

DISSEMINATION SELECTIVE DE L'INFORMATION ET
RECHERCHE RETROSPECTIVE: EN DIRECT OU EN
DIFFERE. PERSPECTIVES D'AVENIR. (SDI AND
RETROSPECTIVE SEARCHING: ON-LINE OR BATCH.
PERSPECTIVE FOR THE FUTURE)

Jean-Yves Fortin, Chef du service CAN/SDI
Institut canadien de l'information scientifique et technique
Ottawa, Ontario, K1A 0S2

RESUME

Il y a deux types de recherche bibliographique: la recherche rétrospective et la recherche SDI (Dissémination sélective de l'information). Le présent document montre l'évolution de ces deux modes de recherche, leur complémentarité en évoquant leurs différences, leur implantation sur des systèmes d'ordinateur, et conclut sur les perspectives d'avenir concernant la recherche SDI. (Two forms of bibliographic searches exist: retrospective and SDI. This paper describes their beginning, their differences, their implementation in computer systems and the future of SDI.)

LA MASSE DOCUMENTAIRE ET SES FILTRES

De nos jours, la recherche, particulièrement dans les domaines scientifiques, se traduit par une courbe exponentielle de sa masse d'information disponible (GOTTSCALK, 1963; ANDERLA, 1973). Ceci a même amené certains savants à proposer un moratoire sur les publications scientifiques (MAUERHOFF, 1974).

Si le flot d'information n'est pas encore endigué, on a vu du moins surgir des instruments pour en contrôler l'écoulement. Il s'est formé tout un réseau de "filtres" (GARFIELD, 1976), en commençant par l'éditeur de périodique qui accepte ou rejette un article proposé pour fin de publication. Si on considère les articles de périodiques comme une source primaire de l'information, on voit que l'accroissement de ces sources primaires a créé la nécessité de sources secondaires, à savoir les répertoires d'analyses et d'indexation. Un autre groupe de filtres exercent ici encore leur fonction en sélectionnant les articles à analyser et à indexer.

DSI ET RECHERCHE RETROSPECTIVE

LA RECHERCHE RETROSPECTIVE ET LA RECHERCHE DSI: FILTRES DE L'USAGER

L'introduction de l'ordinateur dans le processus de photocomposition des publications des sources secondaires a permis d'en arriver à une troisième étape dans le contrôle de l'information: la recherche DSI et la recherche rétrospective. Si les bibliothèques ont eu jusqu'à un certain point le rôle de disséminer l'information, il est intéressant de noter que c'est à l'extérieur de ce milieu que la fonction "dissémination par ordinateur" s'est développée (MAUERHOFF, 1974). Luhn (1958), considéré comme le point de départ de la dissémination sélective de l'information, a oeuvré en dehors des cercles de bibliothécaires.

Dans cette troisième étape, l'utilisateur devient lui-même filtre de l'information. Il détermine les critères d'acceptabilité des articles qu'il désire recevoir. Deux options s'offrent à lui: la recherche rétrospective et la recherche DSI. Il serait bon de définir ces termes et les définitions proposées sont celles normalement acceptées. La recherche rétrospective est une recherche UNIQUE de l'ENSEMBLE DE LA LITERATURE sur une PERIODE DE TEMPS pré-déterminée. La recherche DSI est une recherche PERIODIQUE d'une PARTIE DE LA LITERATURE, celle qui VIENT DE PARAÎTRE.

COMPLEMENTARITE DE LA RECHERCHE RETROSPECTIVE ET DSI

Banques d'information bibliographique

Attardons nous sur ces définitions, puisqu'elles dévoilent la complémentarité des modes de recherche. La littérature qui nous occupe est celle apparaissant sous la forme de sources secondaires (c.-à-d. Chemical Abstracts, BIOSIS Previews, PASCAL, etc.). Pour la recherche rétrospective, on constitue des banques d'information bibliographique ou de larges fichiers bibliographiques ordinaux à partir des données reçues sur bandes magnétiques. Ces fichiers sont appelés à grossir et certains atteignent actuellement 1,5 à 2 millions de références bibliographiques.

Pour la recherche DSI, le nombre de références atteint rarement plus de 10 000 références. Le rapport entre les deux fichiers se situe donc entre 1/150 et 1/200. Cet ordre de grandeur a un impact sur la formulation des stratégies de recherche. On ne peut vraisemblablement utiliser la même stratégie de recherche pour une recherche rétrospective que pour une recherche DSI comme on le verra plus loin.

Le nombre de références du fichier DSI est relativement stable et réfèrent généralement à des documents qui viennent de paraître.

DSI ET RECHERCHE RETROSPECTIVE

C'est donc dire que le fichier est en constante évolution puisque les documents n'y apparaissent pas deux fois et qu'il est témoin de l'activité actuelle dans les domaines du savoir. Ce fichier qui paraît de façon mensuelle, bimensuelle, ou hebdomadaire est ensuite intégré au fichier rétrospectif. Il constitue alors un infime accroissement par rapport à l'ensemble du fichier rétrospectif déjà existant.

Le facteur temps

Une deuxième distinction est la période de temps couverte. Le fichier rétrospectif est l'ensemble de ce qui s'est publié depuis l'an X jusqu'à nos jours. Si l'an X peut être fixé, "jusqu'à nos jours" est modifié périodiquement. Le fichier DSI est la partie "jusqu'à nos jours" prise pour elle-même et qui est utilisée avant son intégration au fichier rétrospectif.

Le facteur recherche

Une troisième distinction existe dans le type de recherche effectuée. La recherche rétrospective est une recherche unique. Cette dernière constitue une étape dans le besoin d'information de l'utilisateur. Il serait illogique pour le même utilisateur de répéter maintes fois la même recherche. Par contre, la recherche DSI, est une recherche périodique. L'utilisateur ne veut plus savoir ce qui s'est fait, mais ce qui se fait. Un élément nouveau s'insère, à savoir la continuité.

Le profil d'intérêt

Cet élément de continuité que l'on retrouve dans la recherche DSI nous amène à reconnaître une quatrième distinction: la nécessité du profil d'intérêt. Il n'y a aucun besoin de conserver de quelques façons que ce soit, les demandes de recherches rétrospectives. Chaque demande est unique. Il en découle qu'il existe en recherche rétrospective, une variété de questions uniques. Par contre, une recherche DSI doit être renouvelée périodiquement. Une stabilité dans le sujet de recherche est requise, d'où la nécessité d'en conserver les éléments constitutifs dans un profil d'intérêt. Il faut avoir une connaissance des besoins de ou des utilisateurs. Le chercheur qui lit périodiquement quinze revues a, par exemple, son profil d'intérêt en tête.

Comme on a pu le constater, il y a complémentarité lorsqu'on distingue les deux formes de recherche. Au point de vue du contenu, l'interrogation d'un fichier rétrospectif est une suite logique de l'utilisation du fichier DSI. Au point de vue de l'utilisation, c'est le contraire: la recherche rétrospective précède la recherche DSI.

L'ORDINATEUR ET SES MODES D'OPERATION

La recherche bibliographique existe depuis fort longtemps et ce

DSI ET RECHERCHE RETROSPECTIVE

n'est que depuis quinze à vingt ans seulement que l'ordinateur a fait son apparition dans ce domaine. On s'aperçoit, en outre, que la recherche DSI a été la première à être envisagée. Ceci s'explique par le genre de recherche et par le genre d'ordinateur en opération. On peut distinguer deux modes d'opération de l'ordinateur: opération en direct et opération en différé.

Dans une opération en direct, l'utilisateur à l'aide d'un terminal, dirige le travail accompli par l'ordinateur. Il exerce un contrôle complet sur les étapes à suivre ce qui lui permet de modifier à tous moments le cours des événements.

Dans une opération en différé, l'utilisateur, après avoir pré-déterminé les étapes à suivre, confie à l'ordinateur le contrôle des opérations. Aucune intervention n'est possible et c'est seulement sur réception des résultats que l'utilisateur pourra modifier sa stratégie et recommencer. Ces deux modes d'opérations peuvent s'appliquer aux deux modes de recherche.

RECHERCHE	MODE D'OPERATION	
	DIFFERE	DIRECT
DSI	DIFFERE	DIRECT
RETROSPECTIVE	DIFFERE	DIRECT

Les ordinateurs qu'on a conçus pour le travail en différé n'avaient pas, du moins au début, la possibilité d'opérer en direct. Cette fonction a été et est encore une fonction surimposée. Par contre, les ordinateurs conçus pour le travail en direct ont aussi la possibilité d'opérer en différé.

RECHERCHE RETROSPECTIVE EN DIFFERE

Les premières applications de recherche bibliographique ont été faites sur des ordinateurs opérant en différé. En d'autres termes, on définissait la ou les questions que l'on soumettait à l'ordinateur et on analysait ensuite les résultats. Il n'y avait aucune interaction possible.

Si on examine l'impact de ce mode d'opération sur la recherche rétrospective on voit la nécessité de bien définir, au point de départ, la ou les questions à soumettre à l'ordinateur. On ne peut en aucun temps modifier la stratégie de recherche lorsque l'ordinateur est en opération. Si l'on soumet chaque question séparément, on s'aperçoit que les coûts d'opérations sont prohibitifs. De par son fonctionnement, l'ordinateur opérant en différé devient rentable lorsqu'on lui fait faire une longue série d'opérations répétitives. Pour augmenter la rentabilité, une meilleure approche est donc de soumettre plusieurs questions à la fois. Or en recherche rétrospective le nombre de questions va varier grandement d'un jour à l'autre, d'une semaine à l'autre, d'un mois à l'autre. Il faut en quelque sorte "différer" la recherche pour

DSI ET RECHERCHE RETROSPECTIVE

accumuler un nombre suffisant de questions permettant de rentabiliser l'utilisation de l'ordinateur.

D'un autre côté, étant donné que le fichier bibliographique grossit périodiquement, il faut comparer un plus grand nombre de documents chaque fois qu'une ou plusieurs questions sont posées. Si un grand nombre de documents permet une meilleure rentabilisation, une augmentation du fichier par rapport au nombre de questions posées crée une saturation augmentant les coûts d'opération plutôt que les réduisant. En effet, on considère comme normal dans un système en différé la lecture séquentielle des documents. En d'autres termes, pour lire la 800ième référence, il faut lire les 799 premières. Or on considère actuellement des fichiers rétrospectifs comprenant plus d'un million de documents. L'ordinateur est rapide, mais il a des limites.

Les solutions proposées par Thiel et Heaps (THIEL, 1972) pour améliorer la construction des fichiers et la formulation des questions, éminemment valables avant l'avènement de la recherche rétrospective en direct, semblent avoir perdu aujourd'hui leur valeur. Au fur et à mesure que le fichier grossira, il faudra être de plus en plus sélectif et formuler des questions de plus en plus précises, si on ne veut pas être submergé par un déluge de références.

Skov (1973) en conclut que l'opération en différé n'est pas la meilleure approche pour la recherche rétrospective, dû aux contraintes imposées. Wilman (1973) après un essai de recherche rétrospective en différé, en arrive aux mêmes conclusions.

RECHERCHE DSI EN DIFFERE

En recherche DSI, on ne retrouve pas une série de questions uniques mais bien une base stable de questions qui doivent être posées périodiquement. On retrouve ici un élément essentiel du fonctionnement en différé: la répétition. Malgré que la mise à jour peut-être fréquente au niveau du profil d'intérêt, l'ensemble des profils constitue une masse relativement stable. L'augmentation du nombre de profils vient même grossir la masse de questions nécessaires au bon fonctionnement d'un ordinateur en différé.

Le fichier bibliographique en recherche DSI est relativement stable et petit, ce qui signifie qu'il se prête facilement à une lecture séquentielle. Ceci n'entraîne donc aucune saturation des opérations et partant aucune augmentation des coûts. Il en résulte que la recherche DSI est à toute fin pratique, une opération en différé.

RECHERCHE RETROSPECTIVE EN DIRECT

L'avènement de la recherche en direct n'est pas dû à une nécessité de l'interaction mais bien à une réduction dans les coûts d'opération des

DSI ET RECHERCHE RETROSPECTIVE

ordinateurs et de la disponibilité de réseaux de télécommunication efficaces (McCARN, 1974). La recherche rétrospective en a tout simplement profité. L'existence de disques de grande capacité a permis de créer des fichiers bibliographiques qui sont, par définition, immenses. Le coût de revient a tellement baissé qu'on peut envisager la disponibilité de ces fichiers en permanence. Un autre aspect de la recherche en opération directe est l'accès aux fichiers. Pour lire, par exemple, la 800ième référence, il n'est pas nécessaire de lire les 799 premières sur un fichier en accès direct. Il en résulte une utilisation beaucoup plus rationnelle des ressources de l'ordinateur.

L'interaction permet de mieux planifier une stratégie de recherche et de dévier de la stratégie établie, si en cours de route, on n'a pas les résultats escomptés. La difficulté d'accumuler un grand nombre de questions pour l'opération en différé, se trouve remplacée par la nécessité d'avoir un nombre d'utilisateurs assez élevé pour qu'il y ait une utilisation constante de l'ordinateur. Les coûts de télécommunication décroissent rapidement et permettront de rejoindre de plus en plus d'utilisateurs. L'avènement de DATAPAC est un exemple d'une rationalisation et d'une bonne gestion des ressources de télécommunication.

RECHERCHE DSI EN DIRECT

La recherche rétrospective a profité certainement de la nouvelle technologie en direct même s'il y a encore beaucoup d'inconnu, en particulier au niveau de l'interaction homme/machine (ARTANDI, 1975). Or on voudrait en faire profiter la recherche DSI.

Il faudrait d'abord examiner les avantages. Il y a d'abord l'accès direct au profil d'intérêt. Pouvoir modifier le profil en accès direct éviterait le délai inhérent au système actuel (ici je pense particulièrement au système CAN/SDI): c'est-à-dire délai de poste, interprétation des corrections demandées, perforation et vérification.

Le changement serait immédiat. On pourrait en faire l'essai sur le fichier rétrospectif et déterminer la valeur des changements introduits. Mais examinons sérieusement cet avantage, particulièrement dans notre contexte canadien.

COMPOSANTES DE LA RECHERCHE DSI

Contrôle des profils

Actuellement tous les profils et modifications aux profils sont vérifiés par un rédacteur de recherche au centre CAN/SDI. Cette vérification permet de s'assurer que le profil pourra repérer les documents pertinents et que le taux de bruit ou de documents non pertinents, sera bas.

DSI ET RECHERCHE RETROSPECTIVE

Les centres CAN/SDI, possédant toute la gamme des instruments de travail, peuvent vérifier les profils plus adéquatement que tout usager ou rédacteur de recherche à l'extérieur des centres.

Fonctions des profils

Lorsqu'on parle de vérifier le profil à l'aide du fichier rétrospectif, il faudrait penser au but du profil. Cole (1976) a défini quatre fonctions d'un profil d'intérêt:

- 1) La mise à jour continue sur un sujet donné
- 2) La prise de décision
- 3) L'interprétation et l'assimilation de l'information repérée
- 4) La satisfaction de la curiosité scientifique

Si la première fonction est vérifiable par une recherche rétrospective, il n'en va pas de même des trois autres. La formulation de question d'un profil d'intérêt DSI, admet une certaine tolérance au bruit et aux documents périphériques au sujet principal de recherche. Le fait de recevoir cette bibliographie par tranche, ce qui est de la nature périodique du DSI, donne à l'utilisateur, l'impression qu'il peut absorber non seulement ce qui se fait dans ce domaine, mais aussi et peut-être surtout ce qui se fait dans les domaines connexes.

La recherche rétrospective impose la nécessité de déterminer le sujet de façon précise, à cause du nombre continuellement grandissant de documents du fichier rétrospectif. La vérification ne doit pas seulement se faire pour les documents spécifiques mais aussi pour les documents périphériques à la recherche. Une recherche rétrospective ne peut répondre à une question générale, dû à la pauvreté, il faut bien le dire, des techniques de repérage des systèmes en accès direct. De ce côté, la recherche DSI particulièrement en ce qui concerne CAN/SDI est encore en avance sur les systèmes de recherche rétrospective. Que l'on pense tout simplement à la troncature à gauche et au système de pondération. La technique de la recherche rétrospective est à l'heure actuelle une copie, et une bien pauvre copie, de la recherche DSI. Une évaluation et une recherche approfondie des techniques de repérage en recherche rétrospective vont bientôt devenir une nécessité.

Rapidité de la recherche DSI

D'un autre côté, la mise à jour du profil d'intérêt en accès direct n'augmente nullement la rapidité de la mise au courant. Par définition, la recherche DSI est périodique. La périodicité ne sera pas accrue parce que le profil est mis à jour en accès direct. Engineering Index, METADEX, GEOREF, pour ne nommer que ceux-ci, demeureront des services mensuels. Là où il y aurait avantage dans la mise à jour du profil en accès direct serait pour des fichiers hebdomadaires. Et encore là, les centres CAN/SDI en particulier, acceptent les modifications par télé-

DSI ET RECHERCHE RETROSPECTIVE

phone ce qui est pratiquement aussi rapide. A la rigueur, on y gagnerait une semaine en modifiant le profil d'intérêt en accès direct.

Lorsqu'on parle de rapidité, il faut la voir non pas du point de vue du pourvoyeur mais bien du récipiendaire. Le plus grand délai à l'heure actuelle est le délai postal et il n'y a aucun lien entre celui-ci et le mode d'opération de l'ordinateur.

Traitement de la recherche DSI

En opération directe, il y aurait deux façons de traiter les recherches DSI. Ou bien on crée un système en différé parallèle, ce qui existe déjà, ou bien on fait les recherches en différé de la même façon que les recherches en direct en y ajoutant l'élément automatique ou répétitif. Dans ce dernier cas, en plus de perdre, du moins actuellement, certaines techniques de recherche, on améliore nullement la périodicité et la rapidité de la recherche DSI. Il en résulte seulement l'illusion d'avoir un plus grand contrôle du profil d'intérêt puisque de l'avis des auteurs (LANCASTER, 1973; WILLIAMS, 1974) la mise à jour directe est le seul avantage d'une recherche DSI. Il faut encore en conclure que la recherche DSI est une opération en différé.

RECHERCHE RETROSPECTIVE ET DSI A L'ICIST

La recherche rétrospective et la recherche DSI sont complémentaires et chacune remplit des fonctions différentes. Il me paraît nécessaire d'avoir une plus grande collaboration dans ce domaine. A l'Institut canadien de l'information scientifique et technique (ICIST) cette collaboration est assez évidente. Depuis novembre 1976 les ateliers CAN/SDI et CAN/OLE ont lieu l'un à la suite de l'autre pour permettre aux usagers de participer aux deux ateliers. Lors d'ateliers spécialisés sur les différentes banques d'information (BIOSIS, INSPEC) l'utilisation des deux services y est discutée.

La recherche DSI a devant elle plusieurs avenues plus ou moins bien exploitées particulièrement concernant les profils de groupe ou les profils standard. On a souvent tendance à voir plutôt les profils d'intérêt comme des profils personnalisés. Or au CAN/SDI, il y a un grand nombre de profils de groupe. On peut les diviser en deux catégories: le profil partagé, rédigé par un seul mais utilisé par plusieurs, et le profil collectif où chaque utilisateur a ses propres questions et reçoit une partie des références repérées.

Un autre aspect qui n'a pas été abordé ici mais qui a son importance pour la recherche bibliographique est le rôle des intermédiaires. En recherche DSI, le rôle des intermédiaires est reconnu depuis longtemps (GARFIELD, 1969). CAN/SDI par son réseau de rédacteurs de recherche à travers le Canada (BROWN, 1970; MAUERHOFF, 1970) est un maillon important de l'éducation des intermédiaires qui apprécient aujourd'hui

DSI ET RECHERCHE RETROSPECTIVE

les avantages de la recherche rétrospective en direct. L'éducation des rédacteurs de recherche ne doit pas se limiter seulement à la "mécanique" de la rédaction des profils (GAFFNEY, 1973), mais doit s'étendre à une connaissance approfondie des banques de données, de l'interview des usagers et du repérage de l'information. Il y a encore beaucoup à faire dans cette voie en vue d'améliorer les recherches DSI.

L'avenir de la recherche DSI est intimement liée à la recherche rétrospective. Il faut voir dans leur complémentarité non pas une lutte de l'une contre l'autre, mais bien une interaction bénéfique pour celui concerné: l'utilisateur.

REFERENCES BIBLIOGRAPHIQUES

- ANDERLA, Georges. Information in 1985. A Forecasting Study of Information Needs and Resources. Paris, Organisation for Economic Cooperation and Development, 1973. 132p.
- ARTANDI, Susan. "On-Line Information Systems in Perspective", dans Journal of Chemical Information and Computer Sciences. Vol. 16, No. 2 (1976), pp. 80-81.
- BROWN, Jack E. "The CAN/SDI Project: Training of Users", dans Users of Documentation. Fédération Internationale de Documentation. Conference and International Congress of Documentation, 35th, Buenos Aires, 21-24 September 1970. (FID 472).
- COLE, Elliot. The Uses of SDI: A Behavioral Approach Applying Concepts from Mass Communications. Texte préliminaire reçu de l'auteur. 1976.
- GAFFNEY, Inez M. "CAN/SDI: Experience with Multi-Source Computer Based Current Awareness Services in the National Science Library, Ottawa", dans Bulletin of the Medical Library Association. Vol. 61, No. 3 (July 1973), pp. 309-313.
- GARFIELD, Eugene. "The Role of the Medical Librarian in SDI Systems", dans Bulletin of the Medical Library Association. Vo. 57, No. 4 (October 1969), pp. 348-351.
- . "Gatekeepers in Science", dans Current Contents. Vol. 16, No. 2 (January 12, 1976), pp. 5-7.
- GOTTSCHALK, Charles M. et DESMOND, Winifred F. "Worldwide Census of Scientific and Technical Serials", dans American Documentation. Vol. 14, No. 3 (July 1963), pp. 188-194

DSI ET RECHERCHE RETROSPECTIVE

- LANCASTER, Frederic Wilfrid et FAYEN, Emily Gallup. Information Retrieval On-Line. A Wiley-Becker & Hayes Series Book. Los Angeles, Melville Publishing Co., 1973. 597p.
- LEGGATE, P. "Progress in Documentation: Computer-Based Current Awareness Services", dans Journal of Documentation. Vol. 31, No. 2 (June 1975), pp. 93-115.
- LUHN, Hans Peter. "A Business Intelligence System", dans IBM Journal of Research and Development. Vo. 2, No. 4 (October 1958), pp. 314-319.
- MAUERHOFF, Georg. "NSL Profiling and Search Editing", dans American Society for Information Science. Western Canada Chapter Annual Meeting, 2nd, Vancouver, B.C., September 14-15, 1970; Proceedings. Information Systems, The University of Calgary, December 1970. pp. 32-53.
- _____. "Selective Dissemination of Information", dans Advances in Librarianship. Vol. 4 (1974), pp. 25-62.
- MCCARN, Davis B. "Trends in Information", dans ZUNDE, Pranas, ed. Information Utilities; Proceedings of the 37th ASIS Annual Meeting, Atlanta, Georgia, 13-17 October 1974. Vol. 11. Washington, D.C., American Society for Information Science, 1974. pp. 145-150.
- SKOV, Helge J. "SDI-Services - Work and Organization", dans International Association of Technological University Libraries. IATUL Conference, 5th, Copenhagen, Denmark, 1973.
- THIEL, L.H. et HEAPS, H.S. "Program Design for Retrospective Searches on Large Data Bases", dans Information Storage and Retrieval. Vol. 9, No. 1 (January 1972), pp. 1-20.
- WILMAN, H. et HALL, Angela M. An Experiment in the Batch Processing of Retrospective Searches. INSPEC Report No. R.73/13. London, The Institution of Electrical Engineers, February 1973. 26 p.