

Une expérience de recensement de ressources accessibles par Internet : « Les Signets de la BnF »

par Jean-François Vincent*

La BnF a entrepris un travail de signalement, de vérification et de mise à disposition sur son site bnf.fr de Signets. Ces adresses de sites Web sont éminemment utiles aux chercheurs, mais aussi aux bibliothécaires qui souhaitent orienter mieux leurs usagers. Pouvez-vous nous présenter cette opération, les moyens qu'elle implique, vos méthodes de travail pour aboutir à un résultat rigoureux dans la veille et la mise à jour ?

Dans quelle mesure utilisez-vous les ressources d'autres établissements ? Dans quelle mesure associez-vous des « experts » de chaque discipline à vos recherches ? Associez-vous également d'autres bibliothèques à votre recherche, ou aux tests des sites que vous signalez ?

Les Signets de la Bibliothèque nationale de France, une sélection encyclopédique et commentée de ressources¹ accessible par Internet, sont consultables depuis octobre 1997 dans les salles

* BnF, Département de recherche bibliographique

de lecture de la BnF, et depuis octobre 1998 sur Internet².

Il faut y voir indissociablement un service rendu au lecteur, qu'il s'agit de guider domaine par domaine vers les ressources majeures disponibles sur Internet, et un travail expérimental en cours, par lequel la BnF se donne les moyens de juger comment il est possible de recenser les ressources gratuites d'Internet, avec quelles méthodes et à quel prix.

Règles et méthodes des Signets

Nous sommes partis à la mi-1997 de règles très simples, qui convenaient à notre absence d'expérience : choisir de bonnes ressources pour nos lecteurs, les décrire, les classer, les tenir à jour. Il s'agissait alors principalement de fournir aux usagers des salles de lecture une aide dans leur recherche par Internet, en visant une complémentarité entre les ressources sélectionnées et nos collections. Selon les disciplines, divers partis documentaires ont été pris, ce qui était tout à fait normal et acceptable dans une première phase d'expérimentation.

Après quelques mois de travail, le passage des *Signets* sur la Toile nous a paru se justifier, même en laissant provisoirement le répertoire en l'état (l'année 1998, celle de l'ouverture du rez-de-jardin, n'était guère favorable à des remaniements de fond !) Cependant, il était clair aussi que l'élargissement du public potentiel des *Signets* et leur « délocalisation » sur la Toile devaient nous conduire à repenser leur fonction et leur contenu. C'est notamment cette tâche que nous nous sommes donnée en 1999, et que nous poursuivons.

Les lignes suivantes décrivent les résultats de nos réflexions qui sont stabilisés ; elles correspondent d'ores et déjà à la réalité de beaucoup des pages des *Signets*, mais pas de toutes, et ne sont pas encore publiées sur la page d'accueil du répertoire, qui s'en tient pour l'instant à une description très générale de son contenu.

Public et contenu

Le public que nous cherchons à servir est le public académique (étudiants, universitaires, chercheurs) ou « professionnel » dans certains cas (bibliothécaires notamment), ainsi que le grand public partageant les mêmes centres d'in-

térêt. Nous cherchons à sélectionner des ressources de référence pour chaque discipline, avec l'ambition de constituer l'étagère de travail de notre public : la collection de ressources de qualité disponibles par Internet et qu'il a souvent besoin d'avoir sous la main.

On trouve donc par exemple dans *Les Signets* : des métasites³ ; des collections de ressources primaires (textes numérisés, bases de données factuelles diverses...) ; des catalogues de bibliothèques en ligne et des bases bibliographiques ; des

périodiques en ligne utiles pour la discipline concernée ; des sites d'institutions importantes ; des ouvrages de référence divers (encyclopédies, annuaires, par exemple) ; certaines ressources interactives (listes de discussion, par exemple) ; des ressources monographiques particulièrement importantes... Cette liste n'est nullement limitative. Qu'on pense à ce que l'on s'attend à trouver dans les collections d'usuels d'une bibliothèque, qu'on s'efforce de le transposer dans le monde documentaire d'Internet, et

on se fera une idée de ce que nous essayons de rassembler.

La langue des ressources n'est pas un critère majeur dans notre sélection, vu le public concerné, même si une certaine préférence doit être donnée aux ressources en français. Les ressources entièrement payantes ne sont pas retenues.

Environ cent cinquante thèmes sont abordés dans la section principale, et sont accessibles par une liste alphabétique. *Les Signets* proposent en outre trois sections à

vocation généraliste. La section « Catalogues de bibliothèques » recense un certain nombre de catalogues qu'il nous a paru commode de rassembler, notamment pour nos propres besoins en service public (mais il y a d'autres catalogues de bibliothèques dans les pages thématiques). La section « Outils de référence » rassemble des généralités (annuaires généralistes, répertoires généralistes de bases de données, répertoires de périodiques, ressources pluridisciplinaires, etc.). La section « Outils de recherche par Internet » décrit les grands instruments de recherche d'Internet (répertoires comme *Yahoo!* et *dmoz*, « moteurs de recherche » comme *Voila* ou *Google*, répertoires sélectifs comme *SOSIG* ou *DutchESS*, etc.) et s'efforce de suivre leur évolution, qui est extrêmement rapide.

Il est important de bien comprendre que ce travail se fait en principe au niveau des généralités des domaines abordés, qui sont eux-mêmes assez larges. Cherchez-vous à connaître les outils les

plus importants d'Internet pour la connaissance de la musique? *Les Signets* sont un répertoire qui devrait vous être fort utile. Cherchez-vous une information immédiate sur le compositeur Kurt Weill? *Les Signets* ne vous la fourniront pas. En revanche, ils vous signaleront *Worldwide Internet Music Resources*, métasite entretenu par la William and Gayle Cook Music Library (Indiana University School of Music), qui vous permettra de trouver le site de *The Kurt Weill Foundation for Music*.

AltaVista ou *Google* vous auraient permis de le trouver aussi? En l'occurrence, c'est exact. Mais ils ne vous donnent pas de garantie sur sa qualité. Or, il y a plus d'un milliard de pages accessibles sur la Toile, sans compter les bases de données dont le contenu est généralement inaccessible aux « moteurs de recherche ».

L'enjeu des répertoires de ressources d'Internet entretenus par des bibliothèques ne peut être de concurrencer, et encore moins de

remplacer, les colossaux outils de recherche commerciaux (dont notre profession aurait d'ailleurs beaucoup à apprendre). Il est de présenter des ressources consultées, approuvées et périodiquement contrôlées par des bibliothécaires qui connaissent leur domaine et son offre documentaire. Ils ne proposent ni plus de données, ni même autant, mais ils les proposent autrement et, pour certaines recherches, mieux.

Notre parti pris de ne sélectionner que des « références » exclut nombre de documents excellents (articles individuels, monographies très spécialisées, etc.). En partie dicté par des contraintes matérielles, il nous semble avoir sa cohérence et son utilité. De toute façon, ne vaut-il pas mieux chercher à faire relativement peu, et bien, que prétendre tout faire sans le pouvoir? Il faudra par ailleurs un certain temps pour que chaque section se stabilise, et pour que cette collection de « références » devienne un jour, peut-être, une « collection de référence », c'est bien clair. Les compétences variées présentes à la BnF peuvent nous laisser espérer que ce n'est pas impossible.

Exemples de notices : ci-dessous, extrait de la section « Discographie » de la page « Musique », entretenue conjointement par le Département de l'audiovisuel et le Département de la musique ; version du 27 janvier 2000.

Discographies

● Gramofile

Ce site est une base de données discographiques classiques unique en son genre. Elle est riche de plus de 24000 notices critiques de disques parues dans la revue anglaise *Gramophone* depuis mars 1983. Une recherche multicritère très performante permet de croiser les titres d'œuvres, compositeurs, interprètes, éditeurs, collections. Les critiques portent en particulier sur l'interprétation et listent également des enregistrements de référence pour chaque œuvre. La consultation est gratuite, mais un formulaire doit être rempli à la première visite.

<http://www.gramofile.co.uk/>

● All-music

Une très grosse base de données pour toutes les musiques non classiques. On y trouve des biographies et discographies d'interprètes et de compositeurs en abondance et la base est enrichie continuellement. Elaboré en Amérique du Nord, ce site est moins performant pour les musiques de traditions non occidentales mais constitue un outil remarquable pour les recherches discographiques en jazz et en rock.

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

Mises à jour

Les contraintes que nous nous imposons pour la mise à jour des *Signets* sont élevées : chaque page doit être actualisée tous les deux mois⁴. Cela recouvre la vérification des liens, mais aussi le contrôle de l'exactitude de la description. C'est ce contrôle qui est le talon d'Achille de tout projet de recensement, et c'est lui qui en fait une opération coûteuse (et souvent fastidieuse !) : la mobilité du contenu des ressources d'Internet, qui est un problème nouveau pour les bibliothécaires, oblige ceux qui veulent les décrire à suivre au long cours leur évolution et à sans cesse reprendre leur ouvrage.

Nous tâcherons dans l'avenir d'affiner les règles de mise à jour pour essayer de gagner du temps, y compris en automatisant ce qui peut l'être (pas grand-chose !) Mais cela restera de toute façon la charge principale, à moins d'accepter une chute de qualité considérable.

Veille et méthodes de sélection

Pour ce qui est des méthodes de veille, le discours que nous aurions à tenir n'est pas original. Je me contenterai ici de renvoyer à l'indispensable *DESIRE Information Gateways Handbook*⁵, et en particulier à sa section « Resource discovery », qui me paraît décrire adéquatement l'état de l'art. C'est dans ce sens que les participants aux *Signets* sont invités à travailler.

La sélection elle-même repose essentiellement sur la connaissance de chaque domaine par le bibliothécaire qui s'en occupe. Nous n'utilisons pas de « grilles de sélection », qui nous apparaissent comme des outils lourds ou peu fiables : en fin de compte, n'est-ce pas toujours la compétence scientifique préalable de celui qui les manie qui détermine le résultat ? Une liste de critères auxquels il faut être attentif

a été établie, notamment pour attirer l'attention des rédacteurs sur certaines spécificités techniques des ressources d'Internet. Mais, pour le reste, les documents d'Internet sont comme les autres et font appel aux mêmes réflexes professionnels que les autres.

Formation

Notre tâche implique en revanche une formation assez lourde, correspondant à l'apprentissage d'une nouvelle facette du métier, complexe et évolutive : la connaissance du réseau et de ses pratiques techniques. Nous avons fait des efforts en ce sens, et nous devons les poursuivre. Mais il est plus facile et bien moins coûteux pour la collectivité de former à Internet des bibliothécaires compétents dans un domaine scientifique que de former à un domaine scientifique des personnes compétentes à propos d'Internet.

Or, c'est la compétence scientifique et documentaire qui peut seule justifier que les bibliothèques se lancent dans cette aventure. On ne voit guère qui pourrait à leur place réaliser des répertoires scientifiquement satisfaisants (ce qui n'exclut pas, loin de là, qu'elles puissent coopérer sur ce point avec les universitaires ou les chercheurs).

Organisation du travail

Les Signets s'appuient sur l'organisation du travail déjà existante dans la Direction des collections : des acquéreurs ou des catalogueurs des services compétents s'occupent donc aussi bien de la documentation traditionnelle que de la documentation en ligne sur Internet. Cela se fait, il est vrai, au prix d'un problème d'accroissement de leur charge de travail, difficulté à laquelle la plupart des bibliothèques intéressées par cette

nouvelle tâche seront sans doute confrontées.

Chaque participant est invité à consacrer environ une journée par mois à ce travail. Cela peut paraître peu, mais il faut bien voir que quatre-vingt-dix personnes sont concernées par cette expérience, l'équivalent de quatre pleins temps et demi, répartis dans presque tous les départements de collections.

Dans chaque département, une personne au moins consacre une part de son temps à la coordination du travail. En outre, l'équivalent d'un gros mi-temps me permet de coordonner l'ensemble. Un comité de rédaction réunit chaque mois les coordonnateurs, pour instruire les évolutions nécessaires et pour régler les problèmes de fonctionnement quotidiens. Cette instance s'est révélée indispensable pour le bon fonctionnement de l'organisation très déconcentrée qui est celle des *Signets*.

La responsabilité du contenu scientifique de chaque page des *Signets* appartient au département qui l'entretient, pour tout ce qui n'est pas défini par des règles de politique documentaire communes.

Les Signets et les missions de la BnF

Dans l'immédiat, *Les Signets* enrichissent l'offre documentaire des salles de lecture, où ils servent toujours de « page d'accueil » aux postes donnant accès à Internet, et constituent une aide et un point de départ pour les lecteurs. Ils tendent par ailleurs, sur la Toile, à fournir un service ayant sa propre cohérence, et ce en français. Par ailleurs, ils développent dans l'établissement une certaine expertise, qui peut lui permettre de jouer en ce domaine un rôle dans l'animation du réseau des bibliothèques françaises.

À plus long terme, lorsque la vocation patrimoniale de la BnF

devra s'exercer relativement à Internet, l'expérience du recensement thématique, plus simple à mener à bien que la conservation ou le recensement sur une base nationale, pourra être réutilisée pour une part.

Coopération

Les Signets recensent aujourd'hui moins de deux mille ressources. À moyens constants, l'ordre de grandeur ne devrait pas changer considérablement vu la charge que représentent les mises à jour. Comment faire alors pour que soient tout de même signalés à nos lecteurs les documents de valeur que notre politique documentaire exclut ? La seule solution possible se nomme coopération.

La question de la coopération en la matière dépasse largement le sujet de cet article. Je soulignerai seulement que l'effort à accomplir pour maintenir en état une collection de ressources d'Internet interdit à n'importe quel établissement individuel d'espérer viser l'exhaustivité, sinon dans des domaines thématiques très réduits. Si les

bibliothèques veulent s'y engager, elles devront unir leurs efforts pour proposer à leurs usagers des « guichets communs » donnant accès à des données réunies par un certain nombre d'établissements. En des termes plus traditionnels et sans doute plus clairs, la forme du catalogue collectif s'imposera. Différentes solutions techniques existent (base de données unique ou bases de données multiples et interrogées simultanément, échanges de notices...), peu nous importe ici.

En ce qui concerne l'utilisateur, le résultat doit être en tout cas, pour une recherche donnée, une fenêtre d'interrogation unique et une page de résultats unique. Il ne sera pas disposé longtemps (si tant est qu'il l'ait jamais été) à errer de sites de bibliothèque en sites d'université à la recherche de recensements minuscules et souvent redondants.

En France, presque tout reste à faire pour organiser cette indispensable coopération. D'autres pays (Grande-Bretagne, Pays-Bas, Allemagne, Suède, Finlande, Danemark, pour s'en tenir à l'Europe)

sont plus avancés. Il faut signaler cependant chez nous l'existence de *Sitebib*⁶, qui organise la coopération dans le domaine des sciences de l'information et de la bibliothéconomie. La cellule MathDoc (Grenoble) et les bibliothèques mathématiques d'Orsay et de Strasbourg participent au projet européen *EULER*⁷ dans le domaine des mathématiques. Enfin, le rapport de Bruno Van Dooren⁸ sur les bibliothèques universitaires et les nouvelles technologies préconise, dans son objectif 7, un découpage thématique de l'effort de recensement entre les établissements universitaires compétents, avec une mise en œuvre envisagée « à brève échéance⁹ ».

Au niveau international, une organisation informelle, *IMesh*¹⁰, s'est réunie pour la première fois en juin 1999 à Warwick (G.-B.). Son objectif est de favoriser cette coopération. En Europe, le projet *Reynard*¹¹, auquel la BnF participe, a débuté en janvier 2000. En termes rapides, il vise à établir les bases techniques, bibliothéconomiques et organisationnelles d'un

« catalogue collectif » par sujets à l'échelle européenne, destiné à un public académique. Il doit aboutir, en deux ans et demi, à la réalisation d'un service pilote qui sera mis à la disposition du public.

Les Signets coopèrent actuellement dans le cadre de *Sitebib*. L'implication de la BnF dans le projet *Reynard*, même si les visées de ce projet dépassent de beaucoup l'échelle de notre répertoire, est une nouvelle étape, qui sera sans doute très importante pour *Les Signets*. L'un de nos objectifs en 2000 est en effet de mettre en place un dispositif technique (base de données) et bibliothéconomique (format de notices, normalisation de la classification, clarification et publication de la politique documentaire) permettant la coopération en accord avec les spécifications qui émaneront de *Reynard*. Ces aménagements devraient se traduire par une amélioration de la qualité du contenu, en conduisant à des pratiques mieux harmonisées, et par une amélioration des accès.

Enfin, le projet d'établissement sera cette année pour la BnF l'occasion de préciser ses objectifs en matière de coopération. On me permettra de ne pas anticiper sur un débat en cours.

Je dirai pour finir qu'il y a place, sur Internet, pour d'autres entreprises que celles qui se battent sous nos yeux à coups de milliards de dollars. Place notamment pour des recensements fondés sur le souci non pas de s'assurer des parts de marché mais de promouvoir le savoir, la culture, le goût des arts et des lettres, enfin ces valeurs qui nous ont fait choisir ce métier plutôt que d'autres. Il n'est pas tout à fait vain de pousser les institutions publiques à investir quelques moyens humains et financiers dans cet effort, qu'elles sont sans doute les seules capables de mener à bien.

En second lieu, il me paraît également important d'assurer un signalement bibliographique en français, c'est-à-dire de construire un *point de vue français* sur les ressources d'Internet. Non par vanité nationale, mais simplement parce que nous parlons cette langue et partageons une certaine histoire. L'activité bibliographique n'est pas seulement la besogne ennuyeuse qui consiste à remplir des champs normalisés, elle est aussi l'humble accès vers la connaissance par les documents. Ainsi, pour une part, elle définit par sa structure les conditions d'exercice de la connaissance tout court. Le point de vue dont je parle, déterminé par l'histoire, est une des composantes de cette structure.

Enfin, si les institutions documentaires françaises ne se chargent pas de recenser, et par là de mettre en valeur, les produits de la recherche et de la culture de leur pays, personne d'autre ne le fera à leur place de façon extensive. En regardant les répertoires existants, il apparaît nettement que même ceux qui affichent une couverture internationale sélectionnent majoritairement ce qui est produit par leurs compatriotes. On peut penser qu'il y a là une demande et un besoin qui se feront sentir rapidement. Ne faut-il pas les devancer, ou du moins s'y préparer ?

1. Un terme pleinement satisfaisant nous manque pour désigner globalement les objets accessibles par Internet et intéressants d'un point de vue documentaire. « Document » n'est pas assez général : une liste de discussion est-elle un document ? « Site » n'est pas plus correct, parce que trop partiel également, etc. Faute de mieux, j'emploierai ici le mot « ressource », comme un terme passe-partout et assez fréquent dans cet usage.

2. [<http://www.bnf.fr/web-bnf/liens/>]

3. « Métasite » : site contenant de nombreuses références à d'autres ressources ; il peut s'agir de répertoires de

ressources accessibles par Internet, ou d'ouvrages divers mais caractérisés par l'abondance des liens hypertextes externes.

4. Voir John Kirriemuir : *A Brief Survey of Quality Resource Discovery Systems: An overview of resource discovery systems and associated database*. Juillet 1999 (en ligne). In *Resource Discovery Network, section Publications*. [7.2.2000] Accessible par Internet : [<http://rdrn.ac.uk/publications/>]. La base de données associée à cette étude montre que la périodicité des mises à jour des descriptions dans les répertoires sélectifs est généralement bien moins grande, et assez souvent indéfinie ou « peu claire ».

5. DESIRE = Development of a European Service for Information on Research and Education. *DESIRE Information Gateways Handbook* (en ligne), 1999 [modifié le 10 novembre 1999]. [7.2.2000] Accessible par Internet : [<http://www.desire.org/handbook/>]. Ce manuel est une ressource essentielle pour qui s'intéresse au recensement de ressources Internet.

6. *Sitebib: coopération entre sites Web en bibliothéconomie et sciences de l'information* (en ligne). Accessible par Internet : [<http://www.abf.asso.fr/sitebib/>]

7. *European Libraries and Electronic Resources in Mathematical Sciences: The Euler Project* (en ligne). [12.2.2000] Accessible par Internet : [<http://www-irma.u-strasbg.fr/EMIS/projects/EULER/>]

8. Bruno Van Dooren : *Bibliothèques universitaires et nouvelles technologies*. Juin 1999 (en ligne). Ministère de l'Éducation nationale, de la Recherche et de la Technologie. [10.2.2000] Accessible par Internet : [<http://www.education.gouv.fr/rapport/van-dooren/default.htm>]

9. Je serais vivement intéressé de connaître toute autre entreprise collective de recensement.

10. *IMesh: International Collaboration on Internet Subject Gateways* (en ligne). [12.2.2000] Accessible par Internet : [<http://www.imesh.org/>] Ce site a pour le moment très peu de contenu. On se référera surtout au compte rendu de la première réunion : Lorcan Dempsey, Tracy Gardner, and Michael Day (UKOLN) ; Titia van der Werf (Koninklijke Bibliotheek) : *International Information Gateway Collaboration: Report of the First IMesh Framework Workshop* (en ligne). *D-lib Magazine*, décembre 1999. [12.2.2000] Accessible par Internet : [<http://www.dlib.org/dlib/december99/12dempsey.html>]. - J'ai représenté la BnF dans cette réunion.

11. Site en cours de construction à la date du 10 février 2000, mais dont le contenu devrait rapidement croître. Accessible par Internet : [<http://www.renardus.org>]