

LA BANQUE DE DONNÉES D'HISTOIRE LITTÉRAIRE

Michel BERNARD

Vous voulez savoir qui a écrit *Les Deux timides* ou *L'École des Maris* ? Rien de plus simple : une encyclopédie sérieuse, une histoire de la littérature munie d'un index ou le *Dictionnaire des œuvres* de Laffont et Bompiani pourront vous indiquer les noms de Labiche et de La Chaussée.

Et si vous vouliez savoir quelle était la profession de Labiche ou le nom du premier éditeur de *L'École des Maris* ? Avouez que la recherche devient plus ardue...

Plus encore : vous voulez savoir la profession des auteurs qui, comme La Chaussée, ont écrit des comédies en vers entre 1720 et 1750. Ou vous voulez savoir quels étaient les écrivains qui auraient pu assister à la représentation des *Deux Timides* en 1860. Cette fois-ci, aucun livre ne pourra vous renseigner. Il n'y a que la Banque de Données d'Histoire Littéraire de l'Université de Paris III (Sorbonne Nouvelle) qui puisse vous y aider, à savoir une banque de données sur l'histoire de la littérature française, créée sous l'impulsion du professeur M. Béhar, avec l'aide de M.-D. Fragonard, de J.-P. Goldenstein et d'A. Viala, et avec l'assistance d'un ingénieur en informatique, M. Sanson. Cette banque toutefois, n'existe encore qu'à l'état de maquette prototype, et n'est pas encore diffusée.

1 -CE QUE L'ON TROUVE DANS LA BANQUE DE DONNÉES D'HISTOIRE LITTÉRAIRE

La BDHL, comme toute banque de données, se présente tout d'abord comme une collection d'indications liées entre elles par des relations logiques. Elle est composée de plusieurs listes : liste de 500 auteurs (date et lieu de naissance, date et lieu de décès, profession), liste de 2 000 œuvres (auteur, date de première publication, genre, éditeur), liste des académies et salons, liste des groupes et mouvements littéraires

LE BULLETIN DE L'EPI N° 51 LA BANQUE DE DONNÉES D'HISTOIRE LITTÉRAIRE

(dates, membres), liste des bibliothèques (date de création), liste des lauréats de prix littéraires, liste des auteurs figurant dans les programmes scolaires, liste des œuvres censurées, liste des grandes dates de l'histoire de France.

Il est à noter que ces listes, par elles-mêmes, constituent une mine de renseignements que l'on trouve difficilement ailleurs. L'indication du genre des œuvres littéraires, en particulier, suivant une codification à trois niveaux, constitue une innovation. On précise d'abord le genre du texte : prose, vers ou mixte. Au second niveau, on indique s'il s'agit d'une œuvre de fiction, de théâtre, de poésie... Au troisième niveau de codage, on donne le genre précis, roman, comédie, tragédie, nouvelle, journal, histoire... Cette méthode, très souple, permet à l'utilisateur et à l'indexateur de s'accommoder des poèmes en prose, des comédies en prose ou en vers, etc.

2 - FONCTIONNEMENT DE LA BDHL

Conçue initialement avec des visées pédagogiques, la BDHL devait pouvoir tenir sur deux disquettes de 360 Ko au format IBM pour être utilisée dans les établissements dotés du matériel IPT. Bien vite, les concepteurs se sont rendu compte qu'elle avait besoin de beaucoup plus de place pour pouvoir contenir toutes les données qui la rendraient réellement utilisable. La BDHL est actuellement implantée sur un compatible IBM doté d'un disque dur.

Au plan du logiciel, les informations de la BDHL sont gérées par Multilog, un produit français agréé par l'Éducation nationale. Il ne s'agit pas du meilleur logiciel du marché mais il a permis de mettre en place une série de menus arborescents qui proposent à l'utilisateur la grande variété des recherches possibles. On ne peut reprocher à Multilog qu'une certaine lourdeur dans les procédures et son manque de souplesse.

L'utilisateur, en suivant les menus proposés par les auteurs, peut opérer des recherches sur tous les fichiers, y compris les croisements, les tris et les statistiques. Les choix proposés sont arborescents et constituent des trajectoires, imposées mais très variées après avoir indiqué le nom d'un auteur, on pourra accéder à des informations chronologiques, à la liste des auteurs appartenant à la même classe d'âge, à l'indication de son appartenance à des groupes organisés, etc.

3 -PROBLÈMES RENCONTRÉS PAR LES CONCEPTEURS

Traduire des faits, quels qu'ils soient, en vue de leur utilisation informatique pose toujours des problèmes. Le type d'information géré dans la BDHL n'échappe pas à la règle. Le codage constitue un lit de Procuste dont il faut bien s'accommoder, à défaut de s'en contenter.

Les fichiers ont (les champs de longueur fixe : les titres qui dépassent trente lettres ou les auteurs comme Barbey d'Aurevilly voient leur nom tronqué à l'affichage. Attention au clavier : tapez Molière à la place de MOLIERE et l'auteur des *Femmes savantes* n'existe plus pour la BDHL...

Plus ennuyeux : l'uniformisation des données. Quelques exemples. Pour les oeuvres du Moyen-Age, on ne connaît évidemment pas d'éditeur. Les concepteur de la BDHL ont choisi dans ce cas là (le faire figurer le nom du destinataire dans cette rubrique. Les œuvres anonymes ont été attribuées à un auteur appelé ANONYME. La date d'édition de quelques textes dramatiques a été remplacée par la date de leur première représentation, faute de pouvoir tout mettre. Pour les œuvres échelonnées (*Les Causeries dit lundi* de Sainte-Beuve par exemple) il a été créé un fichier distinct où figurent les dates de début et de fin de publication. En ce qui concerne la codification des genres, le parti pris a été celui d'adopter la dénomination annoncée par l'auteur. La BDHL permet d'attribuer deux professions à chaque écrivain : c'est insuffisant pour Beaumarchais et surabondant pour Bossuet...

Cependant, ce genre de contrainte est aussi un avantage, dans la mesure où il impose une systématisation des connaissances. Aucun dictionnaire, par exemple, n'est conséquent avec lui-même : on indique la profession d'un auteur mais pas celle d'un autre, on précise ou l'on ne précise pas si une œuvre a été censurée... Dans la BDHL, on s'est systématiquement posé toutes les questions, et l'apparition (du mot "inconnu" sur l'écran indique clairement que l'on n'a pas su y répondre.

Il faut rappeler ici une évidence que l'on perd souvent de vue en utilisant un outil (le ce type : les constatations que permet d'établir la BDHL ne sont valables qu'à l'intérieur du domaine qu'elle s'est fixée. LA BDHL n'est pas le *Catalogue des imprimés* de la B.N. et si vous ne trouvez pas *Le Marquis de Fayolle* de Nerval en consultant la liste des romans parus entre 1840 et 1880 c'est tout simplement que cette œuvre ne figure pas dans la liste... Si la BDHL vous dit que 32 % des œuvres de

Victor Hugo sont des romans il faut entendre : "32% (les œuvres de Victor Hugo figurant dans la BDHL)".

Une fois que l'on a pris cette précaution, il faut bien reconnaître que la BDHL, par son volume et les choix qui ont été faits, est bien représentative de cet étrange Panthéon que l'on est convenu d'appeler "Littérature française". Les œuvres ont été choisies aussi objectivement que possible parmi les plus célèbres, c'est-à-dire les plus citées par les manuels, les programmes scolaires, les encyclopédies. On n'y trouvera guère de lacunes choquantes.

4 - QUELS UTILISATEURS POUR LA BDHL ?

Il convient tout d'abord d'indiquer que la BDHL n'est à l'heure actuelle qu'un prototype. Les listes ne sont pas aussi complètes qu'on le voudrait et le programme d'interrogation est contraignant puisqu'il prétend imposer des parcours à l'utilisateur. On prévoit néanmoins les usages qui pourraient en être faits.

La BDHL est d'abord faite pour être utilisée dans une classe. Les élèves (ou les étudiants) pourraient être amenés à découvrir l'histoire littéraire d'une manière beaucoup plus active que par un cours systématique sur les écoles littéraires ou tel ou tel auteur. La simple promenade que constitue l'interrogation de la BDHL peut se transformer en une stimulante aventure pour peu que l'enseignant ait fixé au départ un thème de recherche. L'utilisation des classements chronologiques permet, sans charger la mémoire de l'élève, de susciter sa réflexion : quels sont les auteurs qui, comme Baudelaire, vivent sous le Second Empire ?

La BDHL présente aussi bien des avantages pour le chercheur, ne serait-ce que par sa valeur heuristique. Pour ne citer que ces exemples, il peut être très intéressant de connaître tous les faits littéraires d'une période donnée, ou d'avoir la liste des auteurs publiés par tel éditeur, ou de répertorier les livres qu'un écrivain pouvait avoir lus à une date donnée.

Comme pour beaucoup de banques de données, les utilisations les plus fructueuses peuvent d'ailleurs ne pas être celles auxquelles avaient pensé les concepteurs. Nul doute que, comme pour FRANTEXT, la demande des utilisateurs créera d'autres fonctions à l'intérieur de la BDHL.

5- PERSPECTIVES D'AVENIR

Une banque de données n'est jamais complète. En matière d'histoire littéraire pas plus qu'ailleurs, l'exhaustivité n'existe pas. Cependant, il est apparu aux concepteurs de la BDHL que l'on pourrait ajouter avec profit un certain nombre de rubriques nouvelles. Déjà, dans cet esprit, on y a introduit la liste des écrivains figurant dans les programmes scolaires, ce qui fournit une indication sur les modalités de consécration.

Par ailleurs, l'équipe de la BDHL travaille actuellement à une indexation thématique des œuvres. Il est apparu qu'il serait très intéressant de pouvoir obtenir, par exemple, la liste des romans du XX^e siècle qui ont parlé de la forêt, ou de la séduction. Une telle indexation, portant sur 2 000 œuvres, doit être conduite à partir d'un thésaurus de mots-clés et pourrait faire l'objet d'un traitement automatique sur la base de résumés.

La BDHL, depuis qu'elle s'est affranchie des limites techniques imposées par les supports magnétiques, peut s'ouvrir à des utilisations variées et puissantes. A terme, il faudra poser la question d'une utilisation aussi facile et aussi large que possible. Les perspectives ouvertes par les serveurs Vidéotex et les CD-ROM ne peuvent aujourd'hui que rendre très optimistes les concepteurs de banques de données.

Michel BERNARD