

Les portails de bibliothèque :

Nouvelles fonctionnalités, nouveaux défis

Clément Arsenault
Professeur adjoint, EBSI,
Université de Montréal
clement.arsenault@umontreal.ca

François-Xavier Paré
Bibliothécaire,
Direction des bibliothèques - Bureau
des systèmes, Université de Montréal
francois-xavier.pare@umontreal.ca

Les auteurs présentent un survol des logiciels de portail de bibliothèque. La notion de portail de bibliothèque est d'abord définie, puis sont présentées et expliquées les principales fonctionnalités de ce type de produit (métarecherche, personnalisation, authentification). On donne ensuite un aperçu du marché des logiciels de portail avant de soulever certaines questions liées à la fiabilité des résultats et à la formation des utilisateurs.

Introduction

Au cours des dernières années, des logiciels spécialisés appelés *portails de bibliothèque* ont fait leur apparition. Ces outils, déjà adoptés par plusieurs bibliothèques, proposent de nouvelles fonctionnalités visant à faciliter le processus de recherche d'information des usagers. Le présent article vise à clarifier le concept de portail de bibliothèque en décrivant les principales fonctionnalités de ce type de produit. Nous aborderons également certains aspects plus controversés des portails, car malgré le succès que connaissent ces logiciels, en particulier en milieu académique, plusieurs remettent en question certains des concepts fondamentaux qui sous-tendent ce type d'outil ou, à tout le moins, nous mettent en garde contre les possibles conséquences de leur utilisation. Nous tenterons donc de mettre en lumière les impacts potentiels de l'implantation d'une telle technologie dans les bibliothèques.

■ Qu'est-ce qu'un portail?

Le terme *portail* a été employé à toutes les sauces depuis quelques années. Zhou (2003) présente un historique détaillé des portails Web et de leur développement en bibliothèques. Il convient donc de retracer les divers usages de ce mot afin de bien comprendre ce qu'il signifie dans le contexte particulier des bibliothèques. On a d'abord utilisé le mot *portail* pour désigner un site Web regroupant dans une seule interface plusieurs ressources portant sur un même sujet ou

destinées à un public en particulier. À titre d'exemples, mentionnons :

- Portail régional de Montréal (<http://www.montreal.gouv.qc.ca/>)
- Portail du droit français (<http://droit.org/>)
- Afrik.com – Le portail de l'Afrique et du Maghreb (<http://www.afrik.com/>)

En fait, assez peu de choses distinguent les portails ci-dessus de n'importe quel autre site Web, si ce n'est leur aspect centralisateur. On y retrouve surtout des liens vers d'autres sites, si bien que ce type de portail se veut avant tout un point de départ pour la recherche sur un sujet donné. Certains préfèrent donc employer le terme *passerelle* (*gateway* en anglais) pour désigner ce type de site.

Par la suite sont apparus les premiers portails dits personnalisés, permettant à l'utilisateur de paramétrer l'aspect et le contenu du site afin de mieux répondre à ses besoins d'information. L'exemple le plus connu est sans doute Yahoo! avec sa fonction My Yahoo!. En y entrant, l'utilisateur tape un nom d'utilisateur et un mot de passe, ce qui permet au site de le reconnaître. L'interface du site est alors adaptée en conséquence : l'utilisateur y retrouvera la météo de sa région, les nouvelles locales et celles liées à ses intérêts, le cours de ses actions en bourse, etc. Bref, le site est modélisé en fonction de l'utilisateur. Pour ce faire, il faut bien sûr que l'utilisateur ait, a priori, indiqué ses préférences lors de son inscription. De nombreuses entreprises ont également décidé d'implanter de tels portails pour leurs employés afin de leur

Library portals: new functionalities, new challenges

The authors present a brief review of software for library portals. The concept of library portals is defined and their main functionalities subsequently presented and explained (meta search engines, personalization, authentication). A general survey of the market for portal software is given before addressing certain questions about the reliability of their results and user training.

donner accès à des informations personnalisées, à un historique de leur paie, etc. Ce type de portail, appelé *portail institutionnel*, est notamment utilisé dans de nombreuses institutions d'enseignement afin de fournir de l'information sur mesure aux différentes clientèles que représentent les étudiants, les professeurs et les employés¹.

■ Le portail de bibliothèque

Plus récemment est apparu un nouveau type de portail destiné aux bibliothèques. Ce dernier reprend certains aspects des portails présentés plus haut : il s'agit d'un site Web ayant une fonction centralisatrice et qui permet un certain niveau de personnalisation. Ce qui différencie ce portail de ses prédécesseurs, c'est qu'il met l'emphase sur une nouvelle fonctionnalité, la métarecherche, à laquelle s'ajoutent diverses fonctionnalités visant à faciliter la recherche de l'utilisateur. Plus précisément, les principales fonctionnalités qui caractérisent le portail de bibliothèque se situent à quatre niveaux :

- **Métarecherche.** Cette fonctionnalité permet de chercher simultanément dans plusieurs bases de données documentaires. La requête de l'utilisateur est envoyée à chacune des bases de données sélectionnées et les résultats provenant de toutes les bases sont présentés dans une seule et même interface, celle du portail. En plus des bases de données bibliographiques et plein texte, on peut envoyer des requêtes à des catalogues de bibliothèque, des dépôts institutionnels ou toute autre ressource similaire.
- **Personnalisation par la bibliothèque.** La bibliothèque peut habituellement contrôler l'apparence du portail (couleurs, logos, textes) ainsi que certains paramètres de recherche par défaut (par exemple le tri des résultats). Il est aussi possible de définir à l'avance des groupes de ressources destinées à être interrogées simultanément (par exemple les catégories « Ressources en médecine » ou « Données statistiques »).

- **Personnalisation par l'utilisateur.** Selon les produits, les portails de bibliothèque permettent aux usagers de conserver un historique de leurs recherches, de créer leurs propres listes de bases de données préférées ou encore de créer des alertes de type DSI (diffusion sélective de l'information²).

- **Authentification.** Les usagers doivent généralement s'identifier au moment d'entrer dans le portail en saisissant un nom d'utilisateur et un mot de passe. Ceci est nécessaire afin d'avoir accès aux fonctionnalités de personnalisation décrites au paragraphe précédent. Qui plus est, le portail permet la recherche dans des bases de données dont l'accès est souvent restreint par licence. La fonction d'authentification permet ainsi de restreindre l'accès aux bases de données à certains usagers ou groupes d'usagers.

Ces fonctionnalités seront examinées plus en détails dans les prochaines sections. Précisons d'emblée que les différents produits n'offrent pas tous les mêmes fonctionnalités. Nous tenterons donc simplement de donner un aperçu de ce qui est possible à l'heure actuelle.

■ Fonction : métarecherche

La métarecherche³ est un processus qui permet essentiellement d'interroger, de façon simultanée et par le biais d'une interface unique, une multitude de ressources hétérogènes telles que des catalogues, des bases de données, des dépôts institutionnels, des services agrégatifs de périodiques, des encyclopédies et d'autres sources d'information en format numérique. Bien qu'elle soit connue depuis longtemps et disponible depuis plusieurs années sur les gros systèmes commerciaux tels que Dialog, Lexis Nexis et Ovid, cette technique de recherche semble susciter, parallèlement à la mise en marché des produits commerciaux de portails de bibliothèque, un intérêt marqué.

En effet, la fonction fédératrice du portail permet de capter, par le biais d'une interface commune, la requête d'un utilisateur, puis de redistribuer celle-ci vers les serveurs des bases qui auront été préalablement sélectionnées, soit à la carte par l'utilisateur, soit par des regroupements préalablement établis par les bibliothécaires, soit par des regroupements automatiques liés au profil de l'utilisateur. Les réponses provenant de chaque serveur sont ensuite reçues et présentées à l'utilisateur dans une interface unifiée et un format passablement uniformisé. Cette forme de recherche s'apparente, d'une certaine façon, à celle effectuée par les métamoteurs de recherche tels que Metacrawler, Kartoo et Ixquick : plutôt que de se constituer un index local, ces outils redistribuent les requêtes vers un ensemble déterminé de moteurs de recherche. La métarecherche propre aux portails présente toutefois l'avantage de lancer des requêtes vers des bases de données propriétaires, dont l'accès est évidemment impossible aux grands métamoteurs publics. En ce sens, le portail devient une sorte de métamoteur du Web invisible (ou en tout cas d'une partie du Web invisible, celle à laquelle la bibliothèque a légitimement accès).

L'intérêt que présente pour l'utilisateur la mise en place d'un service de métarecherche est incontestable : il n'a plus à se familiariser avec les interfaces natives des différents systèmes qu'il veut consulter; il contrôle l'interrogation de ces systèmes par l'intermédiaire d'une seule interface qu'il peut apprivoiser à son rythme et paramétrer à son goût et selon ses besoins. Outre le fait d'offrir un point d'accès unique vers un ensemble hétéroclite de ressources documentaires, la fonction fédératrice du portail permet d'accroître la visibilité des bases et des ressources disponibles dans une institution donnée. C'est, par conséquent, un excellent outil pour mettre en valeur les ressources disponibles et pour en faire la promotion. Le regroupement des ressources sous diverses rubriques, par sujet, par type de contenu ou de documents, permet souvent à l'utilisateur d'interroger des bases dont il ignore l'existence et qui peuvent quelquefois

s'avérer très riches. Ainsi, le service de métarecherche offert par l'intermédiaire du portail permet non seulement la découverte d'informations, mais également la découverte de bases de données et de ressources documentaires (Sadeh, 2004, p. 105).

La métarecherche permet, dans une certaine mesure, de contourner le problème du choix des ressources à interroger puisque celles-ci sont présélectionnées par le bibliothécaire professionnel responsable des collections. Fait à noter, l'étape de la sélection des sources est reconnue, depuis fort longtemps, comme une étape cruciale du processus de recherche documentaire (Harter, 1986, p. 125). Pourtant, des études récentes démontrent que les utilisateurs ont bien du mal à choisir la source appropriée, ne serait-ce que pour retrouver un type de document donné. Ainsi, selon l'étude de Mittermeyer et Quirion (2003, p. 49) portant sur les connaissances documentaires des étudiants entrant au 1^{er} cycle, seulement 28,5% des répondants ont su identifier le type de source adéquat pour retracer un article de périodique. Selon Tallent (2004), la ségrégation des types de documents en bases distinctes constituerait l'un des problèmes les plus frustrants pour les usagers qui ne comprennent généralement pas pourquoi il est impossible de retrouver des références d'articles de périodiques dans le catalogue de la bibliothèque :

Students have little patience for (or knowledge of) library distinctions regarding databases and e-journal collections. [...] For years, reference librarians have been telling students that journal articles are not included in the OPAC and for years students have been confused as to why; it never made sense to them.

(Tallent, 2004, p. 71)

L'expérience tend à démontrer que dans leurs recherches documentaires, les utilisateurs font généralement preuve d'un haut degré de pragmatisme et tendent à réutiliser les sources qu'ils connaissent et qui ont bien fonctionné dans le passé. Selon une récente enquête menée auprès

des utilisateurs des Boston College Libraries (Tallent, 2004, p. 72), ils sont également peu enclins à explorer des ressources qui n'offrent pas un rendement certain. Selon la même étude, il est démontré que les fonctionnalités de type *recherche rapide* qu'offre le service de métarecherche sont très populaires et appréciées des utilisateurs (Tallent, 2004, p. 74).

Cette fonction de recherche rapide, qui consiste surtout à présenter directement à l'utilisateur une sélection des « meilleures » bases, vise à simuler le principe de fonctionnement des gros moteurs de recherche que l'on retrouve sur Internet, dont le désormais incontournable Google. Il ne fait pas de doute que la popularité accrue de ces outils réside principalement dans leur simplicité d'utilisation. Les études de comportement informationnel tendent d'ailleurs à démontrer que les utilisateurs apprécient avant toute chose la simplicité d'utilisation de ces outils et font un très faible usage des fonctions de recherche avancée disponibles (Spink et al., 2001). Avec ce type d'outils, l'utilisateur n'a pas à se préoccuper du choix des sources, du type d'information ou du format des documents qu'il recherche; une simple boîte de saisie disponible dès la page d'accueil lui permet de procéder à la recherche à l'aide de quelques mots-clés. L'option de recherche rapide, paramétrable sur la page d'accueil du portail, se veut en quelque sorte une réponse au problème d'érosion de l'utilisation des services d'information traditionnels dans les bibliothèques. Ce phénomène, observé depuis quelques années, indique que de plus en plus d'usagers, en particulier les étudiants, délaissent les services offerts dans les bibliothèques au profit des moteurs de recherche Web (Griffith et Brophy, 2002; Detlor et al., 2003, p. 41). Il semble, comme le rapporte Luther (2003, p. 1), que « pour plusieurs utilisateurs, la qualité des résultats a moins d'importance que le processus [de recherche] – ils s'attendent à ce que le processus soit rapide et facile » (traduction des auteurs).

L'inclusion de fonctions de recherche fédératrice au sein des portails de bibliothèque saura peut-être séduire la « génération Google » et reconquérir ces usagers. Il est toutefois à souhaiter que la mise en place de fonctions de métarecherche permette aux bibliothèques de se repositionner afin de fournir un service compétitif et efficace en procurant davantage d'autonomie à l'usager (Maisonnette, 2003, p. 214), sans compromettre leur notoriété et leur avantage concurrentiel en tant que fournisseur d'information de haute qualité (Jackson, 2002, p. 36).

Ces dernières années, la question de l'interopérabilité des systèmes a été à l'avant plan dans le milieu de la recherche d'information. Il s'agit bien sûr d'une condition essentielle à la mise en place d'un service de métarecherche. Étant donné la diversité et l'hétérogénéité des sources d'information interrogées (tant au niveau de leur structure, du type d'information qu'elles contiennent, du mode d'encodage de cette information que de la configuration des serveurs sur lesquels elles résident), le paramétrage des protocoles de communication permettant le repérage et l'extraction d'informations peut être relativement complexe. Cet aspect semble d'ailleurs se complexifier de plus en plus du fait de la multiplication des formats propriétaires (Dorner et Curtis, 2004, p. 184; Pace, 2004).

Dans leur rapport comparatif des produits commerciaux de portail de bibliothèque, Dorner et Curtis rapportent pas moins de 12 protocoles ou normes différents (les plus populaires étant Z39.50, MARC, XML et SQL), sans compter les formats propriétaires des fournisseurs (Dorner et Curtis, 2004, p. 193). C'est là un argument de taille en faveur de l'achat d'un produit commercial puisque la majorité de ces produits contiennent déjà, en bonne partie, les métadonnées descriptives et techniques nécessaires pour accéder aux principales ressources (nous en reparlerons plus loin). Le paramétrage local s'en trouve grandement facilité, ce qui permet de mettre en place le service plus rapidement.

Outre les problèmes liés à l'interrogation simultanée des sources, la phase d'extraction pose également des défis particuliers. En général, la réponse des divers systèmes s'effectue en deux temps : en premier lieu, seul le nombre (parfois approximatif) de résultats obtenus dans chaque source est présenté à l'utilisateur; ensuite, de façon automatique ou à la demande de l'utilisateur, les résultats sont affichés. L'unification et la présentation des résultats dans l'interface peuvent s'effectuer grâce à certaines métadonnées paramétrables dans le module de présentation. Il est à noter que si l'élimination des doublons s'effectue généralement grâce à l'utilisation d'identificateurs uniques (par exemple l'ISBN et le DOI), on peut également la paramétrer localement selon des critères de comparaison plus ou moins stricts.

Fait intéressant : dans l'interface du portail, chaque résultat peut être accompagné d'un bouton permettant d'accéder au serveur OpenURL⁴ de l'institution, donc à des services supplémentaires tels que l'accès au texte intégral, la recherche dans le catalogue local ou une demande de prêt entre bibliothèques (en utilisant un logiciel comme SFX ou Link FinderPlus). Normalement, ces boutons apparaissent dans les diverses bases de données auxquelles l'institution est abonnée, à la demande de celle-ci. Mais certains fournisseurs ne permettent pas d'ajouter un tel bouton aux notices dans leur interface. Le portail permet donc d'afficher un bouton OpenURL à côté de tous les résultats, même ceux provenant de bases de données qui ne sont pas « compatibles OpenURL ».

Compte tenu de la réduction de la vitesse de transfert due à l'engorgement des réseaux, il est d'usage que, dans un premier temps du moins, seules les 10 ou 20 premières notices de chaque source soient extraites (ce nombre est généralement paramétrable). Cette limitation se fonde sur la prémisse selon laquelle l'utilisateur ne saurait que faire de centaines de références (Sadeh, 2004, p. 109). Une telle pratique peut toutefois entraîner des conséquences négatives, par exemple lors de l'application d'un tri par perti-

nence à l'ensemble unifié des notices retournées. En effet, chaque ressource interrogée aura repêché un nombre différent de notices et n'aura retourné que les 10 ou 20 premières, mais « premières » selon des critères locaux propres à chaque base. Est-il alors raisonnable d'appliquer un classement par pertinence sur des notices préalablement triées par titre sur une ressource, par date de publication sur une autre ressource ou par fréquence d'occurrence des mots-clés sur une troisième ressource? Le problème vient en majeure partie du fait que « l'algorithme de tri par pertinence employé par un système de métarecherche doit être "objectif"; il ne peut se baser que sur les documents extraits, sans pouvoir connaître les caractéristiques des autres documents contenus dans les ressources desquelles ils proviennent » (Sadeh, 2004, p. 110, traduction des auteurs). Par contre, l'environnement intégré qu'offre le portail permet à l'algorithme de classement d'utiliser des paramètres relatifs au profil de l'utilisateur, ce qui peut représenter un certain avantage par rapport à un environnement traditionnel, où tous les usagers sont considérés de la même façon.

La possibilité qu'offre la métarecherche d'interroger une multitude de sources à la fois est sans contredit très séduisante. S'il n'est plus problématique de combiner les sources, pourquoi se limiter à quelques sources alors que l'on peut interroger constamment toutes les sources? En dépit du risque de bruit contextuel qu'entraîne ce genre de comportement informationnel, il semble que ce phénomène ait déjà été constaté depuis un an ou deux suite à la mise en place successive de services de métarecherche dans les grandes bibliothèques, particulièrement aux États-Unis. Cette situation n'est certes pas réjouissante pour les fournisseurs de contenu, qui voient leurs serveurs assaillis par une montée exponentielle de requêtes d'activités de recherche. Afin de pallier ce problème et pour tenter de régulariser les procédures et les protocoles entourant la métarecherche, la National Information Standards Organization (NISO) a mis sur pied un projet réunissant des experts de plusieurs

milieux. En résumé, ce projet se donne comme objectifs de permettre, dans le contexte d'un environnement de métarecherche :

- aux fournisseurs de services de métarecherche d'offrir des services plus efficaces
- aux fournisseurs de contenu d'offrir un contenu rehaussé tout en protégeant le contenu intellectuel
- aux bibliothèques de fournir des services qui les distinguent de Google et des autres services Web gratuits (NISO, traduction des auteurs)

Une telle initiative vise donc à promouvoir l'établissement de protocoles communs pour les éléments de communication, d'authentification, de balisage et de description de l'information nécessaires au bon fonctionnement de la métarecherche, dans le but de créer un terrain d'entente acceptable pour ces trois principaux intéressés que sont les fournisseurs de services de métarecherche, les éditeurs (fournisseurs de contenu) et les bibliothèques (Luther, 2003; Pace, 2004). Souhaitons que cette initiative puisse à tout le moins fournir un environnement plus robuste et plus efficace à ce phénomène grandissant qu'est la métarecherche.

L'établissement de protocoles communs pour le balisage et la description de l'information intéresse grandement les bibliothèques. Ces dernières années, bon nombre de projets de recherche sur les normes de description et d'analyse se sont intéressés au problème de l'interopérabilité sémantique entre les différents systèmes documentaires. Plus particulièrement, la recherche sur le mapping terminologique⁵ a donné naissance à des projets présentant un intérêt certain dans le domaine de la métarecherche. Car le gros défi est là : comment assurer une recherche précise alors que la majorité des sources d'information utilisent des normes de description, des terminologies et des systèmes de classification différents?

En effet, la combinaison des sources, rendue possible par les modules de

métarecherche, limite grandement l'utilisateur dans l'élaboration de ses stratégies, si bien que très souvent, seule l'interrogation non ciblée par mot-clé (dans tous les champs) reste possible. Imaginons un utilisateur cherchant des ouvrages sur la douleur : il lance une requête simultanée dans un catalogue utilisant les vedettes-matière de la Library of Congress (LCSH), un catalogue utilisant les vedettes-matière de Laval (RVM) et la base de données Dissertation Abstracts. S'il choisit de faire une recherche dans le champ *sujet* avec le terme anglais *pain*, il obtiendra des résultats pertinents de la première source sélectionnée, des résultats non pertinents de la deuxième source – puisque les ouvrages repêchés traiteront de boulangerie – et aucune référence de la troisième source – puisque le terme *pain* ne fait pas partie des descripteurs de cette base. L'utilisateur pourrait s'en remettre au plus petit dénominateur commun, c'est-à-dire la recherche non ciblée par mot-clé, qui lui donnerait sans doute plus de résultats, mais aussi beaucoup plus de bruit.

L'utilisateur est-il bien servi par ce type de service? À vouloir rechercher dans tout, l'utilisateur, qui ne peut raffiner ses requêtes adéquatement, finit peut-être, paradoxalement, par trouver moins. D'où la très grande importance de permettre à l'utilisateur de migrer au besoin de l'interface de métarecherche du portail vers les interfaces natives des sources en consultation, interfaces qui lui fourniront des options plus élaborées puisqu'elles sont paramétrées spécifiquement en fonction du type d'information contenu dans leurs sources. Cette forme de « nivellement par le bas », où seule la recherche non ciblée par mot-clé est possible (malgré tous les problèmes qu'elle comporte), est un des principaux reproches que les bibliothécaires et les autres spécialistes de l'information font à la métarecherche.

Afin de contourner ce problème, plusieurs projets d'interopérabilité sémantique ont vu le jour ces dernières années. Ils proposent des méthodes de mapping terminologique entre les langages documentaires employés dans

divers systèmes d'information, notamment dans le domaine de l'accès multilingue (McCulloch, 2004). Le chemin vers l'interopérabilité sémantique des systèmes d'information est semé d'embûches et le travail à accomplir est colossal. En attendant l'implantation de telles fonctionnalités dans les systèmes de métarecherche, il faudrait peut-être s'assurer de proposer aux usagers des regroupements de sources présentant le plus grand nombre possible de caractéristiques communes ou, du moins, de proposer des sources plus compatibles (par exemple en évitant de combiner des sources de langue différente). Il s'agit d'exercer un bon jugement dans le choix des sources, sans délaissier la mission de formation à la recherche documentaire, afin de permettre aux usagers de tirer tous les bénéfices possibles de la métarecherche tout en évitant de lesurrer en leur laissant croire que tout cela est « magique ».

Fonction : paramétrage par la bibliothèque

Lors de l'implantation d'un portail, la bibliothèque doit prendre certaines décisions quant à l'aspect qu'elle veut lui donner. Le paramétrage subséquent comporte ainsi plusieurs aspects. Il est habituellement possible, d'abord, de contrôler certains éléments de l'apparence visuelle du portail (couleurs, logos, polices des textes, etc.). On peut notamment faire apparaître le logo de la bibliothèque bien en vue sur toutes les pages et choisir un jeu de couleurs et des polices d'écriture appropriés. Plusieurs institutions souhaitent harmoniser l'apparence du portail avec le reste de leur site Web afin d'obtenir une meilleure intégration de ce nouvel outil aux ressources existantes. Strauss (2002) remet même en question la pertinence de développer des portails spécialisés pour les bibliothèques et préconise plutôt le « déploiement de pages, canaux et "lucarnes" » au sein du portail institutionnel (traduction des auteurs). Les modifications esthétiques sont souvent réalisées à l'aide de feuilles de style (CSS), ce qui offre une marge de

manœuvre intéressante tout en facilitant les changements. Il est à noter que tous les produits portails n'offrent pas le même niveau de flexibilité quant à la modification de l'interface.

Les différents éléments textuels de l'interface peuvent également être modifiés, que ce soit le nom des champs, les messages d'erreur ou le contenu de l'aide en ligne. Quelques produits commerciaux offrent une version française. Il demeure néanmoins nécessaire de modifier certains des textes pour refléter les particularités locales.

La bibliothèque peut aussi déterminer les différents regroupements de bases de données présentés aux usagers. En effet, bien qu'un usager puisse théoriquement effectuer sa recherche dans l'ensemble des bases de données de la bibliothèque, il est préférable, dans la plupart des cas, de se limiter à un sous-ensemble de ces ressources. Il revient donc à la bibliothèque de bâtir ces ensembles de ressources en déterminant les catégories offertes aux usagers et, pour chacune de ces catégories, quelles bases de données en feront partie. Cette tâche n'est pas aussi complexe qu'il n'y paraît, ne serait-ce que parce que la plupart des institutions ont déjà des listes de leurs bases de données classées par discipline, classement qui peut être reproduit au sein du portail. À titre d'exemple, les catégories peuvent présenter une discipline particulière (Biologie, Musique, Sciences et technologies), un type de document (Thèses) ou encore un type de ressource (Catalogues d'autres bibliothèques). *Qui plus est, le portail permet de rendre des catégories visibles à certains groupes d'utilisateurs seulement (nous y reviendrons plus en détails dans la section sur l'authentification).*

Un autre type de paramétrage concerne certaines options placées par défaut et liées à la métarecherche. On peut par exemple déterminer quel sera le nombre de résultats par page, si les notices s'afficheront en format long ou abrégé, ou en fonction de quels critères les résultats seront triés au moment de la recherche. Certains produits permettent à l'utilisateur

de modifier ces options et d'enregistrer ses préférences dans son profil. Néanmoins, la bibliothèque se doit de déterminer les paramètres de départ. Il convient ici de souligner une distinction importante entre les différents produits offerts en ce qui concerne la capacité de la bibliothèque à effectuer des changements. Certains fournisseurs gardent un contrôle presque complet sur la configuration du portail. Ainsi, lorsque des modifications s'imposent, la bibliothèque doit contacter le fournisseur pour qu'il les effectue. Ceci peut s'avérer intéressant pour des institutions de plus petite taille, ou dont les ressources sont limitées, mais d'autres bibliothèques préféreront avoir la liberté de configurer le portail elles-mêmes. Dans ce cas, un module d'administration est mis à la disposition du client.

Le fonctionnement des portails repose également sur ce qu'on appelle une base de connaissances. Il s'agit en fait d'une liste des ressources dans lesquelles on peut effectuer des recherches. Pour chaque ressource, la base de connaissances contient toutes les informations nécessaires à la connexion : l'adresse du serveur, les champs interrogeables, le mot de passe utilisé pour se connecter, etc. Cette base de connaissances est normalement bâtie par le fournisseur et contient de l'information sur la plupart des ressources les plus répandues. Encore une fois, la bibliothèque peut avoir accès à cette composante via le module d'administration.

Les modifications à la base de connaissances peuvent être de deux ordres. Premièrement, bien que cette dernière contienne déjà plusieurs ressources, la bibliothèque pourrait vouloir ajouter des ressources locales. Deuxièmement, il est possible d'influencer la mécanique de la métarecherche elle-même. Certains produits permettent à la bibliothèque de modifier les profils Z39.50 ou les connecteurs qu'elle veut utiliser pour rechercher tel ou tel champ d'une base de données. Il va sans dire que ce niveau de paramétrage est bien plus complexe que ceux mentionnés plus haut, si bien qu'il requiert des connaissances plus

poussées. Une autre façon d'influencer la métarecherche est d'accorder un poids plus ou moins grand à certaines ressources afin d'influencer le classement des résultats provenant de cette ressource lors du tri par pertinence.

Outre la question des délais, une autre raison susceptible d'encourager une bibliothèque à effectuer elle-même ces changements est que dans certains cas, un coût est associé à chaque modification demandée par la bibliothèque. Si l'on prévoit avoir beaucoup de modifications à faire (par exemple l'ajout de nombreuses ressources locales), il peut donc s'avérer avantageux d'opter pour un module d'administration local. En définitive, l'effort requis variera considérablement selon le degré de personnalisation désiré par la bibliothèque.

Mentionnons finalement que le paramétrage est souvent un processus itératif pour les bibliothèques, notamment en ce qui concerne l'apparence et les catégories de ressources. Les commentaires des usagers et des bibliothécaires contribueront à rectifier le tir au besoin. Il est également utile d'organiser des tests d'utilisation (Allen, 2002) ou des groupes de discussion avec les usagers (Crowley et al., 2002) pour en savoir plus. Par ailleurs, les portails sont généralement en mesure de produire des statistiques d'utilisation qui peuvent donner une idée des recherches effectuées par les usagers, des ressources et des catégories les plus utilisées et, ainsi, de fournir des indices sur le bien-fondé de certaines décisions prises lors du paramétrage.

Fonction : personnalisation

L'un des aspects les plus intéressants des portails est la possibilité pour l'utilisateur de personnaliser son environnement de recherche. D'abord, comme on l'a mentionné plus haut, l'utilisateur est en mesure d'établir des préférences quant à la recherche, surtout en ce qui a trait à l'affichage des résultats. Le portail est d'ailleurs en mesure d'emmagasiner ce type de préférence et de s'en « rappeler »

d'une session à l'autre. L'utilisateur peut par exemple créer son propre regroupement de bases de données préférées et effectuer ensuite des métarecherches dans cet ensemble de ressources. L'utilisateur n'a donc pas à se limiter aux regroupements disciplinaires prévus par la bibliothèque. La personnalisation amène également d'autres possibilités intéressantes. Le portail permet aux usagers de sauvegarder une requête, par exemple dans le but de l'exécuter à nouveau plus tard. Cette requête peut même faire l'objet d'une alerte (DSI). Dans ce cas, la requête est exécutée périodiquement et les résultats envoyés par courriel. Par ailleurs, il est possible pour l'utilisateur de sauvegarder des notices individuelles afin de se créer une banque de documents pertinents.

On le voit, la personnalisation présente des avantages certains pour l'utilisateur. Le fait de pouvoir emmagasiner dans un même endroit des résultats provenant de plusieurs bases de données s'avère fort utile et simplifie considérablement la tâche du chercheur. C'est d'ailleurs là une des forces des portails : ils offrent des services qui existent dans certaines bases de données (recherche, sauvegarde, alertes, etc.), mais d'une façon intégrée, centralisée. En définitive, le portail est en mesure de mieux soutenir le processus de recherche de l'utilisateur que les diverses bases de données prises individuellement. Il en résulte également des bénéfices indirects pour les bibliothèques : à une époque où ces dernières se plaignent de perdre du terrain face au reste du Web, de telles fonctionnalités ne peuvent qu'aider à attirer et, à l'instar des portails commerciaux, à fidéliser la clientèle. La bibliothèque y gagne également en ce qui a trait à l'image qu'elle projette.

Fonction : authentification

Le Grand dictionnaire terminologique définit l'authentification comme la « procédure consistant à vérifier ou à valider l'identité d'une personne [...] lors d'un échange électronique, pour

contrôler l'accès [...] à un système informatique ou à un logiciel ». L'utilisateur doit donc entrer un nom d'utilisateur et un mot de passe pour pouvoir accéder au portail. Ces informations sont ensuite contre-vérifiées auprès d'une base de données appelée *annuaire*, base qui contient des informations sur les usagers « légitimes ». Par exemple, dans le cas d'une université, l'annuaire en question peut être la base de données du registraire, qui contient de l'information sur tous les étudiants, professeurs et employés. Lors du processus d'authentification, le portail va donc consulter les informations du registraire afin d'identifier l'utilisateur. Il existe différents moyens de communication entre le portail et l'annuaire, mais le plus en vue à l'heure actuelle est une technologie appelée Lightweight Directory Access Protocol (LDAP)⁵.

Mais pourquoi veut-on identifier ainsi les utilisateurs du portail? Il existe plusieurs raisons. D'abord, si l'on désire offrir des fonctions de personnalisation aux usagers, il faut pouvoir déterminer à qui l'on a affaire. Ce n'est donc qu'après s'être identifié que l'utilisateur peut accéder à ses requêtes sauvegardées, à ses préférences, etc. L'utilisation d'un mot de passe nous assure aussi que seule la personne concernée aura accès à ces informations sauvegardées, aspect non négligeable lié à la confidentialité. L'annuaire contient souvent des informations sur l'appartenance de l'utilisateur à un ou à plusieurs groupes personnes. Un utilisateur pourrait donc se trouver dans la catégorie « employé », ou encore dans une catégorie plus restreinte comme « étudiant de première année au Département de chimie ». Le portail de bibliothèque est en mesure d'exploiter ces informations pour offrir des interfaces de départ différentes, adaptées à chaque groupe, ou encore pour permettre à des groupes d'accéder à des ressources qui sont inaccessibles à d'autres. C'est donc là un autre niveau de personnalisation, cette fois sur une base collective plutôt qu'individuelle.

La restriction des accès constitue une autre fonction importante de l'authen-

tification. Une partie considérable des bases de données présentes dans un portail de bibliothèque est protégée par licence. En forçant les gens à s'identifier, on s'assure que seuls les usagers légitimes peuvent accéder au portail. Une autre possibilité, s'il l'on ne voulait pas fonctionner avec des mots de passe, serait d'autoriser l'accès au portail en fonction de l'adresse IP de l'utilisateur. De cette façon, les usagers pourraient accéder au portail sans s'identifier, pourvu qu'ils se trouvent dans la bibliothèque (ou sur le campus). Dans ce cas, on ne pourrait toutefois offrir de fonctions de personnalisation puisqu'on ne saurait pas exactement à qui on a affaire.

Un dernier avantage – un peu plus indirect – de l'authentification : améliorer la qualité des statistiques d'utilisation que la bibliothèque peut obtenir. À la base, un portail n'est pas si différent d'une base de données commerciale. On peut donc s'attendre à pouvoir conserver des statistiques sur le nombre d'accès au portail, le nombre d'accès à chaque ressource accessible via le portail, etc. Mais comme l'authentification permet d'accéder à des informations sur l'appartenance à des groupes, il n'est pas exclu que l'on puisse également obtenir des statistiques agrégées pour chacun de ces groupes. Il est ensuite plus facile de cibler ses efforts de marketing auprès de certaines clientèles. Il faut noter cependant que les statistiques, pour des raisons de confidentialité, ne sont pas compilées de manière individuelle.

Le marché des portails

Si les portails connaissent une certaine popularité depuis deux ou trois ans, c'est qu'ils répondent à de réels besoins. Puisque la demande est là, de nombreux fournisseurs se sont précipités pour prendre leur part du gâteau au cours des dernières années. Ainsi, à l'heure actuelle, la plupart des grands fournisseurs de systèmes intégrés de bibliothèque offrent également un produit de type portail ou une quelconque solution de métarecherche. Parmi les gros joueurs, mentionnons IsacSoft,

Endeavor, ExLibris, Fretwell-Downing, Innovative, SIRSI et WebFeat.

Bien qu'il existe souvent des différences notables en ce qui concerne les fonctionnalités, ces fournisseurs offrent des produits somme toute assez semblables dans leur philosophie. Le prix des produits de type portail est tributaire de divers facteurs, ce qui en explique la grande variation. Parmi les critères pouvant être utilisés pour déterminer le prix, on retrouve parfois le nombre d'utilisateurs que dessert la bibliothèque et le nombre de bases de données auxquelles elle a accès (et qui doivent être configurées pour la métarecherche). Il faut savoir qu'outre le prix de départ pour l'achat du logiciel et les coûts annuels de maintenance, certains fournisseurs exigent un certain montant pour ajouter une nouvelle ressource au portail. En bout de ligne, le coût peut donc varier de quelques milliers de dollars pour une petite bibliothèque publique à plusieurs dizaines de milliers de dollars pour une grande bibliothèque universitaire.

Étant donné le prix considérable de ces produits, on peut se demander s'il ne vaudrait pas mieux envisager un développement maison. Malheureusement, la chose est beaucoup plus facile à dire qu'à faire : la nécessité de développer une base de connaissances constitue notamment un obstacle majeur puisque cela nécessiterait non seulement des centaines d'heures de configuration, mais également des relations privilégiées avec tous les fournisseurs de bases de données. En effet, on voudra que ces derniers nous communiquent tous les changements majeurs apportés à leurs produits de façon à ce que ces changements soient reflétés le plus tôt possible dans notre base de connaissances. En bref, il s'agit d'une entreprise considérable que peu seraient en mesure de mener à bien.

C'est sans doute pourquoi jusqu'ici, les bibliothèques ont généralement privilégié l'achat d'un produit commercial. Parmi les très rares exceptions, notons la California Digital Library (CDL) qui a développé, à partir de 1999, son propre logiciel de métarecherche sous l'appellation

SearchLight (<http://searchlight.cdlib.org/cgi-bin/searchlight>). Aucune personnalisation par l'utilisateur n'y est toutefois possible. La CDL annonçait d'ailleurs récemment qu'elle avait fait l'acquisition du logiciel MetaLib de la firme Ex Libris. Une autre initiative de méta-recherche similaire a été développée au Los Alamos National Laboratory (LANL) sous le nom de FlashPoint. En juillet 2004, FlashPoint permettait la métarecherche dans 13 bases de données différentes. Il ne faudrait pas oublier de mentionner également le logiciel MyLibrary (<http://dewey.library.nd.edu/mylibrary/>), un portail de bibliothèque personnalisable lancé en 1998 à la North Carolina State University, et qui a été adopté par la suite dans plusieurs institutions dans le monde. MyLibrary ne comporte toutefois pas de fonctionnalités de métarecherche, si bien qu'on ne peut pas véritablement le considérer comme un portail à part entière ou, à tout le moins, pas selon les critères que nous avons établis dans cet article. Il s'agit tout de même d'un des premiers exemples de personnalisation Web dans le milieu des bibliothèques, ce qui justifie son inclusion ici.

Conclusion

Ces dernières années, le développement des portails d'entreprise a connu un essor fulgurant. Il est maintenant reconnu que la mise en place de tels systèmes permet indéniablement d'accroître la visibilité et la productivité en stimulant les échanges de connaissances et la collaboration au sein de l'entreprise (Detlor, 2004). Parallèlement à ce phénomène, des portails spécialisés pour les bibliothèques sont apparus, surtout dans les bibliothèques académiques et d'autres grandes institutions. Ceux-ci ont d'abord pris la forme de simples pages Web développées localement, par exemple la page d'accueil de la bibliothèque sur laquelle sont présentées les ressources documentaires accessibles, regroupées et organisées sous diverses rubriques. Plus récemment sont apparus des produits commerciaux permettant de mieux intégrer

et d'exploiter les diverses fonctionnalités qu'offre le portail.

Tel que nous l'avons vu dans cet article, les avantages qu'offre généralement le portail sont multiples : visibilité accrue des ressources de la bibliothèque, paramétrage et personnalisation des interfaces, gestion des accès, élargissement et simplification des fonctions de recherche. L'achat d'un produit commercial offre l'avantage de mettre en place les principales fonctionnalités du portail de façon relativement simple et rapide. Évidemment, cela implique une certaine dépendance vis-à-vis du fournisseur, sans compter les coûts élevés d'achat et des contrats de service de ces produits. L'analyse des besoins par rapport aux divers produits disponibles doit être soigneusement préparée car ceux-ci offrent des services et des fonctionnalités diverses et à des coûts variables.

Bien que les avantages qu'offre le portail de bibliothèque soient indéniables, il ne faut pas voir l'acquisition d'un produit commercial comme une panacée aux problèmes récurrents de visibilité, de gestion et, en particulier, de l'utilisation des ressources numériques auxquels les bibliothèques sont régulièrement confrontées. La mise en place et la promotion de programmes de formation des usagers seront toujours essentielles pour assurer une utilisation plus efficace et adéquate des ressources documentaires en bibliothèque.

Plusieurs s'inquiéteront du phénomène de « Googlisation » et de nivellement par le bas de la recherche d'information qui semble se répandre pernicieusement. Avec la mise en place des interfaces de métarecherche des portails, la recherche par simples mots-clés dans une boîte de requête deviendra de plus en plus la norme. S'il est souhaitable que la recherche d'information se simplifie, il est de notre devoir de s'assurer que tous les usagers y trouvent leur compte : ceux qui veulent une réponse rapide à un besoin simple, mais aussi ceux qui ont des besoins plus poussés. Le développement de requêtes complexes n'est bien souvent possible qu'à

partir des interfaces natives des systèmes et il est donc bien important d'établir des liens entre l'interface du portail et les diverses interfaces des systèmes reliés. Il est essentiel de continuer à former les usagers à la recherche documentaire et à fournir des services de référence de qualité qui visent un équilibre acceptable entre l'autonomie des usagers et la qualité des services que nous leur offrons.

Notes

1. Un exemple parmi tant d'autres : le portail institutionnel de la California Polytechnic State University (<http://portaldemo.calpoly.edu/tour/>). Notons par ailleurs que Lefèvre (2001) offre une typologie des portails plus détaillée.
2. DSI : un service qui avise automatiquement l'utilisateur de l'arrivée de nouveaux documents susceptibles de l'intéresser (généralement par courriel). Il est habituellement nécessaire de définir un ou plusieurs profils d'intérêts pour indiquer le genre de documents que nous aimerions recevoir. Une autre possibilité consiste à établir une DSI à partir d'une requête de recherche : celle-ci sera réexécutée périodiquement et tous les nouveaux résultats seront expédiés à l'utilisateur.
3. Le terme *métarecherche* est équivalent aux expressions *recherche fédérée*, *recherche diffusée*, *recherche parallèle*, *recherche intégrée* et *recherche multi-bases* ou *multi-cibles*. En anglais, on retrouvera notamment les expressions *metasearch*, *federated search* et *broad-cast search*.
4. Pour en savoir plus sur OpenURL, on pourra consulter Vézina (2003).
5. L'établissement d'équivalences ou de quasi-équivalences entre les termes et les concepts de deux ou plusieurs terminologies distinctes.
6. Pour une évaluation relativement exhaustive, consulter le rapport de Dorner et Curtis (2003).

Bibliographie

- Allen, Maryellen. 2002. « A Case Study of the Usability Testing of the University of South Florida's Virtual Library Interface Design », *Online Information Review*, vol. 26, no 1, pp. 40–53.
- Bonnett, Monica. 2001. « Personalization of Web Services: Opportunities and Challenges ». *Ariadne*, no 28. [http://www.ariadne.ac.uk/issue28/personalization/intro.html]
- Crowley, Gwyneth et al. 2002. « User Perceptions of the Library's Web Pages: A Focus Group Study at Texas A&M University ». *The Journal of Academic Librarianship*, vol. 28, no 4, (juil.-août), pp. 205–10.
- Detlor, Brian. 2004. *Towards Knowledge Portals: From Human Issues to Intelligent Agents*. Dordrecht : Kluwer Academic Publishers, 250 p.
- Detlor, Brian, Umar Ruhi, Chris Pollard, Dave Hanna, Mihail Cocosila, Wuping Zheng, Eric Fu, Tao Jiang et Demos Syros. 2003. *Fostering Robust Library Portals: An Assessment of the McMaster University Library Gateway*. Hamilton : Michael G. DeGroot School of Business, McMaster University, 47 p. (MeRC Working Paper ; 4) [http://www.business.mcmaster.ca/msis/profs/detlorb/Working_Paper_4.pdf]
- Dorner, Daniel G. et Anne Marie Curtis. 2003. *A Comparative Review of Common User Interface Software Products for Libraries*. Wellington : Victoria University of Wellington, p. iv-63. [http://www.natlib.govt.nz/files/CUI_Report_Final.pdf]
- Dorner, Daniel G. et Anne Marie Curtis. 2004. « A Comparative Review of Common User Interface Products », *Library Hi Tech*, vol. 22, no 2, pp. 182–97. [http://dx.doi.org/10.1108/07378830410543502]
- Griffith, Jill R. et Peter Brophy. 2002. « Student Searching Behaviour in the JISC Information Environment », *Ariadne*, no 33. [http://www.ariadne.ac.uk/issue33/edner/intro.html]
- Harter, Stephen P. 1986. *Online Information Retrieval: Concepts, Principles, and Techniques*. San Diego : Academic Press, 259 p.
- Jackson, Mary E. 2002. « The Advent of Portals », *Library Journal*, no 15 (15 sept. 2002), pp. 36–39. [http://www.libraryjournal.com/index.asp?layout=articlePrint&articleID=CA242296]
- Lasica, J. D. 2001. « The Promise of the Daily Me: An In-Depth Look at the Different Flavors of Personalization ». [http://www.llrx.com/features/promise.htm]
- Lefèvre, Philippe. 2003. « Les portails d'accès à l'information ». *Documentaliste - Sciences de l'information*, vol. 38, nos 3/4 (sept.), pp. 188–96.
- Luther, Judy. 2003. « Trumping Google?: Metasearching's Promise », *Library Journal*, no 16 (1^{er} oct. 2003), pp. 36–39. [http://www.libraryjournal.com/index.asp?layout=articlePrint&articleID=CA322627]
- Maisonneuve, Marc. 2003. « Recherche multibases : de nouveaux outils pour accroître l'autonomie des usagers », *Documentaliste - Sciences de l'information*, vol. 40, no 3 (juin) pp. 214–17.
- McCulloch, Emma. 2004. « Multiple Terminologies: An Obstacle to Information Retrieval », *Library Review*, vol. 53, no 6, pp. 297–300. [http://dx.doi.org/10.1108/00242530410544376]
- Mittermeyer, Diane et Diane Quirion. 2003. *Étude sur les connaissances documentaires des étudiants entrant au 1^{er} cycle dans les universités québécoises*. Montréal : CREPUQ, 108 p. [http://crepuq.qc.ca/documents/bibli/formation/etude.pdf]
- NISO (National Information Standards Organization). 2004. *MetaSearch Initiative*. [http://www.niso.org/committees/MetaSearch-info.html]
- Pace, Andrew K. 2004. « Much Ado About Metasearch », *American Libraries Online*, June/July. [http://www.ala.org/ala/online/techspeaking/techspeak2004/JuneJuly204muchado.htm]
- Sadeh, Tamar. 2004. « The Challenge of Metasearching », *New Library World*, vol. 105, no 1198/1199, pp. 104–12. [http://dx.doi.org/10.1108/03074800410526721]
- Spink, Amanda, Dietmar Wolfram, B. J. Jansen et Tefko Saracevic. 2001. « Searching the Web: The Public and Their Queries », *Journal of the American Society for Information Science*, vol. 52, no 3 (fév.), pp. 226–34.
- Strauss, Howard. 2002. « Library Portals: A Minority Report », *Net Connect* (Fall 2002), pp. 34–36. [http://www.libraryjournal.com/index.asp?layout=articlePrint&articleID=CA251688]
- Tallent, Ed. 2004. « Metasearching in Boston College Libraries: A Case Study of User Reaction », *New Library World*, vol. 105, no 1196/1197, pp. 69–75. [http://dx.doi.org/10.1108/03074800410515282]
- Vézina, Kumiko. 2003. « La technologie OpenURL : vers une intégration accrue des ressources et services documentaires », *Argus*, vol. 32, no 2, pp. 7–11.
- Zhou, Joe. 2003. « A History of Web Portals and Their Development in Libraries », *Information Technology and Libraries*, vol. 22, no 3 (sept.), pp. 119–28.