

LE REPERAGE DES MANUSCRITS A L'AIDE D'UNE BASE DE DONNÉES D'INCIPITS

Jacqueline HAMESSE

Université Catholique de Louvain
Belgique

Il y a dix-huit ans exactement, en 1972 à Madrid, lors du V^e Congrès de la *Société Internationale pour l'Étude de la philosophie médiévale*, M. Gilbert Ouy fit un exposé remarqué dans le cadre de la Commission de travail sur le traitement des textes par ordinateur. Cette Commission nouvellement créée par la S.I.E.P.M. se réunissait pour la première fois. G. Ouy y exposait les difficultés rencontrées par les médiévistes dans le repérage et l'étude des manuscrits faute d'instruments de travail pertinents:

«Au même titre — et plus encore, peut-être — que les autres médiévistes, les historiens de la philosophie sont gênés dans leur recherche par l'insuffisance quantitative et qualitative des répertoires de collections manuscrites. La solution de ce problème, décisif pour l'avenir de nos disciplines, ne saurait être recherchée dans l'amélioration des répertoires actuels, mais dans une organisation entièrement nouvelle, que le développement de l'informatique rend aujourd'hui possible».

Et M. Ouy proposait alors la mise sur pied d'une organisation internationale qui serait chargée d'informatiser toutes les informations concernant les manuscrits, qu'elles soient paléographiques ou codicologiques, ainsi que la description du contenu des différents documents conservés dans les bibliothèques. Ce projet séduisant et ambitieux en est malheureusement resté à ce stade. S'il n'a pas vu le jour, c'est probablement qu'il visait d'un coup la perfection en tentant d'apporter une réponse à *toutes* les questions que se posent les chercheurs à propos des divers manuscrits qu'ils étudient. Il faut bien avouer aussi qu'il y a vingt ans l'informatique ne s'était pas encore introduite partout et que la technologie ne connaissait pas les développements d'aujourd'hui.

Qu'en est-il vingt ans après?

L'usage de l'informatique pour les recherches bibliographiques a fait des progrès considérables pendant la dernière décennie. Bases de données, recherche automatisée de l'information, interrogations directes bibliographiques se sont développées partout et dans tous les domaines. Et cependant, malgré ces progrès considérables dans les recherches historiques et documentaires, un secteur important pour tous ceux qui s'occupent de textes anciens et médiévaux n'a pas connu le même essor. Il s'agit du traitement des manuscrits. Les grands projets théoriques n'ont pu aboutir. Cependant, des entreprises partielles et localisées ont vu le jour. On peut citer par exemple MEDIUM, la base de données élaborée par l'Institut de Recherche et d'Histoire des Textes à Paris pour les manuscrits latins, le Greek Index Project réalisé à Toronto pour les manuscrits grecs, le Hebrew Codicological Project développé à Jérusalem pour les manuscrits hébreux.

Pour les manuscrits latins, qui font l'objet de cette communication, on regrette beaucoup que l'Institut de Recherche et d'Histoire des Textes ait renoncé à inclure dans sa base de données MEDIUM les *incipits* des différentes œuvres, c'est-à-dire les premiers mots des textes, qui sont le seul moyen tout à fait sûr d'identifier les œuvres dans les manuscrits et qui permettent de les localiser dans les diverses bibliothèques qui les ont conservés. Cette décision est d'autant plus regrettable que l'I.R.H.T. possède un immense fichier d'*incipits*, composé d'environ 400.000 fiches rédigées à la main ou dactylographiées et qui constitue une mine de renseignements pour les chercheurs du monde entier. Mais ce fichier doit être consulté sur place. Son informatisation permettra de mettre à la disposition de la communauté scientifique une documentation précieuse et irremplaçable dans toutes les bibliothèques. On sait aujourd'hui que ce qui n'apparaissait encore hier que comme un rêve se réalisera un jour.

Dans le domaine latin toujours, d'autres entreprises informatisées se sont mises en place. On peut citer parmi d'autres: la Hill Monastic Manuscript Library de l'Université de Collegetown dans le Minnesota, qui a informatisé la description de son fonds manuscrit conservé sur microfilms, ce qui représente environ 72.000 recueils, en grande partie monastiques, provenant de tous les pays d'Europe occidentale. Au Canada, plus précisément à Winnipeg, il existe un catalogue automatisé de la Benjamin Collection for the History of Science. De même, à l'initiative de nos collègues américains, des standards ont été mis au point pour l'établissement des différentes zones devant intervenir dans la description des manuscrits, mais ces standards ne semblent pas avoir été acceptés partout. L'Europe a elle aussi donné le jour à des projets dédiés à l'informatisation des manuscrits latins. La Bibliothèque Nationale de Berlin collecte actuellement tous les *incipits* de manuscrits latins conservés dans les

bibliothèques allemandes. A Munich, outre le Computer Catalog of Medieval Scientific Manuscripts, des chercheurs de l'Institut für Geschichte der Naturwissenschaften du Deutsches Museum ont entrepris la description automatisée des manuscrits de mathématiques datant du moyen âge et de la Renaissance. En Allemagne toujours, le Professeur Dolezalek, établi maintenant en Afrique du Sud, s'est attelé à la description des manuscrits juridiques médiévaux. On le voit, poussés par la nécessité, divers chercheurs ont commencé à combler les lacunes dans le domaine de l'automatisation des documents manuscrits du moyen âge. Mais ces projets sont isolés et jusqu'à présent, malgré différentes tentatives, aucun projet d'ensemble n'a réussi à organiser les entreprises partielles, aucune entente n'a été réellement mise en œuvre pour homogénéiser la saisie des données, pour uniformiser les critères de description ni pour planifier les étapes qui permettraient d'automatiser la description de tout ce matériel manuscrit.

L'heure était donc venue de proposer une organisation internationale dans ce domaine. C'est ainsi qu'est né en 1987 le *Repertorium initiorum manuscriptorum*, association destinée à mettre au point une base de données pour les *incipits* de manuscrits médiévaux. Cette initiative devrait également permettre à tous ceux qui sont concernés par l'étude des manuscrits de rassembler leurs efforts — actuellement dispersés — pour arriver à une entente, à des règles de saisie des données et d'interrogation qui pourraient être adoptées partout et par tous.

Tous les médiévistes ici présents qui sont confrontés avec le problème du repérage et de la localisation des manuscrits lorsqu'ils abordent l'étude d'un texte savent à quel point dans ce domaine l'heuristique reste difficile: catalogues de manuscrits inexistantes ou insuffisants pour de nombreuses bibliothèques; répertoires non publiés ou mal faits; informations codicologiques tout à fait sommaires ou le plus souvent passées sous silence. Bref, la recherche historique n'est pas facilitée à ce niveau faute de bons instruments de travail.

Ce projet de création d'une base de données pour les *incipits* de manuscrits du moyen âge a un but essentiellement utilitaire, au moins dans un premier temps: localiser et inventorier les œuvres actuellement conservées dans les manuscrits. L'entreprise a été décrite à deux reprises: lors de la XV^e Conférence Internationale d'ALLC organisée à Jérusalem en 1988, puis à Toronto en 1989, lors du Congrès conjoint d'ALLC et d'ACH. Le texte publié dans les Actes de Jérusalem a paru et est disponible pour ceux qui sont intéressés.

Pour arriver au but fixé, nous avons commencé à rassembler une série d'*incipits*. La base qui existe actuellement en contient environ 50.000, rassemblés de la manière suivante:

— d'une part, Thomas Falmagne, chercheur du Fonds de Développement Scientifique de l'Université Catholique de Louvain, a travaillé sous ma direction pendant une année pour recueillir systématiquement les *incipits* contenus dans les manuscrits de la Bibliothèque Royale de Bruxelles. Le catalogue de Van den Gheyn qui existe pour cette bibliothèque date du début du siècle et est totalement insuffisant, si pas lacunaire au niveau des *incipits*;

— d'autre part, on a dressé une liste des différents fichiers d'*incipits* existants qui ont été repérés. Ces fichiers, établis pour la plupart à usage privé par des érudits, ne sont pas ou peu accessibles aux chercheurs. Ils contiennent cependant une mine d'informations qui, si elles étaient enregistrées sur disquettes, fourniraient à la suite d'une simple saisie des données, des renseignements précieux et inédits à tous les médiévistes. Voici, parmi d'autres, quelques noms d'érudits s'étant attelés à cette tâche: le Père Kaeppli, par exemple; son fichier qui contient quelques 50.000 fiches est conservé actuellement à Sainte Sabine à Rome. Les Dominicains, en la personne du Père Panella, nous ont donné l'autorisation de l'enregistrer. Le Père Bataillon s'est constitué lui aussi un fichier personnel d'*incipits* qu'il a mis très généreusement à la disposition de notre projet. Aux Etats-Unis, chez les Franciscains de Saint Bonaventure, près de Buffalo, le Père G. Mohan, aujourd'hui décédé, a également constitué un énorme incipitaire d'œuvres philosophiques et théologiques. Le contenu de ce fichier a été enregistré. Il en existe d'autres que nous avons repérés. Nous les enregistrons progressivement grâce à une aide généreuse que le Fonds National de la Recherche Scientifique Belge a accordé pour faire ce travail. Nous espérons être en mesure de faire une démonstration et une interrogation de toutes ces données à Ottawa lors du prochain Congrès de philosophie médiévale qui aura lieu en août 1992. Le matériel qui vient d'être mentionné est essentiellement philosophique et théologique. Il intéressera donc en premier lieu les spécialistes de l'histoire de la philosophie médiévale.

Parallèlement à ce travail d'enregistrement qui peut être fait tout de suite, l'entreprise internationale regroupe des collaborateurs dans différents pays, afin que dans les diverses bibliothèques on s'attèle à l'enregistrement des incipitaires existants et à l'inventaire des manuscrits non encore décrits. Petit à petit le travail s'organise. Ainsi en Pologne, le grand incipitaire de la Bibliothèque Jagellonne de Cracovie est en cours d'enregistrement dans ce cadre. En Italie, grâce à l'aide précieuse de M. Leonardi, les différentes bibliothèques de Florence ont été contactées et ont manifesté leur désir de collaborer. Les informations contenues dans les manuscrits des différentes bibliothèques de Saint-Petersbourg vont être enregistrées et entrer ainsi dans la base de données. Une Table Ronde internationale aura d'ailleurs lieu en Belgique en mars 1992 pour organiser la collaboration.

Avant d'aller plus loin et d'évoquer les problèmes techniques, il faut répondre à une objection qui ne manquera pas de surgir à l'issue de cet exposé: les fichiers de ce type qui sont enregistrés doivent certainement contenir des erreurs à côté de tous les renseignements précieux qu'ils procurent. C'est vrai et ce problème est très préoccupant. Mais il faut aller de l'avant et suivre l'adage scolastique bien connu: «*Melius est sic esse quam non esse*». Il vaut mieux avoir une documentation à sa disposition que rien du tout. Les spécialistes qui pourront la consulter aideront à la corriger progressivement. Le fichier d'*incipits* philosophiques constitué au Centre De Wulf-Mansion de Louvain-la-Neuve a déjà été enregistré. Il est à la disposition des chercheurs; il contient environ 15.000 enregistrements différents ce qui constitue une documentation, limitée certes, mais bien utile pour tous ceux qui s'occupent de philosophie médiévale.

Mais après avoir brièvement évoqué l'histoire de notre projet, après avoir montré comment nous concevons la collecte des données, il est temps d'en venir à l'aspect plus technique et de présenter d'abord le programme d'enregistrement des données et ensuite des exemples d'interrogation de la base.

1 — L'ENREGISTREMENT

Le programme d'enregistrement des données est très simple. Il a été conçu de manière à permettre, sans aucune initiation technique, une saisie rapide et facile des données.

Formulaire d'enregistrement

Documents

1—	\$ID	Identification du MS
2—	\$VI	Lieu
3—	\$DI	Dimension
4—	\$PA	Nombre de folios
5—	\$CA	Références
6—	\$RE	Bibliographies
7—	\$PO	Possesseurs
8—	\$IN	Numérotation

Oeuvres

1—	\$RU	Rubrique
2—	\$TA	Tables des matières
3—	\$DI	Division du texte
4—	\$IN	Incipit
5—	\$EX	Explicit
6—	\$PE	Péripcope
7—	\$CO	Colophon
8—	\$LI	Liturgie

9—	\$TI	Titre
10—	\$AU	Auteur
11—	\$TR	Traducteur
12—	\$ED	Edition
13—	\$RI	Références
14—	\$RE	Renseignements
15—	\$CD	Codicologie
16—	\$DE	Décoration
17—	\$PA	Pagination
18—	\$DA	Date
19—	\$LA	Langue
20—	\$AN	Auteur de la notice \$\$\$

Chaque zone est séparée par un \$.

Les informations pour chaque œuvre sont séparées par \$\$\$.

Exemples d'enregistrement

\$RU Incipit passio sanctorum Basilidis Trophidis et Mandalis Martirum

\$IN In ciuitate orientali cum beatus Basilidis esset in officio Platonis prefecti dicit ei dominus noster Iesus Christus ego te uolo tollere de officio Platonis prefecti en in meo officio esse (...)

\$EX (...) collegerunt corpora sanctorum Basilidis Trophidis et Mandalis et sepelierunt in eodem loco. Et fabricata est sanctitatis eorum basilica ubi florent usque in hodiernum diem eorum beneficia.

\$TI PASSIO SANCTORUM BASILIDIS TROPHIDIS ET MANDALIS

\$VI BRUXELLES, BR-98-100

\$PA 4vb-5rb

\$\$\$

\$RU Incipit passio sancti Barnabe apostoli

\$IN Redemptor et saluator noster Dominus antequam de interitu sathane triumphans uictor exultans penetreret regna polorum (...)

\$EX (...) cum ymnis et laudibus mirifice collatum est ad honorem et gloriam Domini nostri Iesu Christi qui cum patre et spiritu sancto uiuit et regnat in secula seculorum. Amen

\$TI PASSIO SANCTI BARNABE

\$VI BRUXELLES, BR 98-100

\$PA 5rb-7va

\$\$\$

\$RU Incipit passio Nazarii discipuli sancti Petri apostoli

\$IN Nasarius genitus patre nomine Affricano et genere et patria Afro matre uero nomine Perpetua ciue romana baptizata scilicet per manus Petri apostoli Simone iam deuicto (...)

\$EX (...) Ceci enim post multos annos receprunt lumen claudi ambulauerunt demonia quicumque a natiuitate sua habuerunt mundati sunt confestim in uirtute et gratia Domini nostri Iesu Christi ipsi gloria in secula seculorum. Amen.

\$TI PASSIO SANCTI NAZARII

\$VI BRUXELLES, BR 98-100

\$PA 7va-8ra

\$\$\$

Le programme mis au point par Paul Bratley de l'Université de Montréal est très simple. Il prend en compte un nombre variable de zones. Chaque zone est séparée par un \$. L'enregistrement complet d'une *incipit* se termine par \$\$\$.

Le programme permet de créer de nouvelles zones si un élément nouveau l'exige en cours d'enregistrement. Ainsi, à la demande de M. Beit-Arié qui

collaborera probablement au projet pour les manuscrits hébreux, nous avons créé une zone spéciale DI pour les textes acéphales et fragmentaires, qui permet de distinguer les divisions principales d'une œuvre et d'indiquer dans ces cas-là la division interne du texte, à savoir l'intitulé du chapitre ou du livre.

La personne qui enregistre les données a sur son écran la fiche d'enregistrement théorique contenant l'intitulé de chaque zone. Elle remplit uniquement celles qui concernent l'*incipit* à enregistrer. Cette fiche comprend toutes les zones qu'on peut rencontrer théoriquement lorsqu'on fait le relevé des *incipits* ou l'enregistrement des données. Il s'agit d'un idéal, mais en théorie toutes ces zones ne sont que rarement remplies. L'important est de les avoir prévues afin qu'on puisse éventuellement les compléter et fournir d'autres informations dans une étape ultérieure du travail. Dans le cas des fichiers d'*incipits* d'érudits que nous enregistrons, en général des zones comme dimensions, processeurs dans le fichier document ne sont pas nécessairement remplies. Nous avons groupé les différentes informations en deux fichiers distincts: l'un regroupe tous les renseignements stables à propos du manuscrit qui contient l'œuvre. Nous parlons de zones fixes dans ce cas parce que les informations restent les mêmes pour les différents *incipits* contenus dans un seul manuscrit. L'autre concerne les textes contenus dans chaque recueil.

Prenons le cas du manuscrit de la Bibliothèque Royale de Bruxelles, cote 98-100 dont l'enregistrement figure à la page ...; on constate que la zone VI est la même pour les trois enregistrements. On peut donc la reproduire automatiquement. Le travail est effectué dans la bibliothèque même sur un IBM ou un Thoshiba portable. Les données sont stockées sur disquettes et recopiées ensuite sur le disque dur de l'IBM PC qui se trouve à Louvain-la-Neuve. Actuellement 256 K de mémoire nous suffisent. Mais nous savons que nous serons bientôt amenés à élargir la capacité de mémoire, ce qui ne pose plus beaucoup de problèmes actuellement.

Afin que chaque bibliothèque et que tous les chercheurs qui collaborent, adoptent les mêmes critères d'enregistrement, nous sommes occupés à mettre au point un manuel qui permettra une uniformisation des critères de description. La première version de ce manuel est terminée. On y trouve notamment l'explication des divers codes utilisés pour identifier chaque zone d'enregistrement. Mais le texte est actuellement provisoire; il ne s'agit pas encore de la version définitive. Nous ne diffuserons celle-ci qu'en 1992, après la Table ronde internationale dont il a été question. Il est indispensable que les médiévistes, spécialistes des manuscrits, acceptent ces règles avant de les mettre en circulation.

Les médiévistes ne seront pas surpris de voir aborder maintenant les problèmes posés par le relevé des divers *incipits* de certaines œuvres.

En effet, certains textes appartenant à des genres littéraires différents peuvent avoir des *incipits* plus ou moins complexes. Dans le cas d'un sermon qui débute par une citation biblique, le véritable *incipit* du sermon ne sera pas cette citation qui peut être la même dans de nombreuses homélies, mais bien le début du texte qui suit immédiatement. Ainsi pour le sermon qui débute par les mots suivants: *Hortamur vos ne in vacuum gratiam Dei recipitatis. / Diffusa est gratia in labiis tuis; proterera benedixit te Deus, Deus tuus in aeternum. / Verbum secundo propositum dirigitur ad Dominum nostrum Iesum Christum...*, le véritable *incipit* de l'œuvre qui doit être enregistré dans la zone IN est: *Verbum secundo propositum dirigitur ad Dominum nostrum Iesum Christum...* Tout ce qui précède est citation biblique et doit donc figurer dans la zone PE.

La complexité de l'enregistrement est très grande aussi pour toutes les œuvres liturgiques qui sont composées de pièces différentes dont il importe de donner l'*incipit*, ce qui n'a pas toujours été fait jusqu'à présent. Dans les relevés qui ont été réalisés pour les manuscrits de la Bibliothèque Royale de Bruxelles, nous nous sommes efforcés de reprendre les divers *incipits* qui s'y trouvent et pour attirer l'attention des chercheurs qui utiliseront la base de données sur ce problème, nous avons introduit une zone spéciale LI qui indique qu'il s'agit d'une œuvre liturgique et qu'il faut donc faire attention. Il est bien évident que, dans les incipitaires déjà constitués par des érudits dont il a été question tout à l'heure, ces distinctions plus spécifiques n'existent pas en général. Rien ne nous empêchera de compléter plus tard cette documentation parfois fragmentaire.

Certaines œuvres médiévales présentent d'autres spécificités dont il convient de tenir compte: les florilèges par exemple, les manuels scolaires, les encyclopédies et bien d'autres. Il n'est pas nécessaire de s'y attarder pour l'instant puisque le problème a été illustré par un exemple.

2 — L'INTERROGATION DES DONNÉES

Il est temps d'en venir à la dernière partie de cette communication consacrée à l'interrogation des données et aux problèmes de comptabilité informatique qui existent entre les diverses bases de données qui sont déjà en cours de constitution dans différents pays.

Le meilleur procédé pour avoir des solutions correctes lorsqu'on procède à une interrogation est de cumuler des éléments complémentaires appartenant à des zones d'enregistrement différentes.

Exemple d'interrogation

RU *Omelia lectionis eiusdem beati Ambrosii episcopi*

IN *Latent diuina mysteria nec facile iuxta propheticum dictum quisquam hominum potest scire consilium Dei (...)*

EX (...) Et ne quo de generi deprauaretur affatu ab angelo salutatur.

AU

VI BRUXELLES, BR 135

PA 137ra-142rb\$\$\$

RU In annuntiatione beate Marie uirginis. (Lectio sancti euangelii) secundum Lucam

PE In illo tempore missus est angelus Gabriel in ciuitatem Galylee cui nomen Nazareth

RU Tractatus sancti Ambrosii de eadem lectione.

IN *Latent quidem diuina mysteria nec facile iuxta propheticum dictum quisquam. Hanc habemus inter omelias sancti Ambrosii super Lucam et est secundus liber et primum capitulum (...)*

AU AMBROSIUS

VI BRUXELLES, BR 218

PA 79va\$\$\$

RU Euangelium

PE Missus est Gabriel

RU Tractatus Ambrosii

IN *Latent quidem*

VI BRUXELLES, BR 218

PA 317ra-va\$\$\$

RU Lextio sancti euangelii secundum Lucam

PE In illo tempore missus est angelus gabriel a Domino (...) et nomen uirginis Maria. Et reliqua.

RU Tractatus sancti Ambrosii Mediolanensis episcopi

IN *Latent quidem diuina mysteria nec facile iuxta propheticum dictum quisquam hominum potest scire consilium Dei (...)*

Ex (...) Et propheta manifestauit dicens ecce uirgo in utero accipiet.

AU AMBROSIUS

VI BRUXELLES, BR 218

PA 319va-320ra\$\$\$

La seconde solution ci-dessus, lorsqu'on est dans l'impossibilité de cumuler des éléments (soit parce qu'on ne les connaît pas, soit parce qu'ils sont sujets à variation), est de formuler dans la zone IN une interrogation courte (la plus brève possible), portant sur un mot latin pas trop fréquent et qui ne donne pas lieu dans les manuscrits à de nombreuses variantes graphiques. Par exemple, si on veut retrouver une œuvre dont l'*incipit* est: *Latent quidem diuina mysteria nec facile iuxta propheticum dictum quisquam hominum potest scire consilium Dei...*, il vaut mieux interroger uniquement à l'aide de *Latent*, plutôt que de donner toute la phrase citée. En ne prenant qu'un mot comme base, on élimine du même coup une série de variations.

Il suffit de consulter l'exemple ci-dessus pour voir que ce système est bon. En effet, si on avait donné toute la phrase d'*incipit* comme base d'interrogation, le premier des quatre manuscrits contenant le texte n'aurait pas été signalé dans la réponse de l'ordinateur, puisque le deuxième mot *quidem* y est

omis. D'autre part, l'*incipit* du troisième manuscrit est très court. Il n'aurait donc pas été signalé non plus. En interrogeant à l'aide de toute la phrase, on aurait donc risqué d'avoir une réponse incomplète.

Un des avantages de ce type de recherches apparaît immédiatement lorsqu'on examine le tableau. On constate que dans la fiche signalétique du premier manuscrit, la zone AU est vide, ce qui veut dire que l'auteur du catalogue des manuscrits de la Bibliothèque Royale de Bruxelles, ignorait que cette œuvre était de saint Ambroise. Grâce à l'interrogation, on constate, en comparant les *incipits* que le premier et le quatrième manuscrit ont un *incipit* identique (à l'exception du terme *quidem* omis dans le premier), ce qui équivaut à dire que l'on peut probablement attribuer l'œuvre mentionnée dans le premier manuscrit (BR 135) au même auteur, c'est-à-dire saint Ambroise. Les médiévistes verront immédiatement tout le profit qu'on peut tirer de ce type d'interrogation, au-delà du repérage des manuscrits.

Si le chercheur veut mieux connaître le contenu d'un seul manuscrit qui l'intéresse plus particulièrement, il suffit qu'il interroge la base en donnant la cote du manuscrit (par exemple BRUXELLES, BR 135). L'ordinateur lui fournira en réponse la liste de toutes les œuvres décrites et repérées dans le codex. On peut procéder de la même manière pour avoir la liste des manuscrits d'une bibliothèque, ou les différentes œuvres d'un auteur. Le système est très souple et permet de multiples exploitations.

Un des avantages du programme est de faire apparaître lors de chaque interrogation toutes les zones descriptives de la fiche d'enregistrement. Si on veut connaître le contenu général du manuscrit avec les autres œuvres qu'il contient, il suffit d'interroger dans la zone VI, la cote du manuscrit. L'ordinateur va chercher automatiquement toutes les fiches portant la même cote et les fait apparaître à l'écran, de même qu'il donne également toutes les indications codicologiques qu'il possède, à condition bien sûr qu'elles aient été préalablement enregistrées, ce qui n'est pas le cas, comme cela a été dit, pour la plupart des fichiers d'érudits que nous enregistrons pour le moment et qui sont généralement dépourvus de ce type d'informations.

Il convient cependant d'attirer l'attention des chercheurs sur les problèmes posés par une interrogation trop rapide ou incomplète. En effet, le problème des graphies ou des variantes doit être souligné. On risque de perdre des informations si on ne procède pas à des interrogations complémentaires, comme on l'a vu dans l'exemple qui vient d'être présenté. Le chercheur doit donc être attentif à cette question et multiplier les interrogations pour être sûr d'avoir obtenu tous les renseignements qui sont en mémoire d'ordinateur concernant l'œuvre à laquelle il s'intéresse.

Autre exemple d'interrogation

- RU Dominica VII [post octauam penthecosten]
RU [Lectio sancti euangelii] secundum Marcum
PE In illo tempore cum turba multa esset cum Iesu nec haberent quod manducarent conuocatis discipulis ait illis. Misericors super turbam quia ecce iam triduo sustinent me nec hanent manducent. Et reliqua.
- RU Omelia uenerabilis Bede presbiteri de eadem lectione
IN *In hac lectione* consideranda est in uno eodemque redemptore nostro distincta operatio diuinitatis et humanitatis (...)
EX (...) Bene quatuor ut ipso etiam numero docerent euangelicis saturatos esse tibiariis.
AU BEDA VENERABILIS
VI BRUXELLES, BR 80-84
PA 96vb-98rb \$\$\$
- RU Dominica VII [post pentecosten]
LI *Capitulum ad uesperas* Fratres humanum dico propter infirmitatem carnis uestre sicut enim exhibuistis membra uestra seruire immunditie (...) *Oremus* Absolue Domine quesumus nostrorum uincula peccatorum et quicquid pro eis meremur auerte. Per Dominum.
- RU [Lectio sancti euangelii] secundum Marcum
PE In illo tempore cum turba multa esset cum Ihesu nec haberent quod manducarent conuocatis discipulis ait illis misereor super turbam quia ecce iam triduo sustinent me nec habent quod manducent. Et reliqua.
- RU Omelia lectionis eiusdem [uenerabilis Bede presbiteri]
IN *In hac lectione* consideranda est in uno eodemque redemptore nostro distincta operatio diuinitatis et humanitatis (...)
EX (...) Bene quatuor milia ut ipso etiam numero docerent euangelicis se pastos esse cybariis.
AU BEDA VENERABILIS
VI BRUXELLES, BR 104-105
PA 58rb-67va\$\$\$
- RU Dominica septima [post trinitatem]
RU [Lectio sancti euangelii] secundum Marcum
PE In illo tempore cum turba multa esset cum Ihesu nec haberent quod manducarent conuocatis discipulis ait illis misereor super turbam (...) Erant autem qui manducauerunt quasi quatuor milia et diuisit eos. Amen.
- RU Omelia lectionis eiusdem uenerabilis Bede presbiteri
IN *In hac lectione* consideranda est in uno eodemque redemptore nostro distincta operatio diuinitatis et humanitatis (...)
EX (...) per penitentiam declinans ad Dominum se in opere in locutione atque cogitatione conuertit.
AU BEDA VENERABILIS
VI BRUXELLES, BR 135
PA 110ra-111ra\$\$\$

Ce tableau montre un autre type intéressant d'interrogation. Un spécialiste de sermons du moyen âge, travaillant sur les homélies de Bède le Vénéralable et ne connaissant que la citation biblique mise en tête de certains sermons et pas le véritable *incipit* du texte, aura immédiatement une réponse intéressante en basant son interrogation sur *In illo tempore* dans la zone PE et en ajoutant l'information Beda Venerabilis en zone AU. Il constatera en examinant les trois textes affichés en réponse que l'*incipit* des trois œuvres est identique. Il verra également que même si les œuvres ont un *incipit* semblable, l'*explicit* (à savoir la fin du texte) diffère. En effet, dans le troisième cas, le

texte est incomplet et l'*explicit* est donc différent. Toutes ces informations sont fondamentales pour un spécialiste.

Ces quelques exemples suffisent à montrer l'intérêt d'une base de données pour les incipits de manuscrits du moyen âge et les différents types d'informations qu'une interrogation intelligemment menée peut fournir. Le système répond parfaitement aux exigences des spécialistes.

Actuellement nous essayons de perfectionner le logiciel afin que les temps de réponse aux interrogations soient plus rapides. Il y a un moyen d'alléger le système en employant la méthode booléenne et en introduisant des codes pour indiquer qu'une lettre peut ou non être absente, ou pour ajouter éventuellement un caractère. Ce système utilisé pour la recherche approximative d'informations notamment bibliographiques donne de bons résultats et s'adapte tout à fait aux variations graphiques que présente le latin.

Il reste deux problèmes à aborder, avant de conclure : celui de l'accès à la base de données et de la diffusion des informations qu'elle contient, et, d'autre part, la compatibilité informatique entre divers projets déjà en cours. A l'origine le projet avait été conçu dans un but utilitaire : l'accès aux œuvres conservées dans les manuscrits et leur localisation rapide. La base de données doit être accessible à tous les chercheurs. Or nous savons tous que si une interrogation en direct est théoriquement possible, ce système à distance s'avère très onéreux et n'est donc pas à la portée de tout le monde. Actuellement la diffusion sur CD-Rom se répand de plus en plus mais on constate que dans bien des cas, seules les bibliothèques peuvent les acquérir, étant donné leur prix. D'autre part, si ce système de stockage de données convient très bien à des informations bien établies qui contiennent peu d'erreurs — puisqu'il est impossible de corriger lorsque le disque est pressé — peut-on vraiment considérer que, dans le cas des *incipit*, ce moyen de diffusion soit actuellement le meilleur ?

Les données que nous enregistrons sont souvent fragmentaires ; elles contiennent des erreurs, des approximations, des points d'interrogation. La plupart des enregistrements seront donc sujets à correction lorsque les spécialistes les auront examinés. C'est surtout vrai pour les incipitaires constitués par des érudits qui n'ont pas adopté des critères uniformes dans leurs relevés et qui les ont commencés il y a parfois plus de cinquante ans, ne pouvant évidemment pas tenir compte à cette époque des progrès et découvertes faits récemment. Le matériel est donc par définition incomplet. Si nous voulons que tous les chercheurs puissent s'en servir pour y puiser des informations, mais aussi pour compléter et corriger la documentation existante, il faut que nous le diffusions sur un support accessible à tous et bon marché. Seules les disquettes répondent à ces deux critères. On pourrait envisager de livrer

chaque année une partie des données et les chercheurs pourraient les acquérir sous forme d'abonnement. Ce système économique serait accessible à tous et les mises à jour et corrections seraient livrées ultérieurement à chaque personne qui aurait souscrit à l'ensemble.

Outre ce problème de diffusion, il faut arriver à une compatibilité entre les projets existants. Or, c'est loin d'être le cas à l'heure actuelle. Les difficultés sont doubles : les critères d'enregistrement et la compatibilité des divers systèmes informatiques utilisés. Il est absurde qu'à l'heure actuelle, on ne puisse pas récupérer des données déjà enregistrées ailleurs. Il reste tant de chose à faire. Pourquoi perdre du temps et de l'argent à recommencer plusieurs fois le même travail. La collaboration reste encore trop souvent un vain mot. C'est à ce niveau que nos efforts doivent porter. Il faut arriver à un consensus qui satisfasse les chercheurs et qui permette de franchir l'obstacle des difficultés techniques.

Le Repertorium initiorum manuseriptorum a été créé dans ce but : rendre service aux chercheurs et arriver à une collaboration internationale qui soit le fruit d'un dialogue entre spécialistes des manuscrits et informaticiens et qui réponde aux critères scientifiques comme aux exigences techniques.

Les avantages de ce projet sont multiples à long terme :

1. au niveau informatique, on pourra introduire le modèle d'enregistrement proposé dans les bibliothèques des différents pays et encourager l'uniformisation de la saisie des données ;

2. à long terme la base de données pourrait être plurilingue. Elle contiendrait non seulement les manuscrits latins, mais aussi les manuscrits grecs, hébreux, arabes et en langues vulgaires. Ceci constituerait un avantage fondamental pour l'étude de la littérature médiévale de traductions.

Chaque base linguistique serait bien sûr autonome. Les manuscrits grecs, hébreux, arabes, orientaux et en langues vulgaires ont tous leur spécificité et demandent donc des traitements adaptés et adéquats. Cependant, il est intéressant de voir pour le futur l'aspect plurilingue de l'entreprise. Les traductions constituent une bonne part de la production médiévale. Il serait donc intéressant d'étudier de manière comparée les textes philosophiques et scientifiques traduits. La comparaison des bases de données linguistiques permettra certainement de repérer une série de traductions non encore identifiées. Mais ne rêvons pas, cet objectif ne sera pas facile à atteindre. En effet, le domaine grec et le domaine hébreu, pour lesquels il existe déjà des bases de données contenant les descriptions de manuscrits, posent moins de problèmes. Les manuscrits ont été repérés et même si ces bases ne contiennent pas encore les *incipits* des textes, leur relevé s'avère plus facile parce que tout le matériel est rassemblé tant à Jérusalem qu'à Toronto.

La situation n'est pas du tout pareille pour le latin. Qui pourrait dire actuellement combien de manuscrits latins du moyen âge sont encore conservés ? Il est impossible de répondre de manière précise à cette question. Nos connaissances sont trop limitées en la matière.

Que dire alors des manuscrits arabes ? On peut parler à ce sujet de *terra incognita*. Lorsqu'on sait que de nombreux manuscrits arabes du moyen âge sont notamment conservés dans plusieurs pays d'Afrique (le Niger, le Mali, le Sénégal, la Côte d'Ivoire, etc.) sans avoir jamais été étudiés, on imagine sans peine le chemin qui reste à parcourir. D'autant que les conditions de conservation de ces manuscrits sont souvent très mauvaises, pour ne pas dire déplorables.

Pour mener ce projet à bien, il est indispensable qu'il se fasse en collaboration internationale. Aussi, lors de la Table Ronde qui sera organisée en Belgique, nous comptons remettre à l'ordre du jour les questions des standards, pour la description des manuscrits latins. Seul un dialogue entre informaticiens et médiévistes pourra aboutir à l'établissement de règles valables. Lors du Congrès conjoint d'ALLC (Association for Literary and Linguistic Computing) et ACH (Association for Computers and the Humanities) qui a eu lieu à Toronto en 1989, une session avait été consacrée à l'aide que l'ordinateur peut apporter dans l'étude des manuscrits. Lors du débat qui a suivi, le Conservateur des manuscrits de la British Library à Londres a expliqué pourquoi il ne voulait pas accepter les standards proposés par ses collègues américains : essentiellement parce qu'ils ne correspondaient pas aux réalités à décrire. Un dialogue existe actuellement à ce niveau. Dans le cadre de la *Text Encoding Initiative* financée par la Commission des Communautés européennes, le National Endowment for the Humanities et le Social Science and Humanities Research Council of Canada, on essaie de mettre au point des règles de saisie des données dans toutes nos disciplines et on tente de renforcer le dialogue entre philologues et historiens d'une part et informaticiens de l'autre. Nous espérons donc qu'en mars 1992, à l'issue de la Table Ronde, nous pourrons arriver à proposer des règles conformes aux besoins des médiévistes et aux exigences spécifiques des manuscrits. Nous espérons que différents organismes internationaux nous aideront à imposer nos critères et nos règles dans tous les pays.

L'intérêt général de l'entreprise a d'ailleurs déjà été reconnu au niveau international. En effet, nous avons obtenu des aides de l'Agence de Coopération Culturelle et Technique dont le siège est à Paris. De même, la Belgique nous a officiellement encouragés en sélectionnant ce projet dans ses accords culturels avec le Canada, Israël, l'Italie et le Québec. Nous espérons donc que les moyens financiers indispensables nous permettront de progresser dans la

collecte des données et que nous pourrons bientôt faire écho de manière concrète à ce cri du cœur poussé en 1938 par Père Henquinet, érudit médiéviste à la recherche des manuscrits franciscains du moyen âge : «*O utilitas catalogorum, etiam saepe errantium, in quibus operum notantur rationabiliter saltem initia et fines* ».