

MODULE 2 – 1 : LA RECHERCHE DOCUMENTAIRE ÉTAPES ET OUTILS

I.DE KAENEL

[Notes et commentaires de A. Cherbuin en helvetica narrow 10 entre crochets]

Sommaire

Étape 1 Définir l'objectif de la recherche

- Recherche documentaire ou recherche d'information
- Une démarche de résolution de problème
- Des techniques et des processus psychologiques

Étape 2 Faire l'inventaire des outils d'orientation et d'accès

- Connaître les sources et les outils nécessaires
- Grille d'analyse des outils de référence
- Caractéristiques et complémentarité des principaux outils de recherche
 - ouvrages de référence
 - catalogues, catalogues collectifs, métacatalogues
 - bases de données
 - moteurs de recherche, métamoteurs
 - annuaires, portails
 - outils collaboratifs
 - outils « hybrides »

Étape 3 Formuler la recherche

- Les principes d'interrogation
 - interfaces de recherche : formulaires, recherche simple et expert
 - choix des points d'accès (feuilletage des index)
 - recherche par langages documentaires (dictionnaires, thésaurus, plan de classement)
 - recherche en langage naturel, recherche multilingue
- La syntaxe de commandes
- Les opérateurs logiques, de proximité, morphologiques, comparatifs
- Filtres et limites
- Recherche par similarité
- Recherche multifichiers
- Tri et affichage des résultats
 - ordre pré-déterminé
 - ordre dynamique : pertinence, popularité, thèmes
 - Types d'affichage

Étape 4 Evaluer les résultats

- Définitions : la notion de pertinence, bruit, silence, rappel, précision
- Consultation des documents
 - localisation des documents imprimés
 - localisation des sites web

- Critères d'évaluation des documents : grilles d'analyse
- Outils d'analyse de corpus de résultats (lien avec la veille stratégique)

Étape 5 Prélever, traiter, présenter l'information

- Sauvegarder les informations : problèmes techniques de téléchargement
- Extraire les caractéristiques des documents retenus. ajouter des commentaires personnels...
- Déterminer les critères de classement / tri des références et informations
- Mettre en forme l'information : les logiciels de gestion de bibliographie

Étape 6 Evaluer la démarche

- Evaluation des résultats obtenus (efficacité)
- Evaluation de la démarche (efficience)
- Evaluation à chaque étape

[La recherche concerne les catalogues et le web. Tendances : Il y a de moins en moins de recherche documentaire (à partir de catalogues ou de bases de données) aboutissant à des listes de références mais de plus en plus de recherche d'information : il est demandé au documentaliste d'extraire l'information.]

Définitions

■ La recherche documentaire

« actions, méthodes et procédures ayant pour objet de retrouver dans les [fonds documentaires les références des documents pertinents](#) ».

■ La recherche d'information

«actions, méthodes et procédures ayant pour objet d'extraire d'un ensemble de documents les informations voulues.»

Vocabulaire de la documentation, 2ème édition, Afnor, 1987

1

Ere de l'information primaire

- L'utilisateur final ne pense qu'à ça [tout doit être accessible tout de suite]
- Mythe de la bibliothèque virtuelle universelle
- Persistance d'une information secondaire massive
- Documents qui ne seront jamais en ligne
- Ressources en ligne payantes

2

La recherche par étapes

- Définir un objectif : sujet, finalité, contexte
- Faire l'inventaire : grilles d'analyse
- Formuler la recherche : principes d'interrogation, syntaxe, opérateurs
- Evaluer les résultats : tri, analyse de la pertinence
- Prélever, traiter et présenter l'information: sélection, présentation, diffusion
- Analyser : le résultat, la démarche

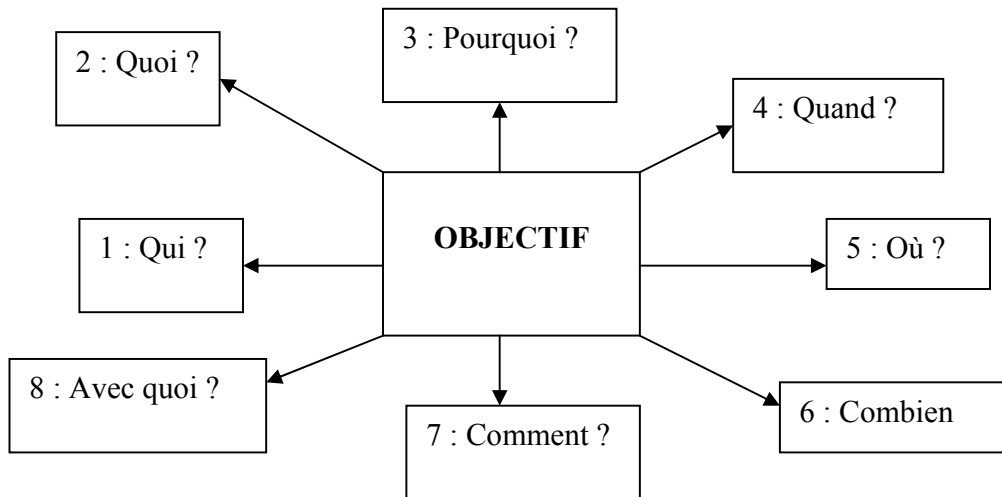
3

ETAPE 1 : Définir l'objectif de recherche

[= mobiliser les idées en vue d'une stratégie de recherche, étape à ne pas éluder !]

Définir l'objectif

4



[L'utilisateur a souvent des difficultés à formuler sa demande !!!

Quoi : Quel est le sujet, que connaît l'utilisateur sur le sujet ?

Pourquoi : dans quel but ? exposé, mémoire...

Quand ? pour quand ? délai ?, quelle période temporelle doit couvrir la recherche ?

Où : aire géographique

Combien : quantité d'information : la recherche doit-elle être exhaustive, synthétique, le coût est-il une limite ?

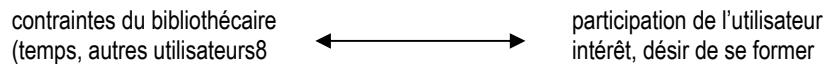
Comment : papier ? électronique seulement ?, quels types de documents ? texte, photos, formules...]

Typologie des demandes

Type de demande	Question	Objectif	Outils
Orientation	Demande brève, simple	Info. pratique immédiate	guides, plans
Factuelle	Question ponctuelle	Info. précise rapide	sources de référence
Bibliographique	Demande de références	Info. précise complète	bibliographies catalogues
Elaborée	Demande complexe	Info. élaborée interprétée	documents tertiaires

[A l'info on pourrait noter nos demandes selon cette répartition

- Niveaux de réponses
 - renseigner, orienter : l'utilisateur se débrouille
 - accompagner l'utilisateur à l'écran
 - montrer, commenter : apprendre à l'utilisateur sur le tas à utiliser catalogue bases de données...
 - former l'utilisateur à la recherches (séances de formation)
 - donner, synthétiser : l'utilisateur ne se prend pas en charge, le bibliothécaire fait pour lui, avec lui



ETAPE 2 : Faire l'inventaire des outils d'orientation et d'accès

Le bibliothécaire de référence doit bien connaître ses sources de référence. La bibliographie est aujourd'hui informatisée et permet une recherche thématique pointue que le bibliothécaire doit maîtriser. L'utilisateur attend du bibliothécaire un engagement professionnel, le bibliothécaire doit savoir s'appuyer sur les catalogues, les bibliographies, les encyclopédies, manuels, thèses, rapports ...]

Connaître les sources quel que soit le support

- **Sources factuelles (informations précises, brèves)**
 - annuaires, dictionnaires, répertoires : adresses, définitions, statistiques, image
- **Sources bibliographiques (documents secondaires)**
 - catalogues : critère de localisation dans bibliothèques ou chez éditeurs/libraires
 - bibliographies : critère intellectuel (sujet, date....)
- **Documents de synthèse (texte + ou - structuré)** [organisés (index, tables des matières), tertiaires],
 - encyclopédies, manuels, articles [de synthèse], thèses, rapports... 6

Connaître ses sources : grilles d'analyse

- Critères de forme
 - Maniabilité, lisibilité, typographie, forme des mises à jour
- Critères de contenu
 - Couverture
 - Type d'informations recensées
 - Domaine de la connaissance
 - Nature du recensement : exhaustif, sélectif
 - Aire géographique, linguistique, temporelle
 - Périodicité des mises à jour
 - Fiabilité, objectivité, technicité du contenu
 - Editeur et auteur
 - Présentation des notices
 - Contenu des notices
 - Organisation des accès
 - classement, sous-classements, tables et index
 - modes d'emploi, table des matières

[Par exemple medline ne contient que des articles. D'autres bases de données contiendront des chapitres d'ouvrages, des proceedings, des rapports...

Est-ce que la base de donnée convient vraiment au sujet ?

Exhaustivité versus sélectivité de la b. de d.

Périodicité des mises à jour

Contenu des notice, fiabilité, technicité du contenu, profondeur de l'information, l'information est-elle abordable pour l'utilisateur ? y a-t-il des résumés ? comment est organisé l'accès : tables des matières, index? présence d'un mode d'emploi utilisable

Outils pour le Web

- Les outils de recherche sont conçus pour une information fortement structurée
- Web
 - espace éditorial ouvert, enrichi et complexe :
- 9 millions de sites, <http://wcp.oclc.org/stats/size.html>
(Ohio College Library catalogue, maintenant groupement de bibliothèques)
 - un des nombreux protocoles diffusés sur Internet
- Conçu comme outil de diffusion
 - 665 millions d'internautes, <http://www.c-i-a.com/>
 - Non conçu comme outil de recherche
 - Développement de nouveaux outils pour rechercher des informations faiblement structurées

8

2.1 Les catalogues

- Définition
- Contenu
- Rôle
- Fonctionnalités : systèmes d'information normalisés, structurés, actualisés 9

Les catalogues

- Dispositif de saisie, stockage : regroupe les notices avec un nombre variable de champs
 - Eléments descriptifs + ou – normalisés : auteurs...(fichiers d'autorités)
 - Eléments textuels : titre, résumés, notes...(champs libres)
 - Eléments d'indexation, normalisés
- codes de classification : situent le document au sein d'une collection organisée
- descripteurs/mots matière : traduisent le contenu dans un langage documentaire donné
 - Eléments repères : date de saisie, nom indexeur

10

L'indexation

- Dispositif d'indexation
 - Indexation : processus destiné à représenter, au moyen des termes ou indices d'un langage documentaire ou au moyen des éléments du langage libre, les notions caractéristiques du contenu d'un document en vue d'en faciliter la recherche, après les avoir identifiés par analyse
 - indexation manuelle (humaine)
 - indexation automatique (de plus en plus fréquente)

11

Points d'accès

■ Les points d'accès sont des index

- issus de la description du document
- issus de l'analyse documentaire
- issus du traitement automatique du texte
 - **TALN : Traitement Automatique du Langage Naturel**
 - **L'informatique a permis la multiplication des index, donc des points d'accès**

Modes d'indexation

■ Indexation manuelle : logique de catégorie

- mots clés ou indices choisis par un être humain en fonction du contenu

■ Indexation automatique : traitement du texte

- traitement en texte intégral
 - indexation intégrale des chaînes de caractères associées au nombre d'occurrences, à la position dans le document [ce sont des logiciels qui sont ajoutés au logiciel de bse, ils permettent d'extraire un sens à partir d'une chaîne de caractères qui est plus ou moins répété. Permet une recherche avec « near »,]
- traitement linguistique, morphologique
 - traitement de la forme des mots : corrections orthographiques, lemmatisation [= ramener chaque terme à sa racine. P ex : il va → aller, chevaux → cheval], traitement phonétique
- traitement sémantique
 - assignation automatique de termes préférentiels

12

On s'achemine vers plus de complexité

- Plus que l'état d'une collection localisée
- Enrichissement du contenu
 - références de documents de nature hétérogène : articles, photos, vidéos, sites web, etc...
- Nouveaux modes de recherche
 - généralisation du principe de navigation : lien, clic, feuilletage, retour en arrière, etc.
- Liens sur les documents primaires

Mimétisme : les catalogues s'organisent comme des outils de recherche du web...

...les fonctionnalités du catalogue se diffusent aux outils de recherche web

13

2. 2 Web invisible

2.2.1 Les moteurs

Moteurs : collecte et index

- Robot d'exploration : spider, crawler
 - logiciel qui explore les réseaux de liens et collecte le contenu dans une base de données

- Module d'indexation : quelle couverture ?
 - nombre de sites, pages indexés : tout ou partie
 - nature des documents indexés (format html, pdf...)
 - champs indexés (TITLE, META, HREF...)
 - rythme de la surveillance, mise à jour des index
 - traitements automatiques
 - pondération des termes selon divers critères
 - traitement sémantique, linguistique
- Ce que les moteurs ne savent pas indexer :
 - certains documents (.pdf., .ppt, .doc, etc.)
 - les pages dynamiques produites à partir de BD
 - le contenu des bases de données [il n'y a pas de recherche sur le contenu des catalogues etc...]
 - les rubriques accessibles par mot de passe
 - les frames
 - les pages interdites de référencement (robots.txt)
- Web caché : ensemble des documents (textes, son, etc.) non indexés par les outils de recherche

14

Moteurs : exemples

	Alta Vista	Lycos	Hotbot	Google	All the Web
Création	1995 Digital	1995 Carnegie	1996 Wired	1998 Stanford	1999 Fast / Dell
Géré par	CMGI	Terra Lycos	Terra Lycos	Google	Fast
Spider	Scooter	Fast	Slurp	Google	Fast
Index*	1,1	2,5 All the Web	2 Inktomi	3	2,5 All the Web
Délai de rafraîchissement de l'index	1-112 j. 12 sem.	1-60 j. 4 sem.	1-59 j. 2 sem.	1-53 j. 4 sem	5-36 j. 3 sem.

* milliard de pages

15

Les métamoteurs (Copernic) couvrent plusieurs moteurs (12 en version gratuite, plus centaines en version payante) il y a perte de précision et de finesse de la recherche même si on gagne en nombre de données.

Métacatalogue de bibliothèques : <http://www.switch.ch/libraries/meta.html> et le métacatalogue européen de Karlsruhe (KVK)

2.2.2 Annuaire de recherche

- Définition (repertoires, directories)
 - catalogues de ressources classées par catégories et organisées hiérarchiquement
- Dispositif de collecte
 - travail de veille des documentalistes du web fait manuellement
 - autoréférencement (payant) par les éditeurs des ressources
 - sélection de sites, pas de pages, pas d'exhaustivité contrairement à un robot (moteur)

■ Dispositif d'indexation

■ indexation humaine [malheureusement les informaticien ont rebâtit leur propre classification au lieu de s'inspirer des bibliothèques ! les thèmes sont en nombre réduit mais il y a un grand nombre de sous-catégories en arborescences. On ne connaît pas les critère de conception de cet arborescence]

- classifications : explosion de nouvelles classifications
- mise à jour des sites, des catégories
- désherbage

16

[Les annuaires de recherche sont associés à des moteurs pour faire rebondir la recherche: ex: ODP (Open Directory Project: référencement par des éditeurs bénévoles) s'associe à Google, Yahoo: Inktomi, va laisser tomber Google)
On peut évoquer un autre type d'outil : les PORTAILS qui intègrent de façon verticale différentes fonctionnalités : sélectionnent personnalisent récupèrent et organisent les données, des ensembles hétérogènes de ressources.
Consulter le catalogue de la Joyner Library (East Carolina) qui offre différents types de documents. S'inspirer pour le Union Catalog de l'EPFL ?]

Annuaire : exemples

	Yahoo	Yahoo F	ODP	Voilà
Création	1994 Stanford	1997 Yahoo	1998	1996 France Telecom
Géré par	Yahoo	Yahoo	Netscape	France Telecom
Sites / Pages	2 000 000	150 000	3 800 000	70 000 15 millions
Thèmes / catégories	14 ?	14 35 000	16 460 000	16 7 500
Index associé	Google	Google	Google	Voilà

17

Analyse des outils web

■ Critères de forme

■ Ergonomie de l'interface : une interface différente peut utiliser un même moteur : MSN et Hotbot avec Inktomi

■ Critères de contenu

- Fournisseur / concepteur : public/privé/individus
- Type d'informations recensées : types de formats, fichiers..
- Organisation du contenu et des accès : organisation statique mais dynamique, selon quels critères ?
- Aides, explications, plan : dans quelle langue ?
- Fiabilité, objectivité : référencement payant ?
- Actualité et mise à jour : délais de rafraîchissement

<http://www.searchengineshowdown.com/>

[www.Abondance.com : analyse les outils web mondiaux. Il faut s'abonner.]

Choix des outils

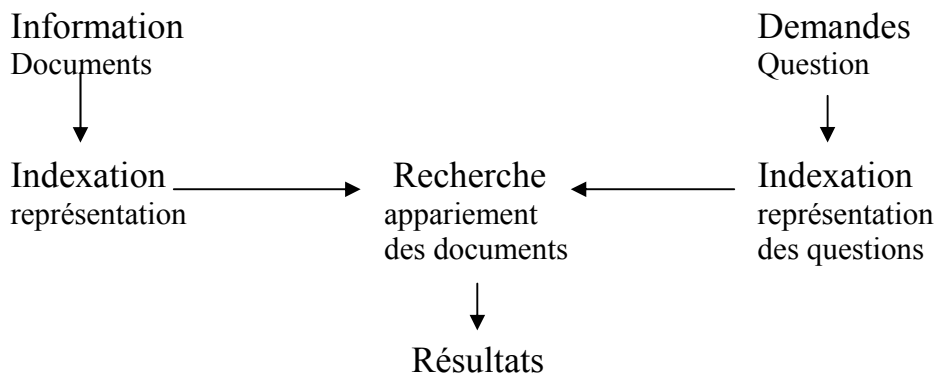
- Par quel type d’outils commencer ?
 - celui qui est susceptible d’offrir le plus de retour
 - il faut une bonne connaissance des modalités d’organisation et de fonctionnement des systèmes pour faire ce choix
- Doit-on commencer par chercher l’information
 - de la question elle-même ?
 - qui mène à l’information ?

ETAPE 3 : La recherche

Question : indexation

- Formuler une question c’est l’indexer !
 - Représenter au moyen de termes, indices d’un langage documentaire, ou au moyen du langage libre les notions caractéristiques d’une question
- Le succès de la recherche exige que le système et l’usager exprime les clés dans les mêmes formes
- Il faut harmoniser l’expression d’un même sujet par l’indexeur et par l’utilisateur occasionnel du système : recherche terminologique souvent indispensable

ETAPE 3 : Formuler la recherche



Papier versus électronique

Info. traditionnelle Support papier Lecture directe Index Sélection a priori	Info. Electronique Fichiers Equipement ad hoc Diversité des accès Ressources illimitées
Ne change pas Connaissance du besoin Connaissance des outils	Change Pas de stockage à priori Accès immédiat à l’info.

La recherche par étapes

- Définir un objectif : sujet, finalité, contexte
- Faire l'inventaire : grilles d'analyse
- **Formuler la recherche : principes d'interrogation, syntaxe, opérateurs**
- Evaluer les résultats : tri, analyse de la pertinence
- Prélever, traiter et présenter l'information : sélection, présentation, diffusion
- Analyser : le résultat, la démarche

Interfaces de recherche

- Interface web pour saisir la requête
 - facilite, uniformise l'accès
 - catalogues, moteurs, annuaires, BD
- Langage de commande (avec syntaxe)
- Formulaire
 - champs libres : langage naturel ou non
 - saisie assistée : liste simple ou structurée
 - liens depuis une liste de résultats (fonction « rebond »)
- Recherche sans saisie (index)
 - navigation sans expression de la requête
 - annuaires de recherche : parcours de hiérarchies
 - essai sur les catalogues

21

Syntaxes de recherche

- Commandes explicites ou implicites (= par défaut)
- Opérateurs explicites ou implicites
 - opérateurs de champs
 - opérateurs logiques ou booléens (ET, OU, SAUF)
 - opérateurs de proximité
 - troncature à gauche, à droite, masque (*, ?...)

*Ni les langages documentaires (thésaurus, notices...), ni les langages d'interrogation (troncature, booléens...) ne sont proposés de manière standard: il n'existe pas de véritable **code universel**, c'est un gros handicap !*

22

Syntaxe des moteurs

Opération	Opérateur	Exemple
Rechercher une expression	« »	«vache folle»
Imposer un terme	+	+vache +folle
Exclure un terme	-	+bretagne -grande
Tronquer un terme	*	vache*
Interrogation spécifique d'un champ	champ	Titre : t : url : u :

voir : <http://www.abondance.com>

[opérateurs numériques : =, <, >, ...

opérateurs booléens

opérateurs de proximité : with (côte à côte) near (dans la même phrase)

opérateurs de champ : t : , a : , s : , u : ...]

Fonctionnement implicite

- Mots vides
 - Filtrés ou non
 - Forcer la recherche des mots vides
- Majuscule / minuscule différenciée ?
- Lettres accentuées différenciées ?
- Opérateur booléen implicite ?
- Troncature : implicite, explicite ou absente

Corrections et reformulation

- Corrections automatiques
 - traitement phonétique, métrique (distance entre le terme erroné et le lexique)
 - traitement linguistique et sémantique
- Reformulation de la recherche
 - manuelle : historique de la requête, utilisation de filtres en recherche avancée (type, langue, date...)
 - automatique : relance à partir d'un champ, voire une référence pertinente, proposition (moteurs...)
 - reformulation interactive: le système soumet un choix denouveaux termes à choisir 24

Synthèse

<i>Modes de recherche</i>	<i>Démarche intellectuelle</i>	<i>Type d'information</i>	<i>Exemples d'outil</i>
Recherche par Navigation arborescente	Figure de l'Arbre Démarche Systématique Du général Au particulier Par menu successifs	Information structurée Organisée en Plan de classement Information secondaire	Tables des matières Classifications documentaires Annuaire thématiques
Recherche par Requête sur les Métadonnées Du document	Principe de l'index Démarche D'indexation de L'information Recherche par champs, logique booléenne	Information Structurée en Champs Information secondaire	Index des livres Banques de données Catalogues de bibliothèques
Recherche par Navigation hypertextuelle	Figure du Réseau Démarche associative , d'une notion à l'autre. Navigation dans un réseau de nœuds et de liens	Information non structurée Information brute primaire	Renvois dans une Encyclopédie Hypertextes Sur CDRom Sites web
Recherche par Requête sur le texte intégral	Parcours de texte Démarche d'analyse Linguistique Recherche Contextuelle sur le contenu	Information non structurée Information brute primaire	Outils de TALN Moteurs de Recherche Outils linguistiques

Changement d'outils

- Changement de catalogues, BD, moteurs
- Métacatalogues
 - systèmes d'interrogation multi-bases
- Métamoteurs
 - interrogent simultanément par le biais d'une interface unique plusieurs outils de recherche
- Moteurs humaines
 - services de réponses par des professionnels aux questions des internautes en direct ou différé
 - <http://answers.google.com/answers/main>

ETAPE 4: Evaluer les resultats

Tri, sélection, évaluation

- Tri prédéterminé sur des critères classiques
 - auteurs, titres, dates , peu liés au contenu
- Tri par pertinence : relevance feedback
 - fréquence, localisation, proximité des termes
- Tri par indice de popularité
 - nombre de liens pointant sur le document (= IP ?)
 - indice de clic (nombre de clics = nombre de prêts ?)
- Tri par lot
 - classement à la volée selon similitude entre documents et la requête

26

Affichage

- Représentation linéaire, textuelle, + ou - figée
 - Représentation partielle du document pour première sélection avec résumés ou pas
 - Affichage du mot ou groupe de mots recherché dans son contexte (surbrillance)
 - Filtres à l'affichage choisis ou non par utilisateur
- distinction par nature d'information : news, vidéo...
- distinction par référencement : liens sponsorisés...
- Représentation graphique
 - Cartes de connaissance établies dynamiquement : des noeuds (les documents) et les liens entre eux

27

Evaluation

- Bruit et silence
 - documents non pertinents parasitent / pertinents n'apparaissent pas, [ce qui oblige à réévaluer l'équation de recherche, p ex combiner des synonymes par « ou », changer de base de donnée...]
- Rappel et précision : documents pertinents
 - leur proportion / ensemble documents pertinents
 - leur proportion / ensemble documents retrouvés
- Pertinence
 - «relevant» : par rapport à la requête
 - «pertinent» : par rapport aux attentes
 - ce qui est conforme à la demande (relevant) pas toujours pertinent pour l'utilisateur (utilité immédiate).

28

ETAPE 5: Prélever. Traiter, présenter l'information

Traitement

- Sauvegarde
 - sauvegarde de la requête pour la suite
 - sauvegarde des adresses (signets)
 - sauvegarde, exportation des informations
- texte, pages, images

- Reformatage, mise en forme
 - traitement informatique
 - traduction : logiciel sur Alta Vista ou Google
- Elaboration et synthèse

29

ETAPE 6: Evaluer la demarche

Evaluation globale

- Information
- Démarche
- Elaboration et synthèse
- Suivi et veille

Quelques signets

Recherche documentaire

Quelques signets....

Annuaire et moteurs

Google : <http://www.google.ch/> (pour la Suisse) et international

Open Directory : <http://dmoz.org/>: annuaire de recherche du web (dmoz. ch pour la Suisse)

Profusion : <http://www.profusion.com/> : meta moteur en ligne

Yahoo : <http://www.yahoo.com/>: annuaire de sites lié au moteur Google
(aussi yahoo.fr pour pages francophones)

Bases de données

exemples de bases de données :

<http://services.inist.fr/public/fre/constl.htm> : catalogue INIST (répertoire d'articles, mémoires...)

Medline Pubmed : <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi> : une des versions de
Medline (gratuite)

Catalogues

Catalogue 1 - RERO : <http://www.rero.ch/eroweb/chameleon.html>

Regarder le format des notices, les possibilités de tri des notices, les liens internes qui pointent sur les index (titres, auteurs)

Catalogue 2 - Bibliothèque Nationale Québec : <http://www.bibliat.gouv.qc.ca:6611/>

2 types de liens internes : liens sur les index (auteur, titre) séparés des liens sur les notices (notices du même auteur)

Catalogue 3 - Joyner Library : <http://systems.lib.ecu.edu/piratesource/>

Formulaires et options adaptés à chaque étape de la démarche

Catalogue 4 : Karlsruhe Virtueller Katalog : <http://www.ubka.uni-karlsruhe.de/kvk.html>

Métacatalogue européen interface Z39.50 ou accès par intervention sur URL

Catalogue 5 - Locator : <http://locatorplus.gov/cgi-bin/Pwebrecon.cgi?DB=local&PAGE=First>

National Library of Medicine ; affichage des résultats par pertinence (logiciel Endeavour)

Catalogue 6 - CCF : <http://www.ccf.bnf.fr/>

Catalogue collectif de France par interface Z39.50

Catalogue 7 : interface SIRSI : <http://www.whittierpl.org/uhtbin/cgisirsi/x/0/49>

Whittier Public library (regarder option Find it fast !)

Répertoires de catalogues

Répertoire de catalogues et métacatalogues suisses: <http://www.switch.ch/libraries/meta.html>

Répertoire de catalogue mondial : <http://www.libdex.com/> : Libdex is a worldwide directory of library homepages, web-based OPACs, Friends of the Library pages, and library e-commerce affiliate links.

Evaluation

Quelques grilles d'évaluation de sites web

Centrale santé / Netscoring : <http://www.chu-rouen.fr/netscoring/>

Outil très complet pour évaluer les sites web surtout dans le domaine de la santé

Cerise : <http://www.ccr.jussieu.fr/urfist/cerise/p361.htm>
Cerise : les questions de base à se poser
Evaluation : identification d'un nom de domaine : <http://www.allwhois.com/>
Base de données d'adresse de serveur qui permet d'identifier un nom de domaine
Evaluation : indice de popularité AV : <http://www.altavista.com/web/webmaster>
Nombre de liens qui pointent sur un site par Alta Vista
Evaluation : indice de popularité Google : http://www.google.ch/advanced_search?hl=fr
Liens qui pointent sur un site par Google
Temesis : <http://www.temesis.com/annuaire/annuaire.php>
Liste des sites évaluant la qualité des sites web

Formation

Sites francophones sur la recherche documentaire
Formation ENSSIB : Formist : <http://formist.enssib.fr/>
Formation à l'usage de l'information dans l'enseignement supérieur.
Formation ENSSIB : Repere : <http://repere.enssib.fr/frontoffice/index.asp>
Ressources électroniques pour les étudiants la recherche et l'enseignement
Formation INSA : SAPRISTI : <http://csidoc.insa-lyon.fr/sapristi/digest.html>
Sentiers d'accès et pistes de recherche d'informations scientifiques et techniques sur l'Internet
Formation UQAM : Infosphere : <http://www.bibliotheques.uqam.ca/InfoSphere/>
Pour apprendre à faire une recherche d'information
Formation URFIST Lyon : RISI : <http://urfist.univ-lyon1.fr/risi/risi.htm>
Recherche d'information sur Internet, JP Lardy
Formation URFIST Paris : Cerise : <http://www.ccr.jussieu.fr/urfist/cerise/>
Conseils aux Etudiants pour une Recherche d'Information Spécialisée Efficace
Formation URFIST Rennes : <http://www.uhb.fr/urfist/index.htm>
Supports de formation et travaux pratiques
Formation UlaVal : GIRI : <http://www.bibl.ulaval.ca/vitrine/giri/>
Guide d'initiation à la recherche dans Internet
Formation Uni. Poitiers : <http://www.sha.univ-poitiers.fr/documentation/tpinternet/listetp.htm>
TP Internet

Statistiques : Statistiques internet

Computer industry almanach : <http://www.c-i-a.com/>
Statistiques sur le web
Searchengineshowdown : <http://www.searchengineshowdown.com/>
Site de veille sur les moteurs de recherche
Statistiques OCLC : <http://wcp.oclc.org/stats.html>
Statistiques OCLC concernant Internet

Suivi et traitement : Outils de suivi et de traitement

Reference Manager : <http://refman.com/>
Outil de gestion bibliographique
Signets (outils de gestion) : <http://www.megagiciel.com/logiciels/313.html/>
Différents outils permettant de gérer des signets

Bibliographie

Recherche documentaires, les outils internet **Quelques ouvrages récents**

- ANDRIEU, Olivier. Trouver l'info sur l'Internet. 4ème éd. Paris : Eyrolles, 2000. XI-425 p.
- CHARTRON, Ghislaine. Recherche d'information sur Internet. In LE MOAL, Jean-Claude, HIDOINE, Bernard (sous la dir. de). La Recherche d'information sur les réseaux. Internet : pour en savoir plus. Cours INRIA, 1996, Trégastel. Paris : ADBS, 1996. p. 43-101
- DALBIN, Sylvie. Instrument de recherche sur le web. In LE MOAL, Jean-Claude, HIDOINE, Bernard (sous la dir. de). La recherche d'information sur les réseaux. Cours INRIA, 30 septembre - 4 octobre 2002, Le Bono (Morbihan). Paris : ADBS, 2002 . p. 11-71
- DENECKER, Claire. Les compétences documentaires : des processus mentaux à l'utilisation de l'information. Villeurbanne : Presses de l'Enssib, 2002 . 208 p.
- FOENIX-RIOU, Béatrice. Recherche et veille sur le web visible et invisible : agents intelligents, annuaires sélectifs, interfaces des grands serveurs, portails thématiques. Paris : Bases, Ed. TEC&DOC, 2001. VI-233 p.
- LAHARY, Dominique. Des bibliothèques traditionnelles aux « bibliothèques virtuelles ». In LE MOAL, Jean-Claude, HIDOINE, Bernard (sous la dir. de). La recherche d'information sur les réseaux. Cours INRIA, 30 septembre - 4 octobre 2002, Le Bono (Morbihan). Paris : ADBS, 2002 . p. 169-203
- LARDY, Jean-Pierre. Recherche d'information sur Internet : méthodes et outils. 7e éd. Paris : ADBS, 2001. 123 p.
- LEFEVRE, Philippe. La Recherche d'informations : du texte intégral au thésaurus. Paris : Hermès, 2000. 253 p.
- MANIEZ, Jacques. Actualité des langages documentaires : les fondements théoriques de la recherche d'information. Paris : ADBS, 2002. 395 p. (Collection Sciences de l'information. Série études et techniques)
- MORIZIO, Claude. La Recherche d'information. ADBS : Nathan, 2002. 126 p. (Nathan université – collection 128 ; 267. Information, documentation)

Des sites web sur la recherche documentaire.

Fourmi, Formist, Cerise, Repère, Risi, Sapristi... Ces sigles ésotériques renvoient à des sites qui hébergent les sites francophones majeurs consacrés à la recherche d'information

Tous ils rappellent deux évidences : d'une part « tout n'est pas sur la Internet », d'autre part l'obtention de résultats satisfaisants passe par une stratégie de recherche cohérente.

Fourmi, Cerise et les documents de l'Urfist de Paris

Les « Unités régionales de formation à l'information scientifique et technique » (Urfist) ont pour mission de sensibiliser aux nouvelles techniques de l'information le personnel des bibliothèques et les étudiants du 3e cycle. À ce titre, certaines d'entre elles (Paris, Lyon, Rennes, Strasbourg, notamment) mettent en ligne des pages traitant aussi bien de la documentation en général que de la recherche documentaire sur l'Internet en particulier

- L'espace « Formations universitaires en réseau aux méthodologies de l'information » également appelé Fourmi, recense les ressources existantes en formation documentaire et crée un fonds de documents pédagogiques dans ce domaine. Au sommaire de ces ressources figurent, entre autres, des cours et des fiches techniques sur la recherche documentaire, des conseils pour la rédaction et l'archivage des références bibliographiques.
<http://www.ccr.jussieu.fr/urfist/fourmi.htm>
- Dans l'espace « Cerise »: conseils aux étudiants pour une recherche d'information spécialisée efficace » deux documents méritent qu'on s'attarde : « pistes sur l'Internet » qui propose une liste d'adresses pour chaque type de ressource et « instruments de recherche » qui, à chaque outil, fait correspondre un type de recherche
- <http://www.ccr.jussieu.fr/urfist/cerise/index.htm>

L'Urfist de Lyon héberge deux espaces originaux : RIsI et DADI

- RIsI : Recherche d'informations sur l'Internet, dresse la liste de tous les outils (en une seule page) et fournit des données actualisées sur les différents types d'outils (répertoires, robots, métamoteurs) et sur les problèmes posés par la recherche d'informations sur Internet.
<http://urfist.univ-lyon1.fr/risi/risi.htm>
- DADI : Répertoire des bases de données gratuites disponibles sur l'Internet, recense les banques de données en accès gratuit. Régulièrement actualisé.
<http://urfist.univ-lyon1.fr/gratuits/index.php>

Cours de l'Urfist de Rennes

- Cet espace accueille de nombreux liens sur des Documents pédagogiques, supports de cours, guides, dossiers consacrés à la recherche documentaire.
<http://www.uhb.fr/urfist/supports>

Les espace FORMIST et REPERE (Ressources électroniques pour les étudiants la recherche et l'enseignement) disponibles sur le serveur de l'ENSSIB permettent également d'accéder à des ressources pédagogiques.

- FORMIST : Formation à l'usage de l'information dans l'enseignement supérieur.
<http://formist.enssib.fr/>
- REPERE (Ressources électroniques pour les étudiants la recherche et l'enseignement)
<http://reperere.enssib.fr/>

Sapristi ou les « Sentiers d'accès et pistes de recherche d'informations scientifiques et techniques sur l'Internet » de l'Insa de Lyon font partie des sites pionniers en recherche d'information sur Internet.

- À signaler : une page « tableau de bord » qui rassemble sous une forme synthétique les principaux outils de recherche avec leurs avantages, leurs inconvénients, leur syntaxe et des liens vers leur module d'aide et (ou) leur fiche technique sur le site «Abondance»
<http://csidoc.insa-lyon.fr/sapristi/digest.html>

Université de Poitiers, mention documentation : cet espace héberge des pages dont l'intérêt est double, car elles constituent à la fois une initiation à la documentation et une tentative (pour certains thèmes) d'enseignement à distance

- Deux thèmes méritent une exploration attentive. Le premier, intitulé « Former à l'Internet, comment et pourquoi ? » est une réflexion sur « faut-il former à l'Internet en limitant cette formation aux outils et aux méthodes de recherche ou, plus largement, faut-il développer et intensifier l'apprentissage de l'univers informationnel ? ». Le second sert de base de travail pendant les séances de travaux pratiques. Il est constitué d'une synthèse sur les outils de recherche et d'une information sur les sites « fédérateurs » (sélection des ressources concernant un domaine). Il est relié à deux modules essentiels : « stratégies de recherche » et « formuler sa question ».
<http://www.sha.univ-poitiers.fr/documentation/formerinternet.htm>

Université de Laval – Canada

- **GIRI - Guide d'initiation à la recherche dans Internet**
<http://www.bibl.ulaval.ca/vitrine/giri/>

La recherche documentaire en ligne - T. BASAGLIA bibliothécaire du service de référence au CERN (pas un scientifique)

1. L'entretien de référence : quelques aspects pratiques

Problèmes :

- L'utilisateur ne sait pas formuler précisément sa demande, il ne la formule pas dans sa langue maternelle mais en anglais, l'anglais n'est pas la langue maternelle du bibliothécaire, culture anglo-saxonne versus culture latine
- le bibliothécaire n'a pas les compétences scientifique de l'usager : en plus de maîtriser ses outils de travail il doit se créer une compétence dans les domaines de la bibliothèque.
- L'utilisateur pense bibliothèque virtuelle et s'attend à avoir tout de suite toute l'information à télécharger ou imprimer, ou bien
- L'utilisateur a un complexe vis à vis du bibliothécaire, est parfois gêné de déranger
- Les formations organisées par la bibliothèques proposées aux chercheurs n'ont pas de succès. Or laisser un utilisateur seul devant un outils qu'il maîtrise mal fait qu'il sera peu efficace dans ses recherches

Conseils :

- Contact direct (l'usager se rend au guichet) : une banque (desk) polyvalente : users' services ou se fait l'accueil, l'information, l'assistance à l'utilisateur, le prêt, quitte à rediriger l'utilisateur vers un reference librarian (Fachreferent)
- Contact virtuel, service « ask a librarian » ou « ask a physician » le mail arrive à la banque unique qui fait le tri.
- Les formations organisées fonctionnent très mal, mais attention, un usager mal formé qui fait sa recherche passe à côté de beaucoup de réponses
- Communication verbale : analyser les questions
 - Questions de l'utilisateur:
 - directionnelles : pratiques, où se trouve quoi
 - informationnelles : comment ?
 - instructionnelles : demande de formation ciblée
 - Questions du bibliothécaire : don pour la communication verbale
 - Reformuler la question, paraphraser, situer (générique – spécifique)
 - Questions ouvertes : est-ce que ? pensez-vous que ? quel aspect vous intéresse ?
 - Questions fermées : suggérer une direction, inclusions, exclusions
- Contraintes de temps, d'argent (banques de données payantes plus performantes que les même gratuites)
- Code de conduite : questions auxquelles le bibliothécaire ne peut pas répondre : éthique, médical, légal
- Code de conduite de l'utilisateur : le bibliothécaire doit rendre l'utilisateur conscient des contraintes financières et du « fair use » des documents électroniques

- Evaluations intermédiaires pour éventuellement recentrer la recherche
- Evaluation finale

2. Le service de référence virtuel

Autrefois le bibliothécaire était le relais indispensable entre l'utilisateur et la banque de données accessible en ligne uniquement

Actuellement les banques de données sont sur internet ou l'intranet (licences de site) d'une école, d'une université ou entreprise ; l'utilisateur y a un accès direct, donc un pouvoir, une influence..

Le bibliothécaire est informateur : il doit rendre l'accès simple et convivial, il doit trier et évaluer la pertinence des informations qu'il met à disposition (impact factor), il doit créer des pages web très soignées, des modes d'emploi des différents catalogues et bases de données , des portails, des catalogues enrichis, voire une plate forme unique (exemple de la bibliothèque de Los Alamos qui a acheté les données brutes d'INSPEC pour les insérer dans son catalogue !)

Le bibliothécaire est également formateur , il crée des modules de formation en ligne.

Le spécialiste, chercheur avancé a besoin de communiquer rapidement avec ses pairs

→ il faut tenir compte du fait que ce type d'utilisateur va privilégier la documentation électronique qu'il peut avoir directement et instantanément . Il est donc nécessaire de privilégier les ressources qui offrent des liens avec le full text. Sinon on peut établir un lien script avec le formulaire de demande de prêt interbibliothèques.

→ Succès des Open archives, preprints, e-publication des centres de recherche au détriment des bases de données bibliographiques traditionnelles basées sur le dépouillement de document issu de la publication éditoriale payante ; l'impact factor de la revue et la peer review de l'article sont mis en balance avec rapidité et la gratuité de l'information distribuée immédiatement et rapidement par et dans un cercle de chercheurs qui se connaissent, et dont on a quand même une information sur la notoriété au vu de leurs publications passées et évaluées par la voie traditionnelle. De plus la notoriété du centre de recherche éditeur est

Bases de données versus moteurs de recherche du web :

Est-ce que google a gagné ? En tout cas les interfaces des bases de données ont tendance à imiter google...

L'avenir du service de référence : le bibliothécaire doit faire face à l'augmentation et à la diversité des ressources, les rechercher et les mettre à disposition de l'utilisateur, former ces derniers et enfin mettre à jour continuellement ses compétences professionnelles...

... et l'avenir des bibliothécaires en général : le bibliothécaire doit être webmaster, promoteur de ses services et négociateur en matière de ressources électroniques (contact avec les éditeurs, regroupement en consortium...)

Bibliographie

- L'entretien de référence:

1. **William A. Katz**, "Introduction to reference work, vol. II: Reference Services and Reference Processes", 3rd ed., New York, McGraw-Hill, 1978, p. 61-119
2. **Jean-Philippe Lamy**, *L'entretien de référence: une notion à introduire dans un enseignement rénové*, BBF, vol. 43, no. 6, 1998, p. 52-58. Version en ligne: <http://www.enssib.fr/bbf/bbf-98-6/07-lamy.pdf>
3. Ohio Reference Excellence Web-based Training: <http://www.olc.org/ore/2open.htm>
4. **Carol Tenopir**, *Virtual Reference Services in a Real World*, Library Journal, vol. 126, no. 12, 2001, p.38,40. Version en ligne: <http://libraryjournal.reviewsnews.com/index.asp?layout=article&articleid=CA90455&display=Online%2BDbsNewsMore&industry=Online%2BDbs&verticalid=151&starting=21>

- Les sources d'information:

1. **Steven J. Bell**, *Has Google won? A Librarian Says Students Have More Data Than They Know What to Do With*, Chronicle of Higher Education, 23.1.2003. Version en ligne: <http://chronicle.com/free/2003/01/2003012301t.htm>
2. **Diane K. Kovacs**, "Evaluating Internet Information". Version en ligne: <http://www.kovacs.com/eval/eval03.html>
3. **Bruce Royan**, *Content creation and rights management: experiences of SCRAN (the Scottish Cultural Resources Access Network)*, Program, vol. 34, no.2, p. 131-142.
4. "La BnF et les nouvelles technologies" (Gallica). Version en ligne: <http://www.bnf.fr/pages/connaitr/num.htm>

- La recherche:

1. **Corinne Verry-Jolivet** (sous la direction de), "Créer et gérer un service de référence", Villeurbanne, Institut de Formation des Bibliothécaires, 1996
2. **Chris Sherman**, "The invisible Web: uncovering information sources search engines can't see", Medford: CyberAge Books, 2001.

3. Le CERN et les archives ouvertes

Voir le site http://library.cern.ch/cern_publications/cern_publication_policy.html