut dorénavant consulter des banques de données de pre-prints, les webs de labos, télécharger les documents de façon gratuite, tout comme se développent des revues électroniques académiques totalement électroniques et gratuites. Les textes sont ouverts aux commentaires et ajouts des lecteurs. Le développement des sites personnels, où chacun peut mettre à disposition sa production, viennent déstabiliser les règles qui avaient cours jusqu'ici : comment par exemple prendre en compte une production collective, dans laquelle un document de base est complété par les commentaires des lecteurs
En matière de contrôle : il n'y a parfois plus de comité de sélection auquel l'auteur soumet un texte à des pairs.

Une initiative récente, les « archives ouvertes », consiste à mettre en place des serveurs où tout chercheur peut déposer ses articles et publications. Cela vient totalement déstabiliser l'édition classique. Des négociations sont actuellement menées avec les éditeurs pour trouver de nouvelles formules et conditions d'abonnement. On passe d'ailleurs de la notion d'abonnement à la notion de contrat, qui spécifie le nombre d'utilisateurs (voire les numéros des machines qui ont le droit de télé télécharger des textes de revues électroniques).

Acteurs de la diffusion

Soulignons pour finir l'importance du serveur universitaire SUNIST (créé en 1984) et de l'ABES, agence bibliographique de l'enseignement supérieur (<http://www.sudoc.abes.fr>). Sans oublier l'INIST (institut de l'information scientifique et technique), issu du regroupement en 1989 des deux centres de documentation du CNRS). Ils permettent de recenser :

- a) laboratoires,
- b) thèses universitaires (SUDOC), thèses en cours, thèses d'ingénieurs de 41 grandes Ecoles
- c) littérature souterraine en Europe : cf SIGLE (INIST)
- d) Les banques de données PASCAL (sciences) et FRANCIS (sciences humaines et sociales) de l'INIST, rendent compte des recherches publiées dans les revues spécialisées

III - Observer la production et les coopérations dans le secteur scientifique

La scientométrie s'attache à mesurer l'évolution de la science et de la production scientifique à l'aide d'analyses statistiques des données concernant la science pour en mesurer le volume, la répartition, les échanges, les pôles disciplinaires, le nombre de publications, de brevets déposés, etc

La scientométrie est née au début des années 1960 autour de Solla Price, avec l'objectif de trouver des lois régissant la production scientifique. Avec la création par Garfield de l'ISI (Institute for Scientific information) qui produit la banque de données "citation index" les mesures statistiques de la science purent se développer. A partir des termes d'indexation des documents on peut dresser des cartes associant des mots-clés entre eux qui indiquent des tendances d'évolution de la recherche.

La revue scientometrics est créée en 1978. Les méthodes bibliométriques (qui prirent le nom de bibliométrie en 1969) appliquent les méthodes statistiques aux publications ; elles permettent d'analyser plus finement le flux croissant d'informations à partir de corpus divers (banques de données bibliographiques, brevets, etc) pour repérer les tendances et les points forts d'une discipline ou pour faire de la veille concurrentielle ou scientifique.

Plusieurs méthodes :

- unidimensionnelles : mesures statistiques simples qui consistent à compter les revues sur un sujet (nombre d'articles ou brevets produits par un laboratoire, nombre de recherches menées en collaboration, etc), comptage des citations (indicateur de l'impact d'un document sur l'ensemble de la communauté scientifique)

- bi-dimensionnelles : permettent d'établir les liens pouvant exister entre deux publications, par exemple lorsqu'elles citent toutes les deux les mêmes auteurs, les mêmes brevets, etc. Deux publications citées ensemble par d'autres sont proches.

a) Méthode de co-citations, de co-occurrence de citations. A et B citant C ont des liens : on cherche à mettre en évidence un aspect qualitatif et non plus seulement quantitatif. Ex : émergence d'un front de recherche, des collaborations sur quels sujets,

b) Méthode des mots associés : liens thématiques : l'évolution des sujets dans le temps, thèmes centraux, termes reliés (ex un procédé toujours indiqué lorsqu'on parle d'un aspect d'un problème) Des outils logiciels ont été développés par le CSI (Ecole des Mines et l'INIST : LEXIMAPPE) pour analyser des corpus de banques de données (brevets notamment)

Les résultats de ces traitements statistiques sont des cartes (outils graphiques associés au traitement logiciel). La question non encore vraiment résolue concerne le formatage des données constituant le corpus (données téléchargées à partir de différentes banques de données qu'il s'agit d'harmoniser), la mise au point d'indicateurs (savoir ce qu'on cherche), l'analyse des cartes (maintenant en trois dimensions). Outils très coûteux et en évolution rapide.

ADEST : Association pour la mesure des sciences et des techniques :

<http://webu2.upmf-grenoble.fr/adest/>

III.1 - quelques indicateurs de la recherche-développement :

A partir de sources statistiques annuelles (faites par l'OST (Observatoire des sciences et des technologies) au Ministère de la Recherche / ministère de l'industrie et du commerce), panorama de la production scientifique à partir de questionnaires recueillis

- auprès d'entreprises qui font de la recherche (mais on comptabilise ailleurs celles qui financent de la recherche extérieure)
- auprès de l'administration : ministères, grands instituts

II.2 - indicateurs de résultats : permettent d'évaluer la productivité de la recherche en tant que secteur d'activité (balance des paiements technologiques, brevets) mesurer et analyser le contenu, les thèmes

III.3 – les relations : La scientométrie permet également de mesurer les relations de collaboration, de coopération entre chercheurs ou laboratoires (index de citations, mesures scientométriques). (cf ch 9 recherches en sciences de l'information)

Le Centre de sociologie de l'innovation (Ecole des Mines, Paris) a développé les premiers logiciels LEXIMAPPE, Candide) qui fonctionnent partir d'un corpus constitué de références bibliographiques par exemple (notamment les champs affiliation, auteurs), ou à partir des index de citations (citation index) faits par l'ISI (institute for science information) on peut voir qui écrit avec qui, qui cite qui, pour en déduire des filiations, notoriété, revues les plus citées..

III.4 - bibliométrie : mesure plus précisément, à partir des références publiées, l'évolution des thèmes de recherche, par couplage (clusters) entre les termes d'indexation relevés dans les références bibliographiques des publications scientifiques.

IV - les chercheurs utilisateurs d'information

La recherche d'information fait partie intégrante du métier de chercheur. Sa relation à l'information est donc tout à fait particulière. Il connaît les membres de sa communauté, avec lesquels il est en relation étroite, assiste régulièrement à des colloques où il les rencontre ; il connaît les revues scientifiques, dans lesquelles il publie. Il lui importe avant tout de connaître tout ce qui se publie (littérature grise surtout) le plus rapidement possible. « faire sa bibliographie » est un exercice courant, pour lequel il existe maintenant des logiciels simplifiés qui respectent les normes de références scientifiques demandées par les revues scientifiques. (voir les logiciels bibliographiques comme end notes)

Les chercheurs sont donc souvent demandeurs d'exhaustivité (tout sur une question), de « bruit » (des documents peuvent ne pas toucher directement leur question, mais les aident à se donner des idées, établir des concordances). Ils trient eux-mêmes les documents pour sélectionner l'information pertinente. Selon la phase dans laquelle ils sont engagés, leurs besoins évolueront, en termes de type d'information recherchée, de précision ou largeur de la demande par exemple.

références bibliographiques

- les chercheurs et la documentation numérique./ sous la direction de Ghislaine Chartron, éd Electre, 20002
- Comportement de communication des chercheurs scientifiques / Andrée Demailly. Documentaliste, vol 15/1, mars 1978
- Le marketing des services d'information / Eric Sutter. Paris, ESF, 1994, notamment p 99 sur le comportement du chercheur dans le domaine bio-médical
- La politique scientifique et l'organisation de la recherche en France, UNESCO, n° 24, (ch 3, p 23) (centre de ressources INTD)
- Quel modèle pour l'édition numérique ? rapport du Wellcome trust, 2004
- La science : une nouvelle marchandise / Georges Ferné. La Recherche, n° 208, mars 1989
- La science et ses réseaux : genèse et circulation des faits scientifiques / sous la direction de Bruno Latour. La découverte, 1989
- - Sciences sociales, expertise et consultance. Numéro thématique. Sciences de la société, n° 32, 1994

- La scientométrie / Callon, Courtial, Penan. PUF que sais-je n° 2727,1993 .
- la science en action / Bruno Latour. La découverte, 1990
- consulter le site de Latour : www.ensmp.fr rubrique recherche

Chapitre 7

Un peu d'histoire...

Suivre l'évolution du secteur longtemps appelé, depuis les années 70, IST (information scientifique et technique) c'est prendre en compte l'évolution sous plusieurs aspects, technique, économique, sociale et politique, et de les considérer comme associés. C'est suivre conjointement l'évolution

- du stock (supports et types de documents),
- des modalités de son accès (outils d'enregistrement de l'information, produits documentaires centralisés ou décentralisés),
- des comportements et besoins d'informations

L'évolution peut se comprendre en identifiant plusieurs périodes, chacune ayant ses caractéristiques propres. Au fur et à mesure que le secteur se complexifie (techniquement, structurellement et économiquement), il y a cohabitation de plusieurs systèmes, superposition, plus que remplacement. Voici quelques fils directeurs....

Repères historiques

Jusqu'en 1900 : les bibliothèques privées

Caractéristiques :

livres (monographies) prédominant. Essentiellement de l'information connaissance pour érudits. A partir de la révolution industrielle, besoin de documents techniques, d'où l'essor exponentiel des revues scientifiques dans le monde :

- 17ème siècle : premières revues scientifiques.
- Milieu 18ème une dizaine,
- vers 1800 une centaine,
- 1850 un millier,
- vers 1900 dix mille

L'accès se fait directement sur les rayons (rangement à l'aide d'une classification, et par les catalogues de la bibliothèque. Il faut se déplacer pour consulter : il s'agit donc d'un système de recherche centralisé.

1451	imprimerie
1537	dépôt légal
1789	mécanisation de la fabrication du papier
1822	invention de la photographie
1839	photogravure (reproduction de photo : très important pour la presse)
1876	Classification DEWEY (USA)
1884	Première linotype (mécanisation de l'imprimerie : on peut fondre une ligne entière)
1895	Création de l'Institut international de bibliographie, par Paul OTLET et Henri LAFONTAINE, dans le but de recenser l'ensemble de la littérature mondiale

1905 CDU (classification décimale universelle) pour organiser ce recensement, décrire et classer le contenu

1900 - 1940 : du document à l'information

Caractéristiques :

Prise en compte des différentes unités d'information contenues dans un même document (chaque article d'une revue par exemple). Avec le *dépouillement* des revues, la notion d'information puis de documentation, apparaissent. Besoin de documents pour la recherche scientifique et technique. Les fichiers sont encore manuels, et recherche centralisée en bibliothèques spécialisées et centres d'information documentaire

1907 création des chemical abstracts (société commerciale américaine en chimie)

1931 l'Institut international de bibliographie devient FID (fédération internationale de documentation) qui existe toujours

1932 UFOD (union française des organismes de documentation) : 1ère formation doc

1936 création d'un Secrétariat d'Etat à la recherche scientifique

1945-1965 : structuration d'un milieu

Caractéristiques :

Développement de la littérature non publiée ou littérature grise (rapports). Notion de recherche fondamentale et de recherche appliquée. Début de la décentralisation (avec la diffusion de produits documentaires comme par exemple les index publiés et les bulletins bibliographiques du CNRS : notion de recherche courante). Centre de documentation. Début de la mécanisation des fichiers (fiches perforées codées) et de la micrographie.

1945 Création des principaux centres publics de recherche :
Centre de documentation du CNRS (créé en 1939), du CEA (énergie atomique),
de l'INRA (recherche agronomique) et d'autres Instituts de recherche
(aéronautique)

1950 création de l'INTD par l'UFOD

1951 premiers ordinateurs opérationnels (1958 1ères utilisations documentaires)

1953 création de l'ANRT agence nationale de recherche technique (qui existe toujours)

1957 premières sociétés de service SSII

1963 création de l'ADBS (scission de l'A-INTD) : des professionnels non diplômés de cette école se sentent aussi documentalistes...

1965 - 1972 : débuts de l'informatique documentaire :

Caractéristiques :

Essor des services de recherche et développement (R&D) et de documentation industrielle. Fichiers mécanisés (Sélecto) et cartes perforées (utilisées ensuite dans les "ordinateurs électroniques"). Développement de la micrographie. L'accès se décentralise, avec la diffusion de produits documentaires

1965 DIALOG, premier serveur US, avec la banque de donnée médicale MEDLARS

1966 N° spécial de la revue de l'ADBS Documentaliste sur l'informatique documentaire
1971 36 centres documentaires français sont automatisés

1972 - 1979 : début d'une politique publique en matière d'information

Caractéristiques :

Information ressource, accent mis sur les brevets industriels. Début des gros systèmes informatiques internes et des bases de données (bibliographiques). Interrogation décentralisée en différé (batch) via un réseau expérimental de télécommunication CYCLADES.

1972 N° spécial de la revue Documentaliste consacré aux logiciels documentaires
Ouverture de FRANCIS, banque de donnée du CNRS en sciences humaines
1973 Dépôt de brevet sur le micro-ordinateur
Ouverture de PASCAL, base CNRS pour les sciences
Création du BNIST (bureau national pour l'information scientifique et technique) au Ministère de l'Industrie
Création des premières ARIST (Agences régionales d'IST au sein des CCI Chambres de commerce)
1974 1er Congrès IDT de l'ADBS (avec le concours de l'ANRT)
1975 CIRCE (centre de calcul qui sert de serveur aux banques de données du CNRS) à Orsay
1976 réseau CYCLADES opérationnel
1978 rapport Nora-Minc sur l'informatisation de la société

1979 : Développement des réseaux

Caractéristiques :

- décentralisation de l'accès (accès à distance via Transpac)
- Mise en ligne des fichiers internes aux gros centres de documentation (bibliographiques). Développement de banques de données factuelles
- Informatisation progressive des centres de documentation avec la micro-informatique (qui débute réellement à partir de 1982)
- On parle alors d'IST (information scientifique et technique)

1979 ouverture du réseau national TRANSPAC par la DGT (dir générale des télécom) : coût indépendant de la distance : on paie au volume de l'information transportée
Ouverture du serveur QUESTEL (financé par l'Etat pendant 10 ans)
lancement du plan télématique (vidéotex)
La MIDIST remplace le BNIST au Ministère de l'Industrie
création des associations de producteurs (GFPB, de serveurs (ACSF), et d'utilisateurs

1980 premiers CADIST (centres d'acquisition et de diffusion de l'IST) au sein de certaines BU

Démarrage d'EURONET (réseau européen de transmission de données), arrêté avec l'interconnexion des réseaux européens

1982 création de la DBMIST au sein du Ministère de l'Education nationale

1983 distribution gratuite des premiers minitels par la DGT

1984 Ouverture du serveur universitaire SUNIST centré sur des applications vidéotex

1985 - 1990 : une offre élargie et diversifiée

Caractéristiques :

- Mise en place de systèmes d'archivage électronique (vidéodisques, disques optiques numériques DON)
- Débuts des produits édités hors-ligne (cédéroms)
- Banques de données en texte intégral
- Début du téléchargement
- On parle d'information spécialisée plus que d'IST

1985 ouverture du Kiosque télématique (paiement à la durée sur la facture téléphonique)

1986 lancement du RNIS devenu NUMERIS

1988 création de l'INIST, qui regroupe les centres de documentation CDSH (sciences humaines) et CDST (sciences) du CNRS
ouverture des kiosques multi-paliers (3615, 16, 17) du vidéotex à haut niveau de facturation

premiers cédéroms sur le marché français

GED gestion électronique des documents, intégrant les références d'un document et le texte intégral numérisé

1990 –1997 : de l'accès généralisé au sur-mesure

Caractéristiques :

- Réseaux locaux internes (RLE ou LAN, et aujourd'hui Intranet)
- Intégration sur le même support du texte, de l'image et du son avec les multi-média ;
- organisation nouvelle de l'information pour naviguer dans les textes avec l'hypertexte (on parle d'hypermédia)
- Notions d'information professionnelle, et de veille technologique (information décisionnelle) : information pour l'action
- Vers des produits très ciblés et de l'information à valeur rajoutée
- Brokers, sociétés de conseils en information, se développent

1989 livre de Martinet sur la veille technologique

1990 Rapport de la commission du Xème plan sur la veille technologique

1991 rapport MAYER "information et compétitivité" pour le Xème plan

1994 Rapport Martre sur l'intelligence économique

1996 réseau universitaire RENATER connecté à Internet

1998 Internet

2000 → Aujourd'hui

Caractéristiques

- Développement du Commerce électronique, développement des Intranet

- Développement des moteurs de recherche (traitement automatique des langues TAL)
- Archives ouvertes,
- Web 02 : production éditoriale de type wikipédia, blog. toute personne peut proposer de l'information, apposer son indexation (folksonomie),
- interconnexion généralisée réseau wifi
- intégration du texte, image, son ; cohabitation de types d'informations différents : banques de données (Internet invisible), sites d'information, de transaction ; production de documents structurés à la source,
- traitement automatique de corpus, web sémantique, ontologies...

Evolutions à suivre :

- les effets se cumulent, ne se remplacent pas, il y a donc cohabitation de plusieurs logiques (économique, gratuite, échange, contrôle, etc)
- il y a passage d'un accès centralisé à un accès décentralisé grâce aux TIC technologies d'information –communication ; passage du macro (ex mini-ordinateurs) au micro (petit, autonome, qui nécessite une mise en réseau pour communiquer avec d'autres, et mise en accès direct. C'est le temps de réponse du système, l'accessibilité du document, qui importent, ainsi que la relation entre la vitesse et le temps (temps réel)
- passage du matériel au virtuel : la bibliothèque lieu de stockage à la bibliothèque virtuelle (ce qui pose la question du stockage des documents physiques)
- la tendance à la marchandisation se retrouve dans les services, les prestations d'information sont désormais payantes (en interne ou en externe)
- Diversification de l'offre de supports, et de moyens d'accès (en mode vidéotex, ASCII puis Internet). La même information devient accessible sur différents supports et par différentes voies (hors ligne ou en ligne), ce qui pose la question du CHOIX, de la validation
- Gestion technique de plus en plus complexe. Lorsqu'on achète du contenu, on achète aussi les méthodes d'accès (ex cédérom) ; la partie logicielle devient partie intégrante de l'information
- Passage l'accès standard (et généralisé) au sur-mesure (à la demande), ce qui nécessite un ciblage des produits et services. Montée de la notion de valeur-ajoutée (conseil, expertise..)
- Les fonctions des spécialistes de l'information se diversifient, vers la production (en amont) ou en aval (vers la relation)
- l'accès direct par chacun reporte sur lui les opérations d'identification, de recherche et de validation

Les avancées techniques récentes, blogs, wikis, netvibe, delicious... (un document est produit à plusieurs, par intervention successive des lecteurs devenus auteurs), les tags transforment l'utilisateur en acteur actif. Sa relation à l'information change, ses compétences aussi. Les unités documentaires participent à l'élaboration de sites documentaires sur les Intranets, obligées de se positionner fortement par rapport à ces nouveaux modes de relation à l'information.

II – Evolution du rapport aux connaissances

On peut également suivre l'évolution du secteur en repérant plusieurs axes en relation :

1 – transformation de la relation aux connaissances.

On passe d'une activité de connaissance à une activité d'information, c'est-à-dire qu'on cherche non seulement à comprendre le monde, mais à agir sur lui. Cela s'est fait en plusieurs étapes :

- Démocratisation progressive de la production et diffusion des connaissances. Le temps des philosophes et des Érudits succède à celui des religieux. Longtemps l'apanage des Monastères, l'érudition se démocratise et passe aux universités dès le Moyen-âge. Création du statut des bibliothèques universitaires à la fin du XIXème siècle

De même, les matières enseignées évoluent, de la métaphysique aux mathématiques : cela est sensible quand on regarde les classifications

- ces connaissances savantes s'enrichissent de connaissances scientifiques (comprendre le monde, comment il fonctionne). Cela s'accompagne d'une forte demande, non plus seulement de savoirs mais des savoir-faire, de procédés ; parallèlement, on assiste à une diversification des savoirs sur le monde (essor des disciplines scientifiques)

Cette dynamique est visible dans l'activité des Société des savants. L'académie des sciences est créée en 1660 (on y fait état des découvertes) qui édite le « journal des savants »

Vulgarisation scientifique (encyclopédie de Diderot et d'Alembert)

1794 création du CNAM, de Polytechnique, de laboratoires de recherche appliquée mise au point de méthodes expérimentales

- Le monde industriel (XIXème). Troisième phase de cette évolution. La révolution industrielle développe des besoins en information technique (sur les machines, les procédés). Les premiers laboratoires de recherche privés sont créés, ainsi que les centres de documentation (technique) industriels : (houillères, soudure, froid...)

- Les informations de gestion ont toujours été nécessaires, elles prenaient la forme de livres de comptes (commerçants), de relevés scientifiques (faits, données).La centralisation des informations pour suivre et contrôler les processus industriels s'est réalisée au XIXème siècle

2 – Evolutions documentaires

21- de la bibliographie à la bibliothéconomie

1870 apparition du terme de documentation qui possède plusieurs sens : rechercher des documents pour appuyer une étude ; identifier ; conserver, ranger, stocker ; organiser pour retrouver ; gérer des collections ; développer des outils (classification Dewey ,1876)

22- du document à l'information

Documentologie : c'est l'étude des supports. Or il y a séparation progressive du traitement du support et du traitement du contenu (l'information). De même, on passe de la monographie aux périodiques (document collectif).Le terme de littérature grise est créé en1960

- *Paul Otlet, un précurseur*

1895 crée l'Institut International de Bibliographie

1905 met au point la CDU (classification décimale universelle) qui organise les savoirs à partir de la Dewey

1903 écrit *sciences de la bibliographie et de la documentation*

Son projet (avec Henri Lafontaine) est de développer la coopération internationale pour recenser l'ensemble des savoirs humains. Pour lui, la science se nourrit des connaissances antérieures « documents de toute nature » dit-il. Google n'invente rien...

La documentation est conçue comme une aide au travail intellectuel : faire des liens, reconstruire l'environnement

Il imagine (conceptualise) le multifenêtrage, la microfiche, l'hypertexte, reprographie...

La naissance de la documentation remonte donc aux travaux d'Otlet. Il s'agit de décomposer les éléments d'un texte, ce qui met l'accent sur la notion *d'unité d'information* c'est-à-dire des concepts d'un texte (un paragraphe, un chapitre, forment une unité d'information, mais aussi une phrase...il convient toujours, selon lui, de garder le sens voulu par l'auteur, de multiplier les accès, de faire les liens entre concepts, de les disperser sous différentes rubriques (pour qu'on les retrouve même si la question est posée différemment).

Notion de fiche mobile, qui permet de décentraliser l'accès aux documents, et de combiner des termes de recherche.)

Enfin, il faut laisser au lecteur le soin d'organiser lui-même les informations recueillies

23- l'accès à l'information : une décentralisation progressive

plusieurs évolutions :

matérielle : du livre enchaîné (moyen Age) aux rayonnages (18 ème)

Des catalogues aux fichiers

Fichiers synthétiques : Taube 1930

Fichiers mécanisés 1950 (sphinxo, sélecto)

Puis aux banques de données (années 60)

Diversification des **accès** au document physique (classement : d'abord par taille de document, puis thématique, auteur) on peut maintenant tout coupler en interrogeant les champs d'une banque de données qui renvoie à la cote de rangement

Le direct : spatial et temporel (catalogue collectif SUDOC qui localise le document) ; accès au texte intégral

Evolution des **langages de description** : Des classifications aux langages combinatoires

24- l'informatique

1951 - Univac, 1er ordinateur

1960 - recherche en ligne aux USA

1965 - début de l'informatique documentaire

1971 - 36 centres de documentation « automatisés »

1973 - brevet du micro-ordinateur

Du traitement séquentiel au traitement inverse

25- Structuration du milieu documentaire

1879 - CAFB (certificat d'aptitude pour les fonctions de bibliothèques) pour les BU

1932 - UFOD union française des organismes de documentation

- FID (Fédération internationale de documentation) qui remplace l'institut de bibliographie

1950 - INTD

1963 - ADBS créée par scission avec l'AINTD (l'association des anciens de l'intd)

III - Petite histoire des outils de mémoire documentaire :

La mise en mémoire de l'information a évolué au fil des temps. Pendant longtemps, le lecteur se promenait dans les rayons pour choisir son document. Puis il y eut les catalogues, bientôt édités (donc transportables dans plusieurs bibliothèques), ensuite les fichiers tels que nous les connaissons encore, et aujourd'hui les banques de données. Cet outil mémorise non pas le texte intégral du document, mais son image (on parle donc d'information secondaire), sa représentation, à l'aide de caractéristiques, tant internes (le contenu) qu'externes (le contenant). Un fichier centralise donc les notices bibliographiques se rapportant à un document physique (dit primaire). Aujourd'hui, la mémoire documentaire est décentralisée en de nombreux points, le document physique se trouvant très souvent éloigné de sa représentation.

Deux modes principaux qui président à la mise en place de fichiers

1 - mode synthétique : une fiche par document, qui centralise toutes ses caractéristiques
exemple : les fichiers de bibliothèques (les références, la cote, sont indiquées sur la fiche)
les fichiers synoptiques peuvent avoir un marquage de couleur sur les bords)
dans les années 50 : le fichier se mécanise, car il faut un outil pour l'utiliser. Dans un fichier à perforation marginale, dans lequel les caractéristiques sont encodées sur le bord de chaque fiche, on effectue la recherche en passant une aiguille qui ne « prend » que les fiches trouées.
en informatique : c'est ce qu'on appelle le fichier séquentiel (qui commence avec le document n°1 avec ses caractéristiques, puis le document n°2, etc). La recherche suppose de consulter l'ensemble du fichier pour comparer les documents en fonction de la caractéristique sélectionnée. C'est long.

2 - mode analytique : une fiche par caractéristique (par ex par mot-clé) : on inscrit le n° des documents possédant ce mot-clé
il y a ainsi un fichier moins lourd (autant de fiches que de mots-clés, et non autant de fiches que de document)

La recherche se fait alors en comparant les fiches concernées (idée de combinaison de mots-clé) : les documents pertinents sont ceux qui figurent sur toutes les fiches. Le résultat, c'est un nombre X de numéros de documents pertinents.
cela suppose alors de consulter un deuxième fichier classé par n° de document, (c'est-à-dire un fichier par document sur laquelle sont portées ses caractéristiques)

Plusieurs générations de fichiers analytiques :

1930 : fichiers manuels améliorés avec les fichiers uniterm (William Taube)

1950 : fichiers mécanisés : une fiche cartonnée par caractéristique (regroupement sur une même fiche tous les documents possédant cette caractéristique : l'intérêt est de n'avoir qu'un seul fichier classé alphabétiquement)

- soit à perforations marginales : le tri se fait par aiguille
- soit à perforations centrales (sélecto) : le tri est optique ; on peut combiner plusieurs termes (en comparant plusieurs fiches) (post-coordination)
- les fichiers automatisés sont organisés sur les mêmes bases (tri sur critères multiples). On les nomme fichiers inverses

Avec le texte intégral, la recherche se fait sur tous les mots pleins du texte (ce qui est évidemment plus long). L'ordinateur établit un index de tous les mots significatifs (mots pleins) du texte, chacun étant suivi de l'adresse (le n°) du document. En amont, aucun travail

d'analyse ou d'indexation. Mais en aval, recherche génère du bruit (trop de documents non pertinents) dû à de mauvaises combinaisons (exemple informatique et documentation peut signifier la documentation dans le secteur de l'informatique, la place de l'informatique dans la documentation, etc).

D'où la mise au point de logiciels d'analyse linguistique qui analysent le texte et le reportent à des formes linguistiques unifiées (lemmatisation, par exemple toutes les formes fléchies : dormir et ses conjugaison, endormissement, etc). C'est l'objet du Tal traitement automatique des langues.

Références bibliographiques :

- Encyclopédie Histoire des bibliothèques, tome IV. Cercle de la Librairie, 1993
 - les années 30 / B. Delmas p 178-193
 - les années 1945-1975 / A. Boulogne p 404-417
 - les années 1975-1991 / B. Guyot p 518-527
 - Les politiques publiques / JP Salaün, p 501-509
- Histoire de la documentation / sylvie Fayet-Scribe, CNRS, 2000
- L'IST et l'entreprise/ B. Guyot. Les cahiers du LERASS n° 19, 1990
- Intelligence économique et stratégie d'entreprise / rapport sous la direction d'Henri Martre. Paris, la Documentation française, 1994
- Micro-ordinateur et traitement de l'information. Paris, A Jour, (1991)
- qu'est-ce que la documentation / suzanne Briet. Tiré à part, 1950
- qu'est-ce que le virtuel / P Lévy. Ed la découverte, 1995
- Histoire de l'information scientifique et technique / Martine Comberousse, 1999. Nathan univ (coll 128)

concernant l'histoire des fichiers

- International organisation and dissemination of knowledge : selected essays of Paul Otlet, by Boyd Rayward. Elsevier, 1990. FID
- La raison graphique : la domestication de la pensée sauvage / Jack Goody. Paris, éd de Minuit, 1979. 275 p
- Traité de documentation : le livre sur le livre. théorie et pratique/ Paul Otlet. Bruxelles, Palais mondial, 1934. 450 p. (réédité en 1989)

Politiques publiques en matière d'IST

1 – les acteurs français

- Structuration et organisation :
- grands centres de documentation d'après-guerre
- années 50 : DGRST et réflexions autour des plans
- politiques publiques : lisibles au travers de plans, rapports au gouvernement, organismes ministériels Plans*

Politiques ministérielles et dispositif d'IST

Les ministères ont un rôle d'observateur (collectent des données sur le secteur IST), de prescripteur (financements) et de mise en place d'un dispositif. Ils lancent des appels d'offre pour développer certains secteurs. Ils insufflent une dynamique. Actuellement par exemple, en matière de développement des technologies (les autoroutes de l'information devenues société de l'information). Ces politiques sont de plus en plus fréquemment liées à l'Union européenne.

Un grand organisme : CNRS : 1945, puis INIST 1989 qui produit les banques de données Francis 1972, et Pascal 1973

Ministère de la culture

Direction de la lecture : gère les BM, les BCP

Ministère de l'Éducation Nationale :

Gère le réseau des BU (bibliothèques universitaires) : Système de repérage, d'accès et de prêt : certaines ont le statut de CADIST (centre d'accès et de diffusion de l'IST) ; PEB, prêt entre-bibliothèques (en Grande-Bretagne, la British lending library), le)

ABES agence bibliographique de l'enseignement supérieur (abes.fr), qui met au point le recensement des thèses, le CCN (1983) devenu SUDOC, catalogue collectif des bibliothèques Les URFIST (unités régionales de formation à l'IST) auprès des BU

Ministère de l'Industrie :

Réseau des 162 CCI, des 21 CRCI (chambre de commerce régionale) réunis dans le réseau APCCI ; une chambre de commerce internationale
65 CCI à l'étranger

ARIST (agence régionale d'information scientifique et technique, appartenant au réseau des chambres de commerce)

Les CRITT (centre régional d'innovation et de transfert technologique), les centres techniques professionnels (ex le CETIM pour les industries mécaniques)

ANVAR 1967 et rattaché en 1979 au M Industrie et M recherche. Aide les PME à développer leur recherche, et finance, notamment, la recherche d'antériorité pour les dépôts de brevets. Vient de fusionner avec la BDPME, et devenue OSEO

Ministère de la recherche :

ANRT (le progrès technique) 1953

Ministère des affaires étrangères

Gère le réseau des postes d'expansion économique des ambassades ; diffusés par le centre français du commerce extérieur et par l'ADIT (agence pour la diffusion de l'information technologique (adit.fr))

La politique publique est lisible au travers des plans (site archivé : <http://www.plan.gouv.fr>)

1990 rapport commission "veille technologique" pour le Xème plan

1991 rapport MAYER "information et compétitivité"

1994 rapport sur l'intelligence économique et stratégie des entreprises. rapporteur Henri Martre

2005 – le commissariat général du Plan devient Centre d'analyse stratégique

<http://www.strategie.gouv.fr> Le Centre d'analyse stratégique est un organisme directement rattaché au Premier ministre. Il a pour mission d'éclairer le Gouvernement dans la définition et la mise en œuvre de ses orientations stratégiques en matière économique, sociale, environnementale ou culturelle.

La société de l'information Politique publique en direction des PME :

http://www.telecom.gouv.fr/fonds_documentaire/men/ticpme/rapportetape1.pdf

La société de l'information, CURIEN Nicolas, MUET Pierre-Alain

FRANCE. Conseil d'analyse économique, Paris; La Documentation française; 2004; 310 pages

<http://www.ladocumentationfrancaise.fr/rapports-publics/044000180/index.shtml>

Site gouvernemental sur la Internet : site portail sur la société de l'information

<http://www.internet.gouv.fr/>

3 - Politique européenne :

Site européen sur la société de l'information :

http://europa.eu/pol/infso/index_fr.htm

EUROPA - http://ec.europa.eu/information_society/index_en.htm

Concepts et recherches en sciences de l'information

I - Quelques concepts et théories qui ont alimenté les sciences de l'information

L'hétérogénéité des concepts et de leur filiation n'est pas évidente, car il y a juxtaposition, appropriation pour enrichir ce qui constitue les bases des sciences de l'information. N'oublions pas que pour nous l'information est à la fois du sens et une forme. Les principaux concepts fondateurs ont été formulés dans la décennie 1947-57.

On peut identifier plusieurs filiations, qui touchent à la distinction de la forme et du sens. La forme peut se découper en symboles et en signaux, constituant un message. Celui-ci se transporte, tout en renvoyant aussi à la nature du raisonnement logique, donc juste (ce qui aboutira à l'algorithme), donc aux conditions de vérité des énoncés. Les choses, on le voit, sont complexes.

Schématiquement, il y a 5 grandes lignées, certaines très techniques (automates, transport et communication de message), d'autres plus philosophiques ou politiques, d'autres encore fondamentalement méthodologiques. Le fait est qu'elles ont convergé vers les années 50 pour construire et aménager ce qu'on appelle encore la théorie de la communication. De cette théorie sont issues les disciplines modernes que nous connaissons aujourd'hui. Les sciences de l'information sont interdisciplinaires car elles empruntent à plusieurs disciplines leurs modèles de réflexion pour développer les leurs.

I.1. Lignée des transports (communication des messages)

Elles sont conduites à la fois par des mathématiciens, des physiciens et des ingénieurs télécommunication, pour lesquels la communication est un acte technique.

Depuis 1924, les ingénieurs de la transmission travaillaient sur une théorie de la transmission du signal (improprement appelée communication). Leur préoccupation était de mesurer la quantité d'information contenue dans un message (signal) pour la faire circuler sur un canal (support physique) dans un temps donné (le débit), avec des contraintes, de frottement, de bruit (entropie). L'information apparaît comme réduite à une donnée mesurable (le bit), qui est codée et décodée pour pouvoir être transmise en fonction de la capacité et des contraintes du support de transmission. La préoccupation majeure est de pouvoir détecter les erreurs pour que le message émis soit identique à l'arrivé chez le récepteur.

1794 : 1er télégraphe mécanique CHAPPE : bras mobiles situés sur des tours : 20 minutes suffisent, par beau temps, pour qu'un signal, émis à Paris, arrive à Toulon

1832 : télégraphe électrique de MORSE : cryptage des signaux en binaire - et .)

1875 : mise au point du téléphone par Bell)

Déjà, en 1927, HARTLEY avait présenté un mémoire sur la transmission de l'information (c'est la première fois que le terme est utilisé par un mathématicien) à un congrès de téléphonie, en proposant une mesure exacte de l'information

En 1947, SHANNON, mathématicien au MIT, publie une "théorie mathématique de la communication" dans la revue "bell telephone system."

Le cadre théorique de la circulation de l'information (définie comme communication) est ainsi produit par un théoricien des sciences dures qui ne considère le message qu'à travers sa mise en forme communicable et pas du tout son contenu. Il s'agit d'un signal et non pas un symbole. On cherche à éliminer les bruits, frottements qui entraînent une déperdition de son amplitude et donc de sa qualité (c'est l'entropie)

Cette théorie est reformulée peu après par WEAVER (MIT), qui définit la communication comme une action complexe résultant de l'interaction entre trois actions : à l'acte technique de base il ajoute deux autres notions, celle du symbole ou de la signification, et celle d'une action qui a des conséquences sur l'émission. Il y a ainsi trois niveaux :

niveau A : acte technique : le signal est rendu transmissible (codage etc)

B : acte sémantique : le signal devient signification (le sens)

C : le signal modifie le comportement du récepteur : il constitue une action (même s'il n'y a pas encore la notion de feed-back / retour) ==> pragmatique

Il intègre ainsi la linguistique, en considérant le langage naturel comme un codage (des lettres assemblées), fait de phonèmes, de morphèmes et de règles syntaxiques dont il convient de tenir compte pour pouvoir coder les significations induites. Il y a ici liaison avec la lignée de la linguistique : sémantique

I.2- La lignée Arts de l'ingénieur, science des systèmes :

Elle est issue de la science des automates, très ancienne, qui cherchait à contrôler les machines et les objets, en analysant les réactions de l'objet face à une sollicitation.

Depuis 1914 le terme de rétroaction ou feedback était utilisé pour indiquer une circularité entre les impulsions et les réponses de l'objet qui venaient réorienter le système : dans un dispositif de régulation, il y a une boucle circulaire (par opposition à linéaire). Le thermostat par exemple offre 2 états :

état 1 : la machine détecte la chaleur, avec pour conséquence la fermeture du chauffage ;

état 2 : le constat de froid aboutit à l'ouverture du chauffage. Les états vont ainsi varier sous l'effet du signal qui provoque une action. L'information-signal fait partie du processus de régulation.

en 1948, Norbert WIENER (mathématicien au MIT) publie "cybernetics or control and communication in the Animal and the Machine", instituant la cybernétique.

Cette discipline est souvent associée à la théorie de la commande (en grec, pilote de navire). Le principe de base avance que tout système transforme une énergie entrante en une action (ainsi les matières premières sont transformées dans l'usine en produit - l'acier). Un certain nombre de variables viennent réguler et contrôler le système pour que l'action puisse se faire (détecter, corriger, signal d'erreur, maintenir l'équilibre de l'ensemble. La cybernétique veut gérer le prévisible et prévoir l'imprévisible par corrections successives, en recherchant des états stables, tout en évitant les pertes d'énergie. Il y a re-bouclage par rétro-action.

Elle étudie donc les relations complexes entre des phénomènes (naturels ou artificiels, aussi bien hommes, animaux que machines). Ces comportements sont qualifiés de comportements d'information, puisque le signal donné provoque une action. WIENER s'intéresse au processus informationnel au cœur de tout système. La communication équivaut à la façon de gérer l'information d'un système dont tous les éléments sont organisés pour réaliser une ou plusieurs fonctions.

Cela explique les apports de la cybernétique aux sciences de l'information et de la communication, puisqu'elle étudie des processus de communication et de contrôle entre des

êtres vivants (hommes, animaux) aussi bien qu'entre des machines, et même au sein de processus sociaux. Elle a notamment tenté d'expliquer les mécanismes de prise de décision. (à l'aide de bouclages successifs et de graphes)

Les notions de boîte noire et d'entropie, sont au centre cette discipline. La boîte noire signifie que les cybernéticiens ne s'intéressent pas à ce qu'il y a dans la boîte mais s'intéressent aux lois d'assemblage des différentes boîtes et aux relations entre ces boîtes (les objets). En faisant varier des éléments externes, on étudie les conséquences pour le système (la boîte). On retrouve aussi cette idée en psychologie expérimentale.

Quant au concept d'entropie, issu de la thermodynamique, il signifie déperdition d'énergie, passage d'un état d'ordre prévisible à un état de désordre aléatoire. Selon Wiener, la nature a tendance à détruire l'ordonné et à détériorer le compréhensible. Ces dégradations successives (maladie, épuisement des ressources naturelles) aboutissant éventuellement à un chaos social pourraient être combattues par la circulation de l'information qui maintient un certain ordre, un équilibre. L'information irriguant tout système, elle provoque des échanges et incite à agir sur le monde extérieur. Selon Wiener, l'information est enrichie lorsqu'elle circule, mais elle se dégrade lorsqu'elle est confinée (le secret) et qu'il n'y a plus de confrontation avec d'autres informations, car la stimulation pousse à l'action. En appliquant cette idée à l'information, il s'agit d'essayer de réduire la perte (déperdition - entropie) d'information lors de sa diffusion.

La cybernétique est donc l'un des précurseurs de la théorie des systèmes, notamment avec la modélisation. Celle-ci est un modèle théorique construit sur papier par analogie avec l'observation des réactions à l'intérieur du système. La modélisation simplifie la complexité de la réalité et en favorise la compréhension. Systémique (science des systèmes) et sciences de gestion fonctionnent beaucoup autour de ces notions.

La cybernétique travaille enfin sur le concept de machine intelligente, d'animaux artificiels (la tortue mécanique de Grey Walter date de 1953), Elle cherche à comprendre le comportement des objets, à établir des relations entre des phénomènes. Entre autres choses, elle voudrait construire une machine qui serait un modèle réduit du cerveau humain. Wiener luttera contre l'utilisation de la technique à des fins militaires (s'isolant du milieu des chercheurs qui travaillent presque tous sur les calculateurs dépendant des militaires)

Cette science des systèmes est donc la racine commune mathématiciens-informaticiens. En 1951, il y aura scission, les informaticiens se concentrant sur l'information linéaire, qui va d'un point à un autre, de l'émetteur vers le récepteur. Ils travaillent sur les conditions de gestion de l'information.

Il y a aura essaimage et enrichissement des éléments de la cybernétique hors du champ qui l'a produite. La notion de système fut notamment reprise en biologie, en sciences du management et en sciences de l'information-communication.

a – la biologie.

Van Foerster en 1959, puis ATLAN en 1972, développent une théorie informationnelle de l'organisation biologique. Tout système est constitué d'une structure rassemblant des éléments autour d'un même but, et reliés entre eux par un système de relations. Toute modification d'un élément modifie l'ensemble du système. Le système, cellulaire par exemple, produit et transmet de l'information Il développe son autonomie en relation avec son environnement (en dépendance, c'est-à-dire en communication).

b - Les sciences de l'organisation,

elles vont donner naissance à la systémique ou théorie des systèmes dans les années 60
Herbert SIMON écrit, en 1969, “ science des systèmes, science de l’artificiel”

Karl POPPER écrit en 1973 la dynamique des systèmes. Il retient quatre directions des sciences de la gestion : concept de modélisation, simulation sur ordinateur, théorie de la prise de décision, et cybernétique des organisations.

Quant à Edgar MORIN, il retient 3 niveaux d’organisation : communicationnel, informationnel et computationnel. Dans sa “méthode”, il développe une vision systémique, qui permet de concevoir et de faire communiquer des systèmes complexes de symboles.

1980 : collège de systémique à l’AFCET, et naissance de la revue internationale de systémique

1989 : 1er congrès européen de systémique

I.3 - La lignée sémantique

A la même époque, le philosophe MORRIS (de la fondation Rockefeller) s’intéresse lui aussi à la question, et plus particulièrement en travaillant au niveau du sens (sémantique), ce qui deviendra une discipline à part entière, la sémiologie, art des signes, des symboles et de leur signification. La communication est ainsi un acte syntaxique, sémantique et pragmatique (pratique) (1946)

I.4 - La lignée des sciences politiques :

Dès 1934, Harold LASSWELL, issu du milieu science politique, formulait de façon linéaire ce qu’était pour lui la communication, avec son schéma devenu fameux :

Qui dit quoi à qui, par quel canal, avec quels effets ?

Destinateur – message – destinataire – usages

L’information est donc liée à l’efficacité (elle transforme)

Il s’inscrivait d’abord dans une communication individuelle de transmission d’information. Ce schéma codifie des interrogations et, surtout, suggère des champs d’études orientés vers les effets. Relation de cause à effet entre les termes, insistant sur l’importance de travailler sur les aspects psycho-sociologiques par observation et enquêtes (telle l’analyse des besoins bien connue des spécialistes de l’information)

Il met en valeur la notion de conséquence, et de fonction (pourquoi s’informe-t-on, comment, y a-t-il une logique rationnelle dans cette démarche, etc). Le « quoi » insiste sur l’analyse de contenu, la façon d’agencer une information (linguistique).

En 1948, il adaptait ce schéma aux mass médias. Il met en évidence la logique de production (l’offre) dans un marché qui la régule. La diversité des acteurs concourant à produire une offre d’information rend difficile aujourd’hui l’application de son schéma, qui reste cependant une référence explicative.

Les sciences de l’information spécialisée (hors mass-médias) sont tout à fait concernées par ces études de production des messages. En fait, leur contexte est soit du type linéaire individuel (Lasswell), soit du type point à masse (un émetteur et plusieurs destinataires).

I.5 - La lignée bibliothéconomique

Dès la fin du XIXème siècle, l’art d’agencer les livres, d’en traduire le contenu et de trouver des moyens d’accès et de stockage. Otlet, puis Briet, Vannevar Bush ou Engelbart en ont écrit les bases et les principes.

En conclusion,

Une grande partie des théories de la communication et de l'information ont été développées dans des milieux scientifiques, qui reconnaissent et cherchent à gérer l'ambigu, le bruit, l'équivoque, l'aléatoire, les fluctuations et redondances.

Elles se sont développées dans deux directions : celle des procédures de validation des machines, ce qui relève du contrôle fonctionnel ; et celle de la simulation des comportements et donc de modélisation.

La théorie de l'information est donc essentiellement envisagée comme régulation des systèmes. Acquérir de l'information est une façon de réduire l'incertitude sur l'état d'un système.

Ce n'est que depuis peu, on va le voir, que se sont développées des réflexions sur les conditions sociales de production et de circulation de l'information.

Ces lignées apportent toutes des réflexions sur les rapports entre science, technique et culture et sur l'intelligence des technologies de communication et d'information.

II – La recherche en sciences de l'information aujourd'hui

Qu'on juge de la diversité des directions de recherche en cours au sein des disciplines que nous venons de nommer (voir aussi chapitre 6)

1 - Les Télécommunications

science du codage (cryptographie),
réseaux-supports (locaux, à hauts débits, etc, téléphonie mobile)

- Organismes de recherche : CNET (centre national d'études des télécommunications, devenu France TELECOM recherche), CCETT (Centre commun d'études de télécomm, Rennes),

2 - L'informatique

Nous ne donnons ici que les principaux intitulés disciplinaires : consulter le CNRS, les universités, les sociétés savantes

<p>1911 : création d'IBM 1931 Bull 1932 Honeywell 1946 premier ordinateur (calculateur) ENIAC (à technologie électronique) UNIVAC fin des années 50 1982: 1er PC 1984 : 1er MAC</p>

plusieurs directions de recherche :

- informatique formelle (automates, robotique) ; théorie de la décision ; algorithmique
- architecture des systèmes informatiques et des réseaux
- Intelligence artificielle IA (1956)
- système d'exploitation
- génie logiciel (1968) : hypertexte, systèmes experts (modéliser le raisonnement de l'expert) lié aux cognitivistes (étude des fonctions de l'esprit humain : langage, apprentissages, raisonnement, etc)

SIAD : systèmes d'aide à la décision

Après une approche techno-centrée (l'homme se plie aux exigences de la machine), se développe des approches anthropo-centrées (la machine essaie de comprendre l'homme).

Deux axes en **Intelligence artificielle** :

- a) traitement automatique des langues, TAL en lien avec la linguistique
- b) l'interaction Homme-machine, en lien avec les cognisciences

a) **traitement automatique des langues (TAL)** : l'histoire commence avec le développement de la logique booléenne (Boole, 1847), qui postule que la logique dépend des mathématiques et non pas de la philosophie. La pensée est de type calculatoire (algorithmique). Les recherches cherchent à développer des machines capables d'imiter la computation (traitement) des symboles du cerveau humain.

Elles cherchent à reconnaître, dans un texte, non pas l'énoncé (c'est-à-dire une suite de termes) mais l'intention qui est derrière et qui se traduit par des concepts. La machine est donc capable d'intégrer le but de l'utilisateur. Avec de nombreuses difficultés, qui tiennent, entre autres, à l'identification des buts à court, à moyen ou à long terme. Par exemple, quand un utilisateur demande la brochure de l'intd pour devenir documentaliste, il y a un but immédiat (l'information), second (la formation) et final (le métier).

Différentes branches :

- analyse morpho-syntaxique
- indexation automatique
- traduction automatique
- reconnaissance de la parole (synthèse vocale)
- langages d'interface (interrogation et transaction)

cf Maret Dominique, Lallich-Boidin G, recherche d'information et traitement de langue, fondements linguistiques et applications, presses de l'enssib, 2005

b) l'interaction Homme-Machine

la machine doit donc être capable d'interagir avec l'utilisateur (dialoguer). Soit elle mène le dialogue (que voulez-vous faire ?, tapez votre question), soit c'est l'utilisateur qui a l'initiative (il inscrit sa demande dans la boîte de dialogue)

l'idée d'une machine « intelligente » provient de sa capacité à intégrer la réponse ou le comportement de l'utilisateur et à caler ses tâches sur lui. Par exemple, s'il ne visualise pas le résultat mais formule une autre requête, cela signifie qu'il n'est pas satisfait de la réponse. Tout ceci suppose d'avoir analysé très finement (voir l'intérêt alors de l'observation minutieuse) les différentes possibilités du comportement de l'utilisateur (il va affiner, élargir, visualiser, recommencer, etc)

La machine se comporte donc en assistant de plusieurs manières : suggérer (un autre terme, plus restreint, plus large) pour aider à la formulation, suggérer d'affiner vu le nombre impressionnant de réponse, proposer le guide, etc. la machine propose des solutions que l'utilisateur confirme ou refuse. Progressivement, elle intègre les réponses pour en dresser un « profil » sur lequel elle se basera pour donner les réponses.

Avec la reconnaissance vocale, elle devra « inférer » d'une phrase comme « je veux des documents récents » les dates qui y correspondent, et le formuler à l'utilisateur pour confirmation (du genre « documents postérieurs à 2000 »)

En matière de Communication homme-machine

- modes de lecture et d'écriture avec les technologies (@books, multimédia)
- modes de recherche via une technologie SRI (systèmes de recherche d'information)

- interfaces
- outils de recherche d'information ; outils de filtrage, outils de représentation...
- outils de traitement (indexation automatique, texte intégral etc)
- outils de présentation de l'information (mapping)
- traitement statistique des données

Organismes : CNRS et INRIA (Institut national de la recherche en informatique et automatique) ;

3 - Les sciences cognitives :

En lien avec l'intelligence artificielle

neurosciences (cerveau)

psychologie cognitive (perception, connaissances, représentations et raisonnements)

linguistique et cognition (sciences du langage)

4 – Les sciences du langage : linguistique, sémantique, sémiologie :

Etude des symboles et leur signification. Analyse de contenu, indexation

5 – Les sciences du comportement

- Psycho-sociologie de la communication

- sociologie des comportements : sociologie des groupes

6 – L'ergonomie :

Concernée par l'adéquation entre l'homme et son environnement de travail, (position des objets, ordinateur) et de la charge pour le maîtriser et s'y sentir bien. Cela concerne la charge cognitive, le corps, les interfaces...

7 - les sciences de gestion :

Ont comme objet de décrire et modéliser les systèmes d'organisation et les systèmes d'information stratégiques comme l'intelligence économique

Prise de décision, la prévision, la planification,

Cf revue française de gestion

8 – la sociologie

Sociologie des organisations

Sociologie du travail

Sociologie des usages

Sont convoquées à des titres divers, notamment pour étudier les relations instaurées entre les personnes par les dispositifs de collaboration, de mutualisation, fortement en lien avec le management.

Conclusion

Ces éléments très schématiquement présentés, montrent la diversité des disciplines mobilisées par le traitement de l'information

Références bibliographiques

Sciences de l'information et philosophie : une communauté d'interrogations. Paris, éd adbs, 1997

Interaction homme-machine et recherche d'information, sous la direction de Céline Paganelli. Paris, Hermès-Lavoisier, 2000

Histoire de la documentation en France : culture, science et technologie de l'information, 895-1937, éd cnrs, 2000

Histoire de la société de l'information / armand Mattelart. Paris, la découverte, 2002

Faire de la recherche en système d'information coord F Rowe. Vuibert, 2002

Assistance intelligente à la recherche d'informations / dir eric Gaussier et MH Stéfani, Hermès-Lavoisier, 2003

Communication personne systèmes informationnels / dir B. Miège, Hermès-Lavoisier, 2003

Sciences de l'information et de la communication : entre sciences et professionnalisation / coordonné par Yves Jeanneret et Bruno Ollivier : revue Hermès, juin 2004

Les ouvrages de la collection Hermès-science (sur les systèmes de recherche d'information)

Des auteurs comme Chaudiron,

Des sociétés savantes comme ISKO

Les compte-rendus des école d'été organisées par l'INRIA

Actes de congrès de la société française de l'information et de la communication (SFSIC)
<http://www.sfsic.org/>

Les archives ouvertes en sciences de l'information – communication
<http://archivesic.ccsd.cnrs.fr/>

Glossaire

Accès direct	recherche, qui utilise des mots-clés, lesquels orientent directement vers le document ou à la partie du document dans laquelle figure ce thème. Cela nécessite, en amont, soit un travail d'indexation (choix des mots-clés) soit une opération automatique qui relève tous les mots d'un texte et les organise en index..
Analyse :	opération intellectuelle qui consiste à réduire un objet en ses éléments constituants pour ensuite recomposer le schéma d'ensemble ou en démontrer la logique. En documentation : analyse documentaire = décrire + apposer des marques + structurer l'information pour la rendre exploitable. Se dit aussi du résultat de cette opération (équivalent à un résumé) Voir aussi l'analyse de l'information qui s'attache au contenu
Back-office	se dit de l'organisation mise en place pour répondre à un service (ant : front-office)
Bruit	trop de documents par rapport à la question posée ; prendre en compte le fait qu'il peut aider à se donner des idées, à établir des concordances, à trouver des éléments auxquels on n'avait pas pensé (ant : silence)
Catégorisation	opération qui consiste à regrouper au sein d'un même ensemble des objets possédant la ou les mêmes caractéristiques. (résultats : typologie, taxinomie, classification)
Collection :	assemblage de documents possédant au moins une caractéristique commune
Info secondaire	information sur l'information (ex bulletins bibliographiques, références, métadonnées) : un document secondaire liste les références bibliographiques (ex bulletin bibliographique, un fichier de bibliothèque) ; il faut ensuite accéder au document primaire
Info primaire	document-source dans lequel se trouve l'information, quel que soit son support
Interrogation par menu ou par arborescence ou encore guidée :	on se rapproche progressivement de l'information recherchée, en suivant la classification, le plan de classement, par des choix successifs
Interrogation par mots clés (accès direct) qui donne immédiatement un résultat.	
Intertextualité	relations entre plusieurs textes
Intra-textualité :	relations à l'intérieur du même document
Métadonnées	éléments /caractéristiques descriptifs d'un document ; elles portent sur la forme, sur le contenu, ou encore sur le type de document, l'auteur...(cf analyse)
Mode analytique	une fiche par caractéristique (par ex par mot-clé) : on inscrit le n° des documents possédant ce mot-clé
Mode synthétique	une fiche par document, sur laquelle sont reportées toutes ses caractéristiques
Ontologie	système de représentation des connaissances., un ensemble structuré de concepts dans un domaine et les rapports entre ces concepts. Les concepts sont organisés dans un graphe dont les relations peuvent être des relations sémantiques ou des relations de subsomption. L'objectif premier d'une

	ontologie est de modéliser un ensemble de connaissances dans un domaine donné. (wikipedia)
Organisation	un système social structuré autour d'objectifs (par exemple vendre un service, un produit, assurer une mission de service public). Elle œuvre dans un certain contexte spatio-temporel
Outil intellectuel	appellations conventionnelles pour langages documentaires : vocabulaires contrôlés, outil linguistique..
Paratexte	éléments qui viennent compléter un texte : notes, références bibliographiques...
Pertinence	correspondance (adéquation) entre la réponse et la demande. Qui répond à la question.
Processus	séquencement d'opérations pour parvenir au but fixé, réalisées selon des procédures
Scientométrie	mesure de la science et de la production scientifique : à l'aide d'analyses statistiques des données concernant la science volume, répartition, les échanges, les pôles disciplinaires, le nombre de publications, de brevets déposés, etc
Sérendipité	« par hasard » : dans le cadre de la recherche documentaire, consiste à trouver des éléments non cherchés au départ ;
Serveur	désigne soit une machine qui abrite la banque de données, soit les sociétés intermédiaires qui assurent sa commercialisation
Système	ensemble d'éléments en interaction, regroupés et agencés pour parvenir à un but.
Taxinomie	(de taxis : arrangement ou ordonnancement fondé sur une loi ou un principe) ; syn : catégorisation, classification

consulter :

Vocabulaire de la documentation, éd Arlette Boulogne, éd ADBS, 2000

Terminologie de base en sciences de l'information, sur le site de l'École des bibliothèques,
Montréal : <http://www.ebsi.umontreal.ca/termino/index.htm>