

Optimisation des accès dans un catalogue plurimédia

par Francis Agostini

Conservateur, médiathèque de la Cité des sciences et de l'industrie

Dans les grandes bibliothèques publiques, l'ampleur des fonds et la variété des médias amènent à élargir les fonctions du catalogue dans plusieurs directions :

– orientation dans les connaissances (autour d'un sujet précis ou dans un domaine large, comment les connaissances sont-elles organisées et reliées entre elles?) ;

– orientation documentaire (comment choisir dans une liste de titres d'un même média ou de médias différents?) ;

– orientation spatiale (comment localiser au plus vite une cote?) ;

– promotion de certains fonds ou de certains titres.

La médiathèque de la Cité des sciences et de l'industrie, qui a pour vocation de dif-

fuser l'information scientifique et technique auprès d'un très large public¹, a entrepris d'améliorer les accès liés aux contenus, sur le catalogue en ligne d'une part, sur le système documentaire d'accès aux audiovisuels, d'autre part.

1. On compte en moyenne 3 500 entrées par jour, et 8 000 entrées les jours de pointe.

Les usages du catalogue

La définition d'objectifs de qualité s'appuie sur l'observation des usages du public, tels qu'ils ont pu être repérés dans plusieurs études de public.

Dès 1991, une étude de public a fourni des indications précieuses sur les usages du catalogue. D'après cette étude, l'intensité d'usage et le type d'usage dépendent de plusieurs facteurs : habitudes de fréquentation², catégorie³, niveau d'étude.

- Les occasionnels et les habitués utilisent assez systématiquement le catalogue, alors que les assidus, étudiants ou actifs, s'en servent relativement peu.
- Les actifs les plus diplômés effectuent des recherches précises alors que les moins diplômés, ou les familles, le manipulent par curiosité.
- Parmi les scolaires qui viennent à la médiathèque sans projet précis, nombreux sont ceux qui explorent un sujet au catalogue comme ils explorent les espaces et les ressources du lieu. D'une manière générale, les scolaires ne sont guère satisfaits de leur utilisation du catalogue.

Les missions de la médiathèque exigent qu'un public profane (profane du domaine scientifique et technique comme du lieu) puisse exploiter les outils disponibles dans une logique de découverte. Le catalogue a d'emblée une double fonction vulgarisatrice : familiarisation avec l'outil informatique, appréhension d'une représentation des connaissances scientifiques et techniques.

De plus, les fonds regroupent une grande variété de médias de la communication scientifique, de la brochure de prévention au traité de spécialité, du film d'enseignement au reportage, du

2. On distingue désormais au sein du public de la médiathèque, les occasionnels, qui viennent pour la première fois ou ne sont pas venus depuis un an : les habitués qui viennent entre deux fois par an et deux fois par mois en moyenne ; les assidus, effectuant plus de deux visites par mois. Les occasionnels représentent 87 % des personnes, les habitués 11 %, les assidus, 2 %.

3. Plusieurs catégories sont identifiées : actifs, étudiants, scolaires, familles.

cédérom ludoéducatif au didacticiel⁴. Comment promouvoir au catalogue les médias favorisant une approche des contenus scientifiques par un public profane ?

Nous nous proposons ici de présenter certains modes d'enrichissement des accès à des parties du fonds susceptibles d'intéresser un large public. Eu égard à la place de l'audiovisuel dans la vulgarisation scientifique et dans la politique documentaire de la médiathèque, nous nous pencherons plus particulièrement sur le traitement des documents audiovisuels. Le souci d'améliorer les accès concerne d'abord l'utilisation du catalogue plurimédia dans une perspective d'intégration des médias. Depuis son installation en décembre 1994, l'étude des accès porte également sur le serveur audiovisuel, qui permet de visionner le programme de son choix au sein d'une base de plus de 3 500 films⁵.

Jouer le jeu du public

La méthode d'étude consiste à simuler des recherches, en s'appuyant sur les requêtes du public (demandes reçues aux points d'information, observations des usagers, statistiques d'exploitation du serveur). Au catalogue (GEAC), on effectuera des recherches sujet sans limitation de fonds ou bien, selon les cas, en limitant la recherche au fonds audiovisuel. Pour ce qui concerne le serveur, l'interrogation se fait par mot. Le logiciel documentaire recherche le mot dans différentes zones de la notice, notamment dans le résumé, qui figure en zone de notes.

À l'instar de l'intégration des médias dans l'espace de la bibliothèque, développer une logique d'intégration au catalogue passe par le rapprochement de références autour d'un même thème, quel que soit le type de document.

4. Ces médias sont désignés par la médiathèque sous le nom de *qualifications*. Ils ont été codifiés pour chaque support et servent en premier lieu pour assurer le développement des collections dans la cohérence.

5. Voir « Médiathèques et médiation », *Bulletin d'informations de l'ABF*, n° 170, 1996.

Dans les espaces, les différentes catégories de documents se signalent par leur aspect physique (format, présentation)⁶. Comment transposer au catalogue en ligne⁷, dans lequel toutes les notices sont traitées sur le même plan et y occupent le même espace à l'écran ?

Dans une bibliothèque dont le fonds est constitué très majoritairement de livres, quelles sont les chances pour un lecteur de tomber sur une notice de film, ou de logiciel, lors d'une recherche thématique ? Derrière cette question se profile celle de l'adéquation des listes d'autorité à l'indexation de tout type de document, et celle de l'infléchissement de la politique d'indexation dans une logique d'accès.

Si, par exemple, le film de Bernard Surugue, *Mara le regard du lion*, est simplement indexé au nom de la maladie qui constitue le sujet du document (l'onchocercose ou cécité des rivières), il aura fort peu de chances d'être proposé à un utilisateur qui ne connaîtrait pas déjà le document. Présent *virtuellement* au catalogue, ne sera-t-il pas en réalité caché au public que la médiathèque vise ? Or il s'agit d'un document d'une rare qualité, dont le propos dépasse largement le cadre d'une pathologie particulière. L'indexeur s'efforcera d'exploiter des entrées plus génériques, comme MALADIE PARASITAIRE ou MALADIE TROPICALE, qui, tout compte fait, sont peu représentées au catalogue. Nous verrons comment la résumé permet d'enrichir encore les accès.

Soumettre le catalogue à une logique d'accès, cela revient à vérifier comment, dans le catalogue public, apparaît l'offre documentaire pour des requêtes jugées particulièrement représentatives des usages du public et des missions de l'institution. Ce travail d'évaluation présente un grand intérêt pour le développement maîtrisé des collections. Mais nous laisserons volontairement de côté les questions de politique d'acquisition, pour concentrer

6. Il faut toutefois remarquer que le lecteur ne peut feuilleter tous les types de document.

7. En utilisant d'abord les techniques traditionnelles, c'est-à-dire sans faire appel au multimédia.

notre attention sur les accès documentaires.

Dès que le bibliothécaire endosse les modes d'approche du public, il doit convenir qu'il existe pour lui un large espace d'intervention pour améliorer les accès, pour réduire le hiatus inévitable entre la collection et sa représentation au catalogue.

Adopter une logique d'accès implique à la fois un travail rétrospectif et un élargissement du traitement courant. Il s'agit d'une part de rattacher le document à une famille d'analogues (les médias ou qualifications); d'autre part d'élargir l'indexation en tenant compte de la catégorie de document, de la structure des accès existants, de la nature et de l'ampleur de l'offre documentaire sur le sujet.

Cet élargissement suppose donc, au moment du traitement, de replacer le document dans le contexte de la bibliothèque : collections et outils documentaires.

Des exemples de situations de recherche, tirés de la médiathèque de la Cité des sciences et de l'industrie, vont nous permettre d'expliciter ces différents éléments.

Commençons par interroger le catalogue à *Sida*. Nous nous trouvons devant une liste de sujets qui couvrent 7 écrans et dont l'ensemble pointe sur 298 documents. La vedette simple, SIDA, renvoie à 88 références dont 6 films et 2 logiciels. Ces titres s'affichent à l'écran triés par ordre alphabétique d'auteur. Ce tri implicite exploite simplement la description bibliographique.

Le bibliothécaire soucieux de structurer l'offre pour en faciliter l'appréhension par un large public, sera amené à faire appel aux caractérisations de genre, de niveau, de support. C'est l'objet du système des qualifications développé à la médiathèque de la Cité des sciences et de l'industrie, qui se présente comme une grille regroupant les médias de la communication scientifique et technique.

Une direction importante pour l'amélioration du catalogue consiste à proposer à l'utilisateur d'exploiter ces caractérisations pour explorer les contenus et l'aider à faire ses choix. Dès que l'offre sur

un sujet se déroule sur plusieurs écrans, le besoin d'une meilleure orientation se fait sentir. De même que les cotes sont ajustées pour répartir des volumes physiques dans des espaces, il nous faut proposer des unités de sens au sein d'ensembles de notices trop vastes pour être appréhendés globalement.

Redescendons d'un ordre de grandeur dans l'échelle des quantités de références rappelées. À partir du moment où la recherche conduit à un corpus réduit (1 ou 2 écrans de notices), on constate que la lisibilité de l'offre au catalogue progresse avec une meilleure intégration des médias.

Offrir un bouquet documentaire plurimédia

Prenons le cas des demandes sur des sujets relativement spécifiques mais susceptibles d'intéresser un public non spécialisé. Une recherche par sujet donne, au catalogue, une liste de titres facilement appréhendable, faisant apparaître des supports variés. Pour choisir les documents qui peuvent répondre à son attente, l'utilisateur a le loisir de consulter les notices complètes qui, dans le cas des films ou des logiciels, comportent un résumé. On peut dire que, sur ces sujets, le catalogue propose une offre *plurimédia*.

Le tableau suivant regroupe un choix de questions posées aux points d'information de la médiathèque, qui rentrent dans cette catégorie.

Sujet	Écrans de titres	Supports non livres
Deltaplane	1 (8 titres)	2 revues 1 film
Arc-en-ciel	1 (4 titres)	2 films
Météorite	2 (16 titres)	5 films
Dauphin	2 (17 titres)	2 films 1 logiciel sur cédérom
Paludisme	2 (18 titres)	3 films
Centrale nucléaire	2 (17 titres)	2 films

Insistons sur le fait que la variété des supports (et plus généralement des médias) recouvre en fait une diversité

dans les moyens d'écriture mis en œuvre et donc dans la mise en forme de l'information. Il s'agit là d'un apport essentiel à l'orientation du public dans l'offre documentaire.

Venons-en à une troisième situation, celle des demandes autour de notions très courantes, qui se retrouvent dans un grand nombre de vedettes-matière composées. Lorsque l'on interroge par des termes ou des locutions larges et que l'on choisit d'abord les vedettes simples, dans une démarche d'exploration, l'offre plurimédia semble bien pauvre si l'on en juge par le nombre de documents audiovisuels rappelés.

Le tableau suivant donne ce nombre, rapporté au nombre total de documents rappelés, pour six vedettes qui se présentent comme des unitermes.

Sujet	Part des notices d'audiovisuels
CONSTRUCTION	0/18
BÂTIMENT	0/35
MAMMIFÈRE	2/78
OcéAN	1/19
FORÊT	3/77
POLLUTION	1/26

L'explication réside dans un traitement « équitable » des différents types de documents qui conduit à indexer un film sur un sujet spécifique exactement comme un livre traitant du même sujet. Si l'on indexe tous les médias « au niveau de finesse du document » et seulement à ce niveau, on court le risque de marginaliser, au catalogue, les médias minoritaires, comme les documents audiovisuels.

Une réponse consiste à indexer à plusieurs niveaux les documents que l'on souhaite promouvoir, comme le film cité plus haut. En suivant ce principe, des films présentant des essences particulières d'arbres méritent d'être « surindexés » par la vedette générique ARBRE. De la même manière, indexer systématiquement tous les films sur le sida, avec la vedette simple (s'ajoutant aux vedettes spécifiques éventuellement nécessaires) permet de rappeler, pour cette recherche, 50 % de titres en plus (9 au lieu de 6).

Ce traitement spécifique, qui heurte quelque peu l'orthodoxie de l'indexation, donne une idée du type de fonctionnalité que nous pourrions demander à un catalogue en ligne, en vue d'améliorer les accès à certains médias.

Des attentes tournées vers le média

Que se passe-t-il lorsque les questions du public incorporent la dimension évoquée plus haut, celle du média ?

Les recherches peuvent être conduites au catalogue en limitant la base consultée aux films. On peut également interroger le serveur audiovisuel par mot.

Commençons par les réponses du catalogue à un choix de questions bien circonscrites, posées par le public :

Énoncé de la demande	Répartition de l'offre
Films sur la conquête de l'espace	13 films liés à 6 vedettes
Films sur les chiens	4 films liés à 4 vedettes
Films dans le secteur de la téléphonie	3 films liés à une vedette
Images de dauphins	2 films liés à une vedette

Dans les deux premiers, une partie des documents sont indexés par des vedettes composées (par exemple, des vedettes mentionnant le nom des planètes ou des pays concernés pour les films sur l'exploration spatiale). Dans les deux autres cas, on a utilisé des vedettes simples (TÉLÉPHONE, DAUPHIN).

Il semble qu'une bonne économie de l'indexation puisse améliorer les accès, d'autant que l'on dispose d'un autre moyen de recherche, qui fournit des réponses plus nombreuses en général : le système documentaire audiovisuel connecté au robot serveur de vidéos.

Recherche sur le serveur audiovisuel

Les films accessibles au public sont signalés au serveur par une notice détaillée, comprenant, outre les éléments figurant au catalogue, la mention de la qualification, qui donne une idée du genre du document. De plus, le résumé fournit des accès supplémentaires, comme le montrent les exemples suivants :

Sujet demandé	Réponse au serveur	Réponse au catalogue (limité aux films)
Dauphin	3 titres	2 titres
Disque laser	2 titres	1 titre
Téléphone	4 titres	3 titres
Chien	14 titres	4 titres
Cœur	25 titres	7 titres

Mais attention au bruit ! Seules 7 réponses sur 14 sont pertinentes sur le sujet CHIEN (4 documentaires et 3 fictions). Quant au sujet CŒUR, il faut retirer 11 références sur les 25 proposées. Ces notices comprennent bien le mot **Cœur** mais dans un sens figuré. On voit bien la nécessité de contrôler l'usage de ce genre de substantif et lors de la rédaction des résumés, de veiller à respecter certaines conventions d'usage.

Cela dit, un système documentaire gérant une base de taille réduite (3 500 titres environ), peut tolérer une part de bruit. Dans ces conditions, nous estimons que ce système apporte un bénéfice réel en termes d'accès pour le public.

Reprenons l'exemple du film produit par l'ORSTOM, « Mara le regard du lion ». Le résumé précise que la maladie, la **cécité** des **rivières**, sévit dans les **savanes** d'**Afrique** de l'Ouest et que son vecteur est une **mouche**. Outre l'enrichissement de l'indexation matière, le résumé fournit plusieurs accès supplémentaires qui ont une pertinence, à des degrés divers.

Il reste une situation assez délicate, celle des questions posées sur des thèmes larges. En général, le serveur va fournir beaucoup plus de réponses que le catalogue, encore que la comparaison soulève certaines difficultés.

Sujet	Notices rappelées
CONSTRUCTION	95
BÂTIMENT	24
MAMMIFÈRE	6
Océan	81
FORÊT	40
POLLUTION	30

Une vérification portant sur la pertinence des notices rappelées et sur les documents du fonds effectivement pertinents conduit à un constat désagréable : on a non seulement du bruit mais aussi, et c'est plus ennuyeux, du silence.

Il ne nous reste plus qu'à analyser les causes et trouver les remèdes.

Le bruit provient souvent du résumé, rédigé en langage libre⁸. L'utilisation non contrôlée de certains mots (termes courants et généraux, polysémies) et l'emploi hors de leur champ propre, de termes caractérisant les grandes disciplines créent du bruit à l'interrogation.

Les statistiques d'exploitation nous fournissent des listes de requêtes ne débouchant sur aucune notice. Ces silences doivent être interprétés. Nombre d'entre eux peuvent donner lieu à un réexamen du fonds et des accès. Pour ce qui concerne les domaines larges, nous comptons « surindexer » les documents à l'aide d'une classification sommaire⁹ adaptée au fonds audiovisuel.

Voilà pour l'aspect *traitement documentaire*. Mais la fonction des résumés déborde largement la multiplication des accès. Il s'agit non seulement de fournir au public une information détaillée sur les contenus d'un film mais de lui donner des indications générales sur la visée du document, ce qui en fait l'originalité, sa construction, les techniques mises en œuvre, la tonalité du discours, la pérennité de l'information¹⁰.

8. Jusqu'à l'installation du serveur, le contenu des résumés n'avait aucun impact sur la recherche documentaire.

9. Cette classification comprendrait au plus 70 intitulés et 3 niveaux hiérarchiques.

10. Les droits des documents audiovisuels sont négociés pour une durée de dix ans et peuvent être renouvelés.

Les outils documentaires à l'épreuve du public

À écouter les remarques des utilisateurs, les catalogues sont loin de répondre à leurs attentes. La succession des écrans, listes et index alphabétiques, leur paraît souvent dénuée de sens. À la perception de l'espace de la bibliothèque comme labyrinthe, répond celle du catalogue. La mise en espace des collections doit donc être transposée au niveau du catalogue dans sa dimension scénographique comme dans sa dimension d'ergonomie cognitive. En partant de la pratique du catalogue au présent, nous avons privilégié le deuxième aspect.

Nous avons cherché à alimenter l'étude des besoins des utilisateurs pour préparer l'évolution du système.

Une première direction de travail concerne l'intégration des médias. Celle-ci passe par la restructuration des données autour de la qualification des documents. Faire de la qualification un critère de même rang que le sujet ou les éléments d'identification permettrait au public, au moment où il le souhaite, de préciser le profil de sa demande. Choisie d'emblée, la qualification (ou un groupe qualifications) pourrait servir à explorer un domaine large. Demandée comme précision en cas d'offre trop vaste sur un sujet, elle fournirait un

moyen rationnel de réduire le corpus de références rappelées. Le tri des notices affichées par type de document (média), dans un ordre favorisant les usages profanes, fournirait certainement au public une aide appréciable.

Nous avons longuement présenté les différents moyens d'adapter la politique d'indexation aux missions de vulgarisation :

- en élargissant l'indexation matière des documents susceptibles d'intéresser un large public (y compris à l'aide d'une classification pour les médias audiovisuels) ;
- en codifiant la rédaction du résumé des documents audiovisuels ou logiciels.

Parmi les voies d'amélioration qui se dessinent à ce stade, citons encore :

- l'intégration des recherches par sujet et des recherches par mot au catalogue (basculer vers la recherche par mot en cas d'échec de la recherche par sujet) ;
- l'incorporation de systèmes d'analyse linguistique (morphologie, syntaxe).

Le renseignement au public comme évaluation

L'observation des usages des systèmes documentaires mis à la disposition du public et la simulation des recherches

sur ces outils font partie de l'évaluation des ressources de la bibliothèque. La qualité de l'accueil et de l'orientation du public, en y incluant la médiation humaine, est conditionnée par ce travail d'évaluation.

Jouer le jeu du public, cela suppose de la part du bibliothécaire :

- de considérer les moyens techniques non comme des outils qui permettraient une communication transparente, mais comme des gisements d'usages, des terrains d'étude des modes d'appropriation de la bibliothèque ;
- de pratiquer, chaque fois que c'est possible, une évaluation en situation de médiation (service au public).

Dans une logique de service, le bibliothécaire accordera une grande attention à la pragmatique de l'échange avec celui qui par sa question ou son attitude a initié l'échange. De la même manière, dans une logique d'accès, il se penchera avec un regard critique sur la pragmatique du dialogue entre la machine et l'utilisateur. Placé devant l'outil documentaire dans le rôle de juge et partie, il pourra s'appuyer sur le contenu des échanges vécus en situation pour les transposer à l'écran.