

Utilisation des instruments de recherche disponibles sur internet

Autor(en): **Pythoud, Christophe**

Objektyp: **Article**

Zeitschrift: **Arbido**

Band (Jahr): **13 (1998)**

Heft 5

PDF erstellt am: **09.06.2024**

Persistenter Link: <https://doi.org/10.5169/seals-770265>

Nutzungsbedingungen

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern.

Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden.

Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

Haftungsausschluss

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

Zielsetzungen und die Umsetzungsstrategien. Das Dokument weist ein Stärken- und Schwächen-Profil auf.

Nachträglich in ein bereits bestehendes Redaktionssystem ein Datenbankmodul einbauen zu wollen, ist ein teures und schwieriges Unterfangen. Das Redaktionssystem darf nicht nur allein zur Produktion der Zeitung konzipiert sein, sondern es muss auch für andere Formen der Publikation wie Internet oder CD-ROM und Datenbanken offen sein.

Informationsbeschaffung ist nicht allein Sache der Journalisten. Einen integrativen Ansatz leistet beim Verfassen des Artikels der Pressedokumentalist, der bei Bedarf die Vielzahl der Informationsquellen im Hinblick auf die konkrete Fragestellung des Journalisten absucht. Das Informationsangebot ist unter Beachtung der Kosten-Nutzen-Relation dem Journalisten aufzuzeigen. Auswahl und Bearbeitung der Information muss Sache des Journalisten bleiben.

Bei der heutigen elektronischen Zeitungsarchivierung ist die Speicherung der graphischen Layout-Form mehr hinderlich als förderlich, da die Wiederverwendungsform von morgen nicht bekannt ist. Reiche, aber neutrale Datenformen wie XML bieten sich dafür an.

Da Recherchen selten ein zweites Mal verwendet werden

können, soll das Gewicht weniger auf die Speicherung der Resultate in house gelegt werden als auf die Kenntnisse der Dokumentalisten, wo was schnell und kostengünstig beschafft werden kann.

Modern konzipierte Dokumentationszentren haben weniger etwas mit der klassischen Nachweisfunktion des alten Zeitungsarchivs zu tun, als mit Stellen der Informationsbeschaffung, vergleichbar mit einer Informationsbeschaffungsdrehscheibe (carrefour d'orientation).

Im Wissen um die Möglichkeiten der Informationsbeschaffung für Foto und Text wird die Zukunft des Pressedokumentalisten liegen; neu wird für ihn die Rolle des Informationsbeschaffungslogistikers sein. Journalisten und Bildredaktoren werden sich vermehrt auf die qualitativen Aspekte der Auswahl und der Bearbeitung der Text- und Bildinformation konzentrieren.

Zyklisch wiederkehrende Informationen zu beschaffen (wie etwa Inhaltsangaben von Filmen), Bilder mit vornehmlich illustrativem und weniger mit Informationscharakter nutzbar zu machen und im richtigen Moment verfügbar zu halten, - dies alles wird zu den neuen Aufgaben der Presse-dokumentation gehören.

UTILISATION DES INSTRUMENTS DE RECHERCHE DISPONIBLES SUR INTERNET

par Christophe Pythoud

Assistant linguistique générale à l'Université de Lausanne

Cet article reprend les éléments d'un exposé donné au mois de février 1998 dans le cadre du séminaire intitulé «Internet für Informationsvermittler». Comme tout ce qui touche à Internet, les informations présentées lors de cette conférence ont probablement évolué depuis lors. Toutefois, l'essentiel de ce qui est présenté ici devrait demeurer d'actualité pendant les quelques mois à venir.

Les informations, comme les jugements de valeur, qui figurent dans ces pages proviennent de trois sources: la presse spécialisée, Internet et mon expérience personnelle. Ces trois sources sont subjectives et j'encourage donc le lecteur à acquérir sa propre expérience des instruments de recherche, et à se forger sa propre opinion, en lisant cet article devant son ordinateur. Les transparents de la présentation originale sont disponibles à l'adresse suivante: <http://www.unil.ch/ling/cp/search.html>.

Les informations disponibles sur Internet sont de nature très diverses: on y trouve bien sûr tous les sujets traités sur le World Wide Web et dans les listes de discussion *Usenet*, mais aussi :

- ◆ des informations financières et boursières;
- ◆ les coordonnées électroniques ou postales de personnes habitant n'importe où dans le monde;
- ◆ l'équivalent des pages jaunes répertoriant diverses entreprises (ce service concerne essentiellement les firmes des Etats-Unis);
- ◆ des services payants donnant accès à des informations

qui ne sont en général pas directement disponibles sur Internet (chapitres de livres, rapports, articles de revues, etc.);

◆ etc.

Pour exploiter rapidement et efficacement la grande quantité d'informations disponibles sur Internet, il existe, grosso modo, trois catégories d'instruments de recherche:

- ◆ les répertoires (en anglais *directories*);
- ◆ les moteurs de recherche (en anglais *search engines*);
- ◆ les outils de personnalisation.

Une brève définition de chaque catégorie est présentée ci-dessous.

LES RÉPERTOIRES (DIRECTORIES)

Les répertoires répartissent les pages WWW en catégories. Par exemple, la version française de *Yahoo!*

(<http://www.yahoo.fr/>) propose le découpage suivant:

- ◆ Actualités et presse
- ◆ Art et culture
- ◆ Commerce et économie
- ◆ Divertissement
- ◆ Enseignement et formation
- ◆ Exploration géographique
- ◆ Informatique et multimédia
- ◆ Institutions et services publics
- ◆ Références et annuaires
- ◆ Santé
- ◆ Sciences et technologies

- ◆ Sciences humaines et sociales
- ◆ Sports et loisirs
- ◆ Sujets de société

Les répertoires sont organisés comme les dossiers que vous trouvez sur votre disque dur: cliquer sur l'un d'entre eux révèle des liens vers divers sites, mais également des sous-répertoires proposant une classification plus fine. Par exemple, si je clique sur «Sports et loisirs», j'obtiens une nouvelle liste:

- ◆ Activités en plein air
- ◆ Animaux
- ◆ Aviation
- ◆ Danse
- ◆ Emploi
- ◆ etc.

Si je clique maintenant sur «Animaux», j'obtiens encore une nouvelle liste («Chats», «Chevaux», «Chiens», «Droits des animaux», «Insectes», etc.) et ainsi de suite. Le découpage en répertoires est effectué par des humains, et non automatiquement par un programme; ceci évite toute incohérence. De ce fait, parcourir un répertoire constitue une bonne manière de commencer une recherche: plus vous descendrez dans l'arborescence des répertoires, plus l'information deviendra spécifique.

Les répertoires les plus connus sont les suivants (il en existe des versions régionales, en français, en allemand, en italien, etc., mais ci-après je donne l'adresse des versions «internationales»):

- ◆ **Yahoo!** (<http://www.yahoo.com/>)
- ◆ **Infoseek** (<http://www.infoseek.com/>)
- ◆ **Lycos** (<http://www.lycos.com/>)

L'utilisation des répertoires est très rentable pour des sujets courants et/ou grand public. Si ce que vous cherchez relève d'un domaine très spécialisé, il est sans doute préférable de commencer vos investigations à l'aide d'un moteur de recherche.

LES MOTEURS DE RECHERCHE (SEARCH ENGINES)

Les moteurs de recherche sont sans doute les outils les plus efficaces pour tirer profit de l'immense quantité d'informations disponibles sur Internet. Ils utilisent des procédés d'indexation sophistiqués pour garder trace des différents sujets auxquels des sites WWW sont consacrés et de l'apparition de ces mêmes sujets dans les listes de discussion *Usenet*. Ils peuvent donc, à partir des informations stockées dans leur base de données, établir la liste de tous les sites traitant peu ou prou des thèmes qui vous intéressent. Les moteurs de recherche actuellement considérés comme les plus puissants sont:

- ◆ **AltaVista** (<http://www.altavista.digital.com/>)
- ◆ **HotBot** (<http://www.hotbot.com/>)
- ◆ **Infoseek** (<http://www.infoseek.com/>)

LES OUTILS DE PERSONNALISATION

Les outils de personnalisation vous permettent de vous créer une «*homepage*» contenant des informations provenant d'un répertoire ou d'un moteur de recherche. Ils vous permettent, lorsque vous vous connectez à un de ces outils, de voir s'afficher, en lieu et place de l'écran d'accueil générique, une page personnalisée contenant les liens les plus récents vers les sujets qui vous intéressent le plus. Les outils de personnalisation sont un moyen simple, et souvent amusant, de se

tenir au courant, sans efforts de recherche, des développements récents dans un domaine précis. Pour utiliser ces services, il faut s'inscrire en obtenant un nom d'utilisateur (*username*) et un mot de passe (*password*). L'inscription est gratuite, mais vous devez fournir quelques indications personnelles comme votre âge, votre lieu de résidence et vos centres d'intérêts. La plupart des services s'engagent à ne pas diffuser ces informations, mais s'en servent pour afficher à l'écran des publicités personnalisées en fonction de votre profil. *Yahoo!* est sans doute le site qui offre les outils de personnalisation les plus complets et les plus conviviaux.

LE FONCTIONNEMENT DES MOTEURS DE RECHERCHE

Avant de présenter plus spécifiquement certains des outils évoqués ci-dessus, j'aimerais fournir quelques précisions sur le fonctionnement des moteurs de recherche. Connaître ces informations vous permettra de mieux choisir vos critères de recherches. Vous aurez aussi des attentes plus réalistes vis-à-vis des moteurs de recherche, car, quelle que soit leur puissance de traitement et l'étendue de leur base de données, il ne faut pas croire aux miracles.

Activités quotidiennes d'un moteur de recherche

Un moteur de recherche occupe son temps aux activités suivantes:

- ◆ il envoie des robots parcourir Internet et collecter le texte des pages WWW disponibles;
- ◆ les informations ainsi obtenues sont traitées par des logiciels d'indexation et organisées dans une très grande base de données;
- ◆ lorsqu'il reçoit une requête, le moteur de recherche va chercher dans cette base de données les adresses de toutes les pages qui contiennent des informations en rapport avec la requête;
- ◆ les réponses obtenues sont triées en fonction de leur pertinence avant d'être présentées à l'utilisateur.

Les robots

Les robots, parfois aussi appelés «araignées» (*spiders*), sont des programmes qui parcourent Internet à la recherche de nouvelles pages ou de pages dont le contenu a changé. Ils copient une partie, ou la totalité, de ce qu'ils trouvent et la renvoient à leur point de départ. Les concepteurs de certains robots ont en effet choisi de ne copier que les premières lignes de chaque page, ou alors la moitié seulement de celle-ci, en partant du principe, discutable, que les informations les plus importantes apparaissent en début de document. D'autres robots privilégient les sites souvent réactualisés. Ici, on part du principe que la fréquence de mise à jour augmente la pertinence ou l'exactitude de l'information, ce qui pourrait aussi être sujet à question.

Si vous ne souhaitez pas que votre site soit indexé, il est possible d'en interdire l'accès aux robots. Vous trouverez comment faire à l'adresse suivante :

<http://info.webcrawler.com/mak/projects/robots/norobots.html>
Inversement, il est possible de leur faciliter la tâche en utilisant les balises HTML <META> dans vos pages WWW.

L'indexation

Nous avons vu ci-dessus que les robots participaient à l'élaboration d'un grand index des pages WWW disponibles: en effet, lorsqu'un moteur de recherche reçoit une requête, il ne

commence pas à parcourir Internet à la recherche de sites satisfaisant aux critères soumis par l'utilisateur, mais il consulte la base de données établie à l'aide des robots afin de fournir une réponse rapide. Les stratégies d'indexation varient d'un site à l'autre. Certains moteurs de recherche, comme *AltaVista* ou *Infoseek* indexent tous les mots de chaque page. D'autres se contentent des premières phrases, ou de la moitié de la page, ou des premières pages du site, selon le type de robots utilisés, comme évoqué plus haut. Il est clair que plus l'index est grand, plus les réponses fournies aux requêtes seront nombreuses. Mais plus les réponses seront nombreuses, plus il sera important d'essayer de les présenter par ordre d'importance.

Le tri des réponses

Un moteur de recherche ne peut se contenter de donner à l'utilisateur la liste des sites correspondant à sa requête: il doit essayer de les présenter par ordre d'importance. Pour ce faire, les stratégies suivantes sont employées, tantôt conjointement, tantôt séparément:

- ◆ on compte le nombre d'occurrences dans le document de chaque mot apparaissant dans la requête, en partant du principe que plus un terme apparaît souvent dans la page WWW, plus le contenu de celle-ci est pertinent pour le sujet considéré;
- ◆ on vérifie que les éléments recherchés apparaissent au début du document, en partant du principe que le thème de la page WWW est indiqué au début de celle-ci;
- ◆ si la requête comporte plusieurs mots, on privilégie les documents où ils apparaissent dans des positions proches les uns des autres.

PRÉSENTATION DE QUELQUES INSTRUMENTS DE RECHERCHE

Nous allons maintenant, dans les lignes qui suivent, examiner quelques-uns des outils mentionnés précédemment, à savoir:

- ◆ *Yahoo!* (<http://www.yahoo.com/>)
- ◆ *AltaVista* (<http://www.altavista.digital.com/>)
- ◆ *Infoseek* (<http://www.infoseek.com/>)
- ◆ *Northern Light* (<http://www.nlsearch.com/>)

Yahoo!

Yahoo! essaie de remplir à la fois toutes les fonctions:

- ◆ c'est le plus grand répertoire disponible sur Internet (nous en avons examiné le fonctionnement plus haut);
- ◆ c'est également un moteur de recherche (en fait, il utilise *AltaVista* pour effectuer ses recherches, mais présente une interface plus conviviale);
- ◆ c'est finalement un des meilleurs outils de personnalisation disponibles, dont je recommande fortement l'utilisation: <http://my.yahoo.com/>.

Yahoo! propose également de nombreuses prestations complémentaires:

- ◆ un service d'informations financières et boursières à l'adresse <http://quote.yahoo.com/>, qui peut fournir des informations intéressantes, mais il ne faut pas oublier que les quotations boursières fournies sont livrées avec vingt minutes de retard;
- ◆ un annuaire, basé sur l'excellent site *Four11* (<http://www.four11.com/>);
- ◆ des pages jaunes (<http://local.yahoo.com/>), peu intéres-

santes pour un Européen;

- ◆ un service de courrier électronique gratuit;
- ◆ etc.

AltaVista

AltaVista (<http://www.altavista.digital.com/>) fut le premier moteur de recherche à rencontrer un large succès sur Internet. Il a été développé dans les laboratoires de recherches de Digital Equipment Corporation (DEC), à Palo Alto, entre 1993 et 1995. Il s'agit toujours du moteur de recherche le plus puissant, tant du point de vue de la quantité d'informations disponibles, quelque cent millions de pages archivées alors que les moteurs concurrents n'en indexent que cinquante ou soixante millions, que du point de vue de la précision des requêtes, qui peuvent être extrêmement élaborées si l'on maîtrise bien la syntaxe des opérateurs booléens.

Parmi les autres qualités d'*AltaVista*, on peut citer la prise en charge des aspects linguistiques:

- ◆ l'interface du moteur est disponible dans vingt-cinq langues différentes;
- ◆ il est possible de choisir la langue dans laquelle on souhaite obtenir des informations;
- ◆ un service de traduction automatique est proposé; il n'est pas très bon, mais se révèle quand même utile lorsqu'on est confronté à une page rédigée dans une langue qu'on ne connaît pas du tout (si vous voulez vous amuser, demandez au système de traduire une page, puis faites-lui retraduire le résultat dans la langue d'origine).

AltaVista distingue deux types de requêtes: les requêtes simples (*simple search*) et les requêtes avancées (*advanced search*). Une requête simple consiste essentiellement en une liste de mots en rapport avec le sujet qui vous intéresse. Cependant, il est possible de préciser des contraintes supplémentaires:

- ◆ pour préciser qu'un mot doit absolument figurer dans les documents trouvés, vous pouvez le faire précéder du signe «+» (par exemple: +mot);
- ◆ on peut également indiquer qu'un mot ne doit jamais être présent en lui préfixant le signe «-» (-mot);
- ◆ on peut aussi demander à *AltaVista* de chercher des bouts de phrase «tels quels» en les faisant figurer entre guillemets (par exemple: "mot ... mot");
- ◆ il faut également savoir que la présence de majuscules dans un mot joue un rôle important; si vous entrez la requête agip, vous obtiendrez les pages contenant agip, Agip ou AGIP; mais si la requête est Agip, vous ne trouverez que Agip; si un mot contient une ou plusieurs majuscules la recherche devient sensible à la casse;
- ◆ il est possible d'utiliser le caractère joker «*»: par exemple Bahn* renvoie à tous les mots commençant par ces quatre lettres et sélectionnera les documents contenant Bahn, mais aussi Bahnhof, Bahnhofstrasse, etc.

Les requêtes avancées offrent, en plus de ce qui précède, quatre opérateurs booléens. Ils sont présentés dans le tableau ci-dessous.

Mot-clé	Symbole	Exemple
AND	&	Antilles AND vacances
OR		Antilles OR Martinique
NOT	!	Antilles NOT Martinique
NEAR	~	Antilles NEAR Martinique

AND permet de préciser que deux termes doivent apparaître conjointement dans le même document: dans l'exemple, je demande tous les documents qui contiennent les mots *Antilles* et *vacances*.

OR permet d'indiquer que l'un ou l'autre terme (ou les deux) doit apparaître dans le document: par exemple tous les documents qui mentionnent les Antilles ou/et la Martinique.

NOT permet de préciser qu'un terme ne doit pas apparaître dans le document: par exemple, je pourrais souhaiter des informations sur les Antilles, mais à l'exclusion de la Martinique.

NEAR signifie que les deux termes de la requête ne doivent pas être séparés dans le texte de plus de dix mots: dans l'exemple, je souhaite obtenir les documents contenant les mots *Antilles* et *Martinique* dans la même phrase (distants de moins de dix mots).

Il est possible de composer des requêtes complexes en utilisant des parenthèses, par exemple: (Antilles OR Martinique) NEAR Vacances.

Lorsque vous effectuez des requêtes avancées, vous perdez la possibilité d'employer les signes plus et moins (+mot et -mot), mais cela est compensé par la flexibilité des opérateurs booléens. (Et réciproquement, vous ne pouvez pas utiliser les opérateurs booléens dans les requêtes simples.)

Dans les requêtes simples comme avancées, il est également possible de préciser l'endroit du texte où l'on souhaite trouver certains mots de la recherche:

- ◆ dans le titre (title:Antilles)
- ◆ dans le texte (text:vacances)
- ◆ dans l'adresse Internet du document (url:antilles)
- ◆ etc.

On peut également limiter la recherche à des sites appartenant à un certain domaine (par exemple uniquement les sites suisses: domain:ch).

AltaVista offre encore bien d'autres facilités de tri et je vous invite à consulter l'aide en ligne de ce moteur pour en savoir plus. Pour les recherches simples, voyez la page:

http://www.altavista.digital.com/av/contents/help_simple.html

Pour les requêtes avancées, consultez:

http://www.altavista.digital.com/av/contents/help_advanced.html

Lorsqu'on maîtrise la syntaxe des requêtes d'*AltaVista*, on peut généralement se débrouiller sans trop de peine avec celle des autres moteurs de recherche qui s'en sont largement inspirés. La seule exception notable est *Infoseek*.

Infoseek

Infoseek (<http://www.infoseek.com/>) est un concurrent de *Yahoo!*: il propose un répertoire, un moteur de recherche et des outils de personnalisation.

Northern Light

Northern Light (<http://www.nlsearch.com/>) est un nouveau venu dans le domaine des instruments de recherche. Sa grande originalité est d'étendre ses recherches à des documents qui ne sont pas disponibles en libre accès sur Internet, comme des rapports, des livres, des revues, etc. Pour ce type d'informations, le service est payant: entre un et quatre dollars par document, avec la possibilité de conclure un abonnement mensuel à cinq dollars, qui permet de réduire un peu les coûts. J'ai obtenu des résultats très satisfaisants

sur différents types de requêtes et il y a fort à penser que *Northern Light* prendra bientôt place parmi les moteurs de recherche les plus en vue sur Internet. La syntaxe employée pour encoder les requêtes est proche de celle d'*AltaVista*.

De quelques autres systèmes...

Je n'ai pas la place pour traiter des nombreux autres instruments de recherche existants. Il faut cependant mentionner les plus performants d'entre eux:

- ◆ *Excite* (<http://www.excite.com/>)
- ◆ *HotBot* (<http://www.hotbot.com/>)
- ◆ *Lycos* (<http://www.lycos.com/>)
- ◆ *WebCrawler* (<http://www.webcrawler.com/>)

Outils spécialisés

Certains sites se spécialisent dans la recherche d'informations d'un type particulier:

- ◆ *Four11* (<http://www.four11.com/>) permet la recherche des adresses électroniques ou postales de personnes vivant à peu près partout dans le monde;
- ◆ *BigBook* (<http://www.bigbook.com/>) offre un service de pages jaunes pour les entreprises américaines (États-Unis et Canada);
- ◆ en Suisse, on dispose de *FirmIndex* (<http://www.firmindex.ch/>);
- ◆ *Deja News* (<http://www.dejanews.com/>) permet de rechercher des informations dans les listes de discussion *Usenet* (*AltaVista* est également très utile dans ce domaine).

Les méta-recherches

Il arrive parfois qu'un moteur de recherche ne trouve aucun site consacré au sujet qui vous intéresse, ou que vous souhaitiez vraiment obtenir une information exhaustive (par exemple la liste de toutes les pages WWW qui possèdent un lien pointant vers le serveur de votre entreprise ou de votre service). Dans ce cas, il est souvent nécessaire d'utiliser un second moteur de recherche, puis un troisième, etc., ce qui ne simplifie pas le tri des résultats. Pour automatiser cette procédure, il existe des moteurs de méta-recherche (*meta-search engines*). Ces sites se chargent pour vous d'invoquer les différents moteurs de recherche et d'unifier les résultats obtenus. Les moteurs de méta-recherche les plus performants sont:

- ◆ *MetaCrawler* (<http://www.metacrawler.com/>);
- ◆ *ProFusion* (<http://www.designlab.ukans.edu/profusion/>).

POUR ALLER PLUS LOIN...

Bien que j'espère que les considérations qui précèdent vous soient utiles, elles ne sauraient remplacer un contact pratique avec les instruments de recherche. Pour prendre en mains un ou plusieurs de ces outils, essayez de trouver les choses suivantes:

- ◆ le site WWW d'un journal publié dans le canton où vous habitez;
- ◆ la recette du Papet vaudois ou du Bami Goreng;
- ◆ si votre nom apparaît quelque part sur Internet;
- ◆ des informations touristiques sur l'endroit où vous prendrez vos prochaines vacances;
- ◆ l'adresse électronique d'un conseiller fédéral;
- ◆ etc.

Pour se tenir au courant

Se tenir au courant des développements récents en matière

de moteurs de recherche n'est pas très difficile. La presse spécialisée publie régulièrement des articles comparatifs sur le sujet et vous devriez trouver, presque chaque mois, dans l'une ou l'autre revue d'informatique, un article sur les moteurs. Qui plus est, il existe un site WWW (<http://search-enginewatch.com/>) qui propose de nombreuses informations sur les moteurs de recherche. On peut également obtenir, sur ce site, un bulletin mensuel consacré aux moteurs de

recherche; il suffit de donner son adresse électronique pour le recevoir.

Finalement, il est possible d'utiliser les moteurs de recherche eux-mêmes pour savoir si de nouvelles informations les concernant sont disponibles sur Internet...

DANS LES TOILES DU WEB UNE NOUVELLE SOURCE D'INFORMATION POUR LES ARCHIVISTES

par Didier Grange
Archiviste de la Ville de Genève

Les institutions archivistiques ayant créé un site WEB sont de plus en plus nombreuses. Si elles pouvaient encore être comptées sur les doigts de la main il y a deux ou trois ans, ce n'est plus le cas aujourd'hui. Le mouvement s'est prodigieusement accéléré. Il ne s'agit certainement pas d'un simple phénomène de mode, comme certains le prétendent, mais bien plus d'un virage fondamental dont nous n'avons pas encore mesuré les conséquences pour notre profession.

Quantité ne rime pas toujours avec qualité. De ce point de vue, de grands progrès ont également été accomplis. La maîtrise technique, la présentation graphique et le choix du contenu ont beaucoup évolué ces derniers mois. La conception des sites s'affine, l'information se diversifie et les perspectives de développement ne semblent avoir comme limites que celles posées par notre propre créativité.

LE WEB EST INCONTOURNABLE

Parmi les nombreuses questions liées au WEB, il en est deux qui reviennent souvent au détour des conversations: quelles informations les archivistes doivent-ils mettre à disposition et quelles informations peuvent-ils trouver sur le réseau?

Bien que la première interrogation mériterait une réflexion et un article en soi, c'est à la seconde que je tenterai de répondre ici brièvement. Car, même si nos ressources et nos compétences ne nous permettent pas toujours de mettre en place un site, rien ne nous empêche d'utiliser le WEB pour rechercher de l'information. L'archiviste ne peut plus se contenter aujourd'hui de lire des monographies, de dépouiller des périodiques spécialisés, d'assister à des conférences ou à des congrès et de recourir à l'aide des collègues avec lesquels il entretient des liens, il doit également rechercher de l'information à travers le réseau. En substance, le WEB constitue déjà une nouvelle source d'information. Notre profession n'échappe pas à ce constat.

UNE INFORMATION ABONDANTE ET EN CONSTANTE ÉVOLUTION

Il est bien difficile de rendre compte en quelques lignes de la richesse du WEB en matière d'archives et d'archivistique. L'information mise à disposition évolue rapidement. Les prestations offertes se diversifient. Les nouveaux partenaires se multiplient. La toile se tisse à un rythme soutenu.

On peut tenter toutefois d'esquisser une typologie reflétant l'état actuel de l'information disponible.

On y trouve:

- des rapports de commissions, de comités spécialisés et d'organismes internationaux (émanant, par exemple, du Conseil International des Archives, de l'European Commission on Preservation and Access et de la Society of American Archivists); des résultats de travaux de groupes de recherche (tels que ceux menés autour des archives électroniques par les Universités de British Columbia et de Pittsburgh);
- des normes et standards (la norme ISAD-G International Standard Archival Description-General, par exemple, est consultable sur un site créé par les Archives Nationales du Canada dans le but de la promouvoir);
- des fiches techniques et des recommandations (semblables à celles que la National Archives and Records Administration à Washington met à disposition et qui ont trait à différents sujets comme la préservation ou la sécurité);
- des périodiques électroniques (notamment celui réalisé dans le cadre de l'Ecole de Bibliothéconomie et des Sciences de l'Information de l'Université de Montréal, baptisé *Cursus*, qui présente des travaux de recherche effectués par les étudiants);
- des sites spécialisés qui se concentrent sur une thématique liée aux archives (tel que le site de l'Université de Stanford, répondant au nom de «CooL», relatif à