



Archives Web Suisse

Une collection commune de sites web patrimoniaux des bibliothèques cantonales suisses et de la Bibliothèque nationale suisse

Atelier de formation à la collecte et à la saisie de sites web patrimoniaux

les 23 et 30 avril 2008 à la Bibliothèque nationale suisse

Partie 1: Collecte

Rédigé par:

Silvia Hofmann (Prestations en ligne, Bibliothèque nationale suisse)

Actualisé par:

Anne-Katharina Weilenmann (Information au public, Bibliothèque nationale suisse)

Avec le soutien de:

Jean-Philippe Accart, Simone Bleuler (Information au public, Bibliothèque nationale suisse),
Marco Vogt (Prestations en ligne, Bibliothèque nationale suisse)

Traduction française par:

Liliane Regamey (Section Utilisation, Bibliothèque nationale suisse)

Version:

Version 1.1

Date:

15 avril 2008

1	Table des matières	
1	Table des matières	2
2	Introduction	3
3	Partie théorique (moteurs de recherche)	4
3.1	Catalogage du web	4
3.2	Fonctionnement des moteurs de recherche, métas-moteurs de recherche et catalogues	6
3.3	Typologie des moteurs de recherche	8
3.4	Quel outil pour quelle recherche ?	9
3.5	Comment sont utilisés les moteurs de recherche ?	10
3.5.1	Utilisation de la recherche avancée	10
3.6	Langage de recherche (syntaxes de recherche)	11
3.7	Quelles sont les stratégies de recherche possibles sur internet?	11
3.8	«Trucs» pour améliorer la recherche	11
3.9	Classement (ranking)	12
3.10	Critères pour l'évaluation de la qualité des résultats d'une recherche	13
3.11	Critères généraux d'évaluation des pages web du point de vue des spécialistes de l'information	16
3.12	Critères de qualité des moteurs de recherche	17
3.13	Ce que les moteurs de recherche ne trouvent pas	17
4	Partie pratique	20
4.1	Méthode de recherche des Helvetica en ligne d'importance patrimoniale	20
4.2	Procédure pour la transmission des Helvetica en ligne d'importance patrimoniale	21
4.3	Exemples	22
4.3.1	Recherche: Cervin et géologie	26
4.3.2	Recherche Basler Fasnacht	31
4.3.3	Recherche: Agriculture dans le canton de Fribourg	34
5	Concepts	38
6	Sources	41

2 Introduction

Internet contient d'innombrables informations reliées entre elles, ce qui permet à chacun d'accéder d'une façon simple au savoir.

Cette simplicité géniale de l'outil peut pourtant conduire à des difficultés.

Accéder aux informations de cet insaisissable réseau, se promener de lien en lien sans prêter attention à son environnement donne une impression de grande facilité à l'utilisateur. Il a l'illusion de pouvoir trouver des réponses à toutes ses questions en un clic. Il lui est toutefois difficile de filtrer dans la masse des informations celles qui sont pertinentes pour lui. Par sa progression de lien en lien dans ce réseau serré, il est exposé sans s'en apercevoir à la linéarité de ce réseau. Il confectionne un tissu (un texte) qui se compose de mailles arbitrairement mises côte à côte. Les composantes de la forme ainsi produite ne s'accordent pas, elles se côtoient sans pouvoir créer un sens. (Weilenmann 2006)

Les moteurs de recherche existent depuis une dizaine d'années. La recherche d'informations, qui est restée longtemps une affaire d'experts, est accessible à l'utilisateur non professionnel.

Ces moteurs ne sont bien sûr pas comparables aux outils «professionnels» de recherche, comme par exemple les banques de données commerciales. Celles-ci se servent de langages de recherche d'une richesse et d'une complexité bien supérieures.

Les moteurs de recherche ne réussissent pas à trouver tous les trésors qui sommeillent dans le web, cette partie qui est nommée depuis longtemps déjà «invisible web» ou «deep web». On estime qu'elle est environ 500 fois plus grande que le «web conventionnel».

La recherche au moyen d'un moteur de recherche est la façon la moins efficace d'aller à la quête d'informations dans internet. L'utilisation de listes de liens, de portails, etc. promet de bien meilleurs résultats.

Ce dossier contient, outre les bases théoriques sur les moteurs de recherche, une partie d'exercices avec des exemples de recherches, une proposition pour les procédures de recherche des Helvetica d'importance patrimoniale en ligne et des règles d'évaluation des résultats de la recherche sur internet.

3 Partie théorique (moteurs de recherche)

3.1 Catalogage du web

Lors du catalogage du web, il est particulièrement important de se souvenir que la quantité de documents du web n'est pas connue et ne peut pas être déterminée ; la saisie est rendue difficile par la structure hypertexte des informations. Ces problèmes n'existent pas dans le cas des banques de données classiques. Là, la quantité de données à saisir est connue, du fait des limites et des règles posées dès la conception de la base de données. Les problèmes pour trouver de nouveaux documents se posent d'une autre façon.

Dans le web, les documents sont potentiellement dans toutes les langues, ce qui pose un problème pour leur description. Etant donné que les moteurs de recherche ne se basent pas sur un vocabulaire d'indexation homogène, mais indexent le texte intégral, seuls les documents qui sont dans la langue du terme de la recherche pourront être trouvés. Dans une banque de données, soit les documents sont tous dans la même langue, soit ceux qui sont dans une autre langue sont décrits au moyen d'un vocabulaire uniforme dans la langue cible de la banque de données. Une aide supplémentaire est apportée par des systèmes de classification et des thesauri multilingues. Les différents types de médias et de formats des données posent un problème supplémentaire, lié à la diversité du web. Le web n'est pas limité aux documents textuels, il contient aussi, par exemple, beaucoup d'informations sous forme de multimédias. La description de ces informations se fera d'une façon différente de celle d'un document textuel. Les très grandes différences de longueur des documents sur le web sont également problématiques. On y trouve aussi bien des documents qui ne contiennent que de peu de mots que des livres entiers, considérés comme un seul document. Certains longs documents sont fractionnés afin de faciliter leur accès.

Le peu de structure des documents est un des gros problèmes de l'indexation du web. Il est vrai que l'on trouve une structure dans les documents HTML ; mais elle n'est pas utilisée avec cohérence par les auteurs, ce qui rend la description de ces documents difficile.

Dans les banques de données, chaque document est unique, la saisie des documents obéit à des critères clairs ; dans le web, du fait de sa structure décentralisée, on trouve une grande quantité de doublets. D'une part, des serveurs complets sont dupliqués (en miroir), d'autre part, les mêmes textes sont intégrés dans différentes offres. (Lewandowski 2005, p. 5-6)

Aperçu des différences entre la recherche d'information dans le web et la recherche d'information classique (Lewandowski 2005, p.8)

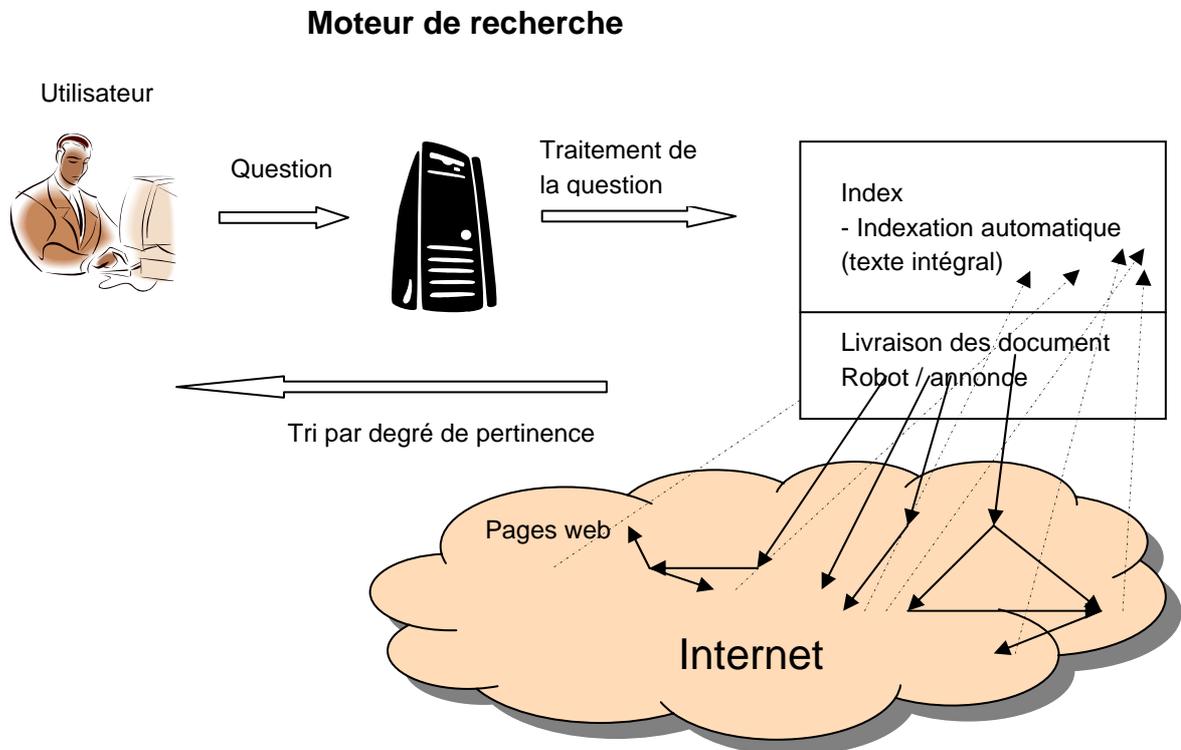
Caractéristiques	Web	Banque de données classiques
Caractéristiques du corpus de documents langue	Les documents sont en différentes langues; l'indexation du texte intégral ne permet pas de différencier les langues	Langues uniques ou langues des documents prédéfinies ; l'indexation des documents en différentes langues est faite au moyen d'une langue unique
Type de documents	Documents de formats différents	En principe, documents d'un seul format
Longueur et « granularité ¹ » des documents	La longueur des documents varie, les gros documents (par ex. les livres entiers) sont souvent fractionnés	La longueur des documents varie dans une mesure raisonnable ; chaque document est une unité

¹ Nombre de subdivisions dans un document

Caractéristiques	Web	Banque de données classiques
Spam	Problème des contenus non désirés trouvés par les moteurs de recherche	Les documents qui seront saisis sont définis au moment de la conception des bases de données
Structure «hyperlien»	Les documents sont liés les uns aux autres	Les documents ne sont en principes par reliés ; pas de nécessité de prouver la qualité du document par l'usage de liens
Contenus Quantité des données / taille du stock de données	La quantité des données ne peut pas être déterminée ; pas d'indexation exhaustive possible	La quantité de données peut être déterminée sur la base de critères formels précis
Exhaustivité des données	Exhaustivité peu certaine	De par la définition du contenu de la banque de données, l'exhaustivité est en principe garantie
Doublets	Les documents peuvent apparaître à plusieurs reprises, quelques fois dans des versions différentes	Contrôle des doublets lors de la saisie des documents. Le contrôle des versions n'est en principe pas nécessaire car il existe une version définitive qui est celle qui est retenue pour la banque de données
Utilisateur Intérêts différents	Les besoins en information de l'utilisateur sont hétérogènes	Groupes ciblés avec précision, besoins en information clairs
Interrogation	L'utilisateur a peu de connaissances des fonctions de recherche proposées	L'utilisateur est familiarisé avec le langage d'interrogation
Système de recherche IR-System Interface	Simple; interface utilisable d'une façon intuitive par des non-professionnels	Interfaces souvent complexes, il est nécessaire d'y être habitué
Classement	Du fait du grand nombre de résultats, un classement par pertinence est indispensable	Le classement par pertinence n'est pas nécessaire du fait que la question est posée avec précision et que les résultats sont peu nombreux
Fonctions de recherche	Fonctions de recherche limitées	Langage de recherche complexe
Modification de la recherche	En général, la seule possibilité est de préciser la recherche dans un 2ème temps	Nombreuses possibilités de modification
Structure des documents indexés	Structure faible; les champs définis dans la recherche avancée sont peu utilisables; pas d'exclusion des spams; pas de possibilité d'affiner les recherches	Structure forte; bonne définition pour la recherche dans un champ précis; les critères de sélection sont définis dès la conception de la banque de donnée

3.2 Fonctionnement des moteurs de recherche, métas-moteurs de recherche et catalogues

Représentation schématique de moteurs de recherche (Griesbaum / Bekavac, 2004, p.38)

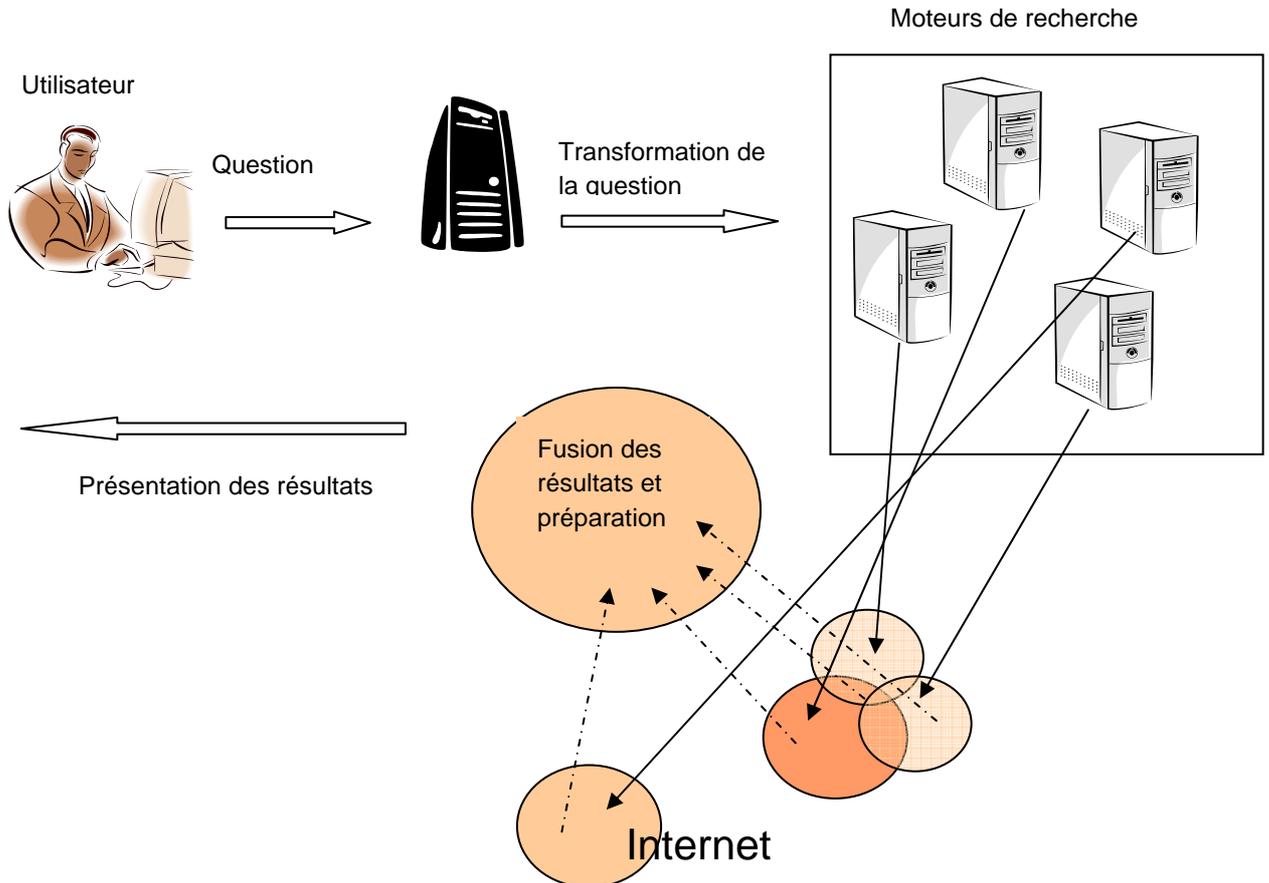


Les moteurs de recherche sont composés essentiellement de trois éléments :

- Acquisition de documents, moissonnage des bases de données à l'aide de robots ou « spiders » (Crawler/Worm/Wanderer). Annonces d'URLs par les usagers
- Indexation automatique à l'aide de logiciels d'indexation – Index
- Traitement de la demande et présentation des résultats à l'aide d'une fonction de recherche qui détermine le nombre de résultat et leur classement – Classement / Ranking



Méta-moteur de recherche

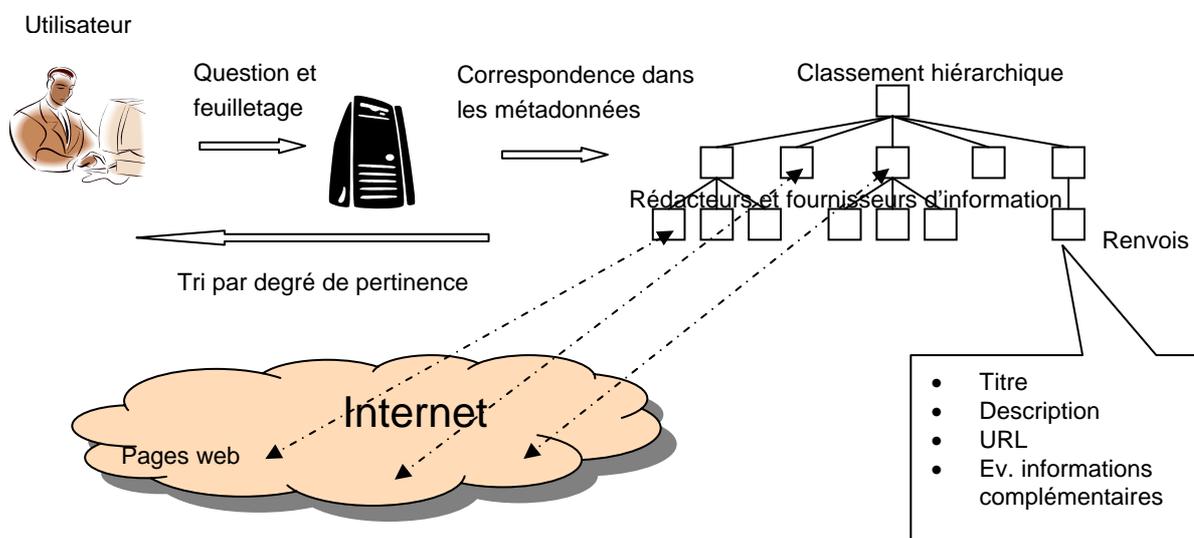


Les méta-moteurs de recherche interrogent plusieurs sources avec une interface, en général des moteurs de recherche et des catalogues, quelquefois aussi d'autres sources comme des dictionnaires

- les fonctionnalités et les opérateurs des différentes sources ne peuvent être en général utilisés que partiellement

Truc: comme les différents services ne livrent en général qu'une partie limitée de leurs résultats aux méta-services, la couverture théoriquement très bonne de ces méta-services n'est en réalité pas optimale.

Catalogue



L'expertise rédactionnelle pour la saisie et le classement hiérarchique des données est cruciale dans les catalogues. Les métadonnées comme le titre, l'URL et un texte descriptif sont saisies. Du fait de leurs coûts de production, les catalogues ne recensent qu'une très petite partie de la toile. De plus, la recherche par mots-clés dans les catalogues ne se base que sur les termes du texte qui décrit les pages web et n'offre pas de lien au texte intégral.

3.3 Typologie des moteurs de recherche

Il existe plusieurs classifications des moteurs de recherche ; les deux qui sont proposées ci-dessous sont les plus connues actuellement :

- Trois groupes (Stock /Lewandowski 2005, p. 1) :
 - **Moteurs de recherche algorithmiques** (ont leurs propres bases de données et système de recherche) comme Google, Yahoo, MSN et Ask.com
 - **Catalogues en ligne élaborés par une intelligence humaine** comme l'Open Directory Project (constitué à partir Google) ou le catalogue Yahoo
 - **Méta-moteurs de recherche**, qui travaillent avec un système de recherche mais n'ont pas de base de données propre ; ils travaillent sur le contenu d'autres outils de recherche (ex. : Kartoo, Metager etc.)
- Quatre groupes (Anonyme) :
 - **Moteurs de recherche généralistes**, comme par ex. Google ou Yahoo
 - **Répertoires internet** : sont des répertoires thématiques, par ex. des dictionnaires. Ce type de moteur est utilisé lorsque l'on veut trouver un service bien défini (par ex. un moteur de recherche spécialisé) ou une collection de liens (portail) sur un sujet précis.
 - **Méta-moteurs de recherche** : interrogent plusieurs moteurs de recherche en même temps, n'ont pas de banque de données propres.
 - **Moteurs de recherche spécialisés** : ils sont spécialisés dans un domaine précis d'internet; ex.: Paperball ou Romso pour la recherche d'articles de journaux.

Un coup d'œil vers l'avenir, le « Semantic web »

Le web sémantique (Semantic web en anglais) est un développement du World Wide Web (WWW) dont le but est de définir formellement les contenus des données lisibles par la machine. Le concept repose sur une proposition du fondateur du web, Tim Berners-Lee.

Les moteurs de recherche sémantiques ne reconnaissent pas que les mots, mais aussi les relations de sens qu'ils ont entre eux. Un moteur de recherche sémantique signale à l'utilisateur que « Java » peut être une île indonésienne, un jeu, un plant de café, une danse à la mode (dans le passé), un langage de programmation. Si la question de l'utilisateur comporte plusieurs mots, le moteur de recherche reconnaît – dans l'idéal – automatiquement de quel domaine il s'agit.

L'utilisateur reçoit au minimum des propositions classées thématiquement.

Pour arriver à cela, il faut que les auteurs de pages web attribuent des identificateurs cohérents à leurs textes. La forme la plus simple d'identificateur existait déjà dans les débuts d'internet. Les auteurs pouvaient décrire le contenu de leurs pages au moyen de métadonnées. De nombreux auteurs n'ont alors pas pu résister à la tentation de manipuler ces métadonnées pour obtenir un meilleur classement de leur page. Les moteurs de recherche ont dû attribuer aux métadonnées une valeur proche de zéro pour le classement. Ce danger se renouvelle avec le web sémantique. La classification des mots doit se faire déjà dans le texte source de la page web. Cela peut fonctionner correctement dans un environnement contrôlé, comme le montrent les intranets d'entreprises. Les auteurs peuvent être formés et être rendus attentifs à leurs fautes. Cela semble impossible à faire dans le vaste monde d'internet. Il faudrait utiliser dans ce cas des services supplémentaires qui apportent leur concours au classement des contenus.(Patzwaldt, online)

Le web sémantique est aujourd'hui encore une vision et un défi majeur.

3.4 Quel outil pour quelle recherche ?

Avant de faire une recherche dans l'internet, il faut se poser la question suivante : quel outil de recherche est le mieux approprié pour trouver ce que je cherche ?

Il y a une différence essentielle entre les moteurs de recherche et les répertoires due à leur mode d'élaboration.

Les « robots », « spider » ou « crawler » des moteurs de recherche fouillent le web systématiquement, ils y « lisent » chaque document et récoltent mot à mot les adresses internet dans une gigantesque base de données. Une liste de mots-clés est établie pour chaque document du web.

Les répertoires ou « catalogues du web » sont établis et tenus à jour par des humains. Une rédaction trouve les documents, décide de leur saisie et les classe dans un catalogue thématique. Les pages d'accueil sont ainsi saisies avec leur titre, une description, l'URL et sont réparties dans un environnement thématique correct (par ex. Yahoo ou Dindo).

Le tableau suivant a pour but d'aider à choisir l'outil approprié au type de recherche (Wolf 2006, on-line):

Outil de recherche	Convient le mieux lorsque...	Exemple
Répertoires	vous désirez avoir un premier aperçu sur un sujet ou chercher une collection de liens sur un thème	Vous cherchez des revues accessibles en ligne sur le thème « internet »
Moteurs de recherche	vous cherchez quelque chose de concret sur un sujet ou vous cherchez des informations actuelles	Vous cherchez des pages internet avec des offres de maisons de vacances dans l'île de Sylt
Méta-moteurs de recherche	vous recherchez un concept très spécifique ou vous voulez couvrir un domaine de recherche particulièrement grand	Vous voulez savoir combien de marches d'escalier a la Tout Eiffel
Moteurs de recherche spécialisés	vous voulez chercher dans un domaine spécifique d'internet (uniquement des pages qui se rapportent à un sujet spécifique ou à une matière)	Vous recherchez des articles récents de quotidiens

3.5 Comment sont utilisés les moteurs de recherche ?

... Les outils de recherche dans internet sont destinés à des utilisateurs non-professionnels. Ils offrent aux usagers professionnels des possibilités de recherche qui ne les satisfont que peu. Les outils actuels utilisent la logique booléenne, c'est-à-dire les opérateurs ET, OU et SAUF. Si l'utilisateur ne saisit pas d'opérateur entre deux termes de recherche, l'outil l'interprète comme un ET, alors que les premiers moteurs avaient le OU par défaut. Les moteurs de recherche n'offrent pas toutes les fonctionnalités du booléen. Ainsi, Google ne permet pas la parenthèse et ne donne pas la possibilité de déterminer un opérateur de proximité, à part la proximité directe (phrase). Impossible d'utiliser un signe de troncature. Si l'on recherche des documents sur le sujet « alimentation » par exemple, il faut lier toutes les variantes grammaticales avec un OU : alimentation OU aliment OU aliments OU alimenter (à la place de la troncature : aliment*). L'usage de la parenthèse facilite les recherches emboîtées comme : je cherche des pages web sur le travail et les loisirs en Belgique et en Hollande, à l'exclusion de celles dans lesquelles la Belgique et la Hollande apparaissent ensemble. Dans Google, il faut formuler cela comme suit : travail loisir Hollande OU Belgique –Hollande OR –Belgique. Aucun utilisateur non-professionnel ne connaît ce langage. [...] (Stock /Lewandowski 2005, p. 2)

3.5.1 Utilisation de la recherche avancée

Selon l'enquête de Machill et al. (2003,), les opérateurs booléens ne sont connus que de la moitié des usagers ; la recherche avancée atteint une notoriété un peu plus grande (59%). Il est toutefois démontré qu'elle est encore plus rarement utilisée que les opérateurs. Seuls 14% des usagers déclarent s'en servir plus souvent (voir Machill et al. 2003, p. 168). Dans les enquêtes expérimentales, le taux est encore plus bas [...]

3.6 Langage de recherche (syntaxes de recherche)

Aperçu des syntaxes de recherche les plus utilisées (Wolf 2006, online):

Combinaison	Opérateur modal	Opérateurs booléens	Champs de sélection	Graphique
et	espace	and	avec tous les mots	
ou	parenthèse	or	avec n'importe quel mot	
sauf	signe « moins »	not	sans les mots	
phrase	guillemets	guillemets	avec le groupe de mots exact	

3.7 Quelles sont les stratégies de recherche possibles sur internet?

Il faut distinguer à la base les stratégies de recherche suivantes :

1. surfer
2. feuilleter (browse)
3. saisie de l'adresse internet
«'Surfer' signifie rechercher d'une manière non structurée, sans plan, dans le monde infini d'internet ; «feuilleter» consiste à utiliser les collections thématiques élaborées, les «Clearing-houses» afin de trouver des liens intéressants.»
(Weilenmann, 2006)

La saisie de l'adresse internet peut aussi donner des résultats (par exemple: <http://www.wetter.de>)

3.8 «Trucs» pour améliorer la recherche

1. Renoncer aux termes très souvent utilisés

Lors de l'utilisation de termes comme «programme», «internet» ou «news», un grand moteur de recherche trouve des milliers de résultats qui n'apportent pas tous une information pertinente pour la question posée. C'est pourquoi il faut renoncer à utiliser des termes aussi couramment utilisés.

2. Travailler avec des synonymes

L'utilisation de synonymes ouvre de nouvelles possibilités au cas où une recherche ne donne pas ou peu de résultats.

Pour une recherche sur la « banlieue », il peut être intéressant d'utiliser aussi le terme « agglomération ».

3. Traduire le terme de recherche dans une autre langue

Au cas où l'interrogation d'un moteur de recherche international avec un terme en français ne donne pas de résultats ou si l'on désire trouver plus d'occurrences, on peut essayer de traduire la recherche en anglais par exemple

4. Trouver des documents en français dans un moteur de recherche international

La plupart des moteurs de recherche proposent la possibilité de filtrer les résultats en fonction de leur langue avec la recherche avancée. En principe, si vous faites une recherche avec un terme en français, vous trouverez des documents en français ou dans d'autres langues mais avec des concepts en français.

5. Apprendre à partir des résultats

L'analyse des résultats peut donner d'autres idées de termes à utiliser pour une recherche. Des concepts utilisables pour une nouvelle interrogation peuvent figurer dans les titres ou les résumés des documents trouvés. Ils permettront de préciser ou de limiter la recherche. Par exemple, utilisez à la place de «insecte» le terme plus spécifique de «moustique».

6. Connaître les fonctions de recherche de plusieurs moteurs de recherche

Aucun moteur de recherche ne propose toutes les fonctions de recherche possibles. Ils se distinguent fortement selon les fonctions qu'ils peuvent offrir. Connaître plusieurs moteurs de recherche et leurs possibilités permet d'obtenir de meilleurs résultats.

7. Travailler avec trois à quatre moteurs de recherche

On devrait en principe travailler avec trois ou quatre moteurs de recherche, y compris les catalogues en ligne et les collections de liens. Cela vaut la peine de bien connaître les syntaxes de recherches des moteurs que vous aurez choisis.

8. Travailler avec les newsletters

Abonnez-vous aux newsletters sur les pages web de votre canton ou de votre commune.

3.9 Classement (ranking)

La méthode de classement des résultats des moteurs de recherche est très importante à connaître. Habituellement ce sont d'autres méthodes que celles utilisées dans les catalogues de bibliothèques ou les banques de données. Les catalogues de bibliothèque, par exemple, ne montrent que les résultats qui correspondent totalement au terme utilisé pour la recherche. Les banques de données offrent la possibilité de formuler plus précisément la question et de définir un choix de sources (c'est-à-dire de définir la ou les banques de données les plus adéquates) afin d'obtenir une quantité de résultats moindre et plus exacte.

Les critères de classement utilisés par les moteurs de recherche (*Lewandowski 2005, online*):

- **La pondération des mots dans un document**
La fréquence d'un terme dans un document
- **L'intervalle entre les termes (proximité)**
On considère, lors d'une recherche avec plusieurs termes, l'intervalle qui sépare ces termes dans le document, donc leur proximité
- **La position du terme de recherche**
Les termes utilisés dans des endroits importants du documents reçoivent une valeur plus haute. Par exemple, la présence dans le titre, dans les têtes de chapitre, dans l'URL
- **La place du terme de recherche dans la question**
Les termes qui apparaissent en premier dans la question sont considérés comme plus importants
- **Les métadonnées**
La présence du terme de recherche dans les métadonnées

- **La position du terme de recherche dans le texte du document**
Le document qui contient le terme de recherche dans ses premières lignes recevra une valeur plus haute que s'il apparaît plus loin dans le texte
- **L'accentuation du terme par des éléments HTML**
Les termes mis en évidence (gras, italique) sont mieux cotés
- **Majuscule / minuscule**
Les documents dans lesquels le terme de recherche apparaît dans la forme exacte de la question sont favorisés
- **La fréquence inverse de documents**
La fréquence relative de la présence d'un terme dans les documents ; les termes rares sont favorisés
- **Les liens**
La présence du terme de recherche dans un lien qui renvoie au document cible ; ce dernier sera favorisé
- **La langue**
Les documents qui sont dans la langue du pays de l'interface utilisé sont favorisés
- **«Geo-targeting» ou «geo-location»**
Les pages qui sont localisées dans une aire géographique proche de l'utilisateur sont favorisées
- **Le niveau dans la liste**
Plus haut le document est placé dans la hiérarchie de sa page web, mieux il est évalué
- **Le nombre de liens qui pointent sur le document**
Plus il y a de liens qui pointent sur le document, plus il est considéré comme important
- **La notoriété des liens**
La structure des liens permet de qualifier la valeur de la page
- **La fréquence de la consultation**
Les documents qui sont consultés par de nombreux usagers sont favorisés
- **L'actualité**
Les documents actuels reçoivent une meilleure cote que les plus anciens
- **La longueur du document**
Les documents compris entre une longueur minimale et maximale sont favorisés
- **Le format des données**
Les documents en format standard HTML sont préférés à ceux qui sont dans d'autres formats (PDF, Word, etc.)
- **La taille du site**
Les documents qui font partie d'une offre riche sont mieux considérés que ceux qui se trouvent dans de petits sites

La plupart des moteurs de recherche utilisent une combinaison de plusieurs critères de classement. Ces dernières années, le critère considéré comme le plus important pour l'évaluation d'un document est la structure des liens. Des procédures complexes ont été développées pour mesurer la notoriété des liens d'un document (par ex., la fréquence de l'utilisation des liens, le nombre de liens qui pointent sur le document).

3.10 Critères pour l'évaluation de la qualité des résultats d'une recherche

Titre

Un premier point de repère pour l'évaluation du résultat d'une recherche est le titre du document. On peut parfois savoir déjà par le titre d'une page web si le document est utilisable ou non. Si l'on fait, par exemple, une recherche sur une personne et que le titre de la page est le nom d'une institution homonyme, il est clair que cette page peut être laissée de côté. Exemple : Goethe – Goethe Institut

[Goethe-Institut - Leitseite](#)

Offizielle Organisation zur Förderung deutscher Kultur.
www.goethe.de/ - 27k - [Im Cache](#) - [Ähnliche Seiten](#)

Extrait du contenu

De nombreux moteurs de recherche donnent dans la liste des résultats un court extrait du texte du document.

Si le titre ne suffit pas pour l'évaluation de la page web, on peut se référer à cet extrait pour juger de la pertinence du document.

Description

De nombreux moteurs de recherche offrent, en plus du titre du document, un brève description de son contenu ; cela peut être sous forme de mots-clés (par exemple dans les répertoires en ligne) qui décrivent plus précisément le contenu de la page ou sous forme de résumé du document.

Résumé

L'opération est encore plus simple lorsque le moteur de recherche livre avec le document un résumé du contenu, créé automatiquement. Un tel résumé reflète le contenu du document, pas d'une manière tout à fait complète toutefois.

URL

L'URL du document donne aussi une information.

L'URL donne des renseignements sur l'aire géographique d'un document ou sur le contexte organisationnel dans lequel se trouve une page web.

Si l'on cherche, par exemple, des informations sur les cochons d'Inde un document dont l'URL est <http://www.cobayesethamsters.com> sera certainement plus pertinent qu'un autre qui aurait comme adresse <http://www.veterinaires.ch/animaux/CochonInde.html> .

On peut voir par l'adresse que le thème « cobaye » est le sujet principal de cette page alors que dans le second cas, le sujet recherché n'est pas le principal sujet du site.

Le nom du serveur

Par exemple: <http://www.hostname.de>

Le nom est choisi librement par l'opérateur. A côté des pages officielles des grandes entreprises, vous trouverez aussi des noms fantaisistes et des noms de projets, films, produits, etc.

Les domaines principaux (Top Level Domain)

Cette liste vous permettra de reconnaître sur la base de la terminaison du nom du serveur, à quel groupe les pages appartiennent.

.com	Utilisée aussi bien par les prestataires de services commerciaux internationaux que par des projets non commerciaux qui ne trouvent pas leur place dans un autre domaine. C'est l'extension la plus utilisée dans le net
.edu	Utilisée aux Etats-Unis par les sites éducatifs, en particulier par les sites des universités. Apparaît aussi de temps à autre dans des institutions internationales en tant que partie de l'adresse
.gov	Utilisée, mais pas toujours, par les autorités gouvernementales américaines
.mil	Utilisée par les institutions militaires américaines
.net	Utilisée par les fournisseurs de services internet, international
.org	Utilisée par les organisations à but non lucratif ; des entreprises peuvent quelquefois se cacher sous cette extension
Codes de pays	Chaque pays a son extension spécifique. Sous l'extension nationale, on peut trouver des institutions gouvernementales, mais aussi des entreprises, des sites de divertissement, etc.
.de	Allemagne
.us	USA
.uk	Grande-Bretagne
.at	Autriche
.ch	Suisse
.nl	Pays-Bas

Catégories

Les répertoires comme Allesklar ou DINO donnent en général, pour chaque recherche par mot-clé, la catégorie du répertoire à laquelle la page web trouvée appartient.

Des exemples:

The screenshot shows the website **allesklar.de** with a search bar containing the text "goethe". Below the search bar, there are two input fields labeled "was" and "wo", and a "suchen" button. The search results are displayed below the search bar, with a red box highlighting the search results section. The search results section is titled "Suche nach: goethe" and shows "1 Kategorien im Webkatalog". The categories listed are: Menschen & Gruppen, Persönlichkeiten, Künstler, Schriftsteller A-Z, and Schriftsteller G. The search results for "Goethe, Johann Wolfgang von" are displayed below the categories. The search results section is titled "Suchergebnis" and shows "1 - 20 | weiter >". The search results are listed below the search results section. The first result is "Goethe - wie Wahrheit wandelt" with a description and a URL. The second result is "Goethe zum Festpreis auf eBay Express kaufen" with a description and a URL. The search results section is titled "Webkatalog (654 Einträge)" and shows "1. Goethe Institut Montreal (Montreal/Kanada)" with a description and a URL.

allesklar.de [Erweiterte Suche](#) [Hilfe](#)

Suche

was wo

[Startseite](#) [Suche](#)

Suche nach: goethe

1 Kategorien im Webkatalog

[Menschen & Gruppen](#) [Persönlichkeiten](#) [Künstler](#) [Schriftsteller A-Z](#) [Schriftsteller G](#) [Goethe, Johann Wolfgang von](#)

Suchergebnis | 1 - 20 | [weiter](#) >

Sponsoren Links

- [Goethe - wie Wahrheit wandelt](#)
Über die Wirksamkeit von Lebenshilfen und Weisheitslehren, über Gemeinsamkeiten heutigen Wissens und einstigen Glaubens, über Vor- und Nachteile der Einsamkeit u. a. m.
<http://www.dietmar-odilo-paul.de>
- [Goethe zum Festpreis auf eBay Express kaufen](#)
Bücher, Hörbücher, Musik, Instrumente, Filme, DVDs, PC- und Videospiele als Neuware zum Festpreis. eBay Express: Alles. Einfach. Neu.
<http://www.ebayexpress.de>

Webkatalog (654 Einträge)

1. [Goethe Institut Montreal \(Montreal/Kanada\)](#)
Das institut informiert über Kulturprogramme und Sprachkurse. Einzelne Rubriken widmen sich der Rechtschreibreform, neuer deutscher Literatur oder Materialien zum Erlernen der deutschen Sprache.
Kategorie: [Kunst & Kultur](#) > [...] > [Einzelne Künstler-Organisationen](#) > [Goethe-Institute](#)
<http://www.goethe.de/uk/mon/deindex.htm>

1 Kategorien im Webkatalog[Menschen & Gruppen](#) > [Persönlichkeiten](#) > [Künstler](#) > [Schriftsteller A-Z](#) > [Schriftsteller G](#) > [Goethe, Johann Wolfgang von](#)**Suchergebnis 1 - 10** von **654**| 1 - 10 | [weiter >](#)**Goethe** Institut Montreal (Montreal/Kanada)

Das Institut informiert über Kulturprogramme und Sprachkurse. Einzelne Rubriken widmen sich der Rechtschreibreform, neuer deutscher Literatur oder Materialien zum Erlernen der deutschen Sprache.

<http://www.goethe.de/uk/mon/deindex.htm>**Goethe** Institut Bibliothek & Information International

Hier kann man Berichte und Publikationen einsehen und sich über Projekte informieren.

<http://www.goethe.de/wis/bib/prj/bii/>**Goethe** Institut Salvador-Bahia

Diese Vertretung des Goethe-Instituts informiert über Veranstaltungen, Sprachkurse, Projekte, Aktivitäten und Informationsdienste.

<http://www.goethe.de/ins/br/sab/>**Goethe** Travel Office (Italien)

Ausgesuchte Hotels und Appartements am Gardasee (Malcesine, Brenzone, Torbole, Riva, Torri). Eintrittskarten für die Opernsaison Arena di Verona, Ausflüge nach Venedig, Verona, Mailand, in die Dolomiten und Seerundfahrten.

<http://www.goethetravels.it/>**Goethe**-Institut Inter Nationes e.V. (München) - Vermittlung der deutschen Sprache und Kultur

Das Zentrum für deutsche Kultur, Sprach und Informationsaustausch ist in zahlreichen Ländern vertreten und bietet dort Sprachkurse, Lehrerausbildung und Fortbildung an.

<http://www.goethe.de/>Fachschaft Jura der Johann Wolfgang **Goethe**-Universität Frankfurt am Main

Die Vertretung der Studierenden bietet aktuelle Informationen, Termine, Veranstaltungen und mehr rund um das Jura-Studium.

<http://www.fachschaft-jura.info/>

3.11 Critères généraux d'évaluation des pages web du point de vue des spécialistes de l'information

(Ne sont pas considérés: les archives FTP, les articles de «new groups», les FAQs ou les banques de données)

Fiabilité

- Caractère plausible de la source
- Compétence de l'auteur
- Indication de l'institution / organisation
- Indication de procédures de contrôle
- Domaine internet

Exactitude et couverture

- Actualité
- Exactitude de la source
- Couverture du sujet
- But et groupes cibles

Pondération

- Objectivité
- Logique

Indication des sources

- g. Possibilité de contact
- h. Répertoire des sources
- i. Certification par des tiers

Convivialité

- j. Structure et navigation
- k. Accès
- l. Possibilités d'utiliser des fonctions interactives et de multimédia

Finalement on fera une estimation de la qualité générale de la page web; cela peut se faire déjà lors de la comparaison avec les autres sources qui sont en relation avec elle.

La checkliste des critères d'appréciation de pages web sous

http://www.lib.berkeley.edu/TeachingLib/Guides/Internet/EvalForm_General_Barker.pdf
apporte des éléments intéressants.

3.12 Critères de qualité des moteurs de recherche

- Nombre de documents indexés
- Actualité des données
- Formats acceptés (Ex. Google: .pdf, .doc, .ppt, images, ...)
- Traitement, maniabilité des pages générées
- Exploitation des documents (totale, partielle)
- Exclusion de domaines
- Profondeur de la recherche et du suivi des liens par les spiders
- Temps de réponse
- Pourcentage de liens inactifs
- Fiabilité et disponibilité
- Interactivité de la recherche (possibilité d'affiner, recherche dans les résultats)

3.13 Ce que les moteurs de recherche ne trouvent pas

Le nombre d'informations que les moteurs de recherche ne trouvent pas est immense.

Le concept de «web invisible» ou de «deep web» désigne cette partie du web que l'on ne peut retrouver avec un moteur.

Le web invisible comprend différents domaines.

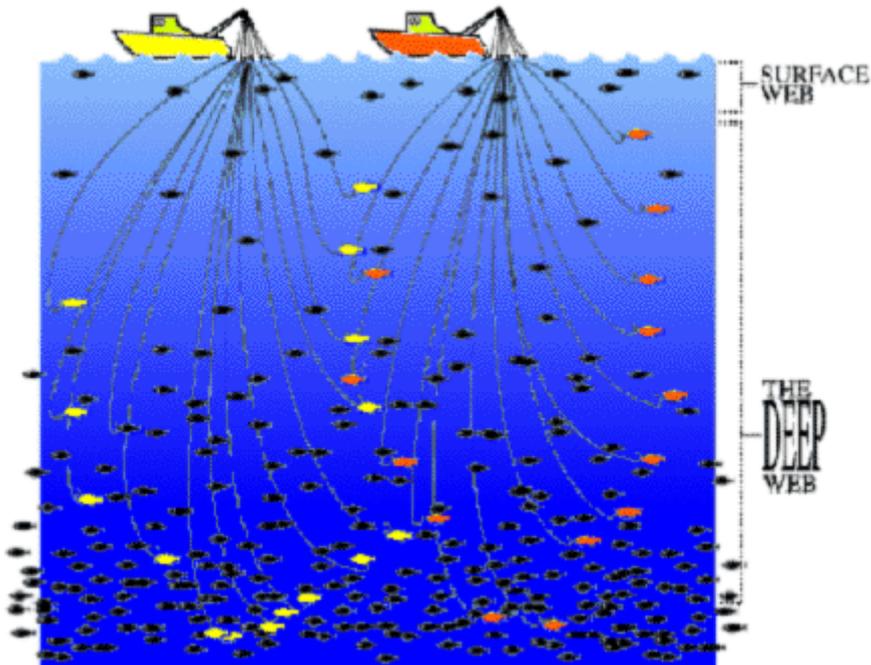


Illustration1: Harvesting the Deep and Surface Web with a Directed Query Engine

En voici une répartition possible entre beaucoup d'autres (Weilenmann, 2006)

Le «**web opaque**» désigne la partie qui pourrait être saisie par les moteurs de recherche mais qui, à cause de restrictions spécifiques, n'est pas indexée ; ce sont par exemple les pages web qui ne sont saisies que jusqu'à un certain niveau (profondeur du crawling), fréquence du crawling (actualité), etc.

Le «**web privé**» contient des pages qui ne sont volontairement pas indexées par les moteurs de recherche ; elles peuvent être désignées au robot par l'obligation d'utiliser un mot de passe pour la recherche et par l'insertion d'une marque d'exclusion (robots.txt).

Le «**web propriétaire**» (le web protégé) exige de l'utilisateur qu'il s'enregistre avant de pouvoir faire une recherche. Ce sont en général des banques de données payantes.

Le «**Truly Invisible Web**» (web vraiment invisible) est constitué de pages qui ne peuvent pas être indexées pour des raisons techniques (par ex. les pages générées dynamiquement).

Les **pages dynamiques** (Karzauninkat, online):

Les pages dynamiques

Internet ne contient pas que des **documents statiques** qui peuvent être saisis par les robots des moteurs de recherche. Une part croissante des informations est livrée par des banques de données. Internet sert dans ces cas de courroie de transmission et non de réservoir de données. Deux exemples simples: les horaires et les annuaires téléphoniques. On ne trouve pas dans internet de longues listes de noms et de numéros, mais un masque de recherche. Sur la base d'une combinaison entre un nom et une adresse, un résultat est livré «on the fly», au moment de l'interrogation. Les pages n'existent qu'un instant sur l'ordinateur de l'utilisateur. Etant donné qu'aucun robot n'interroge ces données, les millions de numéros de téléphone ne peuvent pas être recherchés par un moteur de recherche normal; ils sont donc, aux yeux des moteurs de recherche, invisibles.

Les moteurs de recherche payants

Les banques de données commerciales et les archives professionnelles de tous genres offrent des services payants par internet. Il y a beaucoup de banques de données qui exigent, pour pou-

voir être interrogées, que l'on s'inscrive et que l'on paie une taxe d'utilisation. Les informations contenues dans ces banques professionnelles ne sont pas visibles pour un moteur de recherche normal [...]

Invisibles à cause d'une erreur dans le système

Les banques de données ne sont pas les seules à échapper aux moteurs de recherche. Les serveurs mal installés, les systèmes de gestion de contenu inappropriés, les pages internet mal construites sont aussi des cas que les robots des moteurs de recherche «sautent» et dont ils ne prennent pas les informations dans leurs index. Même si ces informations sont facilement consultables pour un usager normal d'internet, elles restent invisibles.

Invisible car «pas d'accès»

Toutes les informations accessibles par internet ne sont pas destinées au public. Les données d'une entreprise, par exemple, ne sont censées être visibles que par les collaborateurs et sont à cause de cela protégées par un mot de passe. Elles sont donc invisibles pour tous ceux qui ne font pas partie de l'entreprise.

Invisible parce que «beaucoup trop»

Une autre cause de l'absence de certaines pages dans l'index des moteurs de recherche est liée à des raisons de coûts et de capacité. Il existe de nos jours sur chaque serveur de nombreux sites dont la structure et l'organisation sont très compliquées et que les moteurs de recherche ne saisissent pas. A l'extrême, ils se contentent de saisir la page d'accueil. Les moteurs de recherche doivent déjà exploiter d'énormes serveurs afin de rester concurrentiels. Au moment de la rédaction de ce dossier, Google travaille par exemple avec 3000 serveurs.

Invisible parce que «trop lent»

Il ne suffit pas d'indexer une fois une page. De nombreuses pages sont souvent actualisées. Si souvent que les spiders n'arrivent pas à suivre. Ainsi certaines informations actuelles restent invisibles pendant un certain temps.

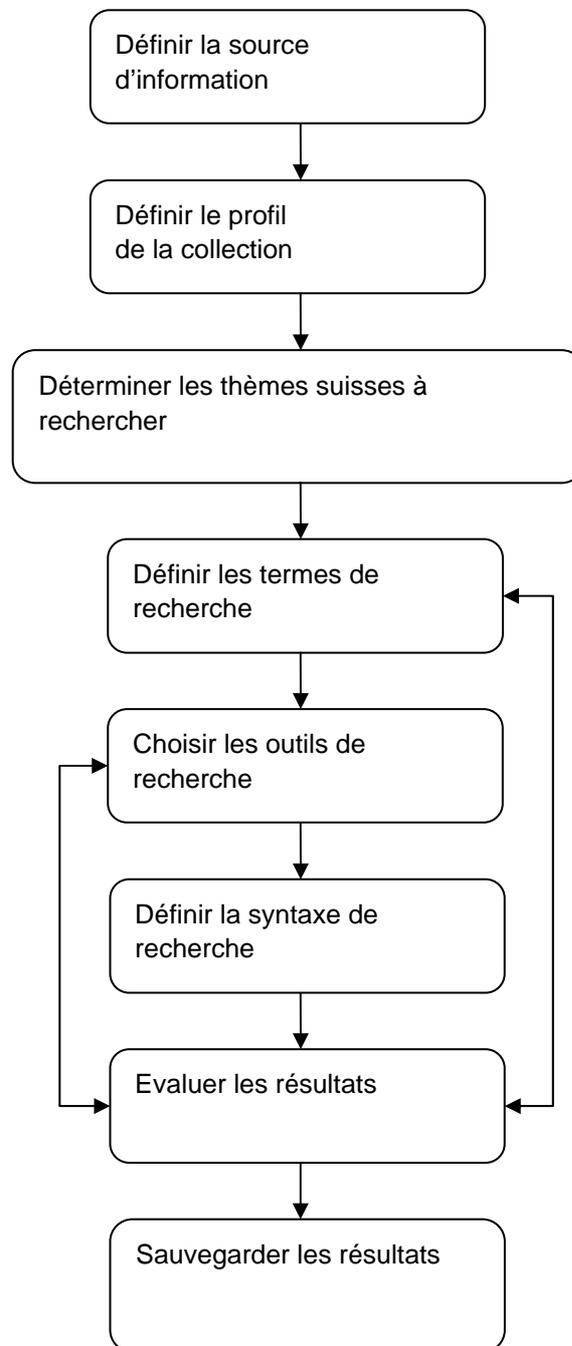
Invisible, au fond, est simplement «non découvert»

Les informations qui ne sont pas trouvées par les moteurs de recherche ne sont pas obligatoirement cachées. Souvent elles peuvent être accessibles tout à fait normalement (elles font parties des données que l'on ne peut consulter que contre paiement ou que si l'on est membre d'une entreprise). Elles n'apparaissent simplement pas dans les résultats donnés par les moteurs de recherche.

4 Partie pratique

4.1 Méthode de recherche des Helvetica en ligne d'importance patrimoniale

Afin d'assurer le succès de la recherche, il faut tout d'abord définir le processus de recherche:



Définir le média d'information

Par exemple : Internet, banque de donnée, etc.

Définir le profil de la collection

Selon les recommandations de la Bibliothèque nationale suisse («Notice collecte») et les directives des bibliothèques cantonales (profil de la collection). On répondra à la question suivante:

«Que collectons-nous, que ne collectons-nous pas, quels domaines sont importants pour notre canton, pour nos usagers, etc.»

Définir les thèmes suisses de la recherche

Par exemple, agriculture, Cervin

Définir les termes de recherche

Quels termes utiliser pour la recherche (par exemple, descripteurs, mots-clés, utilisation de synonymes, d'homonymes, etc.)

Choisir les moteurs de recherche / les répertoires web

Il est judicieux, selon le sujet, de choisir différents moteurs de recherche ou répertoires web.

Définir la syntaxe de la recherche

Important : lire les conseils d'utilisation (l'aide) des moteurs de recherche. Cela permet de définir les opérateurs de recherche et d'appliquer une recherche professionnelle ou une recherche avancée.

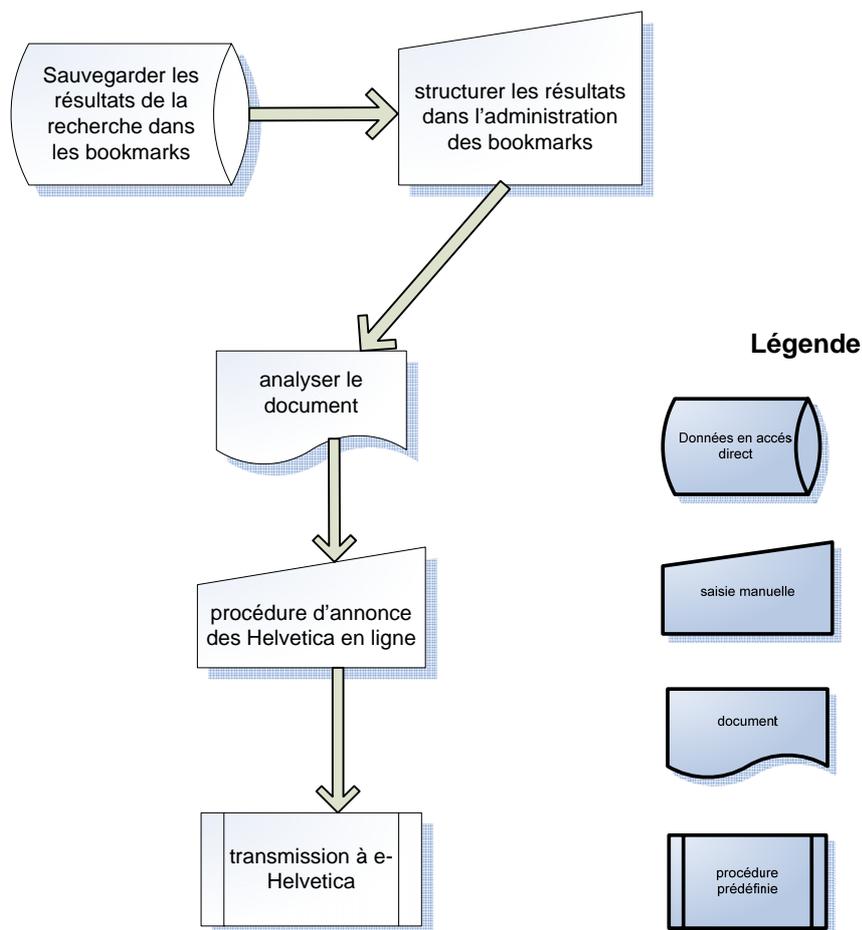
Evaluer les résultats

L'évaluation des résultats permet de décider s'il faut formuler la question autrement ou s'il est préférable de choisir un autre moteur de recherche, un autre répertoire web.

Sauvegarder les résultats

Par exemple, dans les Bookmarks

4.2 Procédure pour la transmission des Helvetica en ligne d'importance patrimoniale



4.3 Exemples

Méta-moteurs de recherche	URL	Description	Evaluation
ETools.ch	http://www.ertools.ch/	Deutsch / English AltaVista, Ask, Bluewin, Entireweb, Google, Lycos, MSN, Search, Seekport, Yahoo ; opérateurs ET/OU ; limitation à la Suisse; Redondances éliminées / classement par pertinence	Convivial, les mots de recherche sont soulignés, l'évaluation de la pertinence n'est pas claire
Moteurs de recherche	URL	Description	Wertung
Search.ch 	http://www.search.ch/	Deutsch / français / italiano / English Opérateurs ET / OU / SAUF, limitation à la Suisse et à tous les cantons	Convivial, les résultats sont pertinents aussi au niveau „canton“ – beaucoup de résultats commerciaux (spams)
Google.ch 	http://www.google.ch	Deutsch / français / italiano / Rumantsch / English Recherche simple / recherche avancée Web, images, actualités, usergroups, etc. Limitation à la Suisse	Convivial, possibilités de recherche puissantes, la publicité trouble la pertinence des résultats
AltaVista Suisse	http://ch.altavista-fr.com/	Sélection du français et anglais ou de toutes les autres langues; recherche simple et avancée Web, images, audio, vidéo, etc. Limitation à la Suisse	Convivial, possibilités de recherche puissantes, publicité, règles linguistiques bizarres
MSN Suisse	http://fr.ch.msn.com/	Deutsch / français ET / OU / SAUF – les «stopwords» sont ignorés Web, images, actualité, etc.	Convivial, pertinence des résultats, pas de limitation à la Suisse, fenêtre de publicité

<p>Yahoo! Suisse</p> 	<p>http://ch.search.yahoo.com/</p>	<p>Deutsch / français / italiano ET / OU / SAUF -> recherche avancée Web, images, vidéo, actualités Limitation à la Suisse et par langues</p>	<p>Convivial, possibilités de recherche puissantes, pas de publicité dans le masque de recherche</p>
<p>Répertoires, collections de lien</p>	<p>URL</p>	<p>Description</p>	<p>Evaluation</p>
<p>Auskunft.ch</p>	<p>http://www.auskunft.ch/</p>	<p>Deutsch Recherche par catégories, les liens pointent sur des sites extérieurs Catégories habituelles</p>	<p>Convivial, les pages commerciales sont mises en évidence</p>
<p>dmoz Open Directory Project</p>	<p>http://dmoz.org/Regional/Europe/Switzerland/</p>	<p>English Recherche possible dans la rubrique Switserzland</p>	<p>Rubriques générales, pas convivial, renvois curieux du fait de la collaboration internationale</p>
<p>Sharelook.ch</p>	<p>http://www.sharelook.ch</p>	<p>Deutsch Méta-recherche -> faite par Apollo7 Recherche Limitation à la Suisse, Bâle et Zurich</p>	<p>Rubriques détaillées, mais presque pas de contenu. Beaucoup de publicité, méta-recherche inutilisable</p>
<p>Webliste</p> 	<p>http://www.webliste.ch/</p>	<p>Deutsch Limitation aux cantons pour certaines catégories</p>	<p>Très riche, collection de liens spécifiques sur la Suisse ; possibilités de recherche dans les listes de liens</p>
<p>Swisslinks</p>	<p>http://www.swissinfo.org/ Chemin: Info Schweiz -> Swisslinks</p>	<p>Deutsch / français / italiano / 6 autres langues Recherche par mot-clé ; pages suisses pertinentes</p>	<p>Convivial, catégories détaillées, richesse et qualité différentes selon les sujets</p>
<p>ch.ch</p>	<p>http://www.ch.ch</p>	<p>Deutsch / français / italiano / Rumantsch / English Portail de la Chancellerie fédérale, en collab. avec les cantons et communes. Classé par «privés», «entreprises» et «autorités» -> guichet virtuel (= réponses aux questions courantes à l'aide d'une liste de liens)</p>	<p>Très convivial, mais pas de liste de liens propre (renvoi sur d'autres sources d'information)</p>

Linker.ch	http://www.linker.ch/	Deutsch Index, résultats classés par pays (dont la CH), pages plutôt commerciales	Agréable, quelques adresses web sur presque tous les su- jets, navigation inhabituelle
Kouik.ch	http://www.kouik.ch/	Français Répertoire et moteur de recherche „Guide de l'internet romand“ ; recherche par rubriques, conte- nus commentés à l'intérieur des rubriques	Convivial, résultats pertinents mais peu nombreux
BigFoot.ch Switzerland Web Di- rectory	http://www.bigfoot.ch/	Deutsch / français / english / italiano Répertoire par thèmes ou cantons / régions Résultats différents selon le choix de la langue Possibilité de recherche à l'intérieur du répertoire	Convivial mais peu de pages suisse indexées
Linkcity.ch	http://www.linkcity.ch/	Deutsch Répertoire des pages web suisses en allemand. Classé par sujet, moteur de recherche pour la col- lection de liens	Résultats pertinents et non pertinents (commerciaux) mêlés
Portails	URL	Description	Evaluation
Z!sch	http://www.zisch.ch/	Deutsch Possibilités de recherche dans les rubriques, ré- pertoires et articles des médias (LZ-Medien) Recherche limitée à la Suisse centrale (zone d'influence de LZ-Medien)	«Agité», beaucoup de publici- té ; les informations sont triées par cantons (Suisse centrale)
Graubünden Portal	http://www.link-gr.ch/	Deutsch Possibilités de recherche dans la banque de liens Répertoire de liens commentés, majorité de pages du canton des Grisons	Convivial ; beaucoup de pa- ges et références commercia- les. Commentaires, en partie messages publicitaires

Basel Linksammlung	http://www.basel.ch/de/links	Deutsch / français / English Possibilités de recherche sur basel.ch uniquement, collection de liens classée par sujets – couvre la région de Bâle (Baseland, Fricktal, Allemagne du Sud et Alsace)	Convivial. Collection de liens sur une petite partie du web, liens pertinents.
AarauVista	http://www.aarauvista.ch/	Deutsch Possibilités de recherche dans les rubriques Entreprise, Associations, pages privées, etc. Rubriques sur le canton d'Argovie	Convivial, parfois peu d'entrées – rubriques comme „Associations“ très utiles; pages commerciales
BL virtuell	http://www.blvirtuell.ch/	Deutsch Rubriques, possibilités de recherche dans les rubriques, contenus en majorité sur le canton de BL, en partie aussi sur le région de Bâle	Convivial, peu de publicité, un peu „vieilli“, la collection de liens n'est pas partout à jour
Appenzellerlinks.ch	http://www.appenzellerlinks.ch/	Deutsch Rubriques, possibilités de recherche dans les rubriques. Liens sur la région Suisse orientale, avant tout sur les deux demi-cantons	Animations et publicité dérangeantes, rubriques sommaires, utile car concentré sur «Appenzell»
Sites web valaisans	http://www.mediatheque.ch/	Deutsch / Français Liste de liens sur le Valais classés thématiquement, sélectionnés par la Médiathèque Valais	
Portail web du Canton de Neuchâtel	http://www.surface.ch/ne/	Français Recherche par commune ou catégorie	
Vaud-Région	http://www.vaud-region.ch/	Français Répertoire avec liens sur le canton de Vaud et alentours, classés par sujet	Convivial, assez peu de résultats



Empfehlenswerte Suchdienste

4.3.1 Recherche: Cervin et géologie

1) Au départ : définir le contenu de la recherche

Quelles informations trouver, collecter et archiver ? [voir «Notice collecte», ch. 2.1, 2.2 et 5.2.1]

Cela signifie concrètement pour cet exemple: examiner les informations avec contenu spécialisé (historique, culturel), proposées par des institutions suisses (nationales, locales) ou des personnes privées (qualité, quantité et importance)

2) Définir le plan de recherche

Définir les autres noms, synonymes, flexions des termes de recherche :

Par ex., recherche dans Wikipedia

„haut valaisan ds'Hore, français Mont **Cervin** ou Le Cervin, italien Monte **Cervino**“

„Synonymes“ et flexions (les troncatures ne son pas utilisables dans les moteurs de recherche)

geologie, geologisch, geologisches, geologischer, geologischen, geologischem

tektonik, tektonischer, tektonischem, tektonisches, tektonischen, tektonisch

The screenshot shows the German Wikipedia article for 'Matterhorn'. It includes a table of contents with sections like 'Geschichte des Namens' and 'Geologie'. A table on the right provides key facts: Height: 4478 m a. M., Location: Grenze Wallis (Schweiz) / Italien, Range: Walliser Alpen, Geographical coordinates: 45°50' N, 7°39' O, First ascent: 14. Juli 1865 durch Edward Whymper, Normalweg: Hörnliweg (Nordostgrat) von Hörnlihütte III+ (25).

Une recherche simple dans un moteur de recherche (comme Google ou Search) donne un nombre non maîtrisable de résultats dont la plupart sont des offres pour des livres ou des hôtels.

-> exclusion de ces termes

Danger de perdre des pages pertinentes dans lesquelles il est question aussi de livres ou d'hôtels

4) Recherche dans un répertoire

Recherche dans un répertoire avec le mot-clé **“Matterhorn“** et le filtre **„Wallis“**.

Les résultats sont présentés par thèmes et triés par rubriques.

Webliste.ch trouve avant tout des pages commerciales dans les rubriques **„Ferienwohnungen & Jugl“**, **„Ausflüge“**, etc

Les recherches dans Sharelook.ch, Linker.ch und Kouik.ch (avec mot clé **„cervin“**) ne donnent que des pages commerciales peu pertinentes.

Swisslinks (Swissinfo) a trouvé, à côté de quelques pages, des extraits d'émissions (Radio DRS, SFDRS, SDA).

57'000
Schweizer
Webverzeichnis

Schnellsuche: matterhorn
Wallis Suche

Firmensuche
Advokaturbüro
Schweiz

Home | Webseite anmelden | Flohmarkt | Lexikon | News Corner | Newsgroups | Witziges | Wetter | Toplisten | Favoriten

Der Begriff **«matterhorn»** wurde gesucht und **50x** gefunden (Suchdauer: 0,02 Sekunden)

Top Rubriken: [Gastronomie & Hotels \(20x\)](#) [Ferienwohnungen & Jugl \(6x\)](#) [Ausflüge \(4x\)](#)

Top Branchen: [Gastronomie & Hotels \(5x\)](#) [Sport & Wellness \(2x\)](#)

Radio Matterhorn
Radio **Matterhorn** war das kleinste, höchstgelegene und meistgehörte Lokalradio in der Schweiz.
Letzter Check: 15.01.2007 - [»Zeige ähnliche Seiten](#) [»Archiv](#)

Hotel Ambassador Zermatt
Ambassador, das Appartement- und Studio-Hotel erster Klasse bietet eine faszinierende Aussicht auf das Matterhorn. Rustikale, komfortable 2- und 3-Zimmer Appartements, geschmackvoll eing...
Letzter Check: 15.01.2007 - [»Zeige ähnliche Seiten](#) [»Archiv](#)

Golf Club Matterhorn
Der 9-Loch Platz vom Golfclub **Matterhorn**, gelegen am Fusse des berühmtesten Berges der Welt, wurde harmonisch in umliegende Bergwelt eingebettet und liegt dabei selber noch auf einer ...
Letzter Check: 15.01.2007 - [»Zeige ähnliche Seiten](#) [»Archiv](#)

Haus Klein Matterhorn
Erholen Sie sich bei uns in Ruhe vom Alltagsstress, lassen Sie Ihre Seele baumeln und verbringen Sie erlebnisreiche Urlaube in einer familiären Ambiente. Wir wünschen Ihnen einen unvergesslichen Aufenthalt.
Letzter Check: 15.01.2007 - [»Zeige ähnliche Seiten](#) [»Archiv](#)

Haus Alpine Zermatt
Eine ganz neue 4-Stern 4 1/2 Zimmer Wohnung (117m2) für 6 Personen mit wunderschönem Blick aufs Matterhorn, zentrale Lage. Nur wenige Schritte zum Skibus oder ins Dorf. Im Winter ...
Letzter Check: 15.01.2007 - [»Zeige ähnliche Seiten](#) [»Archiv](#)

Gemeinde Zermatt
Offizielle Seite, mit Informationen über Sommer- und Wintertourismus.
Letzter Check: 15.01.2007 - [»Zeige ähnliche Seiten](#) [»Archiv](#)

3b) La même recherche en français

Syntaxe de recherche (Google)

cervin geologie OR géologie OR geologique OR géologique OR tectonique -livre -hotel -hôtel -vacance

(différence si l'on cherche avec les caractères diacritiques ou non)

3c)

Syntaxe de recherche (Google)

intitle:cervin geologie OR géologie OR geologique OR géologique OR tectonique -livre -hotel -hôtel -vacance

Curieusement il faut les deux caractères d'espace pour qu'on reçoive des résultats corrects (sinon Google interprète qu'il faut 'cervin' ET 'geologie' dans le titre.)

The screenshot shows a Google search interface with the query 'cervin geologie'. The search results are displayed in French and include several entries related to the Cervin mountain and geological topics. The first result is 'Géologie vivante - Cervin', followed by 'MUSEUM D'HISTOIRE NATURELLE DE LA VILLE DE GENEVE', 'MUSEUM D'HISTOIRE NATURELLE DE LA VILLE DE GENEVE', 'tsr.ch - Visite géologique au Cervin', 'DETEC - Sur les origines du Cervin', 'suisse le cervin une montagne africaine', 'le cervin, une montagne africaine', and 'Le Cervin est-il africain?'. Each result includes a brief description and a URL.

Conclusion

Seuls les moteurs de recherche ont donné des résultats acceptables. La syntaxe de recherche doit comprendre des termes concis pour trouver des pages pertinentes. Une limitation géographique aux pages valaisanne n'est pas pertinente dans cet exemple car de nombreuses pages spécialisées sont produites à l'extérieur. Les répertoires ou portails ne sont d'aucune aide, ils offrent avant tout des pages commerciales (tourisme)

4.3.2 Recherche Basler Fasnacht

1) Au départ : définir le contenu de la recherche

Quelles informations trouver, collecter et archiver ? [voir «Notice collecte», ch. 2.1, 2.2 et 5.2.1]

Cela signifie concrètement pour cet exemple: examiner les informations avec contenu historique ou culturel, proposées par des institutions suisses (régionales, locales) ou des personnes privées (qualité, quantité et importance).

2) Définir le plan de recherche (par ex. recherche dans le Dictionnaire historique, Wikipedia)

Définir les autres noms, synonymes, flexions des termes de recherche :

Par ex., recherche dans Wikipedia

„baseldütsch/baseldytsch“: [basler fasnacht](#), [allemand: Fastnacht](#), [français: Carnaval](#) italien: [Carnevale](#)

„Synonyme“

[baseldütsch: die scheenschte drey Dääg/die drey scheenste Dääg](#),
[deutsch: Karneval, Fasching](#)

Autres termes de recherche possibles : Brauchtum, Volksleben, Volkskunde, Fastenzeit, Folklore/Folklorismus, Fastnachtsbräuche

Termes de recherche spécifiques :

[baseldütsch: Morgestraich/Morgenstraich/Morgenstreich, Bangg, Cortège, Helgen](#)
[deutsch: Schnitzelbank, Fastnachtsmärsche](#)

Une recherche avec „[Basler Fasnacht](#)“ donne, dans un moteur de recherche comme Google 311'000 (Stand 08.02.07) réponses

Basler Fasnacht

Die **Basler Fasnacht** beginnt am Montag nach [Aschermittwoch](#) um 4.00 Uhr mit dem [Morgestraich](#). Sie dauert exakt 72 Stunden und endet am Donnerstagmorgen um 4.00 Uhr. In dieser Zeit wird die [Basler](#) Innenstadt von den Fasnächtlern beherrscht, die in ihren Cliquen durch die Strassen, Kneipen und Geschäfte ziehen. Von Aktiven wird die Basler Fasnacht auch als "die drey scheenschte Dääg" (die drei schönsten Tage) bezeichnet.

Die Basler Fasnacht, im folgenden kurz einfach nur Fasnacht genannt, unterscheidet sich stark vom rheinischen "jecken" [Karneval](#), der vor allem in [Deutschland](#) das Bild von [Karneval](#), [Fastnacht](#) und [Fasching](#) prägt. Dafür sorgt schon die strikte Trennung zwischen Aktiven und Zuschauern.

Jeder der ca. 18.000 aktiven Fasnächtler trägt ein Kostüm mit einer Lave (Maske), unter der der gesamte Körper verdeckt wird. Der Träger des Kostüms ist darunter nicht zu erkennen. Es gehört zum guten Ton, dass ein Fasnächtler sich nicht in der Öffentlichkeit erkennen lässt. In den Cliquen trägt man, ausser beim Morgestraich und am Fasnachts-Dienstag, einheitliche Kostüme, die meist dem "Sujet" (siehe unten) der Clique angepasst sind. Bei den Kostümen kennt die Fantasie keine Grenzen. Oft sind die Masken Personen des öffentlichen Lebens (Politiker und andere Prominente) nachempfunden, manchmal auch Comic-Figuren oder Tieren. Es gibt aber auch ganz traditionelle Laven, die an die [französische Armee der Napoleonischen Kriege](#) erinnern, [Harlekine](#) oder die vor allem bei Einzelpersonen und Wagencliquen beliebten [Waggis-Masken](#).

Am auffälligsten sind die Pfeifer- und Tambouren-Cliquen. Diese spielen bei ihrem Gang durch die Innenstadt Märsche auf ihren [Piccoloflöten](#) begleitet vom Rhythmus ihrer [Basler Trommeln](#). Die Basler Innenstadt wird somit in diesen drei Tagen zum grössten Pfeifkonzert der Welt. Eine Clique besteht meist aus einem Vortrab (Platzmacher), den Pfeifern, dem Tambourmajor und den Tambouren (Trommlern). Die Cliquen folgen dabei keinen bestimmten Routen (ausser am Cortège, siehe unten). Es kommt dabei immer mal vor, dass mehrere Cliquen den Weg kreuzen. In dem Fall bleibt dann eine Clique einfach stehen und wartet, bis die andere passiert hat. Steht man als Zuschauer einer Clique im Weg, wird man vom Vortrab freundlich beiseite geschoben.

Neben den pfeifenden Cliquen gibt es auch [Guggenmusik](#)-Gruppen mit Blechblasinstrumenten. Diese sind allerdings nicht beim Morgestraich anzutreffen, sondern nur bei den Fasnachts-Umzügen (Cortèges) am Montag und Mittwoch und an den Abenden, insbesondere am Dienstagabend bei den Gugge-Konzerten. Weiterhin gibt es zahlreiche Einzelpersonen und kleine Gruppen, die zum Beispiel als [Waggis](#) kostümiert durch die Strassen ziehen und "ahnungslose" Zuschauer mit [Konfetti](#) (im Baseldeutsch [Räppi](#) genannt) bewerfen beziehungsweise es ihnen auch mal in die Jacke stopfen.

Die meisten Cliquen haben sich für die Fasnacht ein so genanntes [Sujet](#) (französisch für "Thema") vorgenommen, das während der Fasnacht präsentiert wird. Es handelt sich dabei um Themen der Zeitgeschichte, an denen in meist satirischer Form Kritik geübt wird. Die Sujets werden beim Morgestraich auf Laternen dargestellt. Beim Cortège sind auch die Costüme und oftmals ein Requisit dem Thema angepasst. Praktisch alle Cliquen verteilen ausserdem passende [Zeedel](#) (Handzettel mit ironischen Versen).



Der [Waggis](#), eines der traditionellen Kostüme bei der Fasnacht

A côté des pages pertinentes, on reçoit aussi beaucoup de résultats avec des arrangements de voyage (voyage en car, hôtel) et des articles de carnaval à vendre. Au lieu d'exclure des termes, la recherche peut être précisée avec d'autres termes liés typiquement au carnaval de Bâle.

3) Recherche avancée dans des moteurs de recherche

Syntaxe de recherche (Google, Yahoo!)

intitle:fasnacht morgestraich OR morgenstraich OR morgenstreich OR bangg OR cortège OR cortege OR helgen "basler fasnacht"

Le terme „Fasnacht“ doit se trouver dans le titre, la phrase „basler fasnacht“ et au moins un des autres termes de recherche doivent se trouver dans le texte. Ceci parce qu'une étude sérieuse sur le carnaval de Bâle doit au moins contenir un de ces termes.

[Google und Yahoo! semblent contenir un thésaurus, Fasnacht donne aussi Fastnacht comme résultat]

Advanced Web Search

You can use the options on this page to create a very specific search. Just fill in the fields you need for your current search.

Yahoo! Search

Show results with

all of these words	<input type="text" value="fasnacht"/>	in the title of the page
the exact phrase	<input type="text" value="basler fasnacht"/>	any part of the page
any of these words	<input type="text" value="morgestraich morgenstraich morgenstrei"/>	any part of the page
none of these words	<input type="text"/>	any part of the page

4) Recherche dans un répertoire

Recherche dans un répertoire avec le terme „Fasnacht“ et une limitation géographique à „Basel-Stadt“. Seule Webliste.ch donne des résultats pertinents (67 résultats 13.2.07). Tous les autres fournisseurs sont décevants : ils donnent de la publicité ou pas de résultats du tout.

Top Rubriken [Fasnacht & Guggen](#) (67x) [Kunst & Kultur](#) (3x) [Transport & Umzug](#) (2x)
Top Branchen [Architektur & Baugeschäfte](#) (1x) [Coiffeur](#) (1x)

[Kottmann Pascal](#) [Kontakt](#)

durchschuss kunst und kommerz muss sich nicht beissen. im gegenteil, werbung oder besser kommunikation muss diese gegensätze auch in verbindung bringen können. darum ist mein tätigkeitsf ...
Letzter Check: 12.02.2007 - [Zeige ähnliche Seiten](#) [Archiv](#)

[Glaibasler Schränzbrieder](#)

Seit 1961 schränzen wir Glaibasler Schränzbrieder an den drei schönsten Tagen durch Basel. Unsere Motivation gründet nicht nur an der Schönheit unserer Fasnacht, sondern auch am gesellig ...
Letzter Check: 12.02.2007 - [Zeige ähnliche Seiten](#) [Archiv](#)

[Guggemuusig Schrupfkepf](#)

Mir sind Wiebli und Männli, bi uns chasch ab 18 Jöhrli mitmache. Unser Hauptziel isch d' Basler Fasnacht
Letzter Check: 12.02.2007 - [Zeige ähnliche Seiten](#) [Archiv](#)

[Guggemuusig Messingkäfer](#)

Am 14.04.1962 treffen sich im Restaurant Viadukt (späteres Happy-Night) 6 Männer zur Gründungsversammlung einer neuen Guggenmusik...
Letzter Check: 12.02.2007 - [Zeige ähnliche Seiten](#) [Archiv](#)

[Charivari s Muusigtheater](#)

Mit dem farblich neu definierten Logo und dem Zusatz: „S Muusigtheater vo dr Fasnacht“ möchten wir unsere Philosophie visualisieren, dass das Charivari vor allem im musikalischen Bereich ...
Letzter Check: 12.02.2007 - [Zeige ähnliche Seiten](#) [Archiv](#)

[Muggedätscher](#)

Die Fasnachtsgesellschaft Muggedätscher ist ein politisch und konfessionell neutraler Verein, der die Tradition der Basler Fasnacht pflegt und Männern sowie Frauen offen steht.
Letzter Check: 12.02.2007 - [Zeige ähnliche Seiten](#) [Archiv](#)

Conclusion

Une recherche avec ,basel' und ,fasnacht' donne un nombre de résultats qui n'est pas gérable. En complétant la recherche avec des termes «typiques» du carnaval, le nombre de résultats diminue et l'on trouve avant tout des pages pertinentes. L'utilisation d'un répertoire local ou d'un portail, dans cet exemple, est plus économique en travail et en temps. Le moteur de recherche local n'est hélas pas encore au point techniquement, mais il donnerait la meilleure couverture du sujet.

4.3.3 Recherche: Agriculture dans le canton de Fribourg

1) Au départ : définir le contenu de la recherche

Quelles informations trouver, collecter et archiver ? [voir «Notice collective», ch. 2.1, 2.2 et 5.2.1]

Cela signifie concrètement pour cet exemple: examiner les informations avec contenu spécialisé (historique, culturel), proposées par des institutions suisses (cantonales, régionales, locales) ou des personnes privées (qualité, quantité et importance).

2) Définir le plan de recherche

Définir le groupe de concepts, les sous-groupes, les sous-thèmes.

Définir les autres noms, synonymes, flexions des termes de recherche :

Définir les langues à utiliser (allemand, français)

Exemple : sous-thèmes

Allemand [Forstwirtschaft](#) [Milchwirtschaft](#) [Tierhaltung](#) [Pflanzenbau](#)

Français [sylviculture](#) [économie laitière](#) [élevage](#) [agriculture](#)

Exemple : synonymes, termes proches, flexions

Allemand [Agrarwirtschaft](#) [Alpwirtschaft](#) [Viehwirtschaft](#) [Nutztierhaltung](#)

[Ackerbau](#) [Weinbau](#)

[Landwirtschaftlich](#) [landwirtschaftliches](#) [landwirtschaftlicher](#) [landwirtschaftlichen](#) [landwirtschaftlichem](#)

Français [exploitation forestière](#) [production animale](#) [production agricole](#) [économie rurale](#) [alpage](#) [viticulture](#) [apiculture](#)

626 [Unbesetzt]
627 Wasserbau
628 Sanitär- und Kommunaltechnik; Umwelttechnik
629 Andere Fachrichtungen der Ingenieurwissenschaften

630

630 Landwirtschaft und verwandte Bereiche
631 Techniken, Ausstattung, Materialien
632 Schäden, Krankheiten, Schädlinge an Pflanzen
633 Feld- und Plantagenfrüchte
634 Obstanlagen, Früchte, Forstwirtschaft
635 Gartenpflanzen (Gartenbau)
636 Viehwirtschaft
637 Milchverarbeitung und verwandte Produkte
638 Insektenzucht
639 Jagd, Fischfang, Naturschutz

640

640 Hauswirtschaft und Familie
641 Essen und Trinken
642 Mahlzeiten, Tischkultur
643 Wohnen, Haushaltsausstattung

Une recherche simple dans un moteur de recherche (comme Google ou Search) donne un nombre de résultats ingérable (36'000 pour agriculture et fribourg en ne comptant que les pages suisses). Dans ces pages on a un mélange de pages commerciales, administratives, politiques et des pages des groupes d'intérêt.

-> L'exclusion de termes choisis ne semble pas pertinente.

Les différents sous-thèmes doivent être recherchés un à un, suivant leur importance.

3) Recherche dans un répertoire ou un portail thématique

a) Recherche dans un répertoire général

Avec le mot-clé [Landwirtschaft](#) et le mot clé ou critère de filtrage [Freiburg](#); ou feuilleter la rubrique [Landwirtschaft](#)

Webliste.ch trouve surtout des pages commerciales.

Kouik.ch a une rubrique [agriculture](#) avec quelques rares résultats

Les recherches dans Sharelook.ch et Linker.ch donnent surtout des résultats commerciaux, étrangers ou pas de résultats du tout.

b) Evaluation des portails et des collections de liens thématiques

La liste de liens de l'Office fédéral de l'agriculture (blw.admin.ch) contient des sites suisses généraux et aussi par exemple les sites des instituts agricoles.

AgroLinks de [www.landwirtschaft.ch](#): la recherche sur le canton de Fribourg donne des références à des pages en français et en allemand

[www.agri-info.ch](#): la recherche avancée sur le canton de Fribourg donne des références à des pages en français et en allemand

-> Lorsqu'ils existent, les répertoires thématiques sont à utiliser car ils donnent accès à des sites web plutôt pertinents.

HOME
SUCHE
KONTAKT

AKTUELL | FACTS | PFLANZEN | TIERE | VOM HOF | SERVICE | SHOP

Home > Service > Agro-Links

Suche nach Links in Agro-Links

Das Suchformular von Agro-Link hilft Ihnen, die gesuchten Weblinks rasch zu finden. Füllen Sie daher die verschiedenen Suchfunktionen möglichst präzise aus. Sie haben ausserdem die Möglichkeit, über Stichworte zu suchen.

[Eintrag ins Adressbuch](#)

Suche

Name

Abkürzung

[Übersicht Kategorien / Gruppen](#)

Kategorie

Gruppe

Sprache d f i e

Kanton oder Land

Stichworte

Alphabetisch

GUT, GIBT'S DIE SCHWEIZER BAUERN.
© 2006 Landwirtschaftlicher Informationsdienst LID

4) Recherche avancée dans les moteurs de recherche

a) Exemple : sous-thème Nutztierhaltung

Recherche dans Search.ch avec différents critères et un filtre sur la région [Freiburg](#)

Recherche [Nutztiere](#) or [Nutztierhaltung](#) or [Tierhaltung](#) or [Viehzucht](#) or [Viehhaltung](#) or [Viehwirtschaft](#) or [Weidewirtschaft](#)

La même recherche avec des termes en français

Filtre région [Fribourg](#)

Recherche [élevage](#) or [bétail](#) or „[production animale](#)“

Les résultats contiennent, outre des pages pertinentes, beaucoup de pages commerciales et de pages qui ne concernent pas le canton de Fribourg.

b) Exemple : sous-thème Nutztierhaltung

Recherche avancée dans [Google.ch](#) et [Yahoo.ch](#) avec une limitation aux **pages suisses**.

Syntaxe de la recherche en allemand :

[Freiburg Nutztierhaltung](#)

[Freiburg Tierhaltung](#)

intitle:tierhaltung fribourg

Syntaxe de la recherche en français :

[élevage fribourg](#)

intitle:élevage fribourg

-> Les quantités énormes de résultats doivent être affinées avec une recherche comme **intitle**.

-> Le problème avec les moteurs de recherche généralistes est le manque de possibilité de limiter la recherche à une aire géographique cantonale ou locale. La recherche avec le mot Fribourg n'est pas concluante puisque des pages qui contiennent les termes Greyerz ou Gruyère mais pas le terme Fribourg ne sont pas trouvées.

[[search.ch](#)]

Die Schweizer Suchmaschine

[Websuche](#) [Telefonbuch](#) [Karte](#) [Wetter](#) [Nachrichten](#) [Assistant](#) | [Anleitung](#)

Anzeigen

- [Scheunentore vom Güller](#)
das bewährte Landwirtschaftstor der Schweizer Bauern, 15% Rabatt
[www.gueller.ch](#)
- [Kostenlose Kleinanzeigen](#)
Gratis inserieren, schnell fündig werden, wenig bezahlen. Los geht's!
[www.Scout24.ch](#)

Hinweis: Kein Dokument enthält alle Stichwörter.

Resultate für 'nutztiere or tierhaltung or viehzucht or nutztierhaltung or viehwirtschaft or viehhaltung or weidewirtschaft' 1-10 von ca. 53 gefundenen Seiten.

1. [Agrarforschung: Veranstaltungen](#)
Agrarforschung ist die Zeitschrift der landwirtschaftlichen Forschungsanstalten: für Nutztiere und Milchwirtschaft (ALP) – für Agrarwirtschaft und Landtechnik, Tänikon – für Agrarökologie und Landbau, Zürich-Reckenholz – für Obst-, Wein- und Gartenbau, Wädenswil
[www.agrarforschung.ch/de/veranst.php?jahr=2005&band=12](#) – 27KB
2. [Agrarforschung: Inhaltsverzeichnis](#)
Agrarforschung ist die Zeitschrift der landwirtschaftlichen Forschungsanstalten: für Nutztiere und Milchwirtschaft (ALP) – für Agrarwirtschaft und Landtechnik, Tänikon – für Agrarökologie und Landbau, Zürich-Reckenholz – für Obst-, Wein- und Gartenbau, Wädenswil
[Mehr Resultate von agrarforschung.ch](#)
3. [NZZ Format: Kühe, Kälber und Kastraten – Superstars der Nutztiere ...](#)
wenn NZZ Format Logo Schmal Reklame-Bild REKLAME Farbbalken mit Navigation LITERATUR ZURUECK ZUR SENDUNG Copyright DTECH,loc=900;grp=[group]
[www.nzz.ch/format/broadcasts/literatures_156.html](#) – 28. Jan 2007 – 15KB
4. [SanaSativa – über Futterhanf](#)
Die SanaSativa®reg; AG föm;l;rdert und unterstützt den Anbau und die Verarbeitung von Bauernhanf
[www.sanasativa.ch/index.php?option=com_content&task=view&id=35&Item...](#) – 27. Jan 2007 – 26KB
5. [SanaSativa – über Futterhanf](#)
Die SanaSativa®red; AG föm;l;rdert und unterstützt den Anbau und die Verarbeitung von Bauernhanf

YAHOO! RECHERCHE [Web](#) | [Images](#) | [Vidéos](#) | [Actualités](#)

tout le Web en français en Suisse

Résultats Web 1 - 10 sur environ 122 pour **intitle:élevage fribourg** - 0.89 s (Ergäbnisse auf Seite)

1. [zu13r Résultats en exposition de l'Elevage Bois des Amourettes Andrea Carrasco-Meissner Moudon Suisse spaniels ...](#)
Groupe d'élevage **Fribourg** 2005. Flamenco Andaluz du Bois des Amourettes ... un excellent caractère, qui d'ailleurs fait la renommée de notre **élevage** ...
[dog-top-info.ch/dog/zuchter/zu13r.htm](#) - 33k - [En cache](#) - [Plus de pages provenant de ce site](#)
2. [Elevage de Berger Allemand du Boidamont à Fribourg en Suisse](#)
Elevage de Berger Allemand à **Fribourg** en Suisse, nous travaillons beaucoup avec l'étranger.
[www.elevageduboidamont.com](#) - 6k - [En cache](#) - [Plus de pages provenant de ce site](#)
3. [Kouik.ch - Suisse romande - Agriculture, élevage - Fribourg, Genève, Jura, Jura bernois, Neuchâtel, Vaud, Valais ...](#)
Guide : Agriculture, alimentation et restaurants " Agriculture, **élevage** " Produits de la ferme ... Formation, aide à la décision, promotion de nouveaux ...
[kouik.ch/search/.../Agriculture,D44_233elevage](#) - 9k - [En cache](#) - [Plus de pages provenant de ce site](#)
4. [Elevage de Cocker, Springer Spaniel Anglais, Syringa, Suisse, Chenil chiens et chats Fribourg Suisse, pension Maison ...](#)
Elevage de syringa, springers, spaniels anglais, cockers. ... Présentation | **Elevage** de Syringa | Springers Spaniels anglais | Cockers ...
[www.chenil-maisonrouge.com](#) - 19k - [En cache](#) - [Plus de pages provenant de ce site](#)
5. [Elevage du Cousimbart - Border Collie - Le Mouret / Fribourg - BOURQUI Daniel et Antoinette](#)
Elevage du Cousimbart - Border Collie - CH-1724 Le Mouret / **Fribourg** - BOURQUI ... **Elevage** du Cousimbart - Border Collie - CH-1724 Le Mouret / **Fribourg**.
[www.elevage-du-cousimbart.ch](#) - 10k - [En cache](#) - [Plus de pages provenant de ce site](#)
6. [Charmey \(Gruyères\) Pays de Fribourg Tiscover Site/Carte élevage de bétail Gastlosen sport d'hiver](#)
Sports d'hiver s'écrit en majuscules ici car la neige tombe en quantité suffisante. ... **Fribourg**/Neuchâtel/Jura/Jura bernois / Pays de **Fribourg** / Charmey ...
[tiscover.ch/chvguide/5ch,fr,SCH1/.../map.html](#) - 17k - [En cache](#) - [Plus de pages provenant de ce site](#)
7. [Charmey \(Gruyères\) Pays de Fribourg Tiscover Bienvenue élevage de bétail Gastlosen sport d'hiver](#)
Sports d'hiver s'écrit en majuscules ici car la neige tombe en quantité suffisante. ... **Fribourg**/Neuchâtel/Jura/Jura bernois / Pays de **Fribourg** / Charmey ...
[tiscover.ch/chvguide/5ch,fr,SCH1/.../home.html](#) - 42k - [En cache](#) - [Plus de pages provenant de ce site](#)
8. [installations destinées à l'élevage des animaux de rente \(PDF\)](#)
Office de la protection de l'environnement du canton de **Fribourg** (OPEN) 1 ... de **Fribourg**, les présentes directives servent à l'élaboration de l'enquête ...
[www.fr.ch/senv/fr/eie/pdf/animaux_rente.pdf](#) - 40k - [Afficher en html](#) - [Plus de pages provenant de ce site](#)

Conclusion

La recherche de pages web pertinentes pour la Suisse, les cantons et les régions pose aux spécialistes des problèmes. A partir de l'exemple de l'agriculture fribourgeoise nous avons fait les expériences suivantes :

- Le thème „agriculture“ est très vaste : il doit être défini précisément, limité et subdivisé. Quelles entreprises doivent être recherchées, quels sont les aspects qui sont importants pour un canton ?
- Contrairement au sujet „Cervin“ (géographiquement bien défini), il faut pour l'agriculture limiter l'aire géographique à un canton lors de la recherche. Cette limitation peut, dans les moteurs de recherche généralistes, donner des résultats incomplets. C'est pourquoi on utilisera de préférence des moteurs et des répertoires qui permettent un filtrage par canton (par ex. Sear.ch)
- On préférera, s'ils existent, les répertoires et portails thématiques aux répertoires généraux, souvent orientés „commerce“. Les recherches dans les répertoires thématiques doivent être faites avec les méthodes définies au début de la recherche.
- Un des problèmes étant l'énorme quantité de résultats trouvés, il faut appliquer les possibilités de recherche avancée, en particulier les possibilités de limitation de la recherche.
- La procédure de recherche doit être refaite pour chaque sous-thème et éventuellement pour chaque entreprise. Cela représente beaucoup de travail et de temps, en particulier lorsque l'on fait la recherche pour la première fois. Il faut en effet pour chaque sujet chercher d'abord quels sont les mots à utiliser et quels répertoires, moteurs de recherche, possibilités de recherche sont susceptibles de donner de bons résultats.

5 Concepts

[Choix de concepts en relations avec la technologie et les services d'internet]

Service d'alerte, de veille

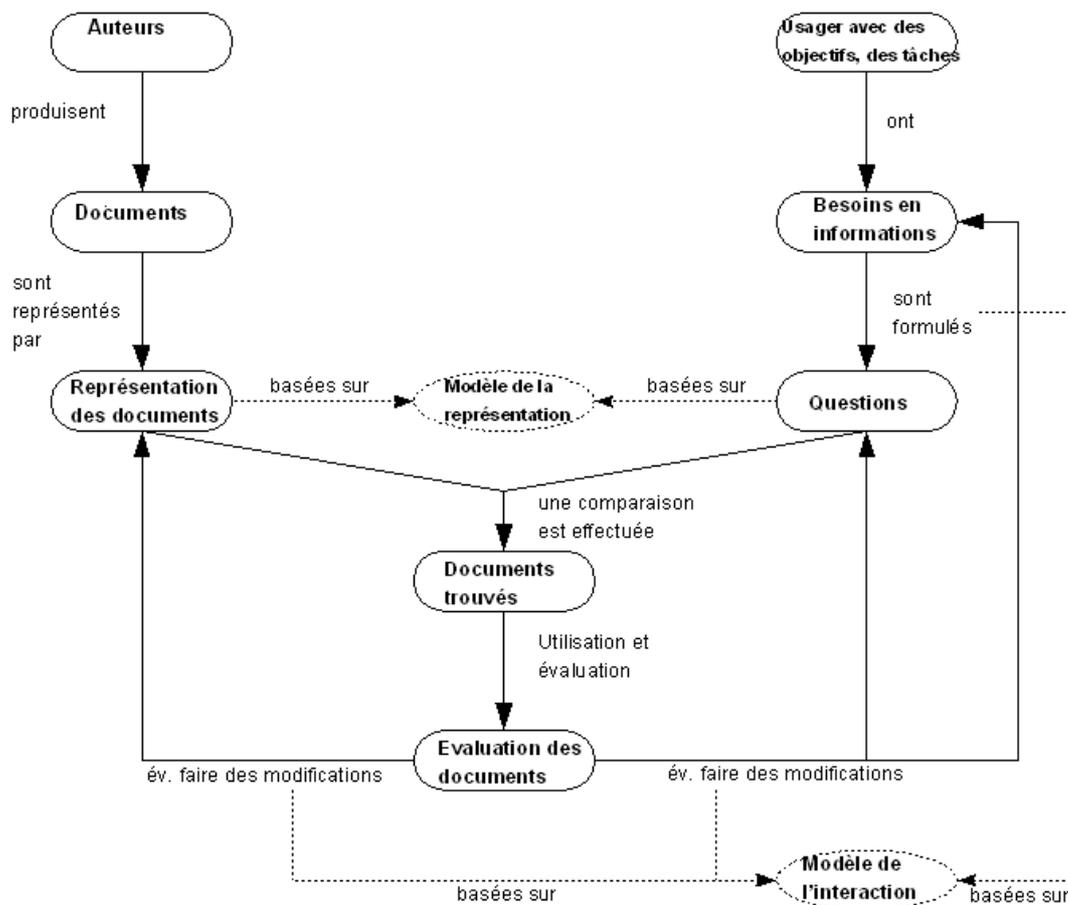
Les services d'alerte, de diffusion sélective de l'information (DSI) ou de Current-Awareness permettent de recevoir régulièrement, après avoir défini un profil de recherche, des informations sur le sujet recherché.

Les informations reçues peuvent être, par exemple, de nouveaux résultats dans un moteur de recherche, des contributions dans un blogue, des tables des matières de revues spécialisées.

Dans un moteur de recherche, les résultats sont cherchés dans une collection de documents déjà présente, suite au lancement d'une recherche par un usager. Les services d'alerte recherchent en continu les nouveaux documents qui répondent à une question prédéfinie et envoient à l'usager un message chaque fois qu'ils en trouvent un. Le message contient soit le document complet, soit une référence au document, soit un résumé.

Système de recherche d'informations (Information Retrieval)

Exemple d'un système de recherche d'information :



Podcasting

Un podcast est une collection de documents audio ou vidéo qui sont offerts sur internet pour être téléchargés et consultés plus tard par l'utilisateur. Par l'entremise d'un abonnement aux flux Really Simple Syndication (RSS), le *podcasting* permet aux utilisateurs d'automatiser le téléchargement d'émissions audio ou vidéo, notamment pour leur baladeur numérique (tel que l'iPod par exemple), sur leur ordinateur personnel, pour une écoute immédiate ou ultérieure. Le terme « podcasting » est une combinaison des mots iPod et Broadcasting (radiodiffusion).

Portails

La notion de «portail» naît dans les années 90 et est utilisée pour désigner de multiples produits comme par exemple: les pages d'accueil de moteurs de recherche, une page d'accueil qui offre un point d'accès dans internet sur un sujet ou des pages web qui proposent des collections de liens. Il n'existe pas de définition précise du concept de portail. Au lieu de cela, on trouve toutes sortes de variantes: Community Portal, Corporate Portal, Enterprise Portal, Enterprise Information Portal, portail d'entreprise, Business-to-Employee-Portal, portail de fournisseur, Knowledge Management Portal, portail scientifique, portail de bibliothèque, pour n'en citer que quelques unes.

La définition du portail du Fraunhofer Instituts für Arbeitswirtschaft und Organisation (IAO) ne considère comme portail que les applications qui ont des caractéristiques bien précises :
«Un portail est une application basée sur les technologies du web, qui offre un accès centralisé à certains contenus personnalisés, ainsi qu'à des procédures établies selon les besoins.
La liaison et l'échange de données entre des applications hétérogènes par le biais d'une plateforme sont les caractéristiques d'un portail. Grâce au Single-Sign-On, il n'est plus nécessaire de se connecter individuellement aux applications intégrées dans le portail ; l'accès se fait au niveau central, par une interface homogène. Les portails donnent la possibilité à des groupes hétérogènes de partager des procédures et de collaborer.»

Un portail, pour mériter cette appellation, doit offrir des possibilités de personnalisation de l'offre d'information. Cette fonctionnalité se base sur des profils, qui peuvent être liés à des rôles spécifiques. Cela signifie que l'utilisateur accède aux informations qui sont en rapport avec le rôle qui a été défini pour lui. (Guba, G. (2006): *Bibliotheken und Portale – ein Überblick*. In: *GMS Med Bibl Inf*. 6(2): Doc 15)

Ce qui différencie les portails des pages web :

1. Un seul point d'entrée vers les ressources à disposition dans un domaine
2. Interaction personnalisée avec les services du portail
3. Accès centralisé aux différents types de données, aux répertoires et aux banques de données
4. Technologies qui encouragent la coopération entre les collaborateurs
5. Intégration des portails avec les procédures et les applications

RSS

Le «RSS» (Rich Site Summary ou really simple syndication) a été développé à la base en tant que système d'échange de données ; il permet d'intégrer d'une façon simple une partie d'une page web (par exemple le chapeau d'un article) dans d'autres pages web. Des développements ultérieurs ont permis d'élargir les possibilités de cette application et de rendre son format plus flexible. Ainsi, le RSS actuel est utilisé par un nombre croissant d'utilisateurs pour la diffusion d'informations diverses, comme les résumés des informations de presse, les dépêches d'agence, les blogues ou la surveillance de l'efficacité de serveurs.

Les contenus préparés par les fournisseurs en format RSS sont appelés (RSS-)Feed.

Web 2.0

L'idée du web 2.0 (les «Wiki» et les blogues appartiennent aussi à cette catégorie) est de déposer dans le web des logiciels d'application ; la seule application dont aura ensuite besoin un utilisateur d'internet est un navigateur web. Les services ne sont pas ouverts uniquement à leurs utilisateurs mais interagissent entre eux. Cela conduit l'utilisateur à une nouvelle conception de l'utilisation du web ; l'élément de coopération devient, grâce à la technologie du web, un aspect essentiel. Le «web 2.0» est donc à comprendre comme une attitude plus que comme une technologie.

Blogue ou weblog

Le terme «weblog» est un assemblage de «web» et «log» (pour Logbuch). La forme courte est «blog» ou «blogue». Un ou plusieurs auteurs peuvent publier sur ce type de pages web leurs contributions. Les blogues se situent à l'intersection entre le journal intime et livre d'hôtes. Les nouveautés sont toujours placées en début de page et relèguent les contributions plus anciennes vers le bas. La saisie est faite en ligne, sans avoir besoin de connaître le format HTML. Les données qui se trouvent sur un blogue sont appelées des «bloggings», les auteurs des blogueurs..

Wiki

Le mot «wiki» est d'origine hawaïenne et signifie «rapide». Les «wiki» sont des collections de pages web libres d'accès qui peuvent non seulement être lues par les usagers mais aussi être modifiées en ligne. Elles fonctionnent avec un «Content Management System».

<http://fr.wikipedia.org/wiki/Accueil>

<http://wiki.org/>

6 Sources

Karzauninkat, Stefan: Die Suchfibel. URL <http://www.suchfibel.de/> (Stand: 13.04.2008)

Lewandowski, Dirk (2005): Web Information Retrieval. In: Information; Wissenschaft und Praxis 56(1), 5-12

Lewandowski, Dirk (2005): Web Information Retrieval. Technologien zur Informationssuche im Internet. URL: <http://www.durchdenken.de/lewandowski/web-ir/> (Stand: 11.01.2006)

Patzwaldt, Klaus (2006). Semantische Suchmaschine Theseus kommt aus Deutschland. URL <http://www.at-web.de/blog/?p=793> (Stand 16.01.2006)

Stock, W.G., Lewandowski, D. (2006): Suchmaschinen und wie sie genutzt werden. In: WISU – Das Wirtschaftsstudium 8-9, 1-7 (Preprint)

Weilenmann, Anne-Katharina (2006): Fachspezifische Internetrecherche. Bibliothekspraxis; 38. München : K.G. Saur

Wolf, Sebastian (2006): „Suche im Internet“. URL: <http://www.ub.uni-bielefeld.de/biblio/search/help/> (Stand: 11.01.2007)