

Clément Arsenault*

Élaine Ménard

École de bibliothéconomie et des sciences de l'information, Université de Montréal,
C.P. 6128, succ. Centre-ville, Montréal (QC) H3C 3J7

Le traitement des articles initiaux dans les catalogues de bibliothèque : une étude de cas

Abstract: This research examines problems caused by initial articles in library catalogues. The problematic records observed are those whose title starts with a word erroneously considered as an article. Many algorithms truncate queries by removing words corresponding to the articles from a stop list, even if the initial word is not an article. Consequently, a certain number of documents remain untraceable.

Résumé : Cette recherche examine les problèmes causés par les articles initiaux dans les catalogues. Les enregistrements problématiques observés sont ceux dont le titre commence par un mot considéré erronément comme un article. Plusieurs algorithmes tronquent des requêtes les mots correspondant aux articles d'une liste d'exclusion, même si le mot initial n'est pas un article. Un certain nombre de documents demeurent donc introuvables.

1. Énoncé du problème

En général, la procédure typique pour le classement alphabétique des titres d'ouvrages dans un catalogue est de ne pas tenir compte des articles initiaux (Bourne 1975, 221). Par exemple le livre intitulé *The earth and its inhabitants* sera classé à la lettre « e » afin d'éviter des trop importants regroupements de titres à « the ». Dans le format MARC, utilisé pour l'encodage électronique des notices bibliographiques, ce procédé est réalisé grâce à l'utilisation d'un indicateur numérique signalant au système le nombre de caractères initiaux à ignorer dans la chaîne de caractères à indexer. Dans l'exemple précédent, le code 4 dans le 2^e indicateur du champ 245 sera utilisé afin d'indiquer que les quatre premiers caractères (t-h-e, plus l'espace) ne sont pas à indexer (LC MARC Standards 2005). Cet indicateur permet de laisser l'article initial dans le champ titre afin qu'il soit affiché, mais sans qu'il soit inclus dans l'index pour la recherche. Au moment du repérage, si l'utilisateur inclut par mégarde l'article initial dans son équation de recherche (p. ex., ti=the earth and its inhabitants), dans certains catalogues un algorithme de détection est activé afin de supprimer automatiquement l'article initial de la chaîne de requête avant de l'envoyer dans l'index. Dans notre exemple, le système présentera donc à l'utilisateur l'index des titres à la lettre « e » et non à la lettre « t ». Ceci étant dit, bien que l'indicateur MARC contrôle la façon dont les articles initiaux sont indexés, il ne contrôle toutefois pas comment s'effectue la recherche (Seeman 2002). La détection automatique des articles initiaux dans les équations de recherche est particulièrement problématique dans les environnements multilingues (Bourne 1975, 222). La décision du catalogueur de désigner un mot initial comme article à ignorer doit tenir compte de plusieurs facteurs, notamment de la langue de l'ouvrage puisque nous pouvons conjecturer qu'un article initial dans une langue correspondra très souvent à un mot « normal » dans une autre langue, par exemple en allemand l'article « die » qui est

* Cette étude est gracieusement subventionnée par le Conseil de recherches en sciences humaines du Canada.

l'homographe du verbe « mourir » en anglais (*to die*). Dans certains cas il est même nécessaire d'analyser grammaticalement les titres car au sein de la même langue il peut y avoir de l'ambiguïté. Prenons par exemple l'article français « la » qui est homographe de l'adverbe de lieu « là », ou encore le mot « un » qui a quelquefois la fonction d'article indéfini *Un destin tragique*, de pronom *L'un d'entre eux*, de nombre *Un, deux, trois, partez !*, ou qui peut faire partie d'une locution adverbiale *Un peu de fatigue*. Il est donc essentiel de faire le traitement documentaire des titres au cas par cas.

Toutefois, au moment de la recherche, les algorithmes de détection inclus dans la grande majorité des systèmes de repérage ne sont pas suffisamment intelligents pour détecter toutes ces subtilités linguistiques. Il en résulte quelquefois des problèmes au moment du repérage. C'est le cas par exemple pour le titre *Las Vegas, the success of excess*. Ce titre sera correctement classé dans l'index sous la lettre « L » puisque le mot « las » dans ce cas-ci fait partie d'un nom propre de lieu. Toutefois, si le mot « las » est interprété par l'algorithme de détection comme un article espagnol lors de la requête, l'utilisateur sera faussement dirigé à la lettre « V » de l'index. Prenons également le cas d'un usager désirant retracer l'ouvrage de Charles-Ferdinand Ramuz intitulé *À propos de tout*. Une requête normale par furetage dans l'index des titres sera « a propos de tout ». Malheureusement, si l'algorithme de détection des articles initiaux est activé, l'usager sera mené à la lettre « P » de l'index puisque le « a » initial de la requête sera considéré, à tort, par le système comme l'article indéfini « a » en anglais, et le texte de la requête sera tronqué à l'insu de l'usager pour devenir « propos de tout ». Et comme le titre a bel et bien été indexé sous la lettre « A », l'usager se retrouvera à une position inexacte dans l'index tel qu'illustré à la figure 1. Il aura par conséquent la fausse impression que le titre ne se trouve pas en collection.



Fig. 1 : Exemple d'une recherche en mode « parcourir index titre ».

Pour contourner ce problème, une « solution » possible pour l'usager est de se positionner dans l'index au début des « A » et de faire défiler la liste complète des titres, ce qui est très peu efficace (voir fig. 2). Dans cet exemple, il faudra cliquer 273 fois sur le bouton « page suivante » pour arriver à localiser le titre en question et qui est bel et bien en collection (voir fig. 3). Mais peut-on raisonnablement supposer que l'usager aura, et le réflexe de procéder de la sorte, et la patience requise pour arriver à ses fins, surtout s'il n'a aucune garantie au départ que l'ouvrage recherché se trouve dans le catalogue.

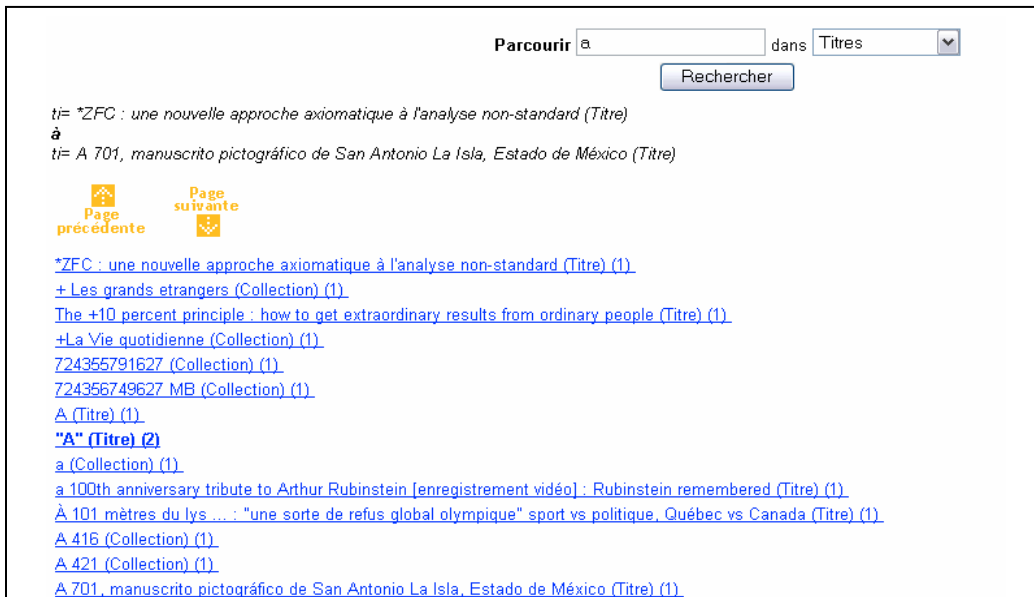


Fig. 2 : Requête pour forcer le positionnement à la lettre « A ».



Fig. 3 : L'item finalement localisé après 273 clics sur le bouton « page suivante » !

Une solution probablement plus efficace serait simplement de ne pas activer de tels algorithmes de détection dans le module de recherche et de fournir des instructions claires indiquant à l'utilisateur d'omettre les articles initiaux lors de la formulation des requêtes. Taylor (2004, 121) rapporte que si les instructions sont clairement positionnées (voir fig. 4) les usagers tiennent généralement compte de cette consigne : « users tend to follow this advice if the instruction is noticeable and can be seen from the search box ».



Fig. 4 : Exemple d'instructions claires sur l'interface de recherche.

En raison de ce qui précède, on peut se demander s'il convient d'utiliser un algorithme de détection basé sur une liste d'exclusion dans un catalogue de bibliothèque puisque celui-ci peut potentiellement causer plus de problèmes qu'il n'en résout. Il est certain que l'utilisation d'une liste d'exclusion offre une forme d'aide aux chercheurs profanes en prenant, en quelque sorte, en charge la formulation de leurs requêtes. Ainsi, l'utilisateur qui obtient une référence sous forme électronique peut faire aisément usage de la fonction copier-coller pour transférer directement sa requête dans la boîte de recherche du catalogue sans se poser plus de question. Si ce titre débute par un mot ou des mots faisant partie de la liste d'exclusion, l'algorithme de recherche amputera la requête des mots non nécessaires, sans que l'utilisateur en ait connaissance. Toutefois, cette même liste d'exclusion, comme nous l'avons vu, n'offre pas que des avantages puisqu'elle peut, dans certaines circonstances, pénaliser l'utilisateur. Par conséquent, on peut se demander quel est le meilleur scénario à adopter.

Il est raisonnable de penser qu'un bibliothécaire professionnel connaisse les moyens de contourner ce genre de problème de repérage, mais il n'en est pas de même pour ce qui est de l'utilisateur final, qui est de plus en plus en situation d'autonomie pour ses recherches bibliographiques. À ce jour, peu d'études se sont penchées sur le problème du repérage d'items connus par titre bien que, comme le mentionne Notess, « the title can often uniquely identify a publication and provide one of the easiest searchable access points for the publication » (Notess, 2001). Il est donc primordial, et même fondamental, de se pencher sur cette question afin de mieux comprendre l'ampleur des problèmes liés à ce phénomène dans les catalogues de bibliothèque.

2. Objectifs de l'étude

Le but de la présente étude est d'examiner l'étendue des problèmes de repérage causés par le traitement erroné d'articles initiaux dans un catalogue de bibliothèque.

Conséquemment, trois objectifs spécifiques ont été définis :

1. Identifier quels articles initiaux posent, par cas d'homographie, un problème potentiel au niveau du repérage des titres et en estimer la fréquence
2. Estimer le nombre total de notices bibliographiques difficilement accessibles par la présence d'un des ces mots/articles en début de champ
3. Établir si ce problème a un impact sur le taux de circulation de ces documents

On peut effectivement supposer que ces « disparitions » auront un impact négatif direct sur la circulation de ces documents. Comment l'utilisateur peut-il utiliser du matériel qu'il ne peut pas (ou très difficilement) repérer ? Comme le mentionnent Nielsen et Pyle, « if you can't find a book in the catalogue, you might as well not have it on the shelf » (Nielsen et Pyle 1995, 292).

La réalisation de ces trois objectifs permettra de mesurer empiriquement la portée de ce problème de repérage. Ces informations seront inestimables pour améliorer la qualité et le taux de succès du repérage dans les systèmes d'information à contenu multilingue, tant au niveau de la formation des usagers que du design et de la conception des systèmes de repérage, notamment au niveau du paramétrage des algorithmes de détection des mots vides dans les requêtes.

3. Méthodologie

Étant donné le caractère exploratoire de cette étude et des moyens limités qui étaient à notre disposition, nous ne pouvions procéder à une étude générale sur un ensemble varié de catalogues. Nous avons par conséquent choisi de réaliser une étude de cas qui pourrait ultérieurement servir de tremplin à une étude plus exhaustive. Nous avons choisi Atrium, le catalogue des bibliothèques de l'Université de Montréal, pour réaliser cette étude de cas <<http://www-atrrium.bib.umontreal.ca:8000/>>. De plus nous avons délimité notre recherche en ne recherchant que les titres de monographies en excluant également les titres de collection de monographies.

Afin de rencontrer les trois objectifs susmentionnés, les trois questions de recherche suivantes ont été formulées :

- Question 1 : Quels sont les articles initiaux les plus susceptibles de poser problème lors du repérage par titre basé sur la liste d'exclusion utilisée dans Atrium ?
- Question 2 : Quelle est la proportion des notices affectées par un classement inexact dans l'index des titres ?
- Question 3 : Quelle est la corrélation entre le taux de circulation de ces ouvrages et le fait qu'ils soient plus difficiles à repérer ?

La collecte des données a été effectuée en prenant au départ les 41 articles identifiés dans la liste d'exclusion utilisée par l'algorithme de détection paramétré dans Atrium.¹ Cette liste s'affiche à l'accès de la version Telnet d'Atrium <atrrium.bib.umontreal.ca>.

| | | | | | |
|-----|------|------|-------|-----|-----|
| a | e | ena | henas | le | um |
| ai | een | enas | het | les | uma |
| al | eene | gli | hoi | li | un |
| an | ein | hai | i | lo | una |
| das | eine | heis | l' | los | une |
| der | eis | hen | la | oi | uno |
| die | el | hena | las | the | |

À noter que cette liste a été développée localement selon les besoins internes de la collection et ne représente qu'une partie des articles initiaux listés dans l'annexe E des Règles de catalogage anglo-américaines (RCAA 2003, 140–143). Nous avons travaillé à partir de la liste locale puisque nous ne pouvions effectuer d'analyse que sur les articles faisant partie de la liste d'exclusion.

Afin de répondre à la première question de recherche chaque article a fait l'objet d'une recherche individuelle en mode parcourir index. La méthode consistait, pour chacun des articles, à se positionner dans l'index des titres à partir du premier titre débutant l'article. L'index était systématiquement et exhaustivement parcouru pour recenser toutes les notices d'ouvrages débutant par un mot homographe n'étant pas un article initial et qui par conséquent serait difficilement repérable. Pour chaque ouvrage identifié la notice MARC a été examinée afin de déterminer dans quel champ se trouvait le mot problématique en question.

Lors de l'examen de la notice MARC le numéro de notice se trouvant dans le champ 001 a également été enregistré afin de pouvoir déterminer le nombre total de notices affectées. À noter que ce nombre était inférieur au nombre d'entrées problématiques retrouvées

dans l'index titre car une notice pouvait contenir plusieurs erreurs et générer deux ou plusieurs entrées problématiques dans l'index. Afin de répondre à la deuxième question de recherche nous avons simplement compilé le total des notices uniques affectées et nous l'avons comparé au nombre total de notices de monographies contenues dans Atrium au moment de la recherche (été 2004). Ces données ont été facilement obtenues du bureau des systèmes.

Les codes à barres de tous les exemplaires reliés à chaque notice en cause ont également été compilés afin de procéder à l'extraction des statistiques de circulation de ces documents. Ces données sont nécessaires pour répondre à notre troisième question de recherche puisqu'il faudra établir des comparaisons entre le taux de circulation des documents affectés et celui des documents en général. Ceci permet de vérifier statistiquement les différences entre les deux groupes en comparant le nombre de prêts annuels moyen obtenus par notice affectée et le nombre de prêts annuels moyen pour toutes les notices de monographie pouvant circuler du catalogue.

4. Observations et analyse

Lors de la collecte de données, qui a eu lieu entre le 5 juillet et le 6 août 2004, nous avons recensé 6 360 entrées problématiques dans l'index des titres. Le tableau 1 présente le nombre d'entrées affectées pour chaque article recherché.

Nous avons identifié six endroits où ces problèmes se trouvent. Il s'agit des champs MARC 240, 245, 246, 700, 740, et 730. L'analyse démontre que le champ dans lequel on retrouve le plus grand nombre d'entrées problématiques est le champ 245 (titre bibliographique). À lui seul il est la source de 5 360 entrées affectées, ce qui représente une proportion 84,3 % du nombre total. À noter que dans cette recherche, nous nous sommes attardés sur les titres de monographies et n'avons pas pris en compte les titres des collections, étant donné que ceux-ci sont extrêmement nombreux et que la période pour effectuer la collecte de données était plutôt courte. Toutefois, comme nous avons examiné les titres de collection pendant une courte durée au début de la période de collecte de données, nous avons pu constater que le nombre d'anomalies dans cette catégorie de titres est extrêmement élevé.

L'examen sommaire des données du tableau 1 indique clairement que certains articles sont beaucoup plus problématiques que d'autres. En fait, près de la moitié des articles de notre liste d'exclusion ne génèrent aucun problème. Par contre il est intéressant de noter que l'article « a » à lui seul génère 4 230 entrées problématiques dans l'index (soit plus de 66 % du nombre total). Ceci s'explique en grande partie par le fait que l'article « a » est très courant en anglais (et même dans certaines autres langues) et est également une préposition utilisée très fréquemment en français. Par exemple au moment de la collecte de données il y avait dans le catalogue 1 205 ouvrages débutant avec les mots « À propos de... ». Toutes ces entrées causent des problèmes de repérage en mode parcourir. À noter aussi quelques problèmes lorsque le titre débute avec l'initial d'un prénom commençant par la lettre « A » (*A.B.C. contre Poirot*) de même que pour de nombreux acronymes (A.A.C.R., A.B.B.), sans oublier les nombreux ouvrages débutant par A.B.C. (*A.B.C. de la lecture*). Il est certain qu'étant donné la forte proportion d'ouvrages de langue française se trouvant dans le catalogue Atrium par rapport à la proportion que l'on retrouverait dans d'autres catalogues d'institutions anglophones, ce problème est probablement exacerbé dans notre cas.

| Champ Article | 240 | 245 | 246 | 740 | 700 | 730 | Total | Prop. (%) |
|--------------------------|------------|--------------|------------|--------------|------------|------------|--------------|----------------------|
| a | 9 | 3519 | 172 | 368 | 157 | 5 | 4 230 | 66,5 |
| e | 29 | 391 | 34 | 101 | 5 | 0 | 560 | 8,8 |
| i | 4 | 377 | 36 | 63 | 53 | 0 | 533 | 8,4 |
| la | 0 | 185 | 5 | 55 | 21 | 3 | 269 | 4,2 |
| le | 0 | 222 | 1 | 5 | 1 | 0 | 229 | 3,6 |
| an | 5 | 121 | 3 | 6 | 28 | 0 | 163 | 2,6 |
| al | 9 | 78 | 1 | 8 | 2 | 1 | 99 | 1,6 |
| el | 0 | 56 | 0 | 0 | 0 | 0 | 56 | 0,9 |
| ai | 0 | 33 | 1 | 19 | 1 | 0 | 54 | 0,8 |
| los | 0 | 40 | 0 | 0 | 0 | 0 | 40 | 0,6 |
| un | 0 | 28 | 0 | 12 | 0 | 0 | 40 | 0,6 |
| um | 0 | 24 | 0 | 0 | 0 | 0 | 24 | 0,4 |
| las | 3 | 11 | 2 | 2 | 5 | 0 | 23 | 0,4 |
| li | 0 | 17 | 0 | 3 | 3 | 0 | 23 | 0,4 |
| the | 0 | 5 | 0 | 0 | 0 | 0 | 5 | 0,1 |
| het | 0 | 1 | 0 | 3 | 0 | 0 | 4 | 0,1 |
| uno | 0 | 4 | 0 | 0 | 0 | 0 | 4 | 0,1 |
| der | 0 | 1 | 0 | 0 | 0 | 0 | 1 | <0,1 |
| ein | 1 | 0 | 0 | 0 | 0 | 0 | 1 | <0,1 |
| eis | 0 | 1 | 0 | 0 | 0 | 0 | 1 | <0,1 |
| les | 0 | 1 | 0 | 0 | 0 | 0 | 1 | <0,1 |
| <i>19 autres</i> | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Total | 300 | 5 360 | 501 | 1 385 | 976 | 739 | 6 360 | 100 |

Tableau 1 : Nombre d'entrées affectées dans l'index par article.

Selon les données présentées dans le tableau 1, on peut voir que le nombre total d'entrées problématiques dans l'index des titres est de 6 360. En recoupant ces données avec les numéros de notices recueillies lors de la collecte des données nous avons trouvé que ces entrées provenaient de 5 111 notices bibliographiques distinctes du catalogue. Ceci s'explique par le fait qu'une notice peut correspondre à plus d'une entrée dans l'index des titres, par exemple une notice qui aurait un titre problématique dans le champ 245 et un autre dans le champ 740.

Nous avons estimé, grâce aux données fournies par le bureau des systèmes des bibliothèques de l'Université de Montréal, que le nombre total de notices de monographies dans Atrium au moment de la collecte des données était d'environ 1 317 741 notices. Nous pouvons donc estimer que la proportion des notices de monographies affectées par l'algorithme de détection des articles initiaux est légèrement inférieure 0,4% (voir tableau 2). Il faut toutefois mentionner que cette proportion tient compte uniquement des titres propres et que ce nombre serait sans aucun doute plus élevé si les titres de collections étaient considérés.

| | |
|---|-----------|
| Nombre d'entrées problématiques dans l'index des titres | 6 360 |
| Nombres de notices affectées | 5 111 |
| Nombre total de notices de monographies dans Atrium | 1 317 741 |
| Proportion des notices affectées | 0,388 % |
| Nombre d'items affectés | 6 101 |

Tableau 2 : Tableau des entrées, notices et items affectés.

Dans le but de trouver une réponse à notre troisième question de recherche concernant le taux de circulation des items affectés, les codes à barres des 6 101 items rattachés aux 5 111 notices identifiées ont été compilés. Cette liste a été transmise au bureau des systèmes des bibliothèques afin d'obtenir les statistiques de circulation pour une durée de 12 mois. Bien que nous puissions en toute logique conjecturer que le taux de circulation des items identifiés sera négativement affecté en raison de leur moins grande visibilité. Toutefois, des circonstances hors de notre contrôle ont fait en sorte que nous n'avons pu pour l'instant obtenir de données fiables pour procéder à notre analyse comparative des taux de circulation des items affectés versus celui de la collection en général. Nous espérons pouvoir procéder à la portion de cette analyse dans un avenir rapproché afin de vérifier notre hypothèse.

5. Conclusion et avenues de recherche

Cette étude nous a fourni des données empiriques essentielles pour comprendre le phénomène du repérage des titres et du classement des articles initiaux dans les systèmes de repérage d'information automatisés. La revue de la littérature effectuée lors de la préparation de ce projet démontre que ce problème est peu étudié et qu'il est important de s'attarder plus à fond sur ce phénomène. La recherche par titre constitue encore le type de recherche le plus fréquent dans les catalogues de bibliothèques. Il est donc souhaitable qu'il soit le plus efficace possible. Les résultats de l'étude démontrent que l'application de l'algorithme de détection des articles initiaux dans les requêtes n'affecte qu'une faible proportion (moins d'un demi pourcent) des notices bibliographiques contenues dans le catalogue en ligne de l'Université de Montréal, Atrium. Bien que cette proportion puisse sembler dérisoire il s'agit toutefois de plus de 5 000 notices qui sont beaucoup moins visibles lors de la recherche en mode parcourir dans l'index des titres. Ce nombre est non négligeable si on considère les coûts d'acquisition et de traitement.

Il est intéressant de constater que sur les 41 articles examinés, l'article « a » est responsable à lui seul des deux tiers des problèmes rencontrés lors de cette étude. On pourrait par conséquent facilement régler une large proportion des problèmes simplement en retirant cet article de la liste d'exclusion. En outre, l'élimination complète de la liste d'exclusion aurait comme conséquence de solutionner la totalité des problèmes observés. Toutefois on peut avancer l'argument que cette élimination risque de créer d'autres problèmes au moment de la recherche, notamment si les usagers incluent par mégarde ou par ignorance les articles initiaux dans leurs requêtes lors de la recherche par titre. En conséquence, le dilemme concernant le maintien ou le retrait de la liste d'exclusion persiste : soit maintenir la liste d'exclusion, ce qui permet de rediriger correctement les requêtes contenant un article initial; soit éliminer la liste d'exclusion, pour éviter de perdre la trace des titres commençant par un mot « normal » mais homographe à un article. Nous espérons que les résultats de notre étude, qui indiquent que l'application d'une liste d'exclusion affecte négativement une petite proportion, mais somme toute non négligeable de notices, quant à leur visibilité dans l'index des titres, fera en sorte que la

décision concernant ce dilemme pourra être prise de façon plus éclairée par les responsables concernés.² Suite aux résultats obtenus dans cette étude, nous préconisons d'éviter d'utiliser une liste d'exclusion sur laquelle se basent les algorithmes de détection d'articles initiaux puisque cette procédure pénalise l'utilisateur lors des recherches par titre. Il nous semble préférable d'inclure simplement des instructions claires directement visibles sur l'interface de recherche spécifiant d'omettre l'article initial au moment de la recherche tel qu'illustré à la figure 4.

Cette étude de cas a permis de chiffrer précisément l'étendue du problème dans le catalogue Atrium. Il sera éventuellement possible d'étendre la recherche à d'autres catalogues et environnements, ou encore de mesurer l'impact direct sur l'utilisateur en milieu de recherche réel. Notre étude nous a permis d'identifier les articles initiaux causant les problèmes, et d'évaluer la proportion d'enregistrements affectés par ce problème mais ne nous renseigne pas sur les comportements de recherche face à ce problème. Par conséquent il serait intéressant, comme complément à notre étude, d'examiner l'étendue des problèmes de repérage du point de vue de l'utilisateur. Une telle étude permettrait d'identifier concrètement de quelle manière l'utilisateur effectue ses recherches dans l'index des titres des catalogues de bibliothèques lorsque les titres débutent par un article initial. En outre, il serait possible dans cette étude de mesurer le taux de succès de l'utilisateur lors du repérage des notices bibliographiques difficilement accessibles par la présence d'un des ces mots/articles en début de champ. On pourrait également établir si le problème en cause a un impact sur le niveau d'efficacité du repérage des monographies (temps et effort requis pour effectuer une tâche de repérage) lors de l'utilisation de l'index de titres. Il serait ainsi possible de mesurer empiriquement la portée de ce problème de repérage en situation réelle. Ces informations seraient fort utiles pour améliorer l'efficacité et le taux de succès du repérage dans les systèmes d'information à contenu multilingue, tels les catalogues, tant au niveau de la formation des usagers que du design et de la conception des systèmes de repérage.

Références

- Bourne, Charles P. 1975. Initial article filing in computer-based book catalogs: techniques, problems, and article frequency. *Journal of Library Automation*, 8 (3), 221–47.
- Library of Congress. 2005. MARC Standards.
Disponible à : <<http://www.loc.gov/marc/>>, visité le 2005-04-01.
- Nielsen, R. et J. Pyle. 1995. Lost articles: filing problems with initial articles in databases. *Library Resources and Technical Services*, 39 (3), 221–22.
- Notess, Greg. 2001. Tracking title search capabilities. ONLINE on the net.
Disponible à : <http://www.onlinemag.net/OL2001/net5_01.html>, visité le 2005-04-01.
- RCAA. 2003. *Règles de catalogage anglo-américaines*, 2^e éd., rév. de 1998, modifications de 2001-2002-2003, Montréal, Asted.
- Seeman, Corey. 2002. RE: Skipping initial articles.
Disponible à : <<http://innovativeusers.org/list/archives/2002/msg02463.html>>, visité le 2005-04-01.
- Taylor, Arlene G. 2004. *The Organization of Information*, 2nd ed. Westport, Conn.: Libraries Unlimited.

Notes

¹ À noter que « ai » est un article contracté en italien et ne devrait pas en principe figurer sur cette liste. À noter également que l'article « l' » n'a pas été étudié dans le cadre de cette recherche, étant donné la difficulté de le retrouver dans le classement d'Atrium.

² À noter que dans le cas qui nous concerne le maintien de la liste d'exclusion est essentiel pour le bon fonctionnement du module d'acquisition dans Atrium. Ainsi, il arrive que dans certaines circonstances la décision de maintenir ou d'enlever cette liste puisse être déterminée par des motifs d'ordre purement technique.