

La difficulté d'accéder aux adresses des chercheurs
français dans les bases de données bibliographiques
L'exemple d'une université lyonnaise

Pascal Bador, Thierry Lafouge

► To cite this version:

Pascal Bador, Thierry Lafouge. La difficulté d'accéder aux adresses des chercheurs français dans les bases de données bibliographiques L'exemple d'une université lyonnaise. Documentaliste - Sciences de l'Information, ADBS, 2006, 43 (1), pp.28-35. <sic_00161084>

HAL Id: sic_00161084

https://archivesic.ccsd.cnrs.fr/sic_00161084

Submitted on 23 Aug 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ÉTUDE

La difficulté d'accéder aux adresses des chercheurs français dans les bases de données bibliographiques L'exemple d'une université lyonnaise

Par PASCAL BADOR et THIERRY LAFOUGE

Note sur les auteurs page 2

Pascal Bador, maître de conférences, et **Thierry Lafouge**, professeur, sont membres de l'équipe **URSIDOC (Unité de recherche en sciences de l'information et du document), EA 3718**.

Auteur à qui la correspondance doit être adressée : Pascal Bador, Centre de documentation pharmaceutique, ISPB-Faculté de pharmacie, Université Claude Bernard Lyon 1, 8 avenue Rockefeller, F-69373 Lyon Cedex 08, France, pascal.bador@univ-lyon1.fr, téléphone +33 (0)4 78 77 70 66

Adresse de Thierry Lafouge : URSIDOC, Batiment Omega, Université Claude Bernard Lyon 1, 43 boulevard du 11 novembre 1918, F-69622 Villeurbanne Cedex, France, thierry.lafouge@univ-lyon1.fr

Les auteurs tiennent à remercier Jean-François Jal, professeur des universités, responsable du projet d'évaluation de la production scientifique des établissements du pôle universitaire de Lyon de les avoir sensibilisés pour mener cette étude.

Chapeau

Version courte

La difficulté d'accéder aux adresses des chercheurs français dans les bases de données bibliographiques est sans doute une des causes de la faible visibilité des universités françaises dans les classements internationaux. À partir de l'exemple de l'Université Claude Bernard Lyon 1, cette étude vise à démontrer que cette situation pourrait en partie s'expliquer par le manque de rigueur des chercheurs dans la rédaction de l'adresse qu'ils font figurer en tête de leurs publications scientifiques. Une autre cause étant l'absence de libellé officiel et normalisé de ces adresses. Des recommandations formulées *in fine* pourront aider à valoriser au plan international les travaux des chercheurs français.

Version longue

La difficulté d'accéder aux adresses des chercheurs français dans les bases de données bibliographiques est sans doute une des causes de la faible visibilité des universités françaises dans les classements internationaux. Fondée sur l'interrogation de la rubrique « adresse » de Medline, Embase.com et Pascal Biomed sur l'année 2003 à l'aide de requêtes permettant de trouver les publications de l'Université Claude Bernard Lyon 1, cette étude vise à démontrer que cette situation pourrait en partie s'expliquer par le manque de rigueur des chercheurs dans la rédaction de l'adresse qu'ils font figurer en tête de leurs publications scientifiques. Une autre cause étant l'absence de libellé officiel et normalisé de ces adresses. Des recommandations formulées *in fine* pourront aider à valoriser au plan international les travaux des chercheurs français.

Texte

METHODE DE MESURE DE L'ACTIVITE SCIENTIFIQUE par le dénombrement des publications des chercheurs, la bibliométrie est désormais bien connue pour ses usages à des fins d'évaluation et de repérage de nouveaux thèmes de recherche ou de partenaires potentiels. Ainsi le « positionnement scientifique » des groupes et équipes de recherche, mais aussi des institutions, évaluées par des études bibliométriques (ou scientométriques), devient un enjeu majeur tant pour les individus que pour leurs organismes de tutelle [1].

Dans ce contexte, le récent classement mondial des cinq cents meilleures universités, réalisé en 2003 puis les années suivantes par l'Université de Shanghai [2], donne une piètre image de la science française. Il apparaît en effet que la première université française figurant dans ce palmarès, celle de Paris 6, n'est placée qu'en 65^e position en 2003, 41^e en 2004 et 46^e en 2005. Et sa consœur d'Orsay (Paris 11) figure en 72^e position en 2003, 48^e en 2004 et 61^e en 2005. Plus inquiétant encore, le classement des cinquante premières universités européennes, toujours selon celle de Shanghai, ne compte que quatre universités françaises et l'École normale supérieure de Paris.

Pour élaborer ce classement, les chercheurs de Shanghai se sont appuyés pour chaque université sur les critères suivants [2] [3] :

- le nombre de prix Nobel issus de cette université ;
- la fréquence de citation des travaux des chercheurs qui y exercent ;
- le nombre d'articles publiés par ces chercheurs dans les deux revues considérées comme les plus prestigieuses du monde, *Science* et *Nature* ;
- le nombre d'articles publiés qui sont indexés dans les bases de données bibliographiques Science Citation Index et Social Science Citation Index de l'Institute of Scientific Information (ISI) ;
- un indicateur de performance qui consiste à diviser les indicateurs précédents par le nombre de chercheurs de l'université concernée en « équivalent temps plein ».

Aussi contestable soit-il, ce palmarès annonce quand même le futur : les organismes de recherche seront dorénavant classés en fonction de leur visibilité dans les bases de données bibliographiques internationales. De plus, ce classement ne fait que confirmer la situation actuelle de malaise dans laquelle se trouve le monde de la recherche en France. Par ailleurs, si la publication de ces résultats peut être l'occasion de faire prendre conscience que la recherche et les universités françaises sont dans une position d'infériorité par rapport à celles d'autres pays (notamment pour des raisons structurelles liées en partie à un sous-financement), alors elle peut être salutaire [4] !

Pour autant, puisque le nombre d'articles publiés par les chercheurs est toujours reconnu comme un critère important d'évaluation de leurs activités, il nous a semblé intéressant d'examiner si la très mauvaise visibilité des articles publiés par nos scientifiques n'était pas due à une insuffisance de précision dans le signalement de l'institution dont il relève. En effet, l'éclatement de la recherche française entre les universités et les autres organismes de recherche comme les EPST (établissements publics à caractère scientifique et technologique : CNRS, INSERM, INRA, etc.) ou les hôpitaux dessert gravement la mesure des activités scientifiques, notamment pour les sciences de la vie et la médecine [4] [5]. Il ne fait pourtant aucun doute que les chercheurs français publient dans les revues habituellement prises en compte dans les systèmes bibliographiques comme les Current Contents, Science Citation Index, Medline, Embase, Pascal, etc., qui sont les outils utilisés pour les analyses bibliométriques.

Dans ces conditions, notre problématique a été d'étudier si effectivement les chercheurs de l'une des plus grandes universités non parisiennes, l'Université Claude Bernard Lyon 1 (UCBL), sont

facilement identifiables à partir de leur université d'appartenance ou si au contraire un certain manque de visibilité de leurs travaux tient par exemple au fait qu'ils appartiennent également à d'autres organismes comme le CNRS, l'INSERM ou le CHU de Lyon. Pour ce faire, nous avons étudié pour l'année 2003 les libellés des adresses signalant l'UCBL dans les notices bibliographiques des articles répertoriés dans les bases Medline, Embase et Pascal Biomed afin de voir si cette université, *a priori* représentative d'une partie de la recherche en France, souffrait ou non d'une sous-visibilité nuisible à sa réputation et à sa reconnaissance au niveau international.

1 Matériel et méthodes

La visibilité des travaux des chercheurs de l'Université Claude Bernard Lyon 1 a été explorée à partir des trois bases de données bibliographiques suivantes qui sont parmi les plus utilisées par les scientifiques du domaine biomédical :

- Medline interface Pubmed qui répertorie environ 14 millions de références extraites de 4.760 périodiques de 70 pays depuis 1953, base produite par la National Library of Medicine (NLM) [6] [7] ;
- Embase.com qui répertorie environ 10 millions de références extraites de 4.524 périodiques depuis 1974, base produite par Elsevier [7] ;
- Pascal Biomed qui répertorie environ 4,5 millions de références extraites de 2.000 périodiques depuis 1987, base produite par l'Institut de l'information scientifique et technique (INIST-CNRS) [8].

Notre objectif étant d'évaluer de façon globale et qualitative la visibilité de l'UCBL en 2003 à travers les libellés des adresses figurant dans les publications des chercheurs, nous avons interrogé dans un premier temps la rubrique « Adresse » ou « Affiliation » des trois bases mentionnées ci-dessus à l'aide de l'équation de recherche la plus large possible telle qu'elle est présentée dans le tableau 1.

Nous avons volontairement choisi une stratégie de recherche très large, avec comme seul critère textuel la ville de Lyon, afin de ne pas cibler la recherche sur un libellé trop spécifique. Les requêtes « *Claude Bernard* » ou « *Lyon 1* », par exemple, auraient entraîné un risque de silence important. Sachant également qu'une recherche sur les composantes de l'université (faculté de médecine, faculté de pharmacie, UFR de chimie, UFR de biologie, etc.) aurait été longue et délicate, cette recherche générale nous a permis d'obtenir un important échantillon de notices bibliographiques (plus de mille pour chaque base de données, voir le tableau 1).

La contrepartie a évidemment été de devoir gérer un bruit documentaire important avec les publications des universités Lyon 2 et Lyon 3, des hôpitaux, des grandes écoles et des organismes de recherche divers établis à Lyon, tels que le Centre international de recherche contre le cancer. Ne pouvant consulter ces presque quatre mille références, nous avons alors décidé de sélectionner manuellement dans chacune des trois bases de données les cent premières références relatives à l'UCBL, ce qui nous a permis d'étudier les libellés des affiliations professionnelles d'un total de trois cents notices bibliographiques.

Afin d'identifier les différents libellés utilisés par les chercheurs de l'UCBL et parmi eux quels étaient les plus pertinents, nous avons ensuite interrogé pour l'année 2003 la rubrique « Adresse » ou « Affiliation » des trois bases de données à l'aide de requêtes utilisant les différentes formulations suivantes :

- « lyon 1 »
- « lyon I »

- « claud bernard »
- « claud bernard » and university
- « universit lyon 1 »
- « universit lyon I »
- « universit claud bernard »
- « faculte de medecine » and lyon
- (« faculte de medecine » and lyon) not « claud bernard »
- (« faculte de medecine » and lyon) not (« claud bernard » or « lyon 1 » or « lyon I »)

2 Résultats

Sur les trois cents notices bibliographiques sélectionnées comme indiqué précédemment et en prenant en compte les différentes écritures possibles (majuscules et minuscules, présence ou non d'accents et de points, abréviations, français ou anglais), nous avons trouvé un total de 44 types de libellés d'adresses différents. Il est évident que ces nombreux libellés nuisent considérablement aux recherches émanant de l'Université Claude Bernard Lyon 1. Le tableau 2 présente les 30 libellés permettant de trouver l'UCBL à partir des deux requêtes les plus intuitives pour un utilisateur, à savoir « lyon 1 » et « claud bernard ». Le tableau 3 présente les 14 libellés ne donnant aucune réponse à ce même type de requête.

Pour aller plus loin dans notre tentative d'explication de la mauvaise visibilité des universités françaises dans les bases de données bibliographiques, nous avons ensuite réparti les trois cents références étudiées selon le caractère complet ou incomplet des adresses qu'elles mentionnaient. Pour ce faire nous avons considéré comme complète toute adresse conforme au modèle suivant :

- Auteur
- Laboratoire et/ou équipe de recherche et/ou composante de rattachement (faculté, institut, etc.)
- Université Claude Bernard Lyon 1
- Rue
- Code postal Ville Cedex

Le tableau 4 est très instructif : il révèle que seules 15 % des adresses rencontrées sont complètes ! Ce tableau montre également qu'une moyenne de 44 % des adresses ne fait aucune mention de l'Université Claude Bernard Lyon 1.

Le tableau 5 présente le nombre de notices bibliographiques obtenues lors de requêtes employant dix formulations qui pourraient être utilisées par un évaluateur extérieur (de Shanghai, exemple pris au hasard !) ou toute personne cherchant un contact ou une collaboration avec l'UCBL.

Dans le but d'explorer notre problématique sous l'angle du fonctionnement des bases de données bibliographiques, nous avons défini dans le tableau 6 les caractéristiques relatives au mode de présentation des adresses des auteurs dans les bases Medline, Embase et Pascal Biomed. Nous indiquons également les caractéristiques propres à leurs logiciels de recherche respectifs en ce qui concerne la prise en compte des majuscules, des accents, des tirets et des points qui sont des caractères pouvant influencer sur les résultats des recherches documentaires.

3 Discussion

La recherche française peu visible dans les classements internationaux

Première université non parisienne par le nombre de ses étudiants (28.000), l'Université Claude Bernard Lyon 1 est une université scientifique, technologique et de santé. Elle apparaît en 338^e position sur 500 dans le classement mondial de Shanghai en 2003, 265^e en 2004 et 266^e en 2005 [2], ce qui fait d'elle la quatorzième des 21 universités françaises de ce classement. Même si les résultats de celui-ci ont été assez peu commentés en France, notre objectif n'est pas de revenir sur les considérations politiques, structurelles ou financières qui peuvent en partie expliquer la quasi-inexistence des universités françaises dans ce type de palmarès. On peut par ailleurs évidemment noter que celui-ci fait la part belle aux universités américaines dont plus d'une trentaine sont présentes parmi les 50 premières de ce classement.

Quels que soient les divers critères utilisés dans l'élaboration des différents classements des universités au niveau mondial, les critères relatifs au nombre de publications des chercheurs ou au nombre de citations dont ils font l'objet sont toujours pris en compte. Ces critères, aussi objectifs soient-ils, utilisent comme élément principal le référencement des articles de recherche dans les bases de données bibliographiques dont la plupart sont en réalité d'origine américaine : bases de l'ISI (Science Citation Index, Current Contents, Journal Citation Reports avec les facteurs d'impact des revues) ; Medline ; Biosis ; etc. Il ne fait aucun doute que ces bases de données prennent en compte prioritairement les revues américaines et d'origine anglo-saxonne au détriment des revues européennes notamment [9].

Pour autant, et même si ce contexte semble bien connu des chercheurs français dont la plupart s'adaptent parfaitement en publiant leurs travaux dans les revues prises en compte dans ces outils bibliographiques, nous avons voulu voir si ce déficit de visibilité et de reconnaissances de la recherche française n'était pas dû à des causes liées à de purs problèmes de forme et de présentation au sein même des articles. Il nous paraît en effet évident que pour pouvoir trouver et localiser les auteurs d'un article, il faut que ceux-ci commencent par indiquer leur adresse complète de façon correcte, tant du point de vue de l'écriture du libellé de leur établissement d'affectation (université en l'occurrence) que du point de vue de la conformité avec les règles d'écriture d'une adresse imposée par la poste.

Pour un libellé officiel unique de l'appellation de chaque université française

Notre étude montre que pour les auteurs de l'UCBL (et sans doute par extrapolation pour un grand nombre d'auteurs français) la réalité est tout autre. Parmi les trois cents références étudiées, l'appellation officielle « Université Claude Bernard Lyon 1 » est loin de faire l'unanimité dans les adresses rédigées par les auteurs. En effet, 43 autres libellés ont été mis en évidence (tableaux 2 et 3) si l'on prend en compte les différents types de caractères utilisés (majuscules ou minuscules, chiffres arabes ou romains, présence ou non de tiret ou de points) ou la formulation du libellé (ordre des termes, abréviations **et** sigles, langue utilisée : français ou anglais). En un mot l'indiscipline et le manque de rigueur règnent en matière d'écriture d'adresses, ce qui pose justement un grave problème pour retrouver l'UCBL à l'aide des outils informatiques.

En effet, ceux-ci ne sont capables de retrouver que des chaînes strictes de caractères successifs, ce qui **empêche une réponse** lorsque la chaîne demandée (« Université Claude Bernard », par exemple) n'est pas strictement identique à la chaîne présente dans l'adresse de la notice bibliographique (« Univ. Claude B. », « Claude Bernard University », « UCBL », « Université Claude-Bernard », par exemple). Le problème est aggravé encore du fait de la double appellation « Université Claude Bernard » et « Université Lyon 1 » pour laquelle on retrouve les mêmes ambiguïtés et à laquelle s'ajoute l'utilisation du chiffre romain (« Université Lyon I »). Une recherche rapide montre d'ailleurs des résultats comparables avec l'Université Pierre et Marie Curie Paris 6 (Paris VI) ou l'Université Paris 11 - Paris Sud Orsay (Paris XI).

Les conséquences de ces différentes appellations de nos universités, pour lesquelles aucune **dénomination** officielle ne semble respectée (si tant est qu'une appellation officielle ait effectivement été donnée ou imposée), sont donc importantes puisqu'elles sont à l'origine d'un important silence documentaire. Cela peut contribuer à expliquer d'une part la sous-visibilité des universités françaises dans les bases de données et d'autre part, par ricochet, la faible fréquence de citations des articles français ; le tout entraînant un mauvais classement lors des évaluations mondiales. Le Comité national d'évaluation de la recherche (CNER) avait d'ailleurs identifié ce problème du caractère imprécis ou incomplet des adresses des auteurs français dans son rapport de décembre 2002 [10].

On peut ainsi facilement comprendre qu'une personne extérieure ou étrangère au système de recherche français ne pousse pas **l'identification** des universités françaises jusqu'à celle de tous les synonymes et libellés possibles. Même si cela peut paraître fastidieux **et contraignant**, un minimum de discipline et de standardisation s'impose donc à nos chercheurs pour l'écriture des libellés de leur université de rattachement s'ils veulent valoriser leurs travaux. Ces exigences paraissent extrêmement simples et faciles à mettre en œuvre pour peu qu'une sensibilisation à l'utilisation stricte et obligatoire d'appellations officielles et non ambiguës des universités soit **assurée** par les instances dirigeantes.

Le manque de rigueur des chercheurs dans l'indication de leur affiliation

Nous avons vu dans le tableau 4 que seulement 15 % des adresses étaient libellés de façon correcte et complète dans les notices bibliographiques. Cela paraît très peu et deux raisons expliquent ce faible pourcentage d'adresses correctes. Tout d'abord les auteurs eux-mêmes semblent peu sensibilisés à l'intérêt de rédiger une adresse selon les règles imposées par la poste (qui garantit dans ces conditions fiabilité et rapidité dans l'acheminement du courrier). Notre étude montre en effet que si l'on considère les cent notices bibliographiques issues de Embase (celle-ci étant la plus fiable des trois bases de données considérées car elle est la seule à utiliser un strict copié-collé des adresses à partir des articles), seuls 26 % des adresses sont complètes. Dans 15 % des cas les auteurs ont omis le cedex, alors qu'il nous apparaît pour le moins surprenant que 13 % des adresses ne mentionnent ni rue, ni code postal. Ces adresses incomplètes sont sans doute pénalisantes pour les personnes ou organismes étrangers souhaitant entrer en contact pour d'éventuelles collaborations avec nos chercheurs, même s'il est vrai qu'aujourd'hui le courriel est très largement utilisé.

Par ailleurs une moyenne de 44 % des adresses étudiées ne fait absolument pas mention de l'Université Claude Bernard Lyon 1, même si certaines d'entre elles sont correctes d'un point de vue postal, et sur ce point les données issues des trois bases Medline, Embase et Pascal se rejoignent parfaitement. Cela veut dire que près de la moitié des articles de notre université ne peuvent être trouvés lors de requêtes simples sur le nom précis de l'établissement (« *Université Claude Bernard* » ou « *Université Lyon 1* ». Nous sommes là au cœur de la problématique étudiée, à savoir la sous-évaluation des universités françaises qui semble due au moins pour partie au fait que nos chercheurs universitaires ne mentionnent pas systématiquement leur université d'appartenance dans le référencement de leurs travaux ; la plupart se contentant d'indiquer leur laboratoire, faculté ou institut (chimie, médecine, pharmacie, etc.), ce qui n'est malheureusement pas **révélateur** de l'appartenance à l'université pour une personne extérieure ou étrangère.

Encore plus difficiles à repérer sont les articles publiés par des universitaires ayant une double ou une triple affiliation. C'est notamment le cas des praticiens chercheurs hospitalo-universitaires qui ne mentionnent que leur appartenance hospitalière ou des chercheurs d'équipes de recherche mixte (CNRS, INSERM, etc.) qui, eux non plus, ne font pas référence à leur université ou à leur EPST. Notre étude montre que 8 % des références bibliographiques sont

dans ce cas, ce qui est très certainement un chiffre fortement sous-évalué puisque nous n'avons pas la prétention de connaître et donc de pouvoir identifier l'ensemble des chercheurs et enseignants-chercheurs de l'UCBL (ce qui représente plus de deux mille personnes !).

Pour compléter nos investigations nous avons voulu voir à travers dix requêtes simples si des différences significatives concernant le nombre de notices bibliographiques trouvées pouvaient être mises en évidence. Là encore les résultats obtenus montrent bien que trouver l'ensemble des articles publiés par les chercheurs de l'UCBL se révèle quasi impossible dans le cadre d'une interrogation rapide faite par une personnalité étrangère connaissant peu l'organisation des universités françaises.

Le tableau 5 nous a en effet montré qu'il est indispensable que les requêtes utilisent une formulation prenant en compte à la fois « lyon 1 » et « lyon I » (ce qui correspond en réalité à l'équation : « lyon 1 » OU « lyon I »), sous peine de passer à côté d'au moins 50 % des articles publiés. Il semble en effet que toutes les universités des grandes villes puissent être trouvées à partir de libellés utilisant soit des chiffres arabes, soit des chiffres romains. La recherche à partir de la requête « *claudes bernard* » semble quant à elle peu pertinente car elle peut concerner tant l'UCBL qu'un hôpital, une rue ou un quai d'où un bruit documentaire rendant les résultats inexploitable (plusieurs centaines de références obtenues pour 2003). Les requêtes « *universite lyon 1* », « *universite lyon I* » semblent pour leur part trop restrictives pour pouvoir être largement utilisées.

En définitive la requête complexe : « *lyon 1* » OU « *lyon I* » OU « *universite claudes bernard* » semble la plus pertinente, sans pour autant offrir la garantie absolue de trouver l'ensemble des articles publiés par l'UCBL. En effet, la requête : (« *faculte de medecine* » ET *lyon*) SAUF (« *claudes bernard* » OU « *lyon 1* » OU « *lyon I* »), qui donne en résultat plusieurs dizaines de références selon les bases, permet de rejoindre la constatation évoquée plus haut selon laquelle de nombreux chercheurs mentionnent la faculté de médecine, par exemple, sans mentionner leur appartenance à l'UCBL.

Le manque de cohérence des bases de données dans le signalement des affiliations des organismes de recherche

Nous avons vu jusqu'à présent que les auteurs étaient largement responsables de la mauvaise visibilité de leur université d'appartenance. Pour autant, d'autres raisons illustrées par le tableau 6 permettent d'expliquer ce phénomène de façon complémentaire. Le mode de fonctionnement et de référencement des adresses par les bases de données doit maintenant être évoqué.

Ainsi, il faut distinguer les bases comme Pascal Biomed qui indiquent les noms de l'ensemble des auteurs d'un article avec mention de leurs adresses respectives. Malheureusement, les adresses en question sont souvent tronquées si la rubrique « Correspondance » de l'article n'est pas reprise intégralement, ce qui est le cas de Pascal Biomed. D'autres bases comme Medline et Embase ne mentionnent que l'adresse d'un seul auteur (le premier de la liste pour Medline, l'auteur responsable de la correspondance pour Embase), ce qui les rend inutilisables si l'on veut obtenir une évaluation objective des publications propres à un organisme donné. De plus, des fautes de frappe lors de la ressaisie des adresses peuvent là encore être à l'origine de biais dans l'évaluation du nombre d'articles publiés. Nous voyons donc que le choix des bases de données bibliographiques en vue d'évaluer le nombre de publications des chercheurs ou des institutions n'est pas anodin et qu'il faut absolument se renseigner au préalable sur leur mode de fonctionnement et de référencement.

Le dernier point que nous voulions évoquer concerne les caractéristiques de fonctionnement des logiciels d'interrogation des bases Medline, Embase et Pascal. Tous les trois autorisent des

requêtes permettant de négliger l'usage des majuscules, des minuscules et des accents ainsi que celui d'éventuels points au sein d'abréviations. Ainsi la requête « *Université Claude Bernard* » donnera les mêmes résultats que « *universite claud bernard* » ; de même « *univ claud bernard* » sera équivalente à « *univ. claud bernard* ». En revanche la prise en compte du tiret peut être différente selon les bases : Medline et Embase le négligent alors que Pascal Biomed le traite comme un caractère à part entière dans les requêtes. Ainsi pour Pascal Biomed la requête « *université lyon 1* » ne donne pas les mêmes résultats que « *université lyon-1* » qui est un libellé effectivement utilisé par certains auteurs. Ces subtilités de fonctionnement peuvent évidemment entraîner elles aussi des biais se traduisant par un silence documentaire plus ou moins important dans l'évaluation du nombre d'articles publiés par une université donnée.

4 Conclusion et recommandation

L'ensemble des causes du manque de visibilité de la science française dans les classements internationaux est certainement multiple et difficile à cerner, et mériterait certainement une étude globale et approfondie des nombreux paramètres pouvant entrer en jeu : politique scientifique au niveau national, structure et organisation des établissements de recherche, statut des chercheurs, mode de financement, choix des thèmes de recherche, partenariats et collaborations, adaptation des modes de publication aux méthodes d'évaluation, etc.

Pour autant, notre étude, bien que ponctuelle et limitée à un cas particulier (mais qui peut sans doute être extrapolé), a mis en évidence le fait que cette faible visibilité pouvait être due pour partie à un manque de sensibilisation des chercheurs à l'importance de mentionner l'adresse de leur(s) établissement(s) de rattachement selon des règles de rédaction strictes et normalisées. On peut noter à ce sujet l'initiative intéressante de l'INSERM qui a diffusé en 2005 à l'ensemble de ses chercheurs sa charte de publication [11]. À l'issue de cette étude, nous pouvons ainsi formuler les recommandations suivantes.

Recommandations à l'intention des universités :

- définir une charte graphique imposant un libellé officiel, unique et normalisé de l'appellation de l'université ;
- diffuser à l'ensemble de son personnel le modèle officiel et normalisé du libellé du nom de l'établissement et de son adresse postale ;
- sensibiliser les chercheurs à l'intérêt du respect de cette charte et de ce modèle normalisé tant pour l'université elle-même, les équipes de recherche, les chercheurs à titre individuel que sur un plan national et international.

Recommandations à l'intention des chercheurs :

- faire preuve d'une extrême rigueur dans l'écriture du libellé du nom de leur établissement ainsi que de l'adresse postale (respect du modèle officiel et normalisé avec usage intensif du copier-coller) sous peine d'une sous-valorisation de leurs travaux dans les classements internationaux ;
- à l'intention des chercheurs à double ou triple appartenance (universitaires, hospitaliers, EPST notamment) : mentionner absolument tous leurs établissements de rattachement selon le libellé officiel et normalisé de chacune des institutions.

REFERENCES

- [1] Cailly B, Goffi E. Colloque Ile de Science. Comment positionner une communauté scientifique : les enjeux de la bibliométrie. *Documentaliste - Sciences de l'information* 2004; vol. 41, n° 3, p. 190-192
- [2] Institute of Higher Education, Shanghai Jiao Tong University. Academic Ranking of World Universities - 2005. <http://ed.sjtu.edu.cn/ranking.htm> , consulté le 15/2/2006
- [3] Orivel F. Pourquoi les universités françaises sont-elles si mal classées dans les palmarès internationaux ? Mai 2004. www.u-bourgogne.fr/IREDU/2004/04006.pdf , consulté le 23/3/2006
- [4] Université René Descartes - Paris 5, Service de la communication. Classement mondial des universités : comparons ce qui est comparable.
www.univ-paris5.fr/fichier/NHRTWYDW5M2PUE8U.PDF, consulté le 23/6/2004
- [5] Haeffner-Cavaillon N. Benchmarking dans le domaine de la recherche bio-médicale. www.eva.inserm.fr/Bibliometrie/EtudeBenchmarking/Benchmarking.pdf, consulté le 14/2/2006
- [6] URFIST Paris. Introduction à Medline (quelques données utiles). www.ext.upmc.fr/urfist/biolo/bioguide2/medline/intro1.htm, consulté le 14/2/2006
- [7] Ratté S. Une analyse descriptive et comparative d'une sélection de bases de données bibliographiques (BDBs) du domaine biomédical. http://bibliothecaire.com/SRatte_BDB_20040604.pdf, consulté le 14/2/2006
- [8] INIST-CNRS. Pascal Biomed. http://bibliosciences.inist.fr/article.php?id_article=25, consulté le 14/2/2006
- [9] Bador P., Ben Romdhane M., Lafouge T. European pharmaceutical journals: Relationship between demand and indexation - the example of the main French document supplier. *The Canadian journal of information and library science*, 2003, vol. 27, n° 2, p. 33-55
- [10] Comité national d'évaluation de la recherche. Évaluation de la recherche publique dans les établissements publics français. Décembre 2002. www.cner.gouv.fr/fr/pdf/bib.pdf, consulté le 15/2/2006
- [11] INSERM. Charte des publications (Adresses - Affiliations). www.eva.inserm.fr/Bibliometrie/Visibilite/Modeles_adresses.pdf, consulté le 15/2/2006

TABLEAU 1 - STRATEGIES DE RECHERCHE DE LA MENTION DE L'UNIVERSITE CLAUDE BERNARD LYON 1 DANS LES TROIS BASES DE DONNEES ETUDIEES

| Base de données | Équation de recherche | Nombre de références obtenues |
|-----------------|--|-------------------------------|
| Medline Pubmed | lyon[ad] Field: All Fields, Limits: Publication Date from 2003/01/01 to 2003/12/31 | 1222 |
| Embase.com | Lyon:ad AND [embase]/lim AND [2003]/py | 1238 |
| Pascal Biomed | (lyon in AF) and (PY=2003-2003) | 1445 |

**TABLEAU 2 - LIBELLES D'ADRESSES REpondant
SOIT A LA REQUETE « CLAUDE BERNARD »,
SOIT A LA REQUETE « LYON 1 »**

| |
|------------------------------------|
| Universite Claude Bernard Lyon 1 |
| Universite Claude Bernard LYON 1 |
| Universite Claude Bernard (Lyon 1) |
| Universite Claude Bernard |
| Universite Lyon 1 |
| Universite de Lyon 1 |
| Universite Claude Bernard Lyon I |
| Universite Claude Bernard Lyon1 |
| Universite Claude Bernard Lyon-1 |
| Universite Claude Bernard-Lyon I |
| Universite Claude Bernard-Lyon |
| Universite Lyon-1 Claude Bernard |
| Univ. Claude Bernard Lyon 1 |
| Univ. Claude Bernard Lyon I |
| Univ. Claude Bernard |
| Univ. Claude Bernard Lyon-1 |
| Universite Claude Bernard-Lyon 1 |
| Universite Claude Bernard-Lyon1 |
| Universite Claude Bernard, Lyon 1 |
| Universite C. Bernard Lyon 1 |
| University Claude Bernard |
| University Claude Bernard of Lyon |
| University Lyon 1 Claude Bernard |
| University Lyon 1 |
| Claude Bernard Lyon I University |
| Claude Bernard University |
| Claude Bernard University Lyon1 |
| Claude Bernard University of Lyon |
| Lyon 1 University |
| UCB-Lyon 1 |

**TABLEAU 3 - LIBELLES D'ADRESSES
NE REPONDANT PAS AUX DEUX REQUETES
« CLAUDE BERNARD » ET « LYON 1 »**

| |
|----------------------------------|
| Universite Claude-Bernard |
| University Claude-Bernard |
| Universite Lyon-1 |
| Universite Lyon I |
| Univ. Claude B. |
| universite Claude-Bernard |
| universite Lyon I |
| universite Claude-Bernard-Lyon-1 |
| Universite Claude-Bernard-Lyon-I |
| UCBL |
| Claude-Bernard University |
| University Claude-Bernard |
| University Lyon I |
| Lyon University of Sciences |

**TABLEAU 4 - CARACTERISTIQUES DES ADRESSES MENTIONNEES
DANS LES NOTICES BIBLIOGRAPHIQUES**

| Caractéristiques des adresses | Medline | Embase | Pascal Biomed | Moyenne |
|---|---------|--------|---------------|---------|
| Adresse complète (avec cedex) | 17 % | 26 % | 1 % | 15 % |
| Adresse « complète » sans cedex | 13 % | 15 % | 26 % | 18 % |
| Absence de rue et code postal | 32 % | 10 % | 15 % | 19 % |
| Mention laboratoire ou faculté avec absence de rue, code postal et Université Claude Bernard-Lyon 1 | 14 % | 3 % | 8 % | 8 % |
| Absence de la mention Université Claude Bernard-Lyon 1 | 39 % | 47 % | 45 % | 44 % |
| Adresses relatives à un chercheur ou hospitalo-universitaire à double affiliation et mentionnant l'autre affiliation que l'Université Claude Bernard-Lyon 1 | 8 % | 8 % | 8 % | 8 % |

**TABLEAU 5 - NOMBRES DE REFERENCES BIBLIOGRAPHIQUES OBTENUES
LORS DE LA RECHERCHE DE L'UNIVERSITE CLAUDE BERNARD LYON 1
DANS LA RUBRIQUE « ADRESSE » OU « AFFILIATION »
DES TROIS BASES DE DONNEES BIBLIOGRAPHIQUES**

| Recherche sur 2003 dans la rubrique « Adresse » ou « Affiliation » | Medline | Embase | Pascal Biomed |
|--|---|--------|---------------|
| « lyon 1 » | 89 | 90 | 93 |
| « lyon I » | 3 | 42 | 62 |
| « claudes bernard » non pertinent (rue, quai, hôpital) | 279 | 336 | 450 |
| « claudes bernard » and university | 228 | 103 | 230 |
| « universite lyon 1 » | 0 142 si recherche sur l'ensemble des rubriques de la notice | 15 | 37 |
| « universite lyon I » | 0 | 8 | 28 |
| « université claudes bernard » | 0 190 si recherche sur l'ensemble des rubriques de la notice | 77 | 143 |
| « faculte de medecine » and lyon | 69 | 49 | 118 |
| (« faculte de medecine » and lyon) not « claudes bernard » | 62 | 47 | 97 |
| (« faculte de medecine » and lyon) not (« claudes bernard » or « lyon 1 » or « lyon I ») | 62 | 47 | 94 |

**TABLEAU 6 – CARACTERISTIQUES RELATIVES A LA PRESENTATION DES ADRESSES
ET AUX LOGICIELS DE RECHERCHE DES TROIS BASES DE DONNEES BIBLIOGRAPHIQUES**

| Caractéristique | Medline | Embase | Pascal Biomed |
|---|---|--|---|
| Adresse des auteurs | <ul style="list-style-type: none"> - Adresse uniquement du premier auteur de la liste des noms d'auteurs - Adresse ressaisie, non copiée-collée, pouvant être tronquée sur certains éléments comme le cedex ou pouvant présenter des fautes de frappe | <ul style="list-style-type: none"> - Adresse uniquement de l'auteur responsable de la correspondance, quelle que soit sa position dans la liste des noms d'auteurs - Adresse complète issue d'un copié-collé de la rubrique « correspondance » de l'article original | <ul style="list-style-type: none"> - Adresses de tous les auteurs de la publication - Adresse souvent tronquée car ne reprenant pas l'adresse complète mentionnée dans la rubrique « correspondance » de l'article original |
| Caractéristiques du logiciel de recherche | Ne tient pas compte des majuscules, des accents, des points et des tirets | Ne tient pas compte des majuscules, des accents, des points et des tirets | Ne tient pas compte des majuscules, des accents et des points Tient compte des tirets |
| Exemples | « Université Claude Bernard » = « universite claud e bernard » | | |
| | « univ claud e bernard » = « univ. claud e bernard » | | |
| | « université lyon 1 » = « universite lyon-1 » | | « université lyon 1 » ≠ « universite lyon-1 » |