

## Les classements internationaux : enjeux, méthodologies et perspectives pour les universités françaises

### Edito

Les classements internationaux sont devenus un enjeu majeur pour comparer les universités au niveau mondial. En quelques années, non seulement les classements (comme celui de Shanghai) ont répondu à de multiples usages (étudiants, politiques publiques, établissements) mais ils ont aussi connu une médiatisation croissante. Face à ce succès, les classements suscitent de nombreuses interrogations : quelles sont les méthodologies utilisées et comment évoluent-elles ? Quelles sont les positions des universités françaises et européennes dans les classements et comment s'expliquent-elles ? Peut-on envisager des méthodologies nouvelles de classements, et si oui sur quelles bases ?

Ce premier numéro de la collection « Résultats et recherches » des e-publications de l'OST vise à contribuer à la réflexion sur les classements. Quatre articles sont proposés pour apporter au lecteur des éclairages différents sur ce sujet éminemment complexe. L'article de G. Filliatreau (OST) retrace les dynamiques de succès des classements, et évoque les défis à venir et les réflexions en cours au niveau européen pour proposer des classements alternatifs. L'article de N. Carayol (GREThA et OST), A. Lahatte (OST) et G. Filliatreau (OST) propose une nouvelle méthodologie de classements des universités, fondée sur une théorie des relations de dominance. Appliqué aux universités américaines, cet article montre comment il est possible de proposer des méthodes

alternatives de classements. L'article de P. Vidal (OST) propose d'analyser le positionnement des établissements français dans 3 classements internationaux, et d'expliquer leurs performances comparées à celles de leurs homologues britanniques et allemands. Le quatrième article de J.A Héraud (BETA) analyse la position des établissements de la région européenne transfrontalière du Rhin supérieur. La diversité des positions est la conséquence de la diversité des institutions au sein de la région, pour lesquelles bien des critères de Shanghai ne sont pas adaptés à capter leur rôle en matière de recherche, d'innovation et de transfert de connaissances sur un territoire.

### Sommaire

- ☞ **Les classements internationaux d'universités : une mise en perspective des défis à relever** ..... 2
- ☞ **Intensité, qualité, excellence et classements des universités** ..... 6
- ☞ **Singularités de la position des établissements d'enseignement supérieur français dans les classements internationaux** ..... 17
- ☞ **L'espace du Rhin Supérieur dans les classements internationaux : réflexions sur un territoire transfrontalier et sur les méthodes d'évaluation** ..... 23

## Les classements internationaux d'universités : une mise en perspective des défis à relever

**Ghislaine Filliatreau, OST**

*ghislaine.filliatreau@obs-ost.fr*

### Shanghai : un succès planétaire sans la moindre publicité

Le "classement de Shanghai", palmarès des universités "de classe mondiale" a été créé à l'origine par trois chercheurs chinois pour leurs besoins, qui était d'évaluer la distance séparant (encore !) les universités chinoises des universités les plus prestigieuses du monde. Visant explicitement à identifier les 500 premières universités mondiales par leur prestige scientifique, ce classement a - sans autre effort de promotion que d'être posté en 2003 sur un site web - rencontré très rapidement un intérêt planétaire.

Comme on le sait, ce classement repose sur 5 mesures liées à la production scientifique de l'établissement. L'idée de se limiter à ces 5 mesures n'était pas là au départ mais l'équipe, après avoir cherché quelles données de référence, portant sur l'ensemble des universités de la planète, étaient accessibles, a dû se rendre à l'évidence : il en existait fort peu, toutes décrivant l'activité de recherche des établissements. A partir de ces données limitées, l'équipe a donc défini les 5 "marqueurs quantitatifs de prestige scientifique international" que l'on connaît : nombre de prix Nobel et médaille Fields parmi les anciens élèves et les professeurs, articles parus dans des journaux prestigieux comme Nature et Science, chercheurs à forte notoriété scientifique ...

Malgré des limites méthodologiques multiples et maintes fois dénoncées, chaque édition annuelle du classement de Shanghai est relayée par les médias dans de très nombreux pays, et utilisée dans un certain nombre d'entre eux - dont la France, pour interpeller les pouvoirs publics sur leur politique.

Comment expliquer ce succès ? Tout d'abord, parce que le classement est gratuit, produit par des universitaires et accompagné d'une méthodologie détaillée. Les mesures sont faites à partir de données de référence produites par des tiers et communément utilisées dans les milieux académiques : ainsi, les observateurs avertis retrouvent bien, aux premières places, la hiérarchie des universités qui sont historiquement connues pour leur prestige scientifique international, ce qui les rassure sur la pertinence de l'exercice dans son ensemble.

La seconde force de Shanghai est qu'il a été le premier à proposer une comparaison internationale, alors même que les préoccupations des établissements d'enseignement supérieur et des décideurs politiques étaient fortement polarisées par la question de la compétition et des coopérations internationales : réunis à Lisbonne en 2000, les chefs d'Etat européens venaient d'affirmer leur volonté commune de mettre en place des réformes profondes de l'enseignement supérieur pour « faire de l'Union européenne l'espace le plus créatif et le plus attractif du monde pour les étudiants et les chercheurs ». Rien d'étonnant à ce que, trois ans plus tard, les résultats du classement aient rencontré un vif intérêt de la part des médias dans beaucoup de pays européens. En effet, le classement montrait la domination des Etats-Unis en termes de nombre et de rangs d'établissements présents : cela apparaissait comme un avertissement bienvenu sur le niveau du défi à relever pour l'Europe. ...

Enfin ce classement, comme de nombreux autres outils de comparaison quantitative, répond à une demande forte de repères factuels : celle des étudiants et de leur famille qui accordent une grande importance aux études et aux diplômes, et plus largement celle des citoyens qui souhaitent avoir des informations précises

**Ghislaine Filliatreau,** directrice de l'Observatoire des Sciences et des Techniques, a participé aux premiers travaux de U-Multirank, financés par l'Union européenne



sur les résultats des établissements qui sont financés sur fonds publics, alors que les ressources limitées des Etats les contraignent à arbitrer durement entre leurs dépenses.

### Sous la pression de la concurrence et des usagers, les classements évoluent rapidement

Les classements internationaux - dont certains sont produits par le secteur privé - sont en concurrence les uns avec les autres, chacun d'entre eux aspirant à devenir un - sinon « le » - classement international de référence.

Aussi sont-ils contraints de s'améliorer pour prendre en compte les vives critiques qui leur sont adressées. C'est ce qui explique qu'ils ne cessent d'évoluer, de telle sorte que ce que l'on regroupe toujours sous le nom générique de "classement" recouvre en fait une large gamme d'outils, depuis le palmarès linéaire et monodimensionnel de type Shanghai première manière jusqu'à des outils sophistiqués comme U-Multirank, que développe actuellement un consortium de laboratoires européens<sup>1</sup>.

Depuis 2003, on peut distinguer trois étapes successives dans cette évolution.

Dans la première génération des classements, l'utilisateur consulte un palmarès fermé, où chaque établissement est positionné par rapport à tous les autres sur la base d'un score global unique, calculé par un indicateur composite. On l'a dit, dans le cas du classement de Shanghai paru en 2003, l'indicateur composite est construit à partir de données portant uniquement sur des performances de recherche, utilisées de telle façon que les établissements anciens, riches et souvent de bonne taille, sont privilégiés. Dans le cas du classement du THES apparu en 2004 par exemple, l'indicateur composite intègre les résultats d'une enquête de réputation auprès d'enseignants sur la qualité de la formation dispensée, et des données factuelles sur les conditions d'étude. Cet enrichissement des critères pris en compte permet à des établissements de prestige mais de petite taille, ou à des établissements très spécialisés, notamment en sciences humaines et sociales, d'être mieux détectés. Depuis, d'autres palmarès unidimensionnels sont apparus. Cependant ces palmarès, séduisants par

leur simplicité mais construits sur des indicateurs de qualité inégale et difficiles à interpréter, sont critiqués pour leur caractère normatif et pour la difficulté à utiliser leurs résultats : qu'apprendre d'une information si globale ?

Ainsi, les auteurs du classement de Shanghai ont très vite proposé plusieurs classements parallèles, déclinés selon différents champs ou disciplines de recherche. Cette évolution permet donc d'observer comment les différents départements disciplinaires d'une université sont positionnés au niveau international. De la même manière, d'autres classements ont enrichi le panel de mesures utilisées, ce qui peut permettre aux responsables d'un établissement d'observer ses points forts et faibles par appréciation des différences du rang obtenu. Cette approche est facilitée par les classements qui, au lieu d'un seul indicateur composite, donnent le détail des différentes mesures qu'ils calculent. De la même manière, cette prise en compte de la réalité multidimensionnelle de la performance des établissements permet une analyse plus riche au niveau des systèmes. Malheureusement, peu de classements sont assez transparents sur le détail de leurs mesures pour permettre aux établissements de mener ce type d'analyse avec précision.

Dans une troisième étape, on change tout à fait la perspective en permettant à l'utilisateur, via une interface interactive, non seulement d'accéder aux différentes mesures calculées, mais aussi de sélectionner lui-même celles qu'il souhaite privilégier pour bâtir sa propre hiérarchie. Comme on le voit, la philosophie est alors radicalement nouvelle : le principe selon lequel il pourrait exister un seul classement dont les critères s'imposeraient naturellement s'efface pour faire place à la notion de "classement pertinent", c'est-à-dire de classement dont la construction est *paramétrée* par les priorités de l'utilisateur. Cette évolution a de multiples conséquences. Notamment, elle incite les établissements à se différencier en fonction de leurs missions vis-à-vis de publics particuliers, et à mettre ainsi en lumière ce qu'ils ont de spécifique. De même, elle permet aux décideurs de faire une analyse systémique

<sup>1</sup> Le rapport final qui rend compte des conclusions du projet U-Multirank est accessible à :

Final Reportec.europa.eu/education/higher-education/doc/multirank\_en.pdf

plus complète, puisqu'elle rend compte de façon non normative d'une plus large variété d'activités et de profils d'établissements.

## U-Multirank, un outil ambitieux de 4<sup>e</sup> génération

Un rapport récemment publié par l'EUA<sup>2</sup> a bien mis en évidence les limites des classements actuels, et les dangers qu'ils représentent s'ils sont utilisés sans esprit critique : le risque d'enclencher une spirale de comportements inadaptés est réel : comparaisons inappropriées, stratification des positions établies, focalisation excessive sur ce qui est mesuré, rétrécissement des marges d'action, renforcement des hiérarchies liées aux ressources, aggravation de la pression pour des résultats immédiats, réduction de la diversité.

L'étude de l'EUA mentionne que l'outil développé depuis près de 15 ans pour l'Allemagne par le *Center for Higher Education (CHE)*<sup>3</sup> est le plus abouti méthodologiquement. Il s'agit d'un classement multidimensionnel (34 mesures différentes sont disponibles, au niveau des départements disciplinaires plutôt que des établissements), qui range les établissements en 3 groupes de qualité (vert, orange, rouge), pour chacune des mesures qui sont sélectionnées par l'utilisateur.

L'étude mentionne U-Multirank<sup>4</sup>, un classement original impulsé par la France, dont la conception a été financée par la Commission européenne.

U-Multirank, qui prend en compte tous les enseignements précédents, propose des solutions originales. Tout d'abord, il combine l'approche du CHE avec un autre exercice appelé U-Map, précédemment développé par le CHEPS<sup>5</sup> sur financement de la Commission européenne. U-Map est un outil de classification horizontale, qui va permettre de sélectionner les établissements qui se ressemblent avant de lancer le classement de type CHE.

Les deux outils sont multidimensionnels, paramétrables par l'utilisateur : leur combinaison fait de U-Multirank un outil d'information ambitieux, permettant à chaque établissement de faire valoir ses spécificités tout en se comparant à celles des autres établissements, et permettant à chaque usager de s'appuyer sur ses propres priorités pour évaluer les établissements et leurs départements.

Le résultat proposé reprend les caractéristiques majeures du rangement du CHE : pas d'indicateur composite et pas de classement, mais une répartition des établissements comparables en trois groupes selon leur adéquation aux spécifications de l'utilisateur.

Dans sa conception, cet outil est certainement le plus abouti et le plus à même de promouvoir l'enseignement supérieur et ceci pour deux raisons :

- d'une part en rendant justice à la diversité des établissements et de leurs activités,
- d'autre part en permettant aux usagers - et d'abord aux étudiants - une exploration contrôlée de l'ensemble des informations disponibles sur ces établissements et leurs départements.

## Une évolution freinée par plusieurs exigences

La plupart des classements internationaux ont pris en compte la nécessité d'évoluer. Par exemple, les concepteurs du classement de Shanghai ont enrichi leur premier classement unidimensionnel par des classements thématiques, et ont récemment lancé un nouvel exercice de comparaison beaucoup plus complet, le « GRUP : Global Research University Profile », qui sera multidimensionnel et paramétrable par l'utilisateur.

Cependant, cette évolution des classements internationaux vers des outils plus sophistiqués et plus inclusifs de la diversité a un prix : il faut pouvoir mobiliser des données quantitatives comparables qui soient perti-

<sup>2</sup> Les classements mondiaux d'universités et leur impact, par Andrejs Rauhvargers, version française à [http://www.eua.be/Libraries/Publications\\_homepage\\_list/EUA\\_Rankings\\_Report\\_french\\_translation.sflb.ash](http://www.eua.be/Libraries/Publications_homepage_list/EUA_Rankings_Report_french_translation.sflb.ash)

<sup>3</sup> L'équipe du CHE a également contribué à définir une grande partie des règles déontologiques et des spécifications de méthode qui sont prônées par l'IREG, une association internationale qui s'est donnée pour objectif de faire connaître et d'encadrer les bonnes pratiques en matière de classement.

<sup>4</sup> Le rapport final de l'étude est accessible à : [http://ec.europa.eu/education/higher-education/doc/multirank\\_en.pdf](http://ec.europa.eu/education/higher-education/doc/multirank_en.pdf)

<sup>5</sup> Le CHEPS est un laboratoire de recherche sur l'enseignement supérieur de l'université hollandaise de Twente : [www.utwente.nl/mb/cheps/](http://www.utwente.nl/mb/cheps/)

nelles, à la fois en signification et en qualité. Il s'agit d'un défi important étant donné la diversité des établissements à l'intérieur de chaque système national d'enseignement supérieur.

Or cet accès aux données est très difficile - et donc coûteux, sachant qu'il existe trois sources possibles de données.

La première est constituée d'un nombre limité de bases qui sont internationales par construction, comme les bases utilisées par les chercheurs pour faire leur recherches bibliographiques, ou par les offices internationaux de brevets pour gérer la propriété industrielle au niveau international, ou par le comité Nobel qui distribue des récompenses de niveau international.

La seconde source est constituée par des bases utilisées au niveau national pour gérer, évaluer ou financer les établissements. Ces bases, qui peuvent contenir des données très pertinentes, sont néanmoins disparates en qualité selon les pays. Surtout elles n'ont pas été construites pour être comparables sur le plan international.

La troisième source possible consiste à recueillir les données directement auprès de chaque établissement. Pour des raisons de coûts, cette méthode est pratiquement irréalisable dans un contexte international par des opérateurs appartenant au secteur public.

De fait, la question des coûts d'investissement et du modèle économique est cruciale : en l'absence de réponse ambitieuse émanant d'opérateurs publics ou sans but lucratif, les classements les plus ambitieux sont et seront développés par des opérateurs privés. Ces derniers sont capables d'investir mais seront ensuite condamnés à rentrer dans leurs frais - qu'il s'agisse de proposer ces classements sur abonnement payant, ou de les faire financer par ceux qui sont classés. Dans tous les cas, ces opérateurs privés n'ont pas intérêt à donner trop d'informations sur leurs méthodes ni à ouvrir leurs données, encore moins à investir pour les marchés peu solvables.

## Disposer de nouvelles modalités de classements : un enjeu majeur pour l'Europe

**C'**est notamment la conscience du danger à abandonner le développement de ces outils au marché qui avait motivé la proposi-

tion de la présidence française de l'Union européenne de promouvoir un classement alternatif, porté par une structure indépendante sans but lucratif, qui a donné lieu au projet U-Multirank.

C'est la raison pour laquelle la Commission européenne avait demandé au consortium choisi pour U-Multirank<sup>6</sup> une analyse du modèle économique préconisé pour un outil qui doit respecter l'ensemble des bonnes pratiques définies par les principes de Berlin, c'est-à-dire se mettre en capacité de prendre en compte la diversité des contextes culturels, linguistiques et économiques de l'ensemble des établissements étudiés. Ce qui représente un énorme défi : si les classements commerciaux actuels ne s'intéressent qu'à quelques centaines d'établissements ayant des caractéristiques communes, U-Multirank se veut représentatif d'une part significative des quelques 16 000 établissements d'enseignement supérieur existants dans le monde.

Dans l'hypothèse d'un dispositif fédératif international, le modèle propose des partenaires dans les différents pays participants, produisant chacun un classement national incluant un tronc commun d'indicateurs comparables qui seront partagés au niveau de la plateforme U-Multirank. Dans chaque pays, les partenaires relaient l'utilisation de l'outil européen, mais les classements nationaux restent diversifiés et servent de banc d'essai pour des développements qui peuvent ensuite être mis en commun pour des aires géographiques pertinentes (la francophonie par exemple). Les coûts centraux européens sont partagés et allégés puisque la production des mesures communes est décentralisée et les développements faits dans chaque pays peuvent être mutualisés.

De plus, ce dispositif peut également être une source d'économies pour les établissements universitaires de chaque pays, qui pourront s'ils le souhaitent partager et contrôler un entrepôt national archivant leurs données, plutôt que d'être acculés à répondre en ordre dispersé

<sup>6</sup> U-Multirank a été réalisé par le CHERPA (*Center for Higher Education and Research Performance Assessment*), qui regroupe cinq institutions : le Centre for Higher Education Policy Studies (CHEPS) de l'université de Twente, le Centrum für Hochschulentwicklung (CHE) de Gütersloh, le Centre for Science and Technology Studies (CWTS) de l'université de Leiden, l'International Centre for Research on Entrepreneurship, Technology and Innovation Management (INCENTIM) de l'Université catholique de Louvain, et l'Observatoire des sciences et des techniques (OST) de Paris.

## Résultats et recherches

à de multiples sollicitations à livrer des données dont ils ne contrôlent pas l'usage.

La Commission européenne a annoncé son intention de lancer un second appel d'offres pour la suite de U-Multirank. Grâce à la participation de l'OST, la France

est en capacité de jouer un rôle pionnier dans ce projet fédératif, et de contribuer ainsi à en concevoir les règles : il appartient aux établissements, comme aux instances nationales qui les rassemblent, de concrétiser leur intérêt pour ces perspectives. ■

### Intensité, qualité, excellence et classements des universités

**Nicolas Carayol, GREThA, et chercheur associé à l'OST**

*nicolas.carayol@u-bordeaux4.fr*

**Ghislaine Filliatreau, OST, Agenor Lahatte, OST**

*ghislaine.filliatreau@obs-ost.fr*

*agenor.lahatte@obs-ost.fr*

#### Introduction

Depuis quelques années, plusieurs classements mondiaux d'universités ont vu le jour, stimulés notamment par une recherche d'information croissante émanant d'étudiants de plus en plus mobiles et désireux de choisir au mieux leur école ou université. Soumis à une concurrence accrue, les responsables de ces institutions accordent désormais une forte attention à ces classements. Si la qualité de l'enseignement est importante pour ces étudiants, le prestige scientifique des établissements est également considéré comme un critère essentiel dans la plupart des classements. Dans un contexte de forte hiérarchisation des établissements, l'enjeu principal des classe-

ments d'universités est de parvenir à rendre compte à la fois de l'intensité de l'activité scientifique des établissements<sup>1</sup> et de leur qualité dans une approche quantitative qui soit intelligible pour les acteurs et les utilisateurs. Pour tenter d'y parvenir, la plupart des classements consistent en un savant dosage de mesures diverses permettant de capturer séparément ces différents aspects. Par

<sup>1</sup> Des problématiques similaires pourraient être développées pour des classements plus orientés vers la qualité de l'enseignement et/ou du placement des étudiants. Dans cet article nous nous limitons volontairement aux approches centrées sur la fonction recherche des universités.



De gauche à droite: **Nicolas Carayol**, (chercheur au GREThA—Université Bordeaux 4 et à l'OST), **Agenor Lahatte** (OST) et **Ghislaine Filliatreau** (OST) ont collaboré sur ce projet de recherche pour proposer une nouvelle méthodologie de classements fondée sur les relations de dominance. Appliquée dans cet article aux universités américaines pour les données de publications et de citations, cette méthodologie pourrait être exploitée pour les universités françaises et européennes sur les mêmes types de données ou sur des données non bibliométriques.

exemple, le classement de ARWU développé par l'Université Jiao Tong de Shanghai, agrège des mesures variées d'excellence (nombre de prix Nobel et médailles Fields, nombre d'articles dans Science et Nature, nombre de chercheurs fortement cités) et des mesures de volume d'activité (nombre d'articles publiés).

Notre méthodologie se distingue de ces approches notamment parce qu'elle permet une grande transparence et une continuité entre les prémisses retenues et les résultats obtenus. Elle permet d'explicitier la hiérarchie des valeurs sur lesquelles on veut établir les comparaisons: souhaite-t-on privilégier la production scientifique? La visibilité globale de l'établissement? Une visibilité sélective? Notre approche permet de répondre à ces questions en adossant les comparaisons d'universités à des jugements de valeur explicites. Dans l'application présentée ici, nous n'utilisons que les informations relatives aux articles publiés et à leur impact (pouvant être mesurés de différentes manières, comme indiqué plus bas) enregistrés dans le *Web of Science* (Thomson Reuters), qui est maintenu et enrichi au sein de l'Observatoire des Sciences et Techniques. Cette approche a l'avantage d'être plus économe en information, s'appuyant uniquement sur des sources contrôlées, homogènes et robustes dans le temps. Il faut aussi noter que nous utilisons cette information de manière très fine, en considérant le nombre d'articles produits par chaque établissement, dans chaque discipline et pour tout niveau possible d'impact de ces articles.

Nous avons modélisé trois types de jugements de valeur. Le premier, traduisant un intérêt pour l'intensité de l'activité de recherche, ne valorise pas particulièrement l'impact des articles scientifiques mais seulement le volume. Le second prend en considération l'impact des articles sans accorder de valorisation particulière aux articles les plus cités dans la discipline de référence. Le troisième donne un poids proportionnellement plus important aux articles les plus cités. Ainsi, en fonction des prémisses retenues, il est possible d'établir des relations de dominance<sup>2</sup> deux à deux au sein d'un ensemble d'universités dans un système de valeur explicite. Nous montrons ensuite comment utiliser ces relations pour construire des classements et des classes de référence.

Après avoir présenté succinctement notre méthode,

nous présentons dans cet article quelques classements portant sur les principales universités de recherche des États-Unis obtenus grâce à cette méthodologie originale.

## Méthodologie pour des classements fondés sur des relations de dominance

Cette section présente la méthodologie fondant les classements et classes de références, dans un abord non technique. Le lecteur pourra se tourner vers Carayol et Lahatte (2011) pour une présentation complète, et vers Carayol et al. (2011) pour une présentation en français. Nous présentons dans un premier temps la question de la mesure de la qualité. Nous introduisons ensuite les fondements des comparaisons bilatérales et leur jugement de valeur associés. Enfin, nous présentons une manière de traduire les comparaisons bilatérales en classements et classes de référence.

### Approximer la qualité des articles

La mesure de la qualité sur laquelle notre méthode est fondée repose, notamment, sur une appréhension de la mesure de l'impact des articles, abordé via trois indices différents, tous basés sur des décomptes des citations reçues. Le premier est le nombre de citations reçues par l'article considéré sur une période donnée après publication. Un tel indicateur, dit de visibilité directe, est basé uniquement sur les références faites aux articles considérés. Il est très intuitif, mais s'avère être une mesure relativement bruitée - les chercheurs se plaignent souvent du fait que leurs «meilleurs» articles ne sont pas les plus cités et il

<sup>2</sup> Nous utilisons le terme de dominance car, en termes généraux, cette théorie se présente comme une généralisation de la notion de dominance stochastique largement utilisée en économie dans la théorie de choix dans l'incertain (Rothschild et Stiglitz, 1970) et dans l'analyse de distributions de revenus ou de richesse (Atkinson, 1970). Elle étend la notion de dominance stochastique en cela qu'elle permet de comparer des profils d'institutions en valorisant aussi la quantité d'items (et non la seule qualité et sa distribution).

arrive que certains articles soient cités bien au-delà de leur contribution scientifique telle qu'estimée par les experts.

Une autre option consiste à utiliser l'impact moyen des articles publiés par le journal dans lequel l'article a paru. Cet indicateur indirect de l'impact d'un article traduit la capacité de ses auteurs à publier dans des journaux plus ou moins cités en moyenne, et donc plus ou moins prestigieux.

Ces deux premiers indices présentent l'inconvénient de favoriser, en termes d'impact, les spécialités ou les sous-disciplines les plus saillantes d'une discipline. L'impact moyen relatif, troisième mesure proposée ici, pallie un tel biais. C'est l'impact moyen du journal rapporté à l'impact moyen des journaux de sa (ou ses) spécialité(s) scientifique(s) de référence. Un tel indice semble être particulièrement utile quand on analyse la production scientifique d'institutions qui sont spécialisées dans des sous-domaines dont la visibilité moyenne n'est pas très élevée. Il permet en outre de prendre en compte la diversité des pratiques de citation dans les différentes spécialités d'une même discipline (par exemple, le fait qu'en mathématiques, la sous-catégorie mathématiques fondamentales est moins citée en moyenne que la sous-catégorie mathématiques appliquées).

Il pourrait sembler naturel d'utiliser l'impact d'un article, défini de l'une des manières présentées ci-avant, comme mesure de la qualité de ce dernier. Nous écartons toutefois cette hypothèse du fait de la variabilité de cet indice d'une discipline scientifique à l'autre (due à des différences de pratiques de citation ou de taux de couverture des bases bibliographiques). Aussi, nous proposons de mesurer la qualité des articles en se basant sur leur position dans la distribution au sein de leur discipline selon l'impact. La qualité  $s_a^k$  d'un article  $a$  dans une discipline donnée  $k$ , sera égale au plus grand  $s$ , tel que son impact est au moins aussi élevé que celui de  $100 \times s\%$  des articles publiés dans cette discipline. Autrement dit, étant donné la distribution de la production scientifique dans la discipline, et pour une mesure d'impact donnée, la qualité d'un article est la probabilité qu'un article tiré au hasard dans la discipline soit d'impact inférieur.

Cette approche permet l'agrégation d'articles de disciplines scientifiques différentes pour tout niveau

de qualité donné mais ne s'avère pas être indispensable dans le cas d'une comparaison intra-discipline.

## Etablir des comparaisons bilatérales

Dans notre modèle, nous distinguons plusieurs types de comparaisons bilatérales entre institutions, chaque type de relation de dominance étant associé à une série d'hypothèses différentes sur la fonction implicite de valorisation de la qualité,  $v(\cdot)$ . Nous parlons ici de relations de dominance. Une relation de dominance de  $i$  sur  $j$  indique que la comparaison est unanime pour toute fonction  $v(\cdot)$  dans une classe de fonctions bien définie. Chaque classe de fonction renvoie à un ensemble de valeurs relatif à une manière particulière de considérer la qualité des articles. Trois types de relations de dominance sont abordées : la dominance forte, la dominance standard et la dominance faible.

La notion de dominance forte requiert la non négativité de la fonction  $v(\cdot)$ , en d'autres termes aucun article ne saurait contribuer négativement à la performance scientifique d'une institution. Cette propriété de non négativité s'avère être nécessaire pour chaque approche différente de la notion de valorisation, elle attribue un rôle très faible aux mesures d'impact. Elle peut ainsi être associée à une valorisation de l'intensité de publication.

Pour la dominance standard, la fonction  $v(\cdot)$  doit être non négative et non décroissante, ce qui veut dire que dans un domaine donné on attribue une valeur la plus élevée aux articles dont l'impact est plus grand. Cette deuxième propriété semble cohérente avec les différentes approches possibles de la valeur de la production scientifique. Elle correspond à une valorisation de la qualité.

En plus des deux propriétés précédentes, l'hypothèse de convexité de  $v(\cdot)$  est retenue dans la dominance faible. Cette dernière hypothèse implique que  $v(\cdot)$  donne un poids proportionnel ou plus que proportionnel aux articles de plus forte qualité. La convexité peut être considérée comme une hypothèse largement partagée puisque les CEO des universités et leur conseil d'administration souhaitent la présence de leur institution dans les segments de forte visibilité académique. Elle correspond à une



valorisation de la qualité.

En plus des deux propriétés précédentes, l'hypothèse de convexité de  $v(\cdot)$  est retenue dans la dominance faible. Cette dernière hypothèse implique que  $v(\cdot)$  donne un poids proportionnel ou plus que proportionnel aux articles de plus forte qualité. La convexité peut être considérée comme une hypothèse largement partagée puisque les CEO des universités et leur conseil d'administration souhaitent la présence de leur institution dans les segments de forte visibilité académique. Elle correspond à une valorisation de l'excellence.

Dans Carayol et Lahatte (2011), un théorème établit les conditions nécessaires et suffisantes à chaque relation de dominance bilatérale, permettant d'établir ces relations de dominance en utilisant les uniques données de publication et d'impact, c'est-à-dire sans spécifier précisément la fonction de valorisation, mais seulement qu'elle doit appartenir à une classe de fonctions bien définie.

Pour la dominance forte, la condition nécessaire et suffisante est que l'institution dominante doit produire toujours plus que l'institution dominée pour tous les niveaux de qualité. Le volume de production joue donc un rôle important en dominance forte puisqu'une institution ne peut dominer fortement une autre si cette dernière a plus de publications pour au moins un niveau quelconque de qualité. Notons que la dominance forte traite tous les niveaux de qualité de la même manière (aucune prime n'est accordée aux articles les plus cités).

La relation de dominance standard attribue une valeur croissante implicite à la qualité. La condition nécessaire et suffisante pour établir une telle relation de dominance est que la performance cumulée de l'institution dominante soit supérieure à celle de l'institution dominée au-dessus de tout niveau minimum de qualité possible. Elle véhicule donc l'idée qu'un article vaut toujours au moins autant qu'un autre article ayant un niveau d'impact plus faible. Ce régime de compensation permet d'établir la domination d'une institution qui a une performance relative plus élevée pour certains niveaux donnés de qualité.

Quant à la dominance faible, elle donne un rôle prépondérant aux articles de qualité élevée. Ici, la condition nécessaire et suffisante est que la performance cumulée de l'institution dominante, pondérée

de la qualité, au-dessus de tout niveau plancher de qualité, soit supérieure à celle de l'institution dominée. Cela implique qu'une institution ne peut pas être dominée par une autre si sa performance en quantité pondérée par sa qualité, pour un segment supérieur quelconque de qualité, est supérieure. La dominance faible dépend ainsi moins de la quantité, puisqu'elle accorde une importance maximale aux segments de plus forte visibilité.

Pour les comparaisons interdisciplinaires, nous agrégeons la valeur des productions des différentes disciplines. Cette opération ne pose pas de problème d'agrégation puisque la mesure de la qualité retenue corrige les différences disciplinaires (cf. supra).

Nous introduisons en outre la possibilité de n'établir ces relations de dominance que sur des segments de qualité haute, dont l'ampleur est ajustable, c'est-à-dire de limiter les comparaisons à des parts variables des meilleurs articles dans une discipline donnée. Un paramètre ( $\emptyset$ ) permet de définir le niveau seuil de qualité au-dessous duquel les publications ne seront pas prises en compte. Si  $\emptyset = 1$  alors tous les articles sont comptabilisés. Plus le segment de part mondiale retenue est sélectif ( $\emptyset$  petit), plus on se limite aux segments de plus forte visibilité en négligeant les articles de qualité plus faible (inférieure à  $1 - \emptyset$ ).

### Construire des classements et classes de référence à partir des relations de dominance

Les relations de dominance sont établies deux à deux. Nous montrons dans cette section que l'on peut utiliser les relations de dominance, telles que précédemment définies, pour classer des universités et pour construire des groupes de référence (ou groupes de pairs).

Une relation de dominance ne peut en général pas être établie pour n'importe quelle paire d'institutions. Lorsque c'est possible (on parle alors de relation complète), il est assez évident et naturel de déduire un classement à partir du complexe de comparaisons bilatérales. Dans le cas général, il s'agit néanmoins de classer des établissements qui ne peuvent pas toujours être comparés pour une relation de dominance donnée. Nous parlons alors de

pseudo-classement, obtenus sur la base de scores attribués aux acteurs sur la base de leur position relative dans le complexe de dominances bilatérales. Le complexe des relations bilatérales de dominance peut être représenté sous la forme d'un réseau dirigé construit en compilant les relations de dominance deux à deux.<sup>3</sup> Nous proposons ici un pseudo-classement construit sur deux critères simples. Le premier, le critère dominant, correspond au nombre d'autres institutions dominées. Le second critère est le nombre d'autres institutions qui dominent l'institution considérée (le score est pris ici par ordre décroissant)<sup>4</sup>.

Les pseudo-classements présentent l'avantage d'être toujours réalisables sur la base de n'importe quelle relation de dominance. Il faut cependant, dans la pratique se garder de construire des pseudo-classements lorsque la proportion de relations de dominance que l'on peut établir sur le nombre de relations de dominance possibles est suffisamment élevée. Nous nommons cette mesure le taux de complétude.

Le fait de ne pas pouvoir toujours établir une relation de dominance entre deux institutions amène à considérer que ces dernières ont l'une vis-à-vis de l'autre le statut de « pair ». Pour une institution donnée, nous pouvons donc constituer sur la base de cette idée l'ensemble de ses « pairs ». Nous parlons de classe de référence de l'institution considérée associée à la relation de dominance utilisée : la classe de référence d'une institution comprend toutes les institutions qui ne peuvent se comparer à elle dans le type de dominance considéré (au sens qu'aucune relation de dominance ne peut être établie, ni dans un sens ni dans l'autre). Cela signifie qu'elles sont à la fois similaires mais dans le même temps différentes puisque cette impossibilité d'établir une relation de dominance indique que chacune des deux institutions a des points forts vis-à-vis de l'autre.

## Application aux universités de recherche américaines

**N**ous illustrons notre méthodologie à partir des données de performances d'un échantillon d'universités américaines. Les données

sont présentées dans un premier temps avant de discuter les résultats obtenus.

### Les données mobilisées

L'échantillon expérimental est constitué de 112 universités américaines, apparaissant dans les meilleures positions du classement mondial des universités produit par l'université Jiao-Tong de Shanghai. Les publications de ces établissements ont été identifiées dans la version "off-line" de la base *Thomson Reuters-Web of Science* maintenue et enrichie par l'OST. Pour cette étude pilote, nous avons choisi d'utiliser les données de publication de 2003, 2004 et 2005, avec une fenêtre de décompte de citations de 3 années. Sur la base de ces options, la production scientifique des 112 établissements considérés dans cette analyse totalise 329 910 articles. Les scores de citations reçues par ces articles vont de 0 à 1 292 et l'impact moyen des journaux dans lesquels ils sont publiés varie de 0 à 27,6.

L'affectation des articles publiés à des domaines scientifiques est basé sur le regroupement des journaux en 9 grandes disciplines. Les 8 premières (Biologie fondamentale, Recherche médicale, Biologie appliquée - écologie, Chimie, Physique, Sciences de l'univers, Sciences pour l'ingénieur, Mathématiques) correspondent à des champs scientifiques distincts alors que la dernière, appelée Multidisciplinaire, regroupe des journaux qui publient des articles pouvant être assignés à différentes disciplines. La non prise en compte des articles de cette catégorie introduirait un biais non négligeable puisqu'il éliminerait une part des articles importants de plusieurs disciplines. Trois de ces journaux (*Proceedings of the National Academy of Sciences USA*, *Science* et *Nature*) sont très prestigieux et leurs articles ont été affectés un à un à leur discipline principale de rattachement (sur une base lexicographique).

<sup>3</sup> Dans les représentations proposées à la section suivante, et par soucis de visibilité, nous avons enlevé les relations de dominance qui peuvent être simplement déduites par l'application de propriété de transitivité des relations de dominance.

<sup>4</sup> Dans Carayol et Lahatte (2011), nous proposons une méthode plus élaborée permettant de classer les établissements sur la base de leur position dans le réseau de dominances bilatérales.

## Classements des universités américaines

Les tableaux 1, 2 et 3 présentent les résultats obtenus pour les 15 premières universités au niveau interdisciplinaire (respectivement lorsque l'impact est mesuré par les citations directes, l'impact moyen et l'impact moyen relatif). Dans chaque tableau et pour chaque type de dominance, sont reportés le rang dans le pseudo-classement ( $\gamma_i$ ) et la performance dans le critère dominant ( $n_i$ ) lorsque toutes les publications sont prises en compte ( $\emptyset = 1$ )<sup>5</sup>.

Bien que les différents classements exposés soient globalement proches, certaines institutions ont des rangs très différents selon le type de relation de dominance associée. Par exemple, l'université Harvard est classée première à la fois pour la dominance standard et la dominance faible (l'impact étant mesuré par les citations, cf table 1, alors qu'elle n'est classée qu'en dixième position lorsque la dominance forte est utilisée, classement dans lequel elle est largement devancée par Seattle (Washington University) qui est alors classée première. De même, le MIT n'est classé qu'en cinquante-cinquième position dans le classement fondé sur la relation de dominance forte, alors qu'il est classé dans les dix premiers pour les deux autres formes de dominance. Ces différences illustrent les effets de la prise en compte différentielle du volume, de la qualité ou de l'excellence dans les classements : comme on l'a vu plus haut, la dominance forte prend pleinement en compte l'effet pur du volume de publication, alors que la dominance standard et surtout la dominance faible font intervenir l'impact des publications, permettant à des institutions de taille éventuellement plus modeste de voir leur niveau d'excellence pleinement valorisé.

En outre, certaines institutions voient leur rang varier sensiblement en fonction de la mesure d'impact utilisée. Par exemple, l'Université Berkeley se classe sixième et cinquième (en dominance et en dominance faible) lorsque l'impact est mesuré par les citations, alors qu'elle se classe en troisième et seconde position (respectivement) lorsque l'impact moyen est utilisé. Ceci indique que les chercheurs de Berkeley sont particulièrement performants dans leur capacité à publier des articles dans les journaux les plus importants de leur discipline.

Par ailleurs, si l'on utilise l'impact moyen relatif, Berkeley se retrouve à la quatrième position révélant ainsi que les chercheurs de Berkeley, ne sont pas seulement excellents dans leur capacité à publier dans les meilleurs journaux de leur spécialité, révélé par le classement selon l'impact moyen relatif, mais aussi qu'ils le font dans des spécialités de plus fort impact de leur discipline, ce qui leur donne un avantage pour les mesures d'impact direct.

## Des réseaux de dominance et des classes de référence pour mieux apprécier le positionnement relatif des universités

Les pseudo-classements ne rendent compte que d'une partie de l'information disponible, et il est souvent utile d'explorer également la structure des réseaux de dominance. La Figure 1 présente le réseau de dominance faible (impact mesuré par les citations) parmi les premières universités dans le pseudo-classement associé. Nous observons qu'en dessous de l'Université Harvard, la structure de dominance fait apparaître une certaine complexité.

En effet, aucune relation de dominance n'a pu être établie entre l'Université du Michigan (à Ann Harbor), l'Université Washington de Seattle, UCLA et l'Université Stanford. Cependant, la structure aval permet de voir que Stanford est mieux classée que les trois autres parce qu'elle domine Berkeley et le MIT, ce que les autres ne parviennent pas à faire.

Lorsque l'on se limite aux articles appartenant aux dix pourcent le plus cités de chaque discipline ( $\emptyset = 0,1$ ), le réseau de dominance entre ces mêmes institutions se modifie naturellement dans le sens d'une plus forte hiérarchisation<sup>6</sup>. Désormais Berkeley domine Michigan et Seattle, et Stanford domine Michigan et le MIT domine UCLA et Johns Hopkins (cf. Figure 2).

<sup>5</sup> Par manque de place, tous les résultats ne peuvent être présentés dans l'article. Les résultats concernant cette étude sont présentés dans Carayol et al (2011). Des résultats plus complets encore sont à la disposition de la communauté sur la page internet suivante :

<http://carayol.u-bordeaux4.fr/ranking.html>

Tableau 1 - Top 15 du pseudo ranking des universités américaines (toutes disciplines)

	Relations de dominance basées sur les citations					
	dominance forte		dominance standard		dominance faible	
	Rang $r_i$	#Dom $n_i$	Rang $r_i$	#Dom $n_i$	Rang $r_i$	#Dom $n_i$
Harvard	10	65	1	111	1	111
Stanford	11	59	2	107	2	107
Seattle	1	82	3	105	3	105
UCLA	7	72	3	105	3	105
UM_Ann_Arbor	7	72	5	104	6	104
Berkeley	17	41	6	104	5	105
Johns_Hopkins	19	41	7	103	7	104
Pennsylvania	28	29	8	101	8	103
VM_Madison	3	78	9	97	11	97
MIT	55	14	10	95	9	102
Columbia	13	48	11	94	10	98
Cornell	5	76	12	94	15	95
Twin_Cities	4	77	12	94	13	95
DE_UCSD	14	48	14	93	12	96
UCSF	48	17	15	91	13	95

Tableau 2 -Top 15 du pseudo ranking des universités américaines (toutes disciplines)

	Relations de dominance basées sur l'impact moyen					
	dominance forte		dominance standard		dominance faible	
	Rang $r_i$	#Dom $n_i$	Rang $r_i$	#Dom $n_i$	Rang $r_i$	#Dom $n_i$
Harvard	1	91	1	111	1	111
Berkeley	7	82	2	106	2	106
Stanford	11	79	2	106	2	106
Seattle	2	89	4	104	4	104
UCLA	5	87	5	103	5	103
UM_Ann_Arbor	2	89	6	102	6	102
Johns_Hopkins	6	83	7	101	7	102
Pennsylvania	13	77	8	100	8	102
MIT	20	62	9	97	9	101
VM_Madison	4	88	10	96	11	96
Cornell	9	80	11	94	12	95
Columbia	16	72	12	94	10	97
DE_UCSD	19	64	13	93	12	95
Twin_Cities	9	80	14	91	15	92
UCSF	31	47	15	91	14	95

Tableau 3 - Top 15 du pseudo ranking des universités américaines (toutes disciplines)

	Relations de dominance basées sur l'impact moyen relatif					
	dominance forte		dominance standard		dominance faible	
	Rang $r_i$	#Dom $n_i$	Rang $r_i$	#Dom $n_i$	Rang $r_i$	#Dom $n_i$
Harvard	1	100	1	111	1	111
Stanford	10	83	2	107	2	107
Seattle	4	95	3	106	3	106
UM_Ann_Arbor	1	100	3	106	3	106
Berkeley	5	89	5	106	5	106
UCLA	3	97	6	105	6	105
Johns_Hopkins	6	89	7	101	9	101
Pennsylvania	13	79	8	100	7	102
MIT	19	72	9	97	8	101
VM_Madison	7	88	10	96	10	96
Columbia	15	79	11	95	10	96
Cornell	9	85	12	95	12	96
DE_UCSD	16	78	13	94	13	95
Twin_Cities	8	87	14	93	14	94
UCSF	30	52	15	92	15	93

En effet, aucune relation de dominance n'a pu être établie entre l'Université du Michigan (à Ann Harbor), l'Université Washington de Seattle, UCLA et l'Université Stanford. Cependant, la structure aval permet de voir que Stanford est mieux classée que les trois autres parce qu'elle domine Berkeley et le MIT, ce que les autres ne parviennent pas à faire.

Lorsque l'on se limite aux articles appartenant aux dix pourcent le plus cités de chaque discipline ( $\emptyset = 0,1$ ), le réseau de dominance entre ces mêmes institutions se modifie naturellement dans le sens d'une plus forte hiérarchisation<sup>6</sup>. Désormais Berkeley domine Michigan et Seattle, et Stanford domine Michigan et le MIT domine UCLA et Johns Hopkins (cf. Figure 2).

En outre, la hiérarchie s'est entièrement clarifiée puisqu'elle est maintenant linéaire depuis l'université de Pennsylvanie jusqu'à l'université du Wisconsin, en passant par l'université de Columbia et l'université de Californie à San Diego (UCSD).

Lorsque l'impact moyen est utilisé pour apprécier l'impact des articles (cf. Figure 3), on observe un déplacement de Berkeley dans la structure de dominance. En effet, Berkeley échappe maintenant à la domination de Stanford et domine désormais le MIT. Cette évolution est partiellement remise en cause lorsque l'impact moyen relatif est utilisé (Figure 4).

En outre, d'autres universités telles que Seattle et Michigan voient leurs positions réévaluées dans ce contexte. Ces deux universités dominent désormais UCLA qui apparaît donc relativement moins performante lorsqu'on se concentre sur sa capacité à placer des articles dans les meilleures revues de chaque spécialité.

<sup>6</sup> Lorsque  $\emptyset$  diminue, plus de relations de dominance peuvent être établies alors qu'aucune ne peut être remise en cause. Le réseau de dominance tend donc naturellement à se hiérarchiser.

Figure 1- Réseau de dominance faible ajusté parmi les meilleures universités américaines (toutes disciplines), quand l'impact est mesuré par les citations directes (avec  $\delta = 1$ )

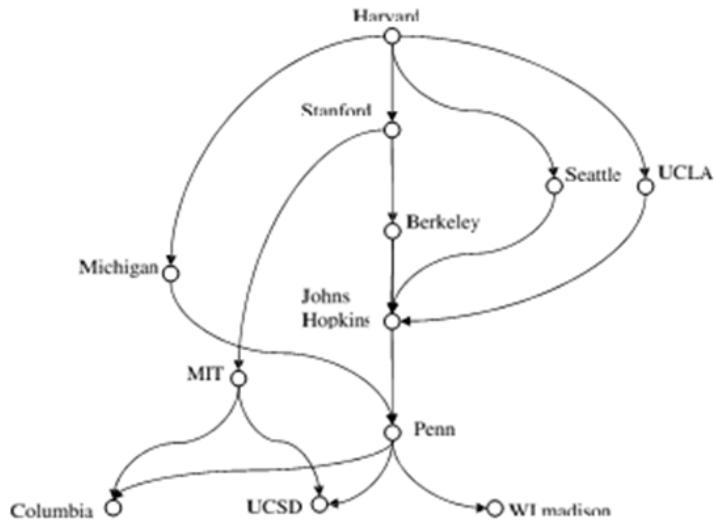


Figure 2 - Réseau de dominance faible ajusté parmi les meilleures universités américaines (toutes disciplines), quand l'impact est mesuré par l'impact moyen, (avec  $\delta = 0,1$ )

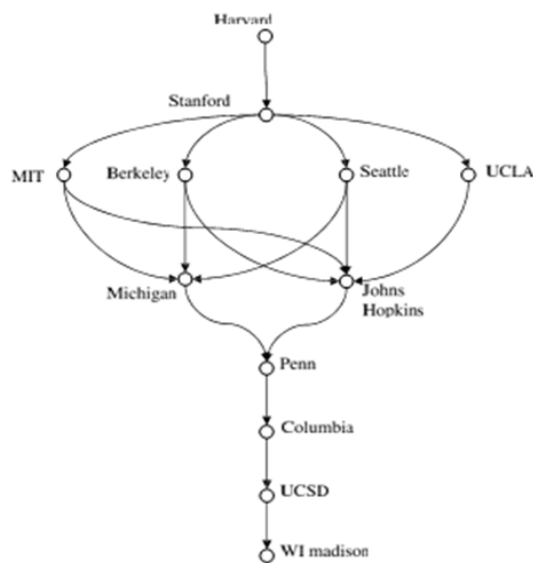


Figure 3 - Réseau de dominance faible ajusté parmi les meilleures universités américaines (toutes disciplines), quand l'impact est mesuré par l'impact moyen (avec  $\varnothing = 1$ )

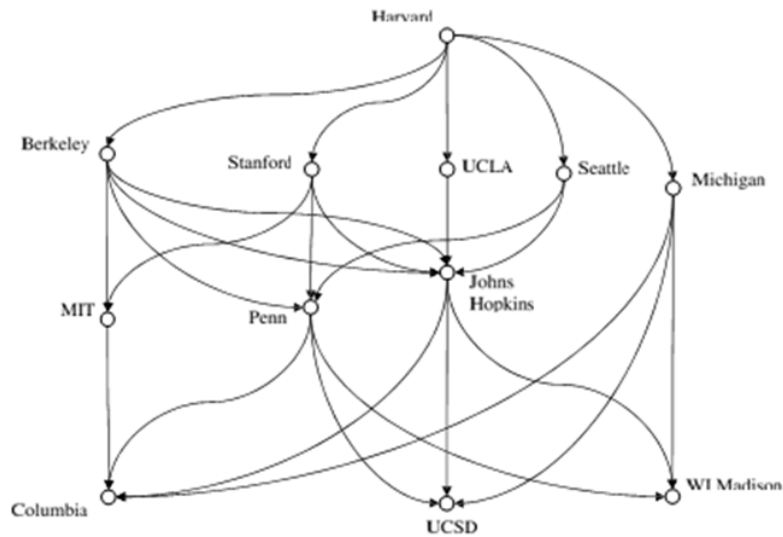
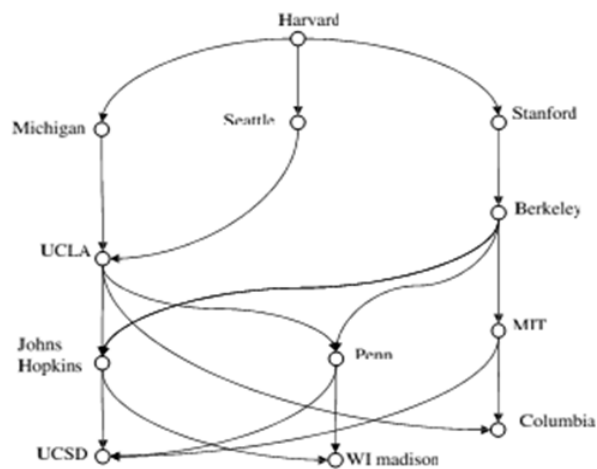


Figure 4 - Réseau de dominance faible ajusté parmi les meilleures universités américaines (toutes disciplines), quand l'impact est mesuré par l'impact moyen relatif (avec  $\varnothing = 1$ )



Pour un type de dominance donné, deux institutions qui n'ont entre elles aucune relation de dominance, appartiennent chacune à la classe de référence de l'autre. Ceci peut être déduit de la lecture des réseaux de dominance ajustés, par le fait qu'il n'y a pas de chemin entre ces deux institutions. La classe de référence d'une institution comprend toutes les institutions qui ne peuvent être comparées à elle, c'est à dire qu'elles sont à la fois proches et différenciées. Par exemple, Stanford et Seattle qui n'ont pas, entre elles, de relation de dominance faible (impact mesuré par les citations), appartiennent chacune à la classe de référence de l'autre. Lorsque l'on étudie la structure de la production scientifique de ces deux institutions, on s'aperçoit que, si ces deux universités ont des productions scientifiques proches leur permettant de se situer toutes les deux au sommet de la hiérarchie universitaire, leur structure de production au regard des citations directes sont effectivement différentes. Stanford a des performances relativement plus élevées parmi les publications les plus citées, alors que Seattle domine significativement dans des segments d'impact plus faible. Ainsi, lorsque l'on se limite aux publications faisant partie des dix pourcent les plus cités de chaque discipline, Seattle est dominée par Stanford : les deux institutions ne font plus partie de leurs classes de référence respectives.

## Conclusion

**N**ous avons proposés différents outils de comparaison des productions scientifiques des universités (classements, classes, réseaux) qui reposent sur une théorie permettant d'établir différentes relations de dominance entre universités sur la base de leur production scientifique et de leur impact. Face à l'expansion récente des classements d'universités, cette approche est originale à plusieurs titres. Tout d'abord elle permet d'affiner les analyses de positionnement des institutions de recherche, en donnant aux utilisateurs potentiels une marge de liberté dans les critères utilisés.

Ensuite, cette approche permet de relier des prémisses explicites sur la valorisation de l'impact scientifique à des comparaisons et des classements. Enfin, ce type d'approche permet à la fois de proposer une méthode de classements et une analyse des structures de réseaux entre institutions en établissant des relations de dominance deux à deux.

Nous avons illustré son application sur un large échantillon d'universités de recherche américaines. Cette théorie pourrait être appliquée à de nombreux contextes dans lesquels un évaluateur, ou une communauté d'évaluateurs, souhaiterait comparer des individus ou des institutions, en fondant leurs jugements sur des principes simples, explicites et ajustables valorisant de différentes manières la quantité et la qualité de leur production scientifique. ■

## Références

- Atkinson, A. B., 1970, On the measurement of inequality, *Journal of Economic Theory* 2, 244-263.
- Carayol, N., Lahatte A., 2011, Dominance relations and universities ranking, *mimeo*.
- Carayol N., Filliatreau G., Lahatte A., 2011, Impact, dominance et classements des universités, *mimeo*.
- Cheng Y., Zitt M., 2009, Towards a global ranking by size-independent bibliometric measures: An experiment on selected US, French and Chinese Universities, *mimeo OST*.
- Garfield, E. 1956, Citation Indexes for Science, *Science* 123, 61-62.
- Merton R.K., 1973, *The sociology of science. Theoretical and empirical investigations*, University Press of Chicago, Chicago.
- Price, de Solla, D.J., 1963, *Little Science, Big Science*, New York: Columbia University Press.
- Quirk, J.P., Saposnick, R., 1962, Admissibility and measurable utility functions, *Review of Economic Studies* 29, 140-146.
- Rothschild, M., Stiglitz, J. E., 1970, Increasing Risk: I. A definition, *Journal of Economic Theory* 2, 225-243.



## Singularités de la position des établissements d'enseignement supérieur français dans les classements internationaux

**Philippe Vidal, chargé de mission à l'OST, professeur émérite à l'Université de Clermont-Ferrand 2**

*Philippe.vidal@obs-ost.fr*

La position de la France dans les classements internationaux des établissements d'enseignement supérieur (EES) n'est pas en rapport avec son niveau de production scientifique. En effet, la France est située entre la 6<sup>e</sup> et la 7<sup>e</sup> place mondiale en terme de production scientifique et à la 13<sup>e</sup> place en terme de visibilité, alors que dans les trois classements internationaux généralistes, l'Academic Ranking of World Universities (ARWU ou classement de Shanghai), le Times Higher Education (THE), et Quacquarelli Symonds (QS), la place des EES français est très en deçà de ce qu'on observe pour des pays comparables en termes de niveau de dépenses de R&D, comme l'Allemagne et le Royaume-Uni. La France a seulement 8 établissements dans ARWU en comparaison de l'Allemagne qui en a 14, ou du Royaume-Uni qui en a 16. De même, la France dispose de seulement 5 établissements dans le classement THE, à comparer à 12 établissements pour l'Allemagne et 32 établissements pour le Royaume-Uni. De même, la France dispose de 5 établissements dans le classement QS à comparer à 11 pour l'Allemagne et 30 pour le Royaume-Uni. Enfin, le premier EES français dans le classement de l'ARWU, l'Université de Paris XI, ne se situe qu'à la 40<sup>e</sup> place.

Les mécanismes de cette situation ont été largement analysés (cf. par exemple l'article de Vidal et Filliatreau, 2011). Il s'agit, ici, de s'intéresser à une comparaison détaillée de ses effets sur les trois classements déjà mentionnés. Le classement de l'ARWU, existe depuis

2003 avec une méthodologie constante. Le classement du THE a modifié plusieurs fois sa méthodologie, ce qui a pour effet de faire varier la place des EES de manière parfois spectaculaire. Cependant, des constantes peuvent être mises en évidence, pour les établissements français au moins. Le classement QS est plus récent, mais il est repris dans celui de l'US News and World Report.

### Les établissements français au sein du classement ARWU

Le classement ARWU est fondé sur l'agrégation de six indicateurs, tous relatifs à la recherche :

- 1) prix Nobel et médailles Fields des anciens élèves (10 %),
- 2) prix Nobel et médailles Fields dans le staff (20 %),
- 3) nombre de chercheurs les plus cités de leur discipline (20 %),
- 4) nombre de publications dans la revue Nature et Science (20 %),
- 5) nombre d'articles indexés dans le Science Citation Index et dans le Arts and Humanities Citation Index (20 %),
- 6) rapport performance/taille (10 %).

L'ARWU classe les 500 premiers EES mondiaux. Il produit un palmarès de 1 à 100, puis des paquets, de 100 à 150, 150-200, 200-300, 300-400, et 400-500.

Trois établissements français (Paris VI, Paris XI, ENS-Paris) se situent dans le Top 100 (respectivement 41<sup>e</sup>, 40<sup>e</sup>, 69<sup>e</sup>), trois dans le groupe 100-150, 2 dans le groupe 150-200, 5 dans le groupe 200-300, 4 dans le groupe 300-400, et 4 dans le groupe 400-500 (tableau 1).

**Philippe Vidal**, chargé de mission à l'OST et professeur émérite à l'Université de Clermont-Ferrand 2, a contribué à de nombreuses réflexions sur les classements.



Tableau 1- Place des EES français dans les classements 2011 ARWU, THE et QS

	ARWU	THE	QS
Paris 11	40		208
Paris 6	41	84	119
<i>Ens Paris</i>	69	59	33
Strasbourg	107	215	227
Paris 7	114	169	238
Aix-Marseille	150	non analysé	non analysé
Paris 5	154	non classé	319
Grenoble 1	163	non classé	254
Toulouse 3	225	non classé	345
Lorraine	230	non analysé	non analysé
Montpellier 2	238	285	329
Lyon 1	239	non classé	350
Paris 9	293	non classé	324
<i>Ecole Polytechnique</i>	307	63	36
<i>Ecole Phys Chim Paris</i>	308	non classé	non classé
Bordeaux 1	345	non classé	non classé
Rennes 1	398	non classé	475
<i>Ecole Mines Paris</i>	426	non classé	non classé
Nice	439	non classé	non classé
<i>Ens Lyon</i>	494	141	133
Versailles St Quentin	500	non classé	non classé
Caen	non classé	non classé	600-700
Cergy	non classé	non classé	600-700
Grenoble 2	non classé	non classé	600-700
Grenoble 3	non classé	non classé	600-700
<i>Iep Paris</i>	non classé	non classé	203
Lille 1	non classé	non classé	346
Lille 2	non classé	non classé	600-700
Lille 3	non classé	non classé	600-700
Lyon 3	non classé	non classé	600-700
Montpellier 3	non classé	non classé	600-700
Paris 1	non classé	non classé	229
Paris 2	non classé	non classé	500-550
Paris 4	non classé	non classé	186
Paris 10	non classé	non classé	600-700
Paris 13	non classé	375	non classé
<i>Ponts ParisTech</i>	non classé	non classé	285
Toulouse 1	non classé	non classé	550-600
Toulouse 2	non classé	non classé	600-700

En italiques : les écoles

A noter qu'un élément nouveau en 2011 est l'arrivée, en assez bonne place, des PRES d'Aix-Marseille et de Lorraine, dont les établissements membres avaient décidé de fusionner.

Les barres d'erreurs inhérentes à l'exercice, et le fait que de très faibles écarts dans les indicateurs peuvent induire de grands changements de rang conduisent à ne considérer qu'avec prudence les fluctuations annuelles. En revanche, les évolutions sur le long terme, c'est-à-dire depuis la création du classement de l'ARWU, sont porteuses d'informations.

On peut constater dans le tableau 2 que :

- les établissements américains et britanniques ont des positions relativement stables et restent dominants au sein du top 100, bien que l'on constate une érosion constante dans les autres tranches du classement, due principalement à l'entrée de pays émergents.

- les établissements allemands sont, eux aussi, marqués par une érosion globale sur la période, sauf en 2011 où on observe, au contraire, un gain qui amène l'Allemagne à dépasser le Royaume-Uni ; si ce résultat devait se confirmer, il mériterait d'être analysé au regard de la politique d'excellence mise en place par le gouvernement allemand au cours des dernières années.

- les établissements chinois montrent une forte montée en puissance dans la tranche 400-500, préfigurant peut-être le début d'un mouvement vers les premières places.

En comparaison, la France apparaît comme globalement assez stable.

**Tableau 2 - Classement ARWU : évolution de la place de la France, de l'Allemagne, du Royaume-Uni, des Etats-Unis et de la Chine depuis 2004**

	2004	2005	2006	2007	2008	2009	2010	2011
<b>France</b>								
1-100	4	4	4	4	3	3	3	3
Total (500)	22	22	21	21	23	23	22	21
<b>Allemagne</b>								
1-100	7	5	5	6	6	5	5	6
Total (500)	43	40	40	41	40	38	36	39
<b>Royaume-Uni</b>								
1-100	11	11	11	11	12	11	11	10
Total (500)	41	40	43	42	43	40	38	37
<b>Etats-Unis</b>								
1-100	51	53	54	54	53	55	54	52
Total (500)	170	168	167	166	157	151	153	149
<b>Chine</b>								
1-100	0	0	0	0	0	0	0	0
Total (500)	13	13	14	17	23	23	23	28

## Les établissements français et les classements THE et QS

Les classements du THE et de QS accordent également une part importante à la recherche, mais ils s'appuient aussi sur des enquêtes de notoriété et sur les taux d'encadrement.

Le **classement du THE** comprend cinq volets :

- 1) recherche (30 %, dont 19,5 % pour l'enquête de réputation,
- 2) citations (32,5 %),
- 3) enseignement (30%, dont 15 % pour la réputation et 4,5 % pour l'encadrement des étudiants en licence),
- 4) proportion d'enseignants et d'étudiants étrangers (5 %),
- 5) relations avec le monde économique (2,5 %).

Le THE produit un palmarès de 400 établissements. Huit EES français sont classés (cf tableau 1, page 18), dont 3 dans le Top 100 : ENS-Paris, 59<sup>e</sup>, Ecole Polytechnique, 63<sup>e</sup>, Paris 6, 84<sup>e</sup>.

Le **classement de QS** comprend six volets :

- 1) réputation en enseignement supérieur et recherche (40 %)
- 2) réputation chez les employeurs (10 %)
- 3) rapport étudiant/staff académique (20 %)
- 4) citations par staff (20 %)
- 5) proportion de staff étranger (5 %)
- 6) proportion d'étudiants étrangers (5 %)

QS produit un palmarès de 1 à 400, puis des paquets de 50 jusqu'à 600, et un paquet final de 100. Trente et un EES français sont classés (cf tableau 1, page 18), dont 2 dans le Top 100 (ENS-Paris, 33<sup>e</sup>, Ecole Polytechnique 36<sup>e</sup>).

## Analyse comparée des trois classements pour les établissements français

Trente deux Universités françaises sont présentes dans l'ensemble des trois classements (cf tableau 1, page 18). Ce nombre est significatif, puisqu'il représente en gros 40 % des universités.

Les rangs des universités classées dans ARWU sont globalement bien corrélés avec les rangs dans THE et QS.

Les Ecoles sont proportionnellement beaucoup moins présentes puisqu'elles ne sont qu'au nombre de 7, soit moins de 4% des Ecoles non intégrées à une université délivrant un diplôme visé par l'Etat. Trois Ecoles (ENS Paris et Lyon, et Ecole Polytechnique) figurent dans les trois palmarès, et deux d'entre elles, l'ENS-Paris et l'Ecole Polytechnique, devancent les Universités dans THE et QS. Les rangs sont bien corrélés entre les trois classements. Quatre autres Ecoles sont présentes, mais de manière mal corrélées.

## La singularité française

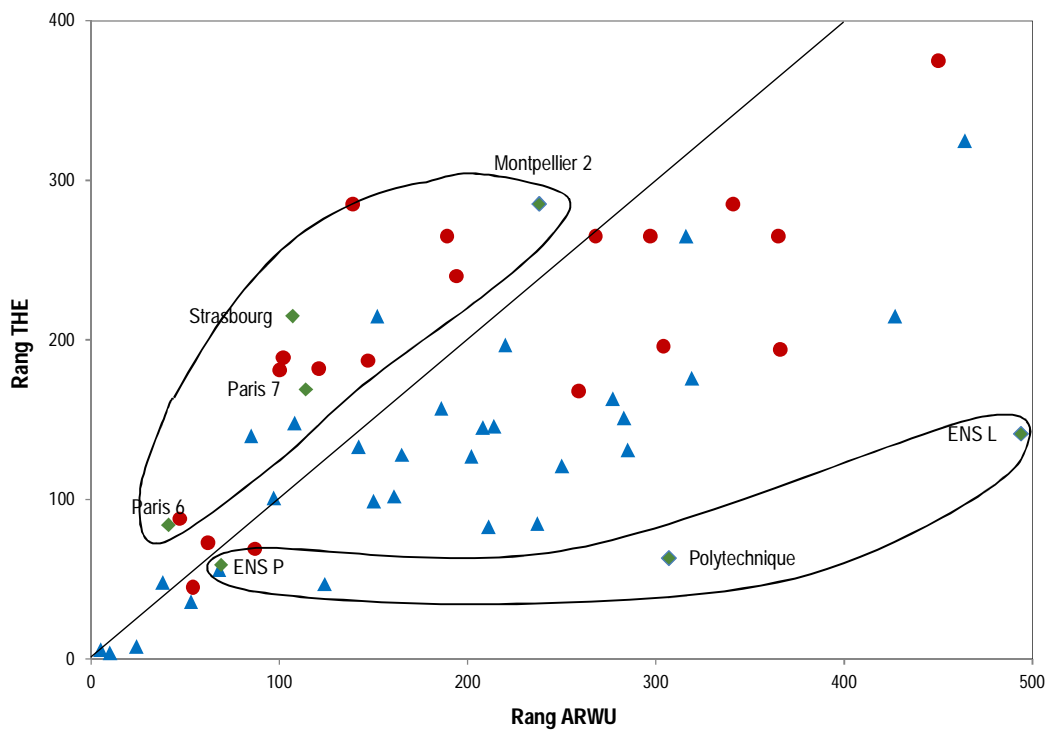
Il est intéressant de comparer la position relative des EES français par rapport aux établissements allemands et britanniques dans les différents classements. La comparaison QS par rapport à ARWU est la plus riche d'informations car portant sur un beaucoup plus grand nombre d'EES. Par ailleurs, on a vu ci-dessus que la méthodologie de QS est proche de celle du THE.

Dans les diagrammes reportant les rangs des classements THE, QS, et ARWU pour les EES français et leurs homologues d'Allemagne et du Royaume-Uni (figures 1 et 2), on voit clairement que les universités françaises sont décalées par rapport à la première bissectrice. Elles occupent des places moins favorables dans les classements THE et QS. On constate également que les trois Ecoles qui apparaissent sont à l'écart de la tendance générale des universités, et se trouvent en position favorable dans THE et QS, et défavorable dans ARWU.

La singularité des positions des EES français par rapport aux EES allemands et britanniques est probablement liée à plusieurs causes. Les pistes suivantes mériteraient d'être explorées et quantifiées :

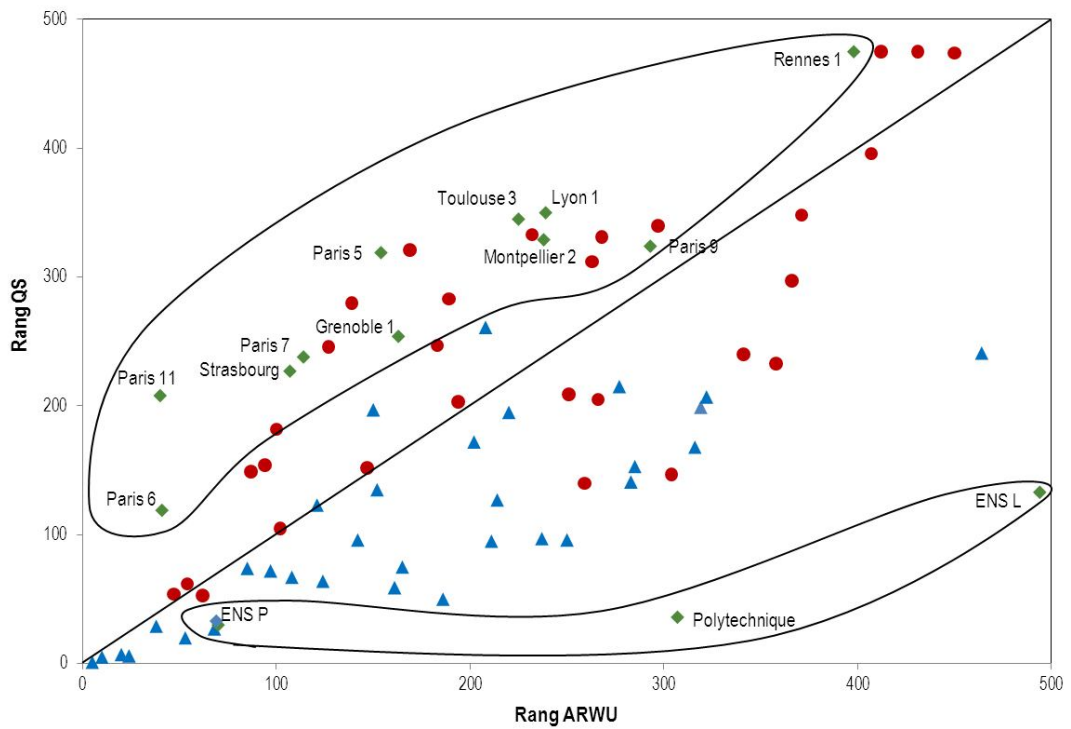
**1) première piste de réflexion liée au ratio entre staff/ étudiant** : si l'on considère l'indicateur du rapport staff académique/étudiant donné par QS, les cinq premiers EES français sont des Ecoles, Paris VI, première université, n'arrivant qu'en sixième position. En revanche, si l'on recalcule les données QS sans cet indicateur, Paris VI passe en troisième position. Ceci n'est pas surprenant car on a vu ci-dessus que cet indicateur pèse pour 20 % dans QS et 4,5 % dans THE.

Figure 1 - Rang THE en fonction du rang ARWU pour les EES allemands, britanniques, et français



◆ EES français ; ● EES allemands ; ▲ EES anglais

Figure 2 - Rang QS en fonction du rang ARWU pour les EES allemands, britanniques, et français



◆ EES français ; ● EES allemands ; ▲ EES anglais

Il n'apparaît pas explicitement dans l'ARWU, mais joue indirectement dans la mesure où les lourdes charges liées à l'enseignement sur les 3 premières années obèrent le potentiel en recherche du staff. Il a donc clairement pour effet d'avantager les Ecoles.

En toute rigueur d'ailleurs, les trois classements discutés ici ne concernent que les universités au sens strict, c'est-à-dire celles où le niveau Licence (Bachelor) est représenté. Le THE stipule d'ailleurs « Universities were excluded from the World University Rankings tables if they do not teach undergraduates ». Les classements THE et QS ne sont par conséquent pas applicables aux Ecoles françaises, sauf celles, telles que les INSA, où les niveaux Bac + 1 et 2 sont représentés. Elles sont pourtant prises en compte jusqu'ici au motif qu'elles possèdent le niveau 3 (admission post-prépa) et peuvent donc exciper de la présence d'étudiants du niveau Bachelor. En conclusion, afin de rendre le paysage français dans les classements plus compréhensible, il serait judicieux de mieux expliciter la structuration duale de l'enseignement supérieur français, et de l'accompagner d'un classement spécifique à ceux des EES qui recrutent

au niveau Bac + 3. Le classement européen U-Multirank (voir ci-dessous) vise justement à ne mettre en comparaison que des EES qui possèdent le même profil.

**2) deuxième piste de réflexion lié aux indicateurs de taille** : il a été largement remarqué que le classement ARWU, ainsi que THE et QS (même si cela est moins explicite) est largement dominé par l'effet de taille, qui favorise généralement les gros établissements (à quelques exceptions près telles que Caltech et MIT). La petite taille des Ecoles françaises pourrait donc expliquer pour partie leurs moins bonnes places dans le classement ARWU par rapport aux classements THE et QS.

**3) troisième piste de réflexion liée aux indicateurs de notoriété** : il a été largement remarqué également que l'ancienneté des EES joue globalement un rôle positif dans leur notoriété, et donc dans leurs rangs dans les classements THE et QS. Les Ecoles françaises classées peuvent à cet égard bénéficier d'un avantage car globalement de fondation beaucoup plus ancienne que les universités.

Tableau 3 - Simulation de la position des PRES dans le classement ARWU

Institutions	Rang mondial
Campus Plateau de Saclay - Ens sup	19
PRES PARIS SCIENCES ET LETTRES ETOILE	30
PRES SORBONNE UNIV	33
PRES UNIV PARIS CITE	43
PRES UNIV GRENOBLE	100-200
PRES UNIV LYON	100-200
PRES UNIV TOULOUSE	100-200
PRES UNIV LANGUEDOC ROUSSILLON	200-300
PRES UNIV BORDEAUX	200-300
PRES UNIV EUROP DE BRETAGNE	200-300
PRES UNIV LILLE	200-300
PRES UNIV NANTES ANGERS LE MANS - UNAM	300-400
PRES LIMOUSIN POITOU CHARENTES	400-500
PRES CENTRE VAL DE LOIRE	400-500
PRES UNIV PARIS EST	400-500
PRES Bourgogne Franche-Comté	400-500

## Perspectives

Comme l'a montré le classement ARWU de 2011, la fusion des EES français, rendue possible par la dynamique des PRES, a permis la création d'universités "comprehensives" de plus grande taille, beaucoup plus visibles au niveau international

Des simulations, réalisées en parallèle pour d'autres PRES ont montré de la même manière, que la plupart des fusions étaient porteuses de gain de visibilité, qui se traduisait par une progression notable de leurs positions dans les classements. Sur ce point voir le tableau 3 (page 22), simulant la position des PRES faite par ARWU, sur la base des mots-clés fournis par l'OST pour le repérage des publications. Les résultats montrent que, si on intégrait ces PRES dans le classement ARWU, les établissements français représentés seraient plus nombreux (21 au lieu de 16) et mieux positionnés que les établissements isolés qui les constituent. Ce résultat n'est pas inattendu, mais il est ici précisé pour ARWU et, en s'appuyant sur la comparaison faite plus haut, on peut faire l'hypothèse qu'il en serait de même dans le classement THE, où l'effet pourrait être amplifié par la présence de certaines Ecoles.

Il convient, enfin, de signaler que le projet de classement U-Multirank, pluridimensionnel, multiparamètres et personnalisé (cf. l'article de Ghislaine Filliatreau dans ce numéro), devrait permettre de mieux mettre en valeur les forces de la France en matière d'enseignement supérieur, ainsi que les avantages qui devraient pouvoir être tirés de sa très grande diversité institutionnelle. Ce projet, dont l'impulsion initiale revient à la France, a fait l'objet d'un pilote dont les résultats sont consultables sur le site de la Commission Européenne ([http://ec.europa.eu/education/higher-education/doc/multirank\\_en.pdf](http://ec.europa.eu/education/higher-education/doc/multirank_en.pdf)). ■

## Références

Vidal P. et Filliatreau G, (2011), Attractivité de la France dans le domaine de l'enseignement supérieur : points forts, points faibles, *Campusfrance, Repères* n° 7.

## Remerciements

L'auteur remercie M. Ying Cheng (directeur du Center for World-Class Universities de l'Université Jiao Tong à Shanghai) de lui avoir aimablement communiqué des données détaillées sur le classement de l'ARWU.

## L'espace du Rhin Supérieur dans les classements internationaux : réflexions sur un territoire transfrontalier et sur les méthodes d'évaluation

**Jean-Alain HERAUD, BETA, Université de Strasbourg et CNRS**

*heraud@cournot.u-strasbg.fr*

Cet article se propose de présenter un territoire de projet transfrontalier, la Région Métropolitaine Trinationale (RMT) du Rhin Supérieur (qui

englobe géographiquement l'Alsace, la Suisse du Nord-Ouest, le Pays de Bade et la partie Sud du Palatinat) du point de vue de sa production scientifique et plus généralement de son poids académique. A cette occasion, nous évoquerons des problèmes méthodologiques, relatifs au choix des indicateurs les plus pertinents. L'enjeu politique est celui d'un projet qui se veut un modèle de coopération pour l'Europe de la connaissance, de l'innovation et de la créativité. Des travaux ont déjà été entrepris (en collaboration entre le BETA et l'OST, avec le soutien de la Conférence du Rhin Supérieur) sur le périmètre de cet espace afin d'en estimer le poids scientifique dans l'ensemble européen pour les

**Jean-Alain Héraud,**  
Doyen de la faculté des sciences économiques et de gestion de l'Université de Strasbourg, professeur d'économie, est très impliqué dans les réseaux transfrontaliers du Rhin Supérieur



principales disciplines scientifiques<sup>1</sup>, mais aussi pour évaluer le degré de connectivité interne. La concentration de capacités scientifiques au sein de la RMT est tout -à-fait significative et son excellence a encore été soulignée en 2011 par l'attribution du Prix Nobel de Médecine à Jules Hoffmann de l'Université de Strasbourg. Ce prix s'ajoute à beaucoup d'autres depuis un siècle (pas seulement quand l'Alsace était française, et pas seulement en Alsace) qui font du Rhin Supérieur un territoire marqué par la créativité scientifique dans la période moderne comme il le fut à la Renaissance en matière de culture, d'arts et de technologie. L'actualité du Prix Nobel nous amène à commencer par appréhender la production et la notoriété scientifiques de la RMT et de ses composantes nationales (dont l'Alsace) par les évaluations du classement dit « de Shanghai ».

Le dernier classement de l'Université Jiao Tong de Shanghai<sup>2</sup>, fait apparaître comme chaque année un certain nombre d'établissements d'enseignement supérieur et de recherche européens dans son palmarès des 500 premiers mondiaux, classés selon des critères reflétant principalement la qualité de la science. Il a souvent été souligné que ce classement comporte, comme tout autre, des biais: accent mis sur la recherche plus que sur la qualité de la formation, importance sans doute excessive donnée à la science du plus haut niveau (prix Nobel) et principalement en « sciences dures », sous-estimation de la recherche appliquée, importance des effets de taille, etc<sup>3</sup>. Malgré tout, ce classement est universellement attendu, scruté et commenté. Il nous a donc paru intéressant d'analyser l'image de la RMT Rhin Supérieur dans un tel référentiel.

## Les établissements de la RMT et le classement de Shanghai

Les sites universitaires à repérer dans le classement ARWU pour ce qui concerne le périmètre de la RMT, sont : Bâle, Freiburg, Karlsruhe, Mulhouse-Colmar (Université de Haute Alsace, UHA) et Strasbourg. Dans la proximité de la RMT se trouve d'autres sites, parfois très bien classés : pour mémoire, indiquons que dans la liste des 100 premières universités au classement général, on trouve Zürich (avec l'ETH au 23<sup>e</sup> rang et l'Université au 56<sup>e</sup>) et Heidelberg (62<sup>e</sup> rang), sans parler d'autres universités qui émergent selon les disciplines, comme Tübingen, Stuttgart, Frank-

furt, ou Konstanz. Mais nous nous en tiendrons ici strictement à la RMT Rhin Sud.

Considérons la liste des 500 premiers établissements mondiaux qui est publiée pour le classement général (toutes disciplines confondues). On peut extraire de cette liste le palmarès « local » suivant :

- **Bâle** est en tête, au 89<sup>e</sup> rang mondial;
- viennent ensuite **Freiburg** et **Strasbourg** dans la classe [102-150]<sup>4</sup> ;
- puis **Karlsruhe** dans la classe [301-400]

Observons maintenant les classements par spécialité. Les listes fournies ici sont seulement celles des 100 premiers établissements mondiaux. ARWU distingue des « sujets » (mathématiques, physique, chimie, informatique, et économie-management) et des « domaines » plus larges qui regroupent des disciplines.

Commençons par les *sujets*, c'est-à-dire des disciplines pures.

- La seule discipline qui émerge est la Chimie. Strasbourg est 14<sup>e</sup> mondial (premier français) et **Karlsruhe** dans la classe [76-100]. Signalons toutefois que Strasbourg se classait encore dans les 100 premières universités pour les Mathématiques en 2009 et 2010.

Dans les « domaines » on trouve un peu plus de citations de la RMT :

- En « Sciences naturelles et mathématiques », **Strasbourg** est dans la classe [51-75]
- En « Sciences de la vie et agriculture », **Bâle** est 34<sup>e</sup> et **Freiburg** dans la classe [51-75]

<sup>1</sup> Voir la note evOREG N°11 de novembre 2010 : *Premiers indicateurs de production scientifique concernant la RMT Rhin Supérieur* <http://www.evoreg.eu/docs/files/shno/Note11.pdf>

<sup>2</sup> « *Academic Ranking of World Universities* » : ARWU, août 2011, <http://www.shanghairanking.com/ARWU2011.html>

<sup>3</sup> Notons qu'une partie des critiques porte en fait sur l'utilisation de l'indicateur synthétique de Shanghai plutôt que sur sa construction : il n'est en rien un outil d'évaluation des établissements et surtout pas en termes de mérite ou d'évolution. C'est un indicateur de résultat, principalement en volume, pointant la notoriété internationale brute de long terme.

<sup>4</sup> Remarque : au-delà du rang 100, les listes sont présentées sur le site de l'ARWU selon des classes de rangs à l'intérieur desquels les établissements sont simplement mentionnés par ordre alphabétique.



- En « Médecine clinique et pharmacie », **Bâle** est 4<sup>e</sup> et **Freiburg** dans la classe [76-100] <sup>5</sup>
- En « Ingénierie, technologie, informatique » comme en « Sciences sociales », aucun établissement de la RMT ne se classe dans les 100 premiers mondiaux.

Comme nous l'avons rappelé plus haut en introduction, ce classement dit « de Shanghai » utilise des critères assez particuliers, mettant l'accent sur la recherche fondamentale d'excellence. Ceci explique sans doute que le KIT de Karlsruhe, une des premières universités d'excellence allemande, mais plus technologique que fondamentale, n'apparaisse pas plus souvent dans le haut des classements. Rappelons aussi que les critères retenus ne donnent que peu de chances d'émerger à des universités de taille moyenne <sup>6</sup>.

Une approche complémentaire a déjà été mise en œuvre, avec la mesure bibliométrique réalisée par le BETA et l'OST.

## Vers une bibliométrie plus approfondie

Les productions de la RMT dans les principales disciplines scientifiques ont été mesurées pour l'année 2008 à partir de la base de publications Thomson-Reuters. Cette évaluation est convergente avec le palmarès de Shanghai sur un point essentiel : la Chimie est la discipline où la région transfrontalière pèse le plus lourd en comparaison européenne, puisque la RMT représente 2,6 % des publications de l'Union Européenne plus la Suisse et la Norvège (UE27+). Sont également confirmées, les performances en sciences de la vie et de la santé (entre 1,5 et 2 % de l'UE27+).

Ce qui n'est par contre pas mis en évidence dans le classement de Shanghai, c'est la contribution notable de la RMT en physique et en sciences pour l'ingénieur (autour de 2%). Il faut croire que le biais du classement joue particulièrement ici : les établissements du Rhin Supérieur ont une production scientifique abondante, mais pas toujours au niveau de l'excellence mondiale – qui est typiquement couronnée par le prix Nobel. Autre facteur explicatif de la divergence des résultats entre notre mesure européenne et le classement mondial ARWU : l'Europe est globalement moins bien représentée au niveau mondial en physique et sciences de l'ingénieur qu'en chimie, ce qui relativise quelque peu

le score européen de la RMT dans ces domaines. Inversement, bien sûr, l'excellent score rhénan en chimie au niveau européen apparaît particulièrement significatif, puisque c'est une spécialisation relative de notre continent.

La question méthodologique qui se pose pour aller plus loin dans la caractérisation d'un territoire comme celui du Rhin Supérieur <sup>7</sup>, c'est de trouver et mettre en œuvre des indicateurs plus orientés vers les applications. L'enjeu est important en Alsace, une région qui se trouve dans une situation particulière soulignée depuis longtemps par les observateurs du système régional d'innovation et les décideurs politiques : le territoire abrite une infrastructure scientifique très développée et productive en science fondamentale (publique), mais peu articulée sur le tissu économique. Ce dernier se révèle assez innovant tout en faisant paradoxalement peu de R&D formelle et en recourant rarement à des partenariats avec le monde académique (sauf dans les pôles de compétitivité). L'explication réside dans une tradition culturelle assez présente dans le Rhin Supérieur, faite de pragmatisme prudent et de méfiance vis-à-vis de la pensée théorique. L'entreprise moyenne alsacienne (ou l'établissement de multinationale présent sur le territoire) valorisent plus les techniciens que les chercheurs, et l'innovation incrémentale que les idées de rupture.

L'ambition politique qui consiste à stimuler la créativité du territoire en l'ouvrant sur de nouveaux schémas de pensée – à renouer avec la tradition de la Renaissance rhénane, d'une certaine manière <sup>8</sup> – devrait s'accompagner d'instruments de mesure appropriés. D'où la nécessité de développer de nouveaux indicateurs statis-

<sup>5</sup> Avec le Prix Nobel de Médecine décerné à Jules Hoffmann, Strasbourg fera certainement un bond dans les prochains classements. En termes de disciplines fondamentales, ce succès montre aussi que la recherche en biologie est très présente et de haute qualité sur le site strasbourgeois.

<sup>6</sup> D'ailleurs, si le rapprochement programmé entre l'UHA et l'UdS se renforce à l'avenir, des centres de recherche alsaciens hors de Strasbourg finiront par s'agréger à la production scientifique strasbourgeoise.

<sup>7</sup> et pour orienter les politiques visant à stimuler l'innovation et la créativité dans des domaines qui ne se limitent pas à la production scientifique fondamentale.

<sup>8</sup> J.-A. Héraud, "Reinventing creativity in old Europe: a development scenario for cities within the Upper Rhine Valley cross-border area", *City, Culture and Society* 2 (2011) pp.65-73.

tiques. En bibliométrie, il serait déjà important d'étendre les bases de données aux disciplines comme les sciences humaines et sociales, aux humanités et aux arts. Pour cela une source comme le *Web of Science* ne suffit pas. De même les indicateurs de créativité technique comme les brevets ne couvrent qu'une partie du champ : peut-on mobiliser d'autres sources statistiques comme les dépôts de marques? Pour appréhender la créativité du territoire, les mesures concernant la proportion d'individus classés dans les diverses catégories de « classes créatives » constitue une autre piste.

## Conclusion

**N**ous concluons sur un des points importants abordés dans cet article, à savoir le fait que les succès des universités techniques (exemple du KIT à Karlsruhe) sont insuffisamment couverts par les

indicateurs habituels, et en particulier minorés dans le classement de Shanghai. Une évaluation plus complète des mérites de tels établissements supposerait de rechercher de l'information sur les relations industrielles, la formation professionnelle, etc. Dans le cas français, un indicateur comme la proportion de financements Cifre dans le total des thèses soutenues apporte une information complémentaire sur le degré de connectivité de la recherche académique avec les secteurs d'application. Il faudrait multiplier la récolte de telles données. Il est clair que l'on perd en comparabilité internationale, voire inter-régionale, quand on commence à considérer des données aussi variées, mais c'est un défi qu'il vaut la peine de relever. ■

**Collection Notes et Comptes-rendus**

- ☀ *Notes et Comptes-rendus n° 1*, octobre 2010  
Ateliers « Innovation et territoires » en collaboration OST - CEA  
« *Exploiter les données et indicateurs pour positionner les capacités de R&D et d'innovation des régions* ».
- ☀ *Notes et Comptes-rendus n° 2*, janvier 2011  
Ateliers « Innovation et territoires » en collaboration OST - CEA  
« *Les indicateurs pour mesurer le poids d'un opérateur de R&D sur un territoire* »
- ☀ *Notes et Comptes-rendus n° 3*, juillet 2011  
Ateliers « Innovation et territoires » en collaboration OST - CEA  
« *Les indicateurs de R&D et d'innovation pour le suivi des SRI* »
- ☀ *Notes et Comptes-rendus n° 4*, décembre 2011  
Ateliers « Innovation et territoires » en collaboration OST - CEA  
« *Les indicateurs de la créativité pour les régions françaises* »

**Collection Résultats et recherches**

- ☀ *Résultats et recherches n° 1*, mars 2012  
Les classements internationaux : enjeux, méthodologies et perspectives pour les universités françaises

Les études de l'Observatoire des Sciences et des Techniques ([www.obs-ost.fr](http://www.obs-ost.fr)) reposent sur les compétences et le travail de l'ensemble de l'équipe.

**Observatoire des Sciences et des Techniques**  
21, boulevard Pasteur - 75015 Paris  
Tél. : 01 44 39 06 80 / Fax : 01 45 48 63 94 / [www.obs-ost.fr](http://www.obs-ost.fr)

**Directrice de publication** : Ghislaine Filliatreau      **Rédactrice en chef** : Valérie Mérindol  
**Réalisation** : Marie-Laure Taillibert

**Point de contact** : Marie-Laure Taillibert ; [marie.laure.taillibert@obs-ost.fr](mailto:marie.laure.taillibert@obs-ost.fr)

ISBN : *en cours*