

A PROPOS D'UN TRAVAIL DE BIBLIOGRAPHIE CRITIQUE
SUR L'ESPACE SCOLAIRE :
APPORTS ET DIFFICULTÉS D'USAGE
DES BANQUES DE DONNÉES INTERNATIONALES

par Marie-Claude Derouet-Besson

Le recours aux banques de données informatisées est désormais une démarche courante lorsque l'on aborde un domaine dont on peut supposer qu'il a déjà été traité. Il est indispensable quand l'objet de la recherche est une étude critique du matériel bibliographique international sur un thème donné. Les fichiers traditionnels, en effet, ne peuvent pas dominer la production écrite internationale à moins d'être assez étroitement spécialisés et la collecte, fiche à fiche, des titres est dévoreuse de temps.

Les avantages des banques de données sont évidents mais parfois trompeurs ; leur utilisation révèle des limites... Celles de toute collecte d'informations, de toute classification, mais aussi celles de l'utilisateur confronté à une masse d'éléments souvent aussi impressionnante qu'hétéroclite.

La rapidité est sans doute l'avantage le plus facilement mesurable. Au C.N.R.S., pour le fichier FRANCIS et les fichiers étrangers comme ERIC, deux ou trois heures pour l'examen du thesaurus avec un spécialiste, pour la détermination de la période de parution et des langues utilisées dans les publications recensées, suffisent pour préparer l'interrogation dont se charge le personnel de la banque. La livraison des fiches demande une huitaine de jours. L'économie de temps est énorme ; les fiches imprimées sont facilement lisibles, maniables, leurs références sont complètes, précises ; une courte analyse présente le document (en français pour FRANCIS, en anglais pour ERIC qui est une banque américaine).

Le second avantage réside dans la masse de titres ainsi recueillis ; plus de deux mille, par exemple, pour l'interrogation de FRANCIS et ERIC sur l'espace scolaire. Une telle masse, sans jamais prétendre à l'exhaustivité peut être considérée comme

représentative des grandes orientations et laisser espérer la connaissance des principaux courants de la recherche internationale dans un secteur.

On touche là, toutefois, la première série des limites des banques de données : celles de la collecte des informations. Elle obéit, pour chaque banque, à des règles précises qui ont des conséquences très importantes sur l'exhaustivité et la représentativité des fichiers par rapport à la production qu'ils sont censés couvrir. FRANCIS, par exemple, ne recense que des productions écrites qui ont été publiées. En outre, son recensement est sélectif : les livres qui y figurent ont été sélectionnés par des documentalistes-analystes, certaines revues sont systématiquement dépouillées, d'autres non. Même si l'éventail est déjà large, chacun doit savoir qu'il y a un tri préalable dont il faut tenir compte. Cette sélection est très accentuée pour les ouvrages étrangers et elle interdit, par exemple, d'accorder un crédit quelconque à des études comparant la fréquence de tel ou tel thème entre les publications françaises et étrangères. Elle élimine donc pour la France tous les textes non édités parmi lesquels se trouvent, en particulier, beaucoup de rapports de recherche.

Pour ERIC la perspective est différente puisque la banque engrange non seulement les publications, mais encore d'autres types de textes : des conférences, des discours prononcés lors de colloques, de congrès..., des rapports administratifs locaux ou fédéraux, des études à très faible diffusion émanant d'associations ou d'organismes de recherche, etc. A priori donc tout document écrit aux Etats-Unis ou dans les pays anglophones dans le domaine de l'éducation peut être répertorié par ERIC. Faut-il croire, pour autant, qu'ERIC est exhaustif pour toute la production écrite en anglais ? D'après l'exemple de l'interrogation sur l'espace scolaire, sûrement pas, même si les titres canadiens, britanniques, australiens, néo-zélandais, etc., recensés sont nombreux, les documents en provenance des Etats-Unis dominant et la couverture des productions des pays étrangers reste incomplète.

Le recours aux banques de données ne supprime donc pas les limites propres à toute collecte d'informations, même s'il élargit énormément le corpus, il ne garantit pas l'exhaustivité. L'étude attentive des principes de constitution de chaque fichier permet, cependant, de réduire considérablement, mais jamais complètement, la part due au hasard dans la quête d'informations et de mieux

apprécier la nature et l'ampleur des lacunes éventuelles. En outre, l'interrogation simultanée de plusieurs banques différentes complète l'étendue du fichier constitué. Sur l'espace scolaire, les informations sur les publications anglosaxonnes données par FRANCIS étaient totalement recouvertes par celles d'ERIC, mais elles ne constituaient qu'environ 15 % des titres fournis par FRANCIS et beaucoup moins encore de ceux fournis par ERIC. Ce recouvrement partiel concrétisait pour l'utilisateur les caractéristiques de la composition des fichiers. FRANCIS apparaissait clairement ne recenser qu'un petit nombre de titres anglosaxons, ERIC redonnait, par comparaison, une idée plus juste de l'importance de la production écrite en anglais.

Si cette masse de titres est, en elle-même, rassurante quant à la couverture du thème, elle est aussi terrorisante. Comment parvenir à se repérer dans cette masse, puis à la dominer ? D'autres limites apparaissent alors, celles qui sont liées à la classification et aux résumés synthétiques proposés sur les fiches.

Les banques de données sont d'énormes fichiers-matières organisés le plus souvent autour de "bases" qui correspondent, en général, à des disciplines universitaires telles que la sociologie, les sciences de l'éducation, l'économie, etc. La première difficulté est parfois d'interroger la bonne base ou plutôt de ne pas se limiter à une seule. Si la recherche en éducation privilégie, bien sûr, l'interrogation de la base "sciences de l'éducation", elle doit aussi avoir recours aux base "sociologie" ou "psychologie", par exemple. Ensuite, seuls les mots-clés permettent de repérer les documents traitant d'un sujet. Or, aussi grande que soit la qualité des documentalistes-analystes qui opèrent le classement des titres, elle n'élimine pas les divergences d'interprétation et le choix du mot-clé, de l'étiquette en quelque sorte de l'ouvrage, reste une opération aussi délicate que capitale. C'est de ce choix que dépend l'efficacité des interrogations. Le thesaurus de la banque qui fournit la liste de tous les mots-clés du fichier propose aussi pour chacun d'eux une définition qui aide le documentaliste-analyste dans son classement et l'utilisateur dans sa recherche. Le thesaurus fournit encore pour chaque mot-clé la liste des autres mots-clés assez proches, auxquels on aurait intérêt à se référer. La difficulté de la classification est, en effet, en grande partie maîtrisée par

l'inscription du même titre dans plusieurs bases et sous plusieurs mots-clés. La multiplication des entrées diminue ainsi le risque de passer à côté d'une référence. L'interrogation, c'est-à-dire le choix des mots-clés, reste une opération à préparer avec beaucoup de soin, thesaurus en main, ou, mieux encore, lors d'un entretien avec un documentaliste-analyste. Il sera d'autant plus facile ensuite de se repérer dans la masse de documents que l'interrogation aura cerner le sujet au plus près. C'est donc grâce au thesaurus de la banque et à la connaissance profonde qu'en a le documentaliste-analyste que l'exposé du thème à couvrir est "traduit" en mots à interroger. Toutes les entrées possibles sont répertoriées et tous les renvois utilisés. De proche en proche, le thème est ainsi délimité et l'on peut considérer que l'approche est terminée lorsque, par le jeu des renvois, on retombe sur des entrées déjà retenues. Sans cette préparation minutieuse, de nombreuses fiches inutiles et nuisibles, par leur présence même, encombreront la livraison et feront perdre du temps, ou bien l'utilisateur passera à côté de références essentielles à sa recherche et indispensables à la qualité de sa documentation.

Les résumés synthétiques, présentés par toutes les banques sont très inégaux. Il faut y prendre garde car certaines distorsions conceptuelles sont fréquentes et s'ajoutent aux difficultés, parfois considérables, de lecture de modes de pensée étrangers. Les synthèses oscillent entre un aperçu des tables des matières ou un résumé de conclusions. La qualité, la longueur diverses des documents mis en fiches expliquent sans aucun doute ces hésitations mais l'utilisateur reste souvent incapable, à partir des renseignements fournis, de déterminer si oui ou non, l'ouvrage intéresse le sujet qu'il veut traiter. Cela n'a guère d'importance si le nombre de fiches imprécises est peu élevé, cela peut être beaucoup plus gênant s'il y en a beaucoup ou si les documents ne sont accessibles que par le service de reproduction de la banque (c'est le cas pour tous les textes, non publiés, répertoriés par ERIC). Le temps perdu et les incidences financières ne sont plus alors négligeables. Toutefois, il faut reconnaître que la tâche demandée aux documentalistes-analystes est presque impossible : comme le classement, la synthèse ne peut être aussi fiable et aussi fine que le souhaite l'utilisateur qui souvent veut se situer lui-même à la pointe de la réflexion dans son secteur. Ainsi, sous la rubrique générale de "perception de l'espace", commune à ERIC

et à FRANCIS, on trouve, par exemple, un regroupement artificiel d'ouvrages de psychologie expérimentale, de sciences de l'éducation, de recherche en éducation, mêlant les fonctions de l'espace scolaire (cadre, outil, objet de l'enseignement ou de l'éducation), ses divers utilisateurs, les modes de contrôle de l'espace, juxtaposant des compte-rendus de recherche, des témoignages ou des points de vue sur tout ou partie de la question. C'est le niveau de généralité du mot-clé et du résumé synthétique qui, la plupart du temps, déçoit l'utilisateur.

L'interrogation des banques de données change, surtout, l'ampleur de l'information mise au service de son utilisateur, elle modifie aussi son approche du matériel bibliographique : pour être plus complète, elle est également plus déroutante parce que plus dispersée. Cependant, et c'est l'essentiel, l'organisation conceptuelle de la masse d'informations recueillies reste toujours à la charge du chercheur. Les entrées thématiques, absolument indispensables, puisqu'elles structurent la banque et permettent son interrogation, ne doivent pas être confondues avec des concepts. Quand les ordinateurs ont trié et imprimé les fiches, soigneusement classées par base et mot-clé, tout le travail reste à faire.

En premier lieu, il faut éliminer toutes les références dont les titres se révèlent trompeurs par rapport au sujet. cette élimination est complexe car elle ne peut pas toujours se faire à partir des fiches elles-mêmes. Lorsque la détermination du contenu n'est pas facile à partir du résumé, il ne reste guère qu'une seule solution : lire le document ! Seule la lecture permet de trancher sans remords... mais le piège du gigantisme se referme bien vite sur l'utilisateur car la masse est telle que la tâche est vite écrasante. La définition des critères d'élimination, malgré son caractère ingrat, est importante, car elle définit aussi, en creux, les critères de l'approche positive, les bases de la constitution du corpus de l'étude. Elle permet, autant que les choix des ouvrages retenus, de réfléchir aux articulations du thème avec ce qui se trouve sur ses marges. Elle est l'occasion de remettre en chantier, à de nombreuses reprises, la délimitation du sujet traité, ce qui favorise sa conception et sa conceptualisation.

La masse des documents peut par elle-même tendre un autre piège : celui de faire croire qu'elle a un sens en elle-même. Or

le nombre de fiches sur tel ou tel thème n'est guère qu'un indicateur de fréquence ou de popularité ; il ne traduit pas automatiquement l'importance d'une question ni même d'un courant de recherche. Il peut montrer, par exemple, la domination d'un type de réflexion mais cela ne le justifie pas, ne le légitime pas pour autant. L'étude du matériel bibliographique sur l'espace scolaire montre, par exemple, que les problématiques qui procèdent par isolement des variables sont les plus répandues. Cela ne signifie cependant pas qu'elles soient forcément les plus appropriées pour déterminer la nature des relations que les usagers entretiennent avec l'espace scolaire. Ce qu'apporte, dans ce cas, l'interrogation des banques de données, c'est la preuve, par la masse de documents qu'elle soumet à la réflexion, que cette problématique aboutit à des résultats très différents, parfois même contradictoires. Elle souligne aussi, en rassemblant quelques cas que la recherche aléatoire dans les fichiers traditionnels et nationaux n'aurait sans doute pas permis de regrouper, qu'une rupture épistémologique s'amorce actuellement entre ces procédures d'isolement des variables et une approche écologique de l'espace scolaire qui tente de prendre tout en compte à la fois.

Un autre piège encore des banques de données réside dans l'aspect bien classé de la documentation qu'elles fournissent. Même si elle n'est jamais parfaite, la classification thématique offre une apparence rationnelle toujours rassurante lorsque l'on est face à une masse de documents. Le danger est de croire que cette organisation du matériel bibliographique recouvre, se confond ou amorce l'organisation conceptuelle du sujet de recherche. Toute classification est, par nature, insuffisante puisqu'elle est réductrice : elle cloisonne, elle étanchéifie en catégories, différentes approches d'un même objet. Face à la classification proposée par la banque, il est indispensable de tenter d'en créer d'autres qui, en se recoupant, restitueront sa dimension à l'objet. C'est donc grâce à une lecture attentive du matériel bibliographique que les intelligibilités d'un même thème seront peu à peu dégagées. Pour l'espace scolaire, par exemple, la classification par mots-clés ("perception de l'espace", "construction", "motricité", "latéralité", "espace personnel", "architecture", "école ouverte", "bâtiment", "environnement", "équipement") a été doublée par de nombreux autres modes d'approche tels que la différenciation des types d'utilisateurs de l'espace scolaire,

des rapports entre l'espace et l'institution scolaire, du discours sur l'école et des pratiques sociales. A chaque fois, le matériel bibliographique interrogé livre ses richesses ou ses lacunes et pose d'autres questions qui font rebondir la réflexion.

Le thème, le sujet choisis pour l'étude bibliographique changent donc, peu à peu, de nature et l'objet de la recherche se construit, en se nourrissant des apports mais aussi des échecs, des interrogations fournis par le matériel bibliographique. C'est à travers ce jeu de propositions et de contre-propositions que l'on parvient à dégager une organisation conceptuelle de la masse des ouvrages. C'est à travers lui que l'on peut passer, pour l'espace scolaire par exemple, d'une description de l'espace à une réflexion sur la sémiologie de l'espace puis à la construction d'un objet de recherche : les pratiques d'espace et au dégagement de quelques pistes pour son étude à venir. L'étude bibliographique a alors joué son rôle, elle a contribué efficacement à la détermination de l'objet de recherche, elle ne s'est pas limitée à une illustration thématique.

L'un des problèmes rencontrés à l'heure actuelle par la recherche en éducation est celui de la capitalisation des résultats. Pour pouvoir progresser en s'appuyant sur ce qui a déjà été fait, il faut capitaliser les recherches antérieures et atteindre une masse critique d'informations qui seule permet de dominer l'état de la recherche et d'avancer sur des bases nouvelles. Faute de cette capitalisation, on risque de reprendre les mêmes canevas, de refaire les mêmes recherches.

Les banques de données sont un instrument pour parvenir à un stade avancé de capitalisation de la recherche... mais leur facilité d'utilisation rend le chercheur fragile face à deux dangers. La première menace est celle de l'alignement sur les recherches déjà existantes. La masse rassemblée peut, en effet, encourager la continuation de recherches sur des sentiers déjà bien balisés. Dans ce cas, la masse attire vers la science hypernormale, elle piège le chercheur car il s'aligne sur elle au lieu de la maîtriser. La masse bibliographique doit servir à capitaliser les résultats antérieurs, à les critiquer, pour aller plus loin, pour aller ailleurs, selon d'autres voies, pour faire autre chose. La seconde menace est celle de la désinformation par la surinformation. L'ampleur de l'information aboutit à un émiettement, à une impossibilité de dominer facilement l'ensemble

qui décourage vite le plus endurci. La tentation est alors de tout laisser de côté et de repartir à zéro sur des bases que chacun croit neuves parce qu'elles sont siennes. Dans l'immense majorité des cas, cela aboutit à la répétition des recherches, à la reprise de pistes, intéressantes, mais déjà explorées. L'une et l'autre menaces, l'une et l'autre tentations, leurrent le chercheur, l'attirent vers le déjà vu, le déjà fait.

L'utilisation de banques de données permet pourtant d'atteindre un niveau critique susceptible de faire de cette masse d'informations un stimulant pour l'esprit. C'est l'étude, l'analyse de cette masse énorme, informe, disparate, qui permet de séparer les genres de production. Certains ouvrages énoncent des thèses, appuyées ou non sur des expérimentations, d'autres racontent des innovations, témoignent de pratiques, d'autres encore n'ont qu'une fonction idéologique, ne sont que discours sur les réalités plus ou moins bien déterminées. Seul un regard extérieur peut donner sens à cet amoncellement, en redonnant à chaque ouvrage sa place sur l'échiquier des connaissances. La recherche des paradigmes du discours sur l'espace scolaire fait apparaître, par exemple, deux tendances organisatrices. Certains structurent l'espace de l'école comme un monde en réduction, un microcosme qui doit reproduire le monde en le simplifiant pour le rendre accessible aux élèves. D'autres conçoivent l'espace de l'école comme devant être totalement perméable au monde pour être en prise directe sur lui. De la même façon, le point de vue que chaque ouvrage adopte vis-à-vis de l'espace scolaire est, selon les cas, globalisant ou particulariste. L'espace scolaire est vu, conçu, selon une approche générale, répétitive et structurelle ou selon l'angle, plus aigu, d'une pratique quotidienne. Là encore, seule la masse de documents a permis de mettre au jour ces approches, qui ne sont pas maîtrisées par les auteurs des ouvrages et de proposer un nouvel objet de recherches : les pratiques d'espace.

La difficulté de l'étude bibliographique est incontestablement augmentée par le recours aux banques de données informatisées qui augmentent le champ de documentation jusqu'à l'insupportable. Pourtant, cette extension est la seule voie possible d'utilisation capitalisatrice et constructive du matériel bibliographique international.