

Denise Malrieu
CNRS-CDSH

A QUAND DES BANQUES DE DONNÉES D'INFORMATION SCIENTIFIQUE EN SCIENCES SOCIALES* ?

DE nombreux commentaires ont été faits de la faible utilisation en France des banques de données bibliographiques (BdB) scientifiques, que ce soit les BdB liées à la recherche fondamentale en sciences exactes ou les BdB en sciences humaines ou sociales ; ce constat a conduit certains membres des instances scientifiques à voir là un investissement difficilement justifiable dans le contexte économique actuel.

Citons les chiffres rapportés par G. Courtois : « *Les 130 bibliothèques universitaires n'ont consommé en 1985 que 3 000 heures d'interrogation environ, soit à peine deux heures par mois et par section. L'enquête menée par l'Association des centres serveurs français le confirme : l'ensemble du secteur enseignement-recherche ne représente que 7,5 % de la consommation totale assurée par les principaux serveurs français. Les chiffres sont donc éloquentes : les universités, c'est le moins qu'on puisse dire, n'abusent pas des banques de données* » (1).

On peut remarquer que, même si ces chiffres sous-estiment l'usage réel des BdB en sciences humaines et sociales sur les serveurs du CNRS et Télésystèmes, les chiffres restent effectivement faibles. Les explications généralement données à ce phénomène dans les milieux de l'information sont soit d'ordre sociologique : marginalité des BdB dans l'ensemble des canaux d'information pratiqués de longue date dans les différents types de communautés professionnelles (5,11), soit d'ordre économique : les publics de la recherche et les universitaires font difficilement appel à un service documentaire payant.

D'autre part, les échanges avec les documentalistes universitaires ou de la BPI confirment un phénomène assez contradictoire : un changement d'attitude positif du public étudiant à l'égard de l'usage des BdB, pourvu que la tarification reste raisonnable, mais sans que cette évolution s'accompagne d'une croissance massive de l'utilisation au cours de ces deux dernières années.

Dans l'autre partie du public universitaire (enseignants, chercheurs), l'engouement récent en sciences humaines et sociales pour l'outil informatique, et ses applications documentaires, s'accompagne non d'un usage accru des BdB existantes (quoiqu'à l'occasion on ne rechigne pas au téléchargement), mais d'activités d'appropriation de l'information (confection de fichiers personnalisés individuels ou de laboratoire). Ces activités documentaires s'accompagnent souvent d'une argumentation sur le caractère peu adapté des BdB de masse¹ aux besoins de la recherche.

Ces activités de constitution de BdB personnalisées se sont multipliées souvent de façon compulsive et artisanale, sans réelle réflexion sur les objectifs et les méthodologies de traitement, et il n'est pas certain que les critiques adressées aux BdB de masse par ce dernier public ne soient pas, à moyen terme, valables pour ces nouveaux produits.

Aussi nous semble-t-il nécessaire de pousser plus avant l'analyse des causes de la crise que traversent les BdB liées à la recherche scientifique. Si crise il y a, celle-ci nous semble due à différents facteurs que nous n'avons pas compétence d'analyser dans leur ensemble : il n'est pas question de faire ici l'analyse historique des décisions et de la politique scientifique menées à l'égard des BdB de masse depuis leur création, ni de l'ambivalence permanente des instances de décision à l'égard d'un outil « vitrine », mais jamais doté des moyens correspondant aux ambitions. Ceci constituerait à soi seul un chapitre de l'histoire du CNRS et de l'Université. Les facteurs économiques demanderaient en particulier à être pris en compte, que ce soit l'emprunt de modèles anciens de production industrielle pour l'organisation de la production des BdB de masse, ou que ce soit l'organisation de la diffusion et, en particulier, l'absence d'études de marché des informations scientifiques.

On pourrait invoquer aussi l'incapacité des BdB — sans doute liée

*Je remercie Robert Nadot pour les encouragements à l'analyse de ces questions, et pour les critiques et l'aide apportées à la rédaction de ce texte.

1. Nous distinguons les BdB très spécialisées de laboratoire, à usage interne, et les BdB de masse moins spécialisées, à large diffusion sur les grands serveurs.

aux facteurs précédemment cités — à bénéficier des résultats de la recherche-développement en sciences de l'information et au peu de développement de cette discipline en France. Cependant l'approche purement technique des problèmes qui y prévaut suppose que l'usage des BdB sera accru, soit en facilitant l'accès (améliorer la convivialité des logiciels d'interrogation), soit en diminuant bruit et silence : à travers le développement de recherches évaluatives sur les logiciels d'interrogation et techniques de classement et indexation (indexation manuelle ou automatique et interrogation de textes intégraux ou pas); ce dernier aspect est de loin le plus développé car il est doublement « motivé » : faire l'économie du travail du documentaliste analyste-indexeur, travail souvent sujet à critique de la part des auteurs. Or cette approche, même si elle apporte quelques améliorations, nous semble présenter un certain nombre de faiblesses :

— elle définit l'offre *a priori*, sans étude préalable, des types de demande d'information scientifique. Elle présuppose l'existence d'un utilisateur unique, alors qu'il existe des classes d'utilisateurs avec des demandes diversifiées et qui sont outillées de façon diverse quant à leur capacité à formuler la demande ;

— elle parle de l'information comme si celle-ci existait en soi, indépendamment de sa source et de sa cible, comme si un même message avait la même valeur d'information quel que soit l'utilisateur, ou quelle que soit la phase de travail où se situe l'interrogation. D'où les polémiques sans fin sur l'évaluation de la pertinence des documents et l'évaluation des BdB ;

— elle ne prend pas en compte les spécificités de l'information scientifique : la fonction de sélection exercée par les BdB scientifiques ne peut être vraiment opératoire que si les dimensions de tri qu'elle met en oeuvre sont définies par rapport au type de questions pratiquées dans le champ scientifique. A ce propos, on peut constater une forte disparité d'utilisation des BdB suivant les secteurs d'activité et leur position sur l'axe recherche appliquée/recherche fondamentale, les banques de données factuelles (que ce soit en économie, gestion ou chimie...) étant, toutes proportions gardées, plus utili-

sées que les BdB traitant de la littérature de recherche.

Il semblerait donc que les BdB les plus opérationnelles sont celles qui ont une forte composante technique, ou dont les objets ou techniques font l'objet d'un fort consensus ; dans ce dernier cas, l'organisation de l'information (la définition des critères de sélection ou de tri) dans la BdB est plus facilement adaptée aux types de question des utilisateurs. A cet égard, le secteur des SHS sur lequel est centrée cette analyse, car c'est celui que nous pratiquons, est sans doute exemplaire par le faible consensus scientifique qui y règne et la faible diffusion des outils auxquels ces sciences ont donné naissance.

Nous voudrions donc préciser ici les axes de recherche qui permettraient une meilleure adéquation de l'offre et de la demande, étant entendu que cette dernière ne peut être définie à partir de l'évaluation de la satisfaction par rapport à l'offre. Seule une double approche de la demande — étude des questionnements sous-jacents aux recherches et étude des besoins sous-jacents aux demandes adressées au système d'information — permettra de définir une méthodologie de description du document qui autorise les tris utiles à l'interrogation.

Les limites du modèle bibliothéconomique

Les BdB scientifiques de masse sont encore élaborées selon le modèle de la documentation *in situ* : selon un schéma *a priori* qui découle directement de la bibliothéconomie et non d'une étude des comportements de recherche d'information ou de questionnement scientifiques.

Plans de classement et indexation restent le pilier de la recherche documentaire, opération de classification du document dans des tiroirs préétablis : le gain opéré par l'informatisation étant d'autoriser des classifications plus riches et plus fines que le fichier manuel et une manipulation d'ensemble plus importante que précédemment.

Une recherche scientifique ne se définit pas seulement par l'objet concret d'étude, mais aussi par le type d'approche, le cadre théori-

que d'interprétation et le choix de la méthodologie mise en oeuvre, l'objet scientifique étant le résultat de cette construction. Or la description bibliothéconomique reste une description unidimensionnelle ; le critère essentiel de sélection est la thématique du document : l'objet social étudié (lecture, sport, aspirations professionnelles etc.) et non le type de problème et de démarche scientifique engagée.

Ce premier point a pour conséquence, vu la croissance de la production scientifique, un pouvoir de sélectivité faible de la BdB. La classification des articles étant thématique et non fonctionnelle par rapport aux procédures mises en oeuvre dans la recherche, les classes ainsi formées comportent des items toujours plus nombreux et non différenciés ; la sélection des documents se fait alors en deux temps : le premier choix effectué par le système ne dispense pas l'utilisateur de trier des listings souvent lourds, listings sur lesquels la deuxième opération de sélection n'est pas toujours évidente à faire, surtout si l'analyse est absente ou succincte et l'indexation pauvre. Le rapport coût-bénéfice n'est pas toujours ce qu'attendait l'usager, compte tenu de la difficulté des BdB à gérer le rétrospectif.

Cette faible sélectivité est d'autant plus sensible que l'interrogation de BdB dans le cadre d'une activité de recherche est le lieu privilégié d'expression d'une ambivalence inhérente à la recherche d'information : désir de connaître ce que les autres ont fait, besoin de se situer ; en même temps, crainte de voir remise en question sa propre problématique ou méthodologie, ou de ne pas trouver les âmes sœurs dans cette abondance de références. Cette dernière provoque soit une réaction de défense chez le non-spécialiste : « *Est-il possible que tant de personnes travaillent sur mon sujet ? Si je lisais toute la littérature triée par la BdB, je ne ferais jamais de recherche* », soit déception chez le chercheur confirmé, qui n'est intéressé que par les documents pertinents dans sa propre optique de recherche.

Ce modèle bibliothéconomique a pour autre conséquence une pauvreté des questions autorisées par le système et une sous-utilisation de l'outil informatique. L'utilisateur, dans son interaction avec le système, moule sa procédure de questionnement sur le schéma

bibliothéconomique sans s'autoriser d'autres schémas sans doute plus économiques, mais soit impossibles, car non prévus par le système, soit possibles mais non utilisés car divergents du schéma habituel : on peut constater par exemple que, dans une phase déjà avancée de repérage du champ et de la littérature scientifique, la démarche qui consiste à partir d'un article bien ciblé et de ses citations peut être une démarche très économique par rapport à une démarche classique.

Cependant cette stratégie n'est généralement pas utilisable sur les BdB. Elle serait d'ailleurs d'autant plus efficace — en particulier pour l'étude de la communication et de la transmission des savoirs — que serait pris en compte, non plus comme dans *Social science citation index* la citation brute de fin d'article, mais l'environnement sémantique de la citation (cf. 13, 26). L'observation, tirée de notre pratique, des questions adressées au système, mais qu'il reste à étayer de façon systématique, montre que ces questions ne sont pas forcément de type bibliographique au sens de recherche d'accès aux documents *via* les catalogues.

Les BdB exercent de fait une double fonction : bibliographique et d'information sur la recherche. Les exemples de questions présentés (cf. encadré, exemples 2 à 6) suggèrent que les BdB sont utilisés comme un ensemble de données à retraiter pour pouvoir suivre l'évolution des champs scientifiques. On peut constater que les logiciels d'interrogation classiques ne sont pas adaptés à ce type de questionnement. Le couplage de traitements statistiques secondaires et de tris supplémentaires s'avérerait utile dans ces cas-là, comme une aide à la visualisation graphique de la structuration des champs de recherche dans la BdB. Aucun serveur en sciences humaines ou sociales n'offre, pour l'instant, l'équivalent des traitements secondaires de l'ESA/IRS ou des traitements de type leximappe², qui aideraient à cette structuration du champ.

2. Méthode d'analyse des cooccurrences de mots-clés d'indexation à l'intérieur d'une BdB d'un secteur scientifique, et donnant lieu à une représentation cartographique des cooccurrences fortes (8).

De plus, l'observation des questions met en évidence la nécessité d'une prise en compte modulable de différentes dimensions de description du document en fonction du type de questions posées. Elle montre aussi que les dimensions de tri à prendre en compte pour s'adapter à l'utilisateur et satisfaire ses questions portent sur :

- l'élaboration dont relève l'article : note de synthèse, bilan de recherche, analyse théorique... ;
- l'approche disciplinaire : étude sociologique, clinique... ;
- le périodique : recherche fondamentale, ou vulgarisation, ou recherche appliquée ;
- la source ou le matériel traité ;
- la méthodologie : enquête, étude longitudinale, observation systématique ;
- les caractéristiques de l'échantillon étudié : âge, sexe, langue maternelle...

La difficulté à adapter la stratégie de recherche au type d'utilisateur lors de l'interrogation des BdB françaises est un des problèmes que soulevait H. Dupuy du service télématique de la Bibliothèque publique d'information (14).

Les BdB dans la communication scientifique

Un deuxième facteur des difficultés rencontrées par les BdB scientifiques de masse nous semble être lié à l'oubli que l'information n'existe pas en soi, mais dans un processus de communication. Ce problème peut être abordé sous trois angles :

— *sous l'angle de l'utilisateur en tant qu'acteur social* : il s'agit de prendre en compte l'implication de la démarche de recherche d'information dans les processus d'action et de communication sociales. C'est la démarche suivie par G. Wersig et G. Windel (28), qui soulignent la nécessité pour les sciences de l'information de décrire le comportement de recherche d'information à l'intérieur du contexte plus large de l'action dont il fait partie : il est un indicateur de l'état du processus de traitement d'un problème ; suivant les étapes de l'action entreprise pour résoudre ou modifier une situation ressentie comme problématique et plus ou moins clairement formulée, l'information recherchée et utile changera de forme et de contenu.

— *suivant le type de rapport à l'activité de recherche* : V.L. Blember et P. Leggate (4), dans leur enquête sur la recherche d'information bibliographique chez les médecins, montrent que les trois groupes (praticiens, médecins chercheurs et chercheurs) n'utilisent pas de façon identique les services d'information : bibliothèques d'hôpital et interrogation de BdB ; suivant le mode d'insertion sociale de cette activité, le rapport à l'écrit comme rapport à une problématique et le mode de communication par l'écrit varieront : F. Poulle (22) pose à propos de la littérature grise le problème de comment caractériser et classer les écrits issus des différents corps sociaux qui composent le monde de la psychiatrie.

Quelques exemples de demandes

— *Articles ou ouvrages théoriques sur l'interactionnisme symbolique.*

— *Analyse sémantique d'une notion : la « qualité de l'éducation ».*

— *Etude sociologique de l'introduction d'une méthode pédagogique : quand, dans quelles matières d'enseignement, dans quels pays s'est diffusée la pédagogie par objectifs ?*

— *Etude des champs de diffusion d'une théorie : dans quels secteurs de la psychologie la théorie de Brunswik a-t-elle diffusé ?*

— *Etude historique des développements d'une méthodologie : les descendants de Lazarfeld.*

— *Ouvrages et articles sur la perception d'autrui.*

— *Enquêtes par budget temps menées entre 1980 et 1984 sur la lecture chez l'adolescent.*

— *Un étudiant de 1^{er} cycle cherche des ouvrages sur les théories de la communication dans le contexte pédagogique.*

— *Un enseignant du secondaire veut connaître les travaux francophones sur l'effet Pygmalion.*

— *Un chercheur désire des articles se référant à la contribution d'A. Kukla, parue dans les mélanges Berkowitz (1978), et celles de J.G. Nicholls (1976, 78, 79).*

Les résistances des milieux de recherche à utiliser un instrument de masse standardisé, en dehors des circuits normaux d'information, nous semblent relever des mêmes processus liés à la communication. Le chercheur se fie plus à sa connaissance des réseaux informels « *Je lis tel article, parce qu'il est de tel auteur et de telle école* », qu'à un repérage systématique tel qu'il est offert par un fichier de BdB. Les documentalistes perdent souvent de vue que l'on cherche à s'informer non seulement pour s'informer (résoudre un problème), mais pour argumenter, définir sa position à l'intérieur d'un espace ou d'un groupe social de référence. Le recours aux BdB dans le cas d'un appel d'offres en est un exemple: il s'agit de vérifier que sa position est bonne, pas trop concurrencée sur le marché de la recherche.

La BdB de masse dérangerait donc certaines habitudes de sélectivité sociale de l'information, liées au fonctionnement de la recherche (c'est ce problème que mentionne J.-C. Chamboredon, dans son article sur les difficultés rencontrées dans l'évaluation des pairs au sein des commissions CNRS) et à certains principes d'économie dans la sélection des informations utiles. Citons J.-C. Demailly: « *Les contacts interpersonnels lui fournissent une information immédiate, déjà sélectionnée et quantitativement limitée, souvent préalablement "digérée" par d'autres et donc simplifiée, sans déranger ses habitudes de travail et de pensée, et tout en bénéficiant d'un feedback qui accélère et simplifie les choses. Alors qu'à l'opposé, les instruments secondaires de rappel de l'information tendent à fournir une information optimale en multipliant souvent le nombre, les origines et la complexité des données disponibles.* »

— *sous l'angle de la source et du canal d'information*: on pourrait dire que les BdB sous-estiment la caractérisation de la source dans le traitement qu'elles offrent. Dans quelle mesure le système documentaire est-il apte à considérer un document scientifique, non seulement comme un objet à archiver, mais comme un acte de communication d'un agent donné, caractérisé par un certain mode d'insertion institutionnelle dans la recherche, véhiculant un certain type de problématique, vers un public plus ou moins proche ?

On peut constater que, dans nombre de BdB en SHS, l'information que représente l'article est nettoyée de son environnement: des éléments tels que le rattachement institutionnel de l'auteur ou le type de communauté émettrice du périodique ne sont pas considérés comme porteurs d'information. Il nous semble pourtant que la prise en compte de cette information permettrait d'analyser de façon plus précise à travers les BdB (qui ont pour elles d'être les seules à offrir une vue globale des champs scientifiques, de leur évolution et de la diversité des pratiques scientifiques en SHS), les modes de diffusion de méthodologies ou modèles entre spécialités ou les modes de communication entre types de recherche ou communautés scientifiques (recherche fondamentale et recherches sur le terrain par exemple). On peut citer l'étude de R. Grinder sur la communication entre les différents secteurs de recherche sur l'adolescence, où il souligne le cloisonnement entre les différents types de recherche: conventionnelles, recherche-action ou centrées sur la politique sociale.

L'étude déjà citée de H. Dolman et H. Bodewitz est un exemple intéressant de l'évolution, selon le schéma de Fleck, de la « sédimentation » différenciée d'un concept (ici le concept de neuro-

peptide de De Wied), suivant les types de recherche et les groupes sociaux: les cercles de recherche clinique et appliquée adoptent le concept plus facilement que le cercle restreint de la recherche fondamentale.

— *sous l'angle des retraiteurs de l'information*: ou de l'insertion des centres de production des BdB dans les circuits de production et de communication de l'information scientifique, et des rapports entre ce type d'insertion et la nature du retraitement de l'information engrangée. Il s'agit de l'étude socio-cognitive du traitement de l'information que l'on pourrait désigner comme socio-ergonomie cognitive.

Suivant le type d'insertion des centres documentaires dans les lieux de production de l'information (de la recherche) et le mode d'organisation du travail, comment s'effectuent le repérage et la collecte de l'information (en temps réel, accès à l'information vive) ? Que retient-on comme étant de l'information ? Quels sont les filtres, comment s'opère l'interprétation dans la fabrication des BdB (ergonomie de la lecture et de la prise d'information, retraduction à travers le langage documentaire, mode d'élaboration et d'évolution de ce dernier) ?

Ce problème a des implications directes quant au choix des struc-

R

tures de production des BdB, problème que nous soulevions dans *Le bilan de la documentation en psychologie* (2). Autant une structure centralisée traitant de plusieurs disciplines, disconnectée des lieux de recherche, paraît peu apte à une collecte vivante de l'information qui suppose des liens informels avec ces lieux de production, autant une structure spécialisée au niveau d'une discipline entière, et sans sous-structures spécialisées par domaines semble ergonomiquement peu viable, car un même individu ne peut prétendre maîtriser l'évolution de l'ensemble des champs d'une discipline ; autant la juxtaposition de minibases très spécialisées sans structure de coordination paraît, dans un système classique de thésaurus et d'indexation, peu apte à contrôler les polysémies issues de la juxtaposition des langages de spécialité.

Améliorer l'information

Les systèmes d'information de grande diffusion ne possèdent à l'heure actuelle aucun dispositif d'évaluation permanente des produits offerts, et de l'évolution de la demande. Améliorer le caractère performant des banques d'information scientifique passe par l'étude des types de questionnement des utilisateurs (et non utilisateurs) du système, de façon à assurer une meilleure adéquation des critères de sélection et stratégies de questionnement offerts par le système à ceux des usagers (20).

La recherche d'une adéquation meilleure de l'*Information retrieval system* (IRS) à l'utilisateur touche à celle-ci aux deux bouts de sa chaîne (3) :

- dans sa capacité à interpréter le type de demande qui lui est adressée et à offrir les filtres adéquats ;
- dans sa méthodologie de traitement de l'information : savoir décrire le document sur les dimensions pertinentes.

Capacité à interpréter correctement la demande

La performance de l'IRS dépend de sa capacité à définir un espace d'interrogation correspondant à la demande et, en quelque sorte, isomorphe aux dimensions de description des documents : les inférences faites à partir des données sur les objectifs de la biblio-

graphie (thèse, appel d'offres, maîtrise, bilan de recherches), sur le type d'utilisateur (degré d'expertise dans le domaine), doivent aider à la sélection sur les dimensions non cognitives (types de document, de communauté émettrice, de périodique, d'élaboration effectuée dans le document) ; l'interprétation de l'aspect purement cognitif de la question devra se faire de façon interactive par confrontation des structures conceptuelles et des choix procé-

dures de l'usager avec ceux de la banque d'information (19). Cette interprétation ne peut être qu'interactive, car l'usager n'est jamais à même d'explicitier d'emblée l'ensemble des contraintes de son problème.

La pratique actuelle des interrogations en SHS effectuées par des étudiants semble suggérer — une étude plus précise reste à faire — que l'étudiant se contente souvent de quelques références appelées par une logique simple, les limi-

tes purement quantitatives étant apportées par le budget disponible pour l'interrogation.

Un système d'information redéfini serait à même d'aider l'utilisateur à expliciter sa demande comme à se situer dans les champs de connaissance, théories ou pratiques scientifiques existantes sur un terrain donné. Citons J.T. Dillon (12), qui s'efforce de définir à des fins didactiques une typologie des questions sous-jacentes à la recherche en éducation : « *La classification est un acte conceptuel. On classe non les questions individuelles, mais les types de questions... Un schème classificatoire est un outil analytique pour comprendre la recherche à trois niveaux : l'étude individuelle, le corpus d'études et l'entreprise entière de recherche dans un champ donné. Ces schèmes différents peuvent aider à montrer aux étudiants la nature, les types et fonctions des questions dans la recherche. Ils peuvent les aider à percevoir le genre de questions à poser, à identifier la question qui peut faire l'objet d'une recherche, à la situer par rapport aux autres questions étudiées dans le champ. Ils peuvent les aider à concevoir le domaine à étudier et à organiser un bilan des recherches.* »

C'est dans la mesure où la BdB peut aider l'utilisateur à se constituer une espèce de cartographie conceptuelle et méthodologique du champ scientifique qu'il aborde, qu'on peut espérer un usage des banques d'information scientifique dans leur spécificité par des publics en contact avec les résultats de la recherche mais non professionnels de celle-ci (praticiens, étudiants), comme une utilisation par les professionnels de la recherche dans les

phases ultérieures à la prise de contact avec un champ (travaux de type méthodologique ou historique par exemple).

Méthodologie de traitement

En ce sens, il nous paraît tout à fait improbable que les banques d'information en sciences humaines et sociales jouent réellement une fonction de vulgarisation scientifique ou didactique et d'accès facilité à la recherche, si aucune réflexion épistémologique ne sous-tend la méthodologie de traitement de l'information scientifique engrangée.

Il est clair qu'un tel objectif demande un nouveau traitement documentaire des textes scientifiques. On peut en particulier faire l'hypothèse que la caractérisation du questionnement, de la problématique et de la procédure de résolution adoptée dans l'article implique une méthodologie qui prend en compte, non plus des listes de descripteurs, mais le réseau sémantique et la structure argumentative du texte.

Il est probable qu'elle pourra à moyen terme s'aider utilement de l'analyse du discours et de l'argumentation, ainsi que des travaux en intelligence artificielle du type de ceux qu'élaborent, à l'aide du logiciel DEREDEC, le Groupe de recherche en analyse du discours de Grenoble et le Centre de sémiologie de Neuchâtel. Il s'agit bien de travailler, non seulement sur l'objet explicite d'étude, mais aussi sur le mode de questionnement lié au type de compréhension recherché.

Il faut remarquer qu'en SHS pour une même discipline l'hétérogénéité de ces démarches est grande suivant le type d'ancrage social de la recherche et cette caractérisation des démarches de recherche peut être d'une grande utilité pour une banque d'information. On peut en effet penser que lorsque le chercheur pose une question au système pour connaître « les travaux proches des siens », ce qui l'intéresse, ce n'est pas forcément de retrouver un discours identique au sien de façon explicite (même objet empirique, même méthodologie), mais des procédures, des modèles de démarches analogues, éventuellement dans des champs très éloignés du sien.

Il semble nécessaire à court terme de préciser les fonctions argumentatives ou descriptives de la démarche scientifique suivie, tel-

les qu'elles sont remplies par les différentes parties de l'article ou du résumé : articulation du discours, insertion des hypothèses, références aux théories discutées, exposé des résultats, etc. (25, 26). Cette analyse du discours faciliterait ainsi efficacement la fonction d'exploration et de sélection au moment de l'interrogation.

Un mode possible d'étude des critères spontanés de sélection utilisés par le chercheur serait d'analyser la façon dont ce dernier classe les références obtenues par l'IRS comme ayant un rapport avec sa recherche, et de faire expliciter quel type de rapport est perçu avant, mais surtout après lecture du document.

En résumé, on peut dire qu'un document scientifique peut être caractérisé par sa source, (éventuellement sa cible), son canal de diffusion (type de périodique par exemple) et par son contenu défini en termes de problématique, de méthodologie et cadre théorique.

C'est là que tout le travail reste à faire pour définir les dimensions de description adéquates de ces problématiques, cadres conceptuels, et procédures méthodologiques, qui caractérisent les démarches scientifiques en sciences humaines et sociales.

Pour être opératoire, la description du document doit passer d'une description unidimensionnelle (liste de descripteurs différenciés quant à leur fonction descriptive) à une description multidimensionnelle fonctionnelle (18). Le document est alors caractérisé par sa situation sur un ensemble de dimensions non cognitives, par une certaine structure conceptuelle et des caractéristiques procédurales et argumentatives de façon à autoriser les tris adéquats selon la demande.

Dans un tout autre domaine, la recherche menée par A.M. Pejtersen et J. Austin (21) pour adapter les dimensions de description de la littérature de fiction aux critères de choix utilisés par les usagers de bibliothèque, nous paraît de ce point de vue intéressante : le système classificatoire multidimensionnel est fondé sur l'analyse de la formulation des besoins par les usagers et prend notamment en compte les thèmes de l'ouvrage, le cadre temporel et géographique ou social, l'expérience émotionnelle recherchée, les intentions de l'auteur, la lisibilité.

3. " *Classification is a conceptual act. What is classified are not the individual questions but kinds of questions... A classification scheme serves as one analytical tool for understanding inquiry at three levels: the individual study, a corpus of studies and the entire enterprise of research in a given field... Schemes of different sorts can serve to instruct the students about the nature, kinds and functions of research questions. They can help him to perceive of the range of questions that might be asked, to identify the question for research... to situate its place in relation to other questions studied in the domain. They can help him to conceive of the domain for study and to organize the review of research.* »

Dernière remarque : vue l'hétérogénéité des méthodes d'étude suivant les disciplines, il nous semble qu'une telle démarche de redéfinition du traitement implique une approche disciplinaire au sens traditionnel du terme. Autant l'idée que les documents scientifiques peuvent être, quel que soit leur champ, passés dans un moule unique de traitement documentaire paraît en gros valable si l'on parle d'élaboration de catalogues, autant elle nous semble être à l'origine de l'inefficacité des BdB lorsqu'elles visent un objectif d'information scientifique.

Conclusion

Les BdB de masse ont acquis au cours du temps et du fait de leur caractère cumulatif une fonction qui n'est plus strictement bibliographique, mais d'information sur les développements scientifiques.

Les outils en cours d'élaboration que sont les catalogues collectifs (CCN, CCO, revues de sommaires) vont, dans un proche avenir, mordre sur la fonction bibliographique traditionnelle des BdB. De ce fait, il semble nécessaire de redéfinir entre ces deux ensembles des fonctions de complémentarité, où les BdB deviendraient à part entière des banques d'information scientifique, porteuses d'information finale.

Les BdB de masse en sciences sociales ne servent bien pour l'instant ni le public des chercheurs ni le public des non spécialistes. Une meilleure adéquation de l'offre et de la demande supposerait un ensemble de recherches pluridisciplinaires. Ces recherches s'articuleraient autour de deux pôles.

L'analyse des demandes, devant aider à l'élaboration de modèles de l'utilisateur, appellerait :

- une étude sociologique des comportements de recherche d'information selon les groupes professionnels et les modalités institutionnelles de l'activité de recherche ;
- une étude pragmatique de l'action de recherche d'information suivant les phases de l'activité professionnelle ou de recherche, les objectifs poursuivis et les types de questionnement correspondants.

Un examen de l'offre appelant :

- une redéfinition de la méthodologie de traitement, apte à mettre en rapport la grammaire de texte et les dimensions de description

de son contenu avec les différents types de questionnement ;

- cette redéfinition du traitement appelle elle-même une étude socio-ergonomique de la production des banques d'information ;
- l'analyse de l'offre et de la demande appelant une étude économique des produits et marchés des informations scientifiques suivant leurs canaux de diffusion.

Septembre 1986

RÉFÉRENCES

1. **Belkin N.J., Oddy R.N., Brooks H.M.**, « Ask for information retrieval », *The Journal of documentation*, vol. 38, n° 2 et 3, 1982, p. 61-71 et 145-164.
2. **Bastide A., Malrieu D.**, *Bilan de la documentation en psychologie : nécessité de la définition de choix sectoriels et propositions de développement, rapport*, 1985.

3. **Belkin N.J., Seeger T., Wersig G.**, « Distributed expert problem treatment as a model for information system analysis and design », *The Journal of information science*, vol. 5, 1983, p. 153-159.
4. **Blember V.L. & Leggate P.**, « Linking a medical user survey to management for library effectiveness: 1. The user survey », *The Journal of documentation*, vol. 41, n° 1, 1985, p. 1-14.
5. **Bornes C.**, *Transfert de l'information scientifique dans un milieu de recherche*, MIDIST, La Documentation française, 1980.
6. **Brooks H.M., Daniels P.J., Belkin N.J.**, « Using problem structures for driving human-computer dialogs », *RIA O* 1985, p. 645-660.
7. **Chamboredon J.C.**, « Les difficultés de l'évaluation scientifique » *Revue française de sociologie*, vol. 22, 1981, p. 291-295.
8. **Callon M., Courtial J.P., Turner W.A., Bauin S.**, « From translations to problematic networks: an introduction to co-word analysis », *Social science information*, vol. 22, n° 2, 1983, p. 191-235.

9. _____
Courtois G., *Le Monde-Campus*, 10 avril 1986.
10. _____
Crooks G., « Towards a validated analysis of scientific text structure », *Applied linguistics*, vol. 7, n° 1, 1986, p. 57-69.
11. _____
Demailly A., « La Recherche et la communication scientifique à la lumière du modèle de la rationalité limitée de H.A. Simon », *Bulletin des bibliothèques de France*, vol. 26, n° 7, 1981, p. 410-418.
12. _____
Dillon J.T., « The classification of research questions », *Review of educational research*, vol. 54, n° 3, 1984, p. 327-361.
13. _____
Dolman H., Bodewitz H., « Sedimentation of a scientific concept: the use of citation data », *Social studies of science*, vol. 15, 1985, p. 507-527.
14. _____
Dupuy H., « Le B.A. BA des BD: deux ans d'interrogation des banques de données françaises à la bibliothèque publique d'information », *Bulletin des bibliothèques de France*, t. 28, n° 5, 1983, p. 468-472.
15. _____
Hapeshi K., Belkin N.J., *Developing and evaluating an automatic on-line information retrieval system based on user problem statements*, RIAO, 1985.
16. _____
Grinder R.E., « Isolationism in adolescent research », *Human development*, vol. 25, 1982, p. 223-232.
17. _____
Kemp D.A., « Relevance, pertinence and information development », *Information storage and retrieval*, vol. 10, n° 2, 1974, p. 37-47.
18. _____
Leplat J., « Les représentations fonctionnelles dans le travail », *Psychologie française*, vol. 30, n° 3-4, 1985.
19. _____
Mc Aleese R., « Some problems of knowledge representation in an authoring environment: exteriorisation, anomalous state, meta-cognition and self-confrontation », *Programmed learning and educational technology*, vol. 22, n° 4, 1985, p. 299-306.
20. _____
Martin T.H., Carlisle J., Treu S., « The user interface for interactive bibliographic searching: an analysis of the attitudes of nineteen information scientists », *Journal of the American society for information science*, March-April 1973, p. 142-147.
21. _____
Pejtersen A.M., Austin J., « Fiction retrieval: experimental design and evaluation of a search system based on user's values criteria », *The Journal of documentation*, vol. 39, n° 4, 1983, p. 230-246.
22. _____
Poulle F., « Les écrits en santé mentale: la littérature grise », Rapport Mire, 1986.
23. _____
Regazzi J.J., « Evaluating indexing systems: a review after Cranfield », *The Indexer*, vol. 12, n° 1, 1980.
24. _____
Remy D., « La recherche documentaire automatisée à la bibliothèque d'Orsay: bilan de trois années de fonctionnement, 1977-1979 », *Bulletin des bibliothèques de France*, t. 25, n° 7, 1980, p. 341-346.
25. _____
Rouault J., « Méthodes linguistiques en informatique documentaire », *Informatics 8: advances in intelligent retrieval*, Oxford, 16-17 avril 1985.
26. _____
Swales J., « Citation analysis and discourse analysis », *Applied linguistics*, vol. 7, n° 1, 1986, p. 39-56.
27. _____
Vickery A., « An intelligent interface for on-line interaction », *Journal of information science*, vol. 9, n° 1, 1984, p. 7-18.
28. _____
Wersig G., Windel G., « Information science needs a theory of "information actions" », *Social science information studies*, vol. 5, 1985, p. 11-23.