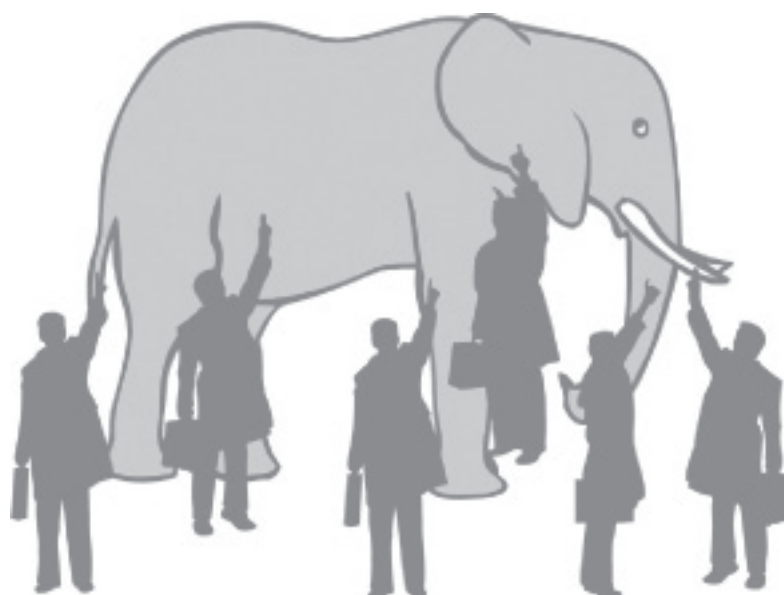


EUA REPORT ON RANKINGS 2011

LES CLASSEMENTS MONDIAUX D'UNIVERSITÉS ET LEUR IMPACT

Andrejs Rauhvargers



« Les aveugles et l'éléphant »

de John Godfrey Saxe (1816–1887)

Six hommes d'Hindoustan,
Très enclins à parfaire leurs connaissances,
Allèrent voir un éléphant
(Bien que tous fussent aveugles)
Afin que chacun, en l'observant,
Puisse satisfaire sa curiosité.

Le *premier* s'approcha de l'éléphant
Et perdant pied,
Alla buter contre son flanc large et robuste.
Il s'exclama aussitôt :
« Mon Dieu ! Mais l'éléphant
Ressemble beaucoup à un mur ! »

Le *second*, palpant une défense,
S'écria : « Ho ! qu'est-ce que cet objet
Si rond, si lisse et si pointu ?
Il ne fait aucun doute
Que cet éléphant extraordinaire
Ressemble beaucoup à une lance ! »

Le *troisième* s'avança vers l'éléphant
Et, saisissant par inadvertance
La trompe qui se tortillait,
S'écria sans hésitation :
« Je vois que l'éléphant
Ressemble beaucoup à un serpent ! »

Le *quatrième*, de sa main fébrile,
Se mit à palper le genou.
« De toute évidence, dit-il,
Cet animal fabuleux
Ressemble à un arbre ! »

Le *cinquième* toucha par hasard à l'oreille
Et dit : « Même le plus aveugle des hommes
Peut dire à quoi ressemble le plus l'éléphant ;
Nul ne peut me prouver le contraire,
Ce magnifique éléphant
Ressemble à un éventail ! »

Le *sixième* commença tout juste
À tâter l'animal,
La queue qui se balançait
Lui tomba dans la main.
« Je vois, dit-il, que l'éléphant
Ressemble beaucoup à une corde ! »

Ainsi, ces hommes d'Hindoustan
Discutèrent longuement,
Chacun faisant valoir son opinion
Avec force et fermeté.
Même si chacun avait partiellement raison,
Tous étaient dans l'erreur !

MORALE :
Trop souvent dans les guerres théologiques,
Les parties en conflit, je crois,
Se moquent dans la plus totale ignorance
De ce que veulent dire leurs opposants,
Et palabrent à n'en plus finir sur un Éléphant
Qu'aucune d'entre elles n'a jamais vu !

It was six men of Indostan
To learning much inclined,
Who went to see the Elephant
(Though all of them were blind),
That each by observation
Might satisfy his mind.

The *First* approach'd the Elephant,
And happening to fall
Against his broad and sturdy side,
At once began to bawl:
« God bless me! but the Elephant
Is very like a wall ! »

The *Second*, feeling of the tusk,
Cried, - « Ho ! what have we here
So very round and smooth and sharp ?
To me 'tis mighty clear
This wonder of an Elephant
Is very like a spear ! »

The *Third* approached the animal,
And happening to take
The squirming trunk within his hands,
Thus boldly up and spake:
« I see, » quoth he, « the Elephant
Is very like a snake ! »

The *Fourth* reached out his eager hand,
And felt about the knee.
« What most this wondrous beast is like
Is mighty plain, » quoth he,
« 'Tis clear enough the Elephant
Is very like a tree ! »

The *Fifth*, who chanced to touch the ear,
Said : « E'en the blindest man
Can tell what this resembles most ;
Deny the fact who can,
This marvel of an Elephant
Is very like a fan ! »

The *Sixth* no sooner had begun
About the beast to grope,
Then, seizing on the swinging tail
That fell within his scope,
« I see, » quoth he, « the Elephant
Is very like a rope ! »

And so these men of Indostan
Disputed loud and long,
Each in his own opinion
Exceeding stiff and strong,
Though each was partly in the right,
And all were in the wrong !

MORAL :
So oft in theologic wars,
The disputants, I ween,
Rail on in utter ignorance
Of what each other mean,
And prate about an Elephant
Not one of them has seen !

TABLE DES MATIÈRES

ÉDITORIAL ET REMERCIEMENTS	7
SIGLES ET ACRONYMES	9
GLOSSAIRE	10
RÉSUMÉ	11
I. INTRODUCTION	18
Objet du rapport	18
Brève histoire des classements	19
Aperçu des implications des classements	20
Principes de Berlin sur le classement des établissements d'enseignement supérieur	21
II. MÉTHODOLOGIES DES CLASSEMENTS MONDIAUX LES PLUS RÉPUTÉS	23
1. Classements internationaux publiés sous forme de palmarès	24
2. Classements axés uniquement sur la performance de la recherche (avec ou sans palmarès)	39
3. Classements multidimensionnels	46
4. Classements Internet	60
5. L'étalement concurrentiel sur la base des acquis de l'apprentissage	62
III. ANALYSE DES RÉSULTATS (IMPLICATIONS POUR LES UNIVERSITÉS)	64
Quelles sont les universités prises en considération dans les classements mondiaux d'université publiés sous forme de palmarès ?	64
Les classements et la mission de recherche des universités	65
Les classements et la mission d'enseignement des universités	67
Biais et défauts	68
Le risque d'accorder trop d'importance aux positions dans les classements	70
IV. PRINCIPALES CONCLUSIONS	72
V. INTERPRÉTER LES RÉSULTATS DES CLASSEMENTS	73
BIBLIOGRAPHIE	79

ÉDITORIAL ET REMERCIEMENTS

Le présent rapport a été établi à la demande de l'EUA en raison du nombre croissant de classements internationaux et nationaux dans l'enseignement supérieur, en réponse aux demandes de plus en plus nombreuses d'informations et de conseils adressées par les institutions membres quant à la nature de ces classements, en raison de l'intérêt manifesté par les gouvernements nationaux à l'égard des systèmes de classement et, enfin, compte tenu de la décision de la Commission européenne d'établir un « classement européen ».

Un groupe de travail ad hoc du Conseil de l'EUA a d'abord été constitué en 2009 afin d'examiner comment l'Association pourrait répondre au mieux aux besoins de ses membres à cet égard. Cette initiative a débouché sur une décision du Conseil de lancer un projet pilote en vue de publier en 2011 la première d'une série d'études de l'EUA sur les classements.

Ce projet a été confié à un Comité éditorial, présidé par le Président de l'EUA, le professeur Jean-Marc Rapp, ancien recteur de l'Université de Lausanne, et composé des membres suivants : le Professeur Jean-Pierre Finance, Président de l'Université Henri Poincaré et membre du Bureau de l'EUA ; le Professeur Howard Newby, Vice-président de l'Université de Liverpool ; le Professeur Jens Oddershede, Vice-président de l'Université du Danemark du Sud et Président de la Conférence des recteurs danois. Le Professeur Andrejs Rauhvargers, Secrétaire général de la Conférence des recteurs lettons, a accepté notre invitation à réaliser cette analyse et est l'auteur du présent rapport.

La Fondation Gulbenkian et la Fondation Robert Bosch nous ont fait l'honneur d'accepter de soutenir ce projet sur une période de deux ans.

Le rapport porte principalement sur les classements internationaux et examine également un certain nombre d'autres projets en cours visant à mesurer la performance des universités. Il présente une description et une analyse des méthodologies utilisées par les principaux classements internationaux sur la seule base d'informations rendues publiques et librement accessibles. Il a pour but de répondre aux besoins des membres de l'EUA, qui subissent souvent des pressions pour figurer dans les classements ou y améliorer leur position d'une manière ou d'une autre.

Il apparaît clairement que, malgré leurs lacunes, des biais et des défauts manifestes, les classements sont désormais bien établis. « Les parties prenantes et le grand public les acceptent facilement car ils proposent des informations simples, prêtes à consommer » (Groupe d'experts sur l'évaluation de la recherche universitaire – AUBR, 2009). C'est pourquoi il est important que les universités prennent conscience du degré de transparence des classements du point de vue de l'utilisateur, de la relation entre ce qui est annoncé comme étant mesuré et ce qui est effectivement mesuré, des méthodes de calcul utilisées pour établir les « scores », et de leur signification.

Il convient toutefois de souligner que les classements internationaux sous leur forme actuelle ne prennent en compte que 1 à 3 % (entre 200 et 500) des 17 000 universités qui existent dans le monde, se désintéressant du reste. Ils ne concernent directement qu'environ la moitié des membres de l'EUA, dans un nombre limité de pays dans lesquels l'EUA compte des membres, et sont étroitement liés à la richesse de ces pays.

Le rapport confirme que la plupart des classements internationaux se concentrent principalement sur des indicateurs liés à la fonction de recherche des universités. Les tentatives de mesure de la qualité de l'enseignement et de l'apprentissage supposent généralement l'utilisation d'indicateurs indirects, peu en rapport avec le processus d'enseignement et rarement pertinentes. L'importance des liens que l'université peut entretenir avec des partenaires externes et avec son environnement en général est largement ignorée. Lorsque des données existantes sont disponibles, elles ne sont souvent pas utilisées de manière systématique, et les facteurs liés à la réputation acquièrent dans de nombreux cas une importance disproportionnée. Ensemble, ces facteurs débouchent sur une représentation simpliste de la mission, de la qualité et de la performance institutionnelle, qui manque de pertinence pour la vaste majorité des établissements, en particulier au moment où la diversification et la caractérisation des profils institutionnels figurent au premier rang des priorités à travers l'Europe.

En revanche, l'apparition des classements mondiaux au cours des dernières années a attiré l'attention sur l'enseignement supérieur et mis en lumière les universités, qui sont de plus en plus souvent comparées à l'échelle nationale et internationale. Les classements ont certainement contribué à promouvoir une plus grande responsabilité et à améliorer les pratiques de gestion. Ils ont encouragé la collecte de données plus fiables et ont été invoqués dans certains pays pour augmenter les investissements dans l'enseignement supérieur. Bien que l'on dise qu'ils peuvent être utilisés pour orienter le choix des consommateurs, peu d'éléments convaincants donnent à penser que ce soit le cas (si ce n'est en Asie ou dans le cas du classement CHE en Allemagne).

Nous considérons qu'il serait actuellement difficile d'affirmer que les avantages offerts par les informations qu'ils fournissent sont supérieurs aux « conséquences indésirables des classements », compte tenu du manque de transparence observé. Il existe en effet un risque que le temps consacré par les universités à la collecte et à l'utilisation de données et de statistiques en vue d'améliorer leur performance au regard des classements le soit au détriment des efforts entrepris dans d'autres domaines, comme l'enseignement et l'apprentissage ou l'engagement régional.

Dans une perspective d'avenir, les producteurs de classements prennent des mesures visant à améliorer les méthodologies utilisées, ce que nous ne pouvons qu'encourager et que nous suivrons dans le cadre de rapports ultérieurs. Le Groupe d'experts internationaux sur les classements (IREG) a annoncé qu'il réaliserait un audit des différents classements. Nous espérons que des experts indépendants seront associés à cette initiative, ce qui renforcerait considérablement sa crédibilité. Parmi les questions en suspens à examiner à l'avenir figure celle de la « démocratisation » des classements et ce que cela suppose, pour que davantage d'universités dans le monde soient en mesure d'y trouver leur place. Certaines initiatives européennes récentes qui visent à étendre la portée des classements afin de prendre en considération les différentes missions de l'université semblent indiquer que le manque de données comparables à l'échelle internationale pose problème. Le débat se poursuivra, et ces questions feront notamment l'objet de futurs rapports de l'EUA.

Remerciements

L'analyse approfondie présentée dans ce rapport a été réalisée par le Professeur Andrejs Rauhvargers. Le Comité éditorial souhaite remercier très sincèrement le Professeur Rauhvargers, pour son engagement et le temps considérable qu'il a dédié à la recherche, à la description précise et à l'analyse des différents classements, évaluations et classifications figurant dans l'étude. Ce fut une entreprise difficile, entre autres parce qu'il a été décidé dès le début de ne prendre en considération que les informations accessibles au public sur les différents classements examinés dans le cadre de cette étude.

Le Comité éditorial souhaite également remercier tous les membres du personnel de l'EUA qui ont contribué à la préparation, à la révision et à la publication du présent rapport.

Bruxelles, juin 2011

Jean-Marc Rapp
Jean-Pierre Finance
Howard Newby
Jens Oddershede

SIGLES ET ACRONYMES

A&HCI	Arts & Humanities Citation Index
AHELO	Évaluation des résultats de l'enseignement supérieur
ARWU	Classement académique des universités mondiales (Academic Ranking of World Universities) – Université JiaoTong Universities by Shanghai (désigné sous le sigle « SJTU » dans les publications plus anciennes).
AUBR	Working group on Assessment of University-Based Research (Groupe de travail sur l'évaluation de la recherche universitaire)
CPP	Nombre de citations par publication
ESI	Essential Science Indicators
FCSm	Mean fields citation score (moyenne des citations par champ disciplinaire) – indicateur bibliométrique
ETP	Équivalent temps plein
HEEACT	Conseil d'accréditation et d'évaluation de l'enseignement supérieur de Taïwan
EES	Établissement d'enseignement supérieur
ISI	Institute for Scientific Information (Institut pour l'information scientifique – actuellement Thomson Reuters)
IREG	International Ranking Expert Group (Groupe d'experts internationaux sur les classements)
JCR	Journal Citation Reports
MNCS	Mean normalised citation score – indicateur bibliométrique
P	Nombre de publications (dans les revues de référence) – indicateur bibliométrique
AQ	Assurance qualité
SSCI	Social Sciences Citation Index (base de données bibliographiques en sciences sociales)
SCI	Science Citation Index (base de données de citations scientifiques)
THE	Classement du Times Higher Education (Times Higher Education World University Ranking, anciennement « Times Higher Education Supplement Ranking », désigné sous le sigle « THES »)
THE-QS	Classement du Times Higher Education (Times Higher Education World University Ranking) en coopération avec Quacquarelli-Symonds jusqu'en 2009
THE-TR	Classement du Times Higher Education (Times Higher Education World University Ranking) en coopération avec Thomson Reuters après 2009

GLOSSAIRE

Classement

Le mot « classement » peut avoir plusieurs significations : un classement « actif » consistant à classer, trier, un ensemble d'éléments selon un ou des critères préexistants (classement alphabétique, par taille, par couleur, etc.) ; une « classification », qui est un classement plus élaboré, c'est-à-dire un système organisé et hiérarchisé de classement ; un classement « passif » consistant à constater le résultat d'une compétition au sens large du terme (classement d'une course, classement d'un concours, classement des meilleures ventes, etc.) ; à la fois le **système d'évaluation** de la force d'un compétiteur et le **classement** qui en résulte².

Évaluation

1. Une position sur une échelle ;
2. Une estimation de la situation, en particulier de la situation financière ;
3. Un nombre, une lettre ou un autre signe qui renvoie à la capacité de quelque chose.

Un système d'évaluation appliqué aux sports analyse les résultats de compétitions sportives afin de fournir une évaluation objective de chaque équipe ou joueur. La classification des évaluations de chaque équipe selon un ordre décroissant permet l'établissement d'un classement dans lequel l'équipe la mieux notée occupe la première position.

Bourse d'études

Une aide financière accordée à un étudiant pour lui permettre de couvrir ses frais de scolarité et d'entretien.

Bourse

Une subvention accordée à un étudiant (ex. par un établissement d'enseignement supérieur ou une fondation).

Assurance qualité³

Un programme de suivi et d'évaluation systématiques des différents aspects d'un projet, d'un produit ou d'un service, afin de garantir le respect de normes de qualité.

Il convient de noter que la qualité est déterminée par le promoteur du programme. L'assurance qualité ne peut malheureusement pas garantir pleinement la fabrication de produits de qualité, mais en renforce la probabilité.

Les deux principes fondamentaux sur lesquels repose l'assurance qualité sont l'adaptation à l'usage prévu (le produit doit être adapté à l'usage prévu) et la perfection dès le premier essai (les erreurs doivent être éliminées). L'assurance qualité comprend notamment la réglementation de la qualité des matières premières, des assemblages, des produits et des composants ; les services liés à la production ; les processus de gestion, de production et d'inspection.

Il convient par ailleurs de noter que la qualité est également déterminée par les utilisateurs ou les clients, et non pas par la société en général. La qualité n'est pas synonyme de « prix élevé » ou de « haut de gamme ». Des biens peu onéreux peuvent également être considérés comme des produits de qualité s'ils répondent aux attentes du marché. L'assurance qualité ne se limite au contrôle des aspects qualitatifs d'un produit ou service, elle analyse la qualité afin de garantir sa conformité à des normes spécifiques et à des plans établis.

Étalonnage concurrentiel

On entend par étalonnage concurrentiel (benchmarking) le processus de comparaison des procédures et des indicateurs de performance d'une entreprise aux meilleurs résultats du secteur et/ou aux meilleures pratiques d'autres secteurs. Il s'agit du processus d'identification des « meilleures pratiques » applicables tant aux produits qu'aux processus sur lesquels reposent la fabrication et la fourniture de ces produits. La recherche des « meilleures pratiques » peut être menée dans un secteur particulier, mais aussi dans d'autres secteurs. Il s'agit de mesurer la qualité des politiques, des stratégies, des programmes, des produits, etc. d'une entreprise et de comparer ces mesures à des mesures normalisées ou à des mesures similaires dans les meilleures entreprises. Les objectifs de cette démarche sont : (1) de déterminer ce qui doit être amélioré et où ces améliorations doivent être apportées ; (2) de comprendre comment d'autres entreprises atteignent des niveaux de performance élevés ; (3) d'utiliser ces informations pour améliorer les performances de l'entreprise⁴.

² <http://fr.wikipedia.org/wiki/Classement>.

³ Une définition de l'assurance qualité applicable à l'industrie a été choisie à dessein dans le contexte des classements.

⁴ <http://www.businessdictionary.com/definition/benchmarking.html>.

RÉSUMÉ

Au cours des dernières années, les classements d'universités ont acquis une importance croissante dans le monde entier. Le nombre de classements mondiaux a augmenté⁵ au cours de la réalisation de la présente étude et devrait continuer à augmenter. Les classements se sont mondialisés en 2003 lors de la publication des résultats du premier classement mondial d'universités par l'Université Jiao Tong de Shanghai. L'importance des classements semble connaître depuis lors une croissance exponentielle.

Les classements ont toujours suscité des controverses. Il peut être utile à la société de disposer d'un classement qui permette de déterminer qui est « le meilleur au monde » (et qui ne l'est pas). Les responsables politiques souhaitent

utiliser des informations pratiques et synthétiques – précisément ce que fournissent les classements. Dans une période caractérisée par de considérables contraintes financières, les dirigeants dans différents pays sont de plus en plus intéressés par des comparaisons de la performance d'établissement d'enseignement supérieur (EES) divers fondées sur des indicateurs objectifs. Cependant, les résultats des classements, et en particulier des classements mondiaux, dépendent fortement du choix des indicateurs et de la manière dont ils sont pondérés. Il est en outre difficile, voire impossible, de mesurer et de quantifier la qualité en tant que telle ; c'est pourquoi les classements utilisent divers indicateurs indirects, dont certains n'ont que peu à voir avec la qualité réelle de l'enseignement ou de la recherche.

Objet et principes de cette étude

La présente étude de l'Association européenne de l'Université (EUA) a pour objet d'informer les universités des méthodologies à la base des classements mondiaux les plus réputés et de leur impact potentiel.

Cette étude de l'EUA sur les classements mondiaux d'universités a été réalisée sur la base d'un certain nombre de principes convenus :

- Elle examine les classements mondiaux d'universités les plus réputés et d'autres tentatives de mesure de la performance (classements, évaluations, classifications) qui présentent un intérêt pour les universités européennes. Cependant, cette étude ne traite pas des classements nationaux ;
- Cette étude ne vise pas à classer les classements eux-mêmes, mais à présenter à l'intention des universités une analyse des méthodologies à la base des classements mondiaux ;
- Cette étude se fonde sur les seules informations publiques et librement accessibles, plutôt que sur des enquêtes ou entretiens auprès des producteurs de classements. Cette approche a été privilégiée afin de mettre en évidence le degré de transparence de chaque classement du point de vue de l'utilisateur ;
- Des efforts ont été déployés pour déterminer ce qui est annoncé comme étant mesuré, ce qui est effectivement mesuré, les méthodes de calcul des scores par indicateur et, le cas échéant, des scores finaux, et définir ce que les résultats signifient réellement.

Nous pensons que cette analyse de la méthodologie des classements fournira aux universités des informations utiles à un moment où les classements exercent une influence

croissante sur les politiques de l'enseignement supérieur et sur l'opinion publique à leur égard.

Classements examinés dans le cadre de cette étude

Conformément aux principes mentionnés plus haut, ont été sélectionnés pour cette étude de l'EUA les classements internationaux d'universités suivants :

- Le classement académique des universités mondiales (*Shanghai Academic Ranking of World Universities – ARWU*) – Shanghai Ranking Consultancy, Chine ;
- Le classement du Times Higher Education (*Times Higher Education World University Ranking*) :
 - en coopération avec Quacquarelli Symonds jusqu'en 2009 ;
 - en coopération avec Thomson Reuters depuis 2010 ;
- Le classement US News & World Report (*World's Best Universities Ranking*) en coopération avec Quacquarelli Symonds, États-Unis ;
- Le classement mondial des universités – Reitor (Рейтор), Russie ;
- L'évaluation de la recherche universitaire européenne - Groupe de travail AUBR, Commission européenne ;
- Le classement de Leiden – Université de Leiden, Pays-Bas ;
- Le classement établi par le Conseil d'accréditation et d'évaluation de l'enseignement supérieur de Taïwan, basé sur la publication d'articles scientifiques (*Performance Rankings of Scientific Papers for World Universities* – Higher Education Evaluation and Accreditation Council) ;

⁵ S. Marginson : entretien publié dans *The Australian*, le 12 janvier 2011.

- Le classement universitaire du CHE – Centre pour le développement de l'enseignement supérieur / Die Zeit, Allemagne ;
- Le classement d'excellence du CHE – Centre pour le développement de l'enseignement supérieur / Die Zeit, Allemagne ;
- La classification U-Map – UE ;
- Le classement U-Multirank – UE ;
- L'évaluation des résultats de l'enseignement supérieur

(AHELO) – Organisation de coopération et de développement économiques (OCDE) ;

- Le classement cybermétrique des universités du monde (Webometrics Ranking of World Universities – Laboratoire de cybermétrique, Centro de Ciencias Humanas y Sociales, Espagne.

Résumé des observations et conclusions

Conclusions générales

1. Les tendances observées au cours des dernières années mettent en évidence **la probabilité d'une augmentation du nombre de classements d'universités**, mais aussi d'une spécialisation accrue.

2. Les responsables politiques et la société en général considèrent souvent les classements mondiaux d'universités comme des instruments de « transparence » pour les universités, bien qu'il soit difficile de prétendre le contraire, c'est-à-dire qu'en l'absence de classements, les universités seraient « non transparentes ».

3. Les classements mondiaux d'universités se caractérisent par une grande diversité :

- Les classements d'universités qui ont pour principal objectif d'établir un palmarès des universités les plus réputées : le classement académique des universités mondiales ou classement de Shanghai (ARWU), qui repose essentiellement sur des indicateurs de la recherche ; le classement du Times Higher Education (THE), au départ en coopération avec Quacquarelli Symonds (THE-QS) et, depuis 2010, en coopération avec Thomson Reuters (THE-TR) ; et, sur la base d'indicateurs différents, d'autres classements tels que le classement établi par l'agence russe Reitor.

- Les classements d'universités qui portent uniquement sur la performance de la recherche – avec ou sans palmarès : le classement de Leiden, sans indice composite ; le classement du Conseil d'accréditation et d'évaluation de l'enseignement supérieur de Taïwan (HEEACT), qui établit son palmarès sur la base d'un indice composite ; l'évaluation de la recherche universitaire européenne (AUBR), qui est une méthodologie d'évaluation de la recherche axée sur la transparence à différentes fins, plutôt qu'en vue d'établir un classement.

- Les classements et classifications d'universités qui reposent sur un certain nombre d'indicateurs et n'ont pas pour objet d'élaborer des indices composites ou des palmarès : le premier classement d'universités établi par le Centre pour le développement de l'enseignement supérieur (CHE) en Allemagne a été conçu pour aider les étudiants potentiels à choisir une université selon leurs besoins ; la classification européenne U-Map pour leur permettre de trouver et de comparer des universités qui présentent des profils similaires ; le classement européen U-Multirank pour comparer la performance des universités quant à différents aspects de leurs activités.

- L'OCDE a mis en œuvre les premières phases de son Évaluation des résultats de l'enseignement supérieur (AHELO), qui vise à évaluer la performance des universités d'après les acquis de l'apprentissage démontrés par les étudiants.

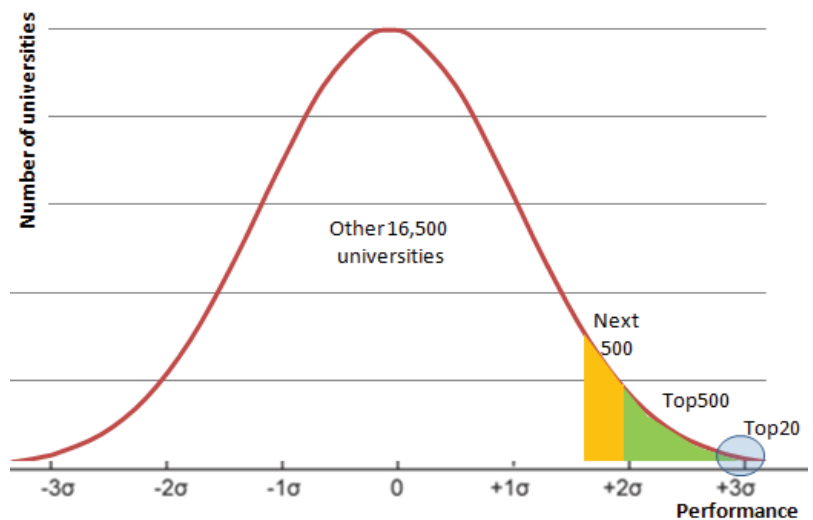
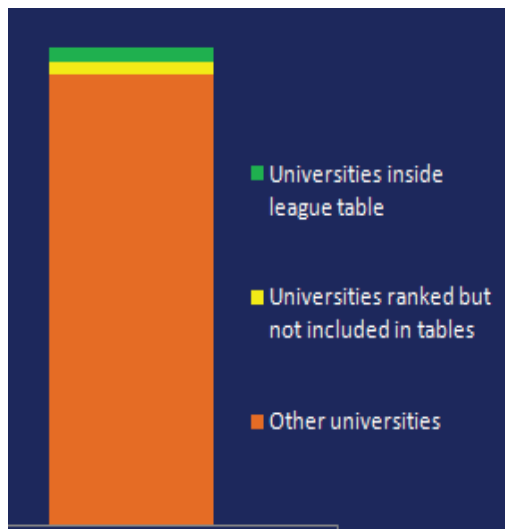
- Les classements d'universités établis sur la seule base de leur visibilité sur Internet : Webometrics.

4. Malgré les nombreux défauts, biais et lacunes des classements, **« les parties prenantes et le grand public les acceptent facilement car ils proposent des informations simples, prêtes à consommer »** (Groupe d'experts sur l'évaluation de la recherche universitaire – AUBR, 2009). Par conséquent, les classements d'universités sont désormais bien établis. Leur nombre devrait ainsi augmenter, mais ils devraient se caractériser par une spécialisation accrue.

Un classement des élites qui ébranle l'ensemble du système

5. Les classements mondiaux de premier plan (ARWU, THE-QS et THE-Thomson Reuters, US News and World Report Ranking [USNWR], HEEACT, Reitor et d'autres) ne concernent que les universités les plus réputées. D'une part, les classements recensent de 1 à 3 % des universités (entre 200 et 500 universités) dans le monde, alors qu'il en existe environ 17 000. D'autre part, il convient de noter que les classements publiés sous forme de palmarès reposent sur des méthodologies qui **peuvent uniquement produire des résultats stables pour 700 à 1 200 universités** dans les classements mondiaux et pour à peine 300 universités dans les classements par champ disciplinaire.

Proportion d'universités prises en considération dans les classements mondiaux par rapport au nombre total d'universités dans le monde.



6. En raison de la conception élitiste sur laquelle se fondent les méthodologies des classements mondiaux, plus de **16 000 universités dans le monde ne figureront jamais dans ces classements**. La question rhétorique de Jamil Salmi (2010) « Combien d'universités peuvent figurer parmi les 500 premières ? » et sa réponse « cinq cent » sont explicites.

7. Un problème, ou une « conséquence indésirable », comme les producteurs de classements dénomment parfois les effets négatifs des classements, est que tant la société que les responsables politiques sont tentés de juger tout l'enseignement supérieur dans le monde à l'aune des critères que les classements utilisent pour détecter les meilleures universités de recherche, plutôt que d'appliquer l'un des principes fondamentaux de l'assurance qualité : le principe de « l'adaptation à l'objectif visé ». En d'autres mots, méritent d'être prises en considération non seulement les universités de recherche mais aussi les universités importantes à l'échelle régionale ou celles qui visent à élargir l'accès à l'enseignement supérieur à un plus large éventail de jeunes. Par conséquent, une « conséquence indésirable » des classements mondiaux est que **les EES avec d'autres missions que celle de se situer à la pointe de la recherche pourraient être obligés de reconsidérer leur raison d'être** à un moment où la différenciation des missions figure au premier rang des priorités de l'enseignement supérieur dans toute l'Europe.

8. Les descriptions des méthodologies de la plupart des classements mondiaux sont simplifiées et permettent rarement au lecteur de comprendre comment les scores de chaque indicateur et le score final de l'indice composite sont établis. Afin d'être en mesure de suivre ces calculs, il est nécessaire d'avoir accès à des sources plus spécialisées, à des informations plus élaborées, peu utiles à un utilisateur curieux qui chercherait uniquement à comprendre d'où proviennent les chiffres.

La combinaison d'indicateurs en vue d'établir un score final – un simple calcul ?

9. Il convient de noter que divers indicateurs ont des dimensions et des dénominateurs différents, tels que le nombre de publications, les effectifs, les citations par enseignant-chercheur⁶, etc. Par conséquent, avant de combiner les scores de chaque indicateur en vue d'établir un score global sur la base d'un indice composite, les scores de ces indicateurs sont traités mathématiquement afin d'éliminer leur caractère dimensionnel. Ceci implique que **les scores publiés par indicateur ne sont généralement pas les valeurs des indicateurs en tant que telles**. Il s'agit dans la plupart des cas du rapport entre la valeur de l'indicateur pour l'université en question et celle de l'université qui affiche la valeur la plus élevée pour ce même indicateur⁷. Les classements ne précisent généralement pas dans les intitulés de colonne que le nombre indiqué n'est pas la valeur de l'indicateur en tant que telle, mais bien le résultat d'autres opérations mathématiques. En conséquence, les scores affichés dans les classements peuvent entraîner des malentendus.

10. Les indices composites comportent toujours des éléments subjectifs. Lors de l'élaboration d'un indice composite, chaque indicateur se voit attribuer une pondération dans le score global. Cela signifie que **le degré d'importance des indicateurs est déterminé par le jugement subjectif des producteurs de classements**. En d'autres mots, l'indice composite reflète le concept de qualité propre aux producteurs du classement. Ceci explique pourquoi les classements publiés sous forme de palmarès ne peuvent en principe pas être « objectifs ».

⁶ Comme équivalent des termes 'academic staff' et 'faculty', l'expression 'enseignant-chercheur' a été retenue dans un sens large, désignant tout personnel, titulaire ou contractuel, chargé d'enseignement ou/et de recherche.

⁷ Par exemple, si la valeur de l'indicateur « publications par membre du personnel académique » d'une université X est de 57, cela ne signifie pas que chaque membre du personnel académique de l'université X publie 57 articles par an. Cela signifie en revanche que la productivité de l'université X en termes de publications représente 57 % de la productivité de la première université quant au nombre de publications.

11. Les indicateurs utilisés dans les classements peuvent être exprimés en valeurs absolues (nombre de publications, de citations, d'étudiants, de membres du personnel, etc.) ou en valeurs relatives (publications par membre du personnel, citations par publication, financement par étudiant, etc.). Ce simple aspect devrait être pris en considération lors de l'analyse des résultats de classements. Si un classement se fonde essentiellement sur des valeurs absolues, les scores obtenus sont liés à la taille de l'établissement, c'est-à-dire que le classement favorise les grandes universités. Si des valeurs relatives sont privilégiées, les universités plus performantes et pas nécessairement plus grandes obtiendront des scores plus élevés. Les classements ARWU et Reitor sont des exemples de classements qui reposent principalement sur des valeurs absolues. Le classement HEEACT, principalement, et les classements THE-QS et THE-TR reposent essentiellement sur des valeurs relatives (à l'exception des enquêtes de réputation). Le classement de Leiden, qui ne combine pas de scores d'indicateurs, repose à la fois sur des indicateurs dépendants et indépendants de la taille des établissements.

12. La position des universités qui figurent actuellement au premier rang des classements ne peut être tenue pour acquise. Les universités les mieux classées doivent au contraire déployer des efforts supplémentaires considérables pour maintenir leur position actuelle, car leurs concurrents évoluent également (CHERPA, 2010).

Dans quelle mesure les indicateurs sont-ils fiables ?

13. Globalement, **les classements mondiaux d'universités reflètent beaucoup plus précisément la performance de la recherche universitaire que la qualité de l'enseignement.** Les indicateurs bibliométriques, utilisés pour mesurer la performance de la recherche dans la plupart des classements, comportent également des biais et des défauts, mais n'en demeurent pas moins des mesures directes.

14. **Les indicateurs relatifs à l'enseignement sont tous des indicateurs indirects**, comme l'est, tout au mieux, leur rapport avec la qualité de l'enseignement.

La mesure de la qualité de l'éducation par le nombre de lauréats du Prix Nobel parmi les diplômés de l'université (ARWU) est un extrême. Cet indicateur peut être considéré comme étant lié à la qualité de l'éducation, mais de façon très particulière et assez indirecte. Juger la qualité de l'enseignement sur la seule base du taux d'encadrement des étudiants sans examiner l'enseignement/l'apprentissage même (THE-QS) est un autre extrême. Il a par ailleurs été prouvé que le taux d'encadrement des étudiants (sur la base du ratio étudiants/enseignants) peut être facilement manipulé. Des indicateurs comme les salaires des enseignants ou la durée des études ne mesurent pas réellement la qualité en tant que telle. L'indicateur relatif à la durée des études examine un aspect important dans certains pays, mais peut être difficilement considéré comme un indicateur valable dans d'autres.

Il est malaisé de déterminer si des cycles d'études beaucoup plus courts ou des taux de réussite élevés sont des signes de grande qualité ou de faibles exigences. Ces indicateurs sont également susceptibles d'être manipulés.

15. Le classement des universités du CHE se fonde sur des indicateurs de l'enseignement sélectionnés en vue d'aider les étudiants potentiels à choisir des EES appropriés. Les informations utilisées sont dans une large mesure tirées d'enquêtes de satisfaction des étudiants – qui sont particulièrement adaptées à cette fin, notamment parce que le classement CHE n'établit pas de palmarès.

16. Quant aux indicateurs bibliométriques utilisés pour mesurer la performance de la recherche, la tendance à privilégier les sciences naturelles et la médecine est manifeste depuis la publication du premier classement ARWU en 2003. Tous les classements fondés sur des indicateurs bibliométriques favorisent les sciences naturelles et la médecine – les 21 disciplines de l'ISI sont essentiellement des sous-domaines des sciences naturelles et de la médecine. Les sciences sociales y sont sous-représentées et les sciences humaines en sont tout simplement absentes. Les traditions en matière de publications et de citations varient par ailleurs selon les champs disciplinaires. Il existe davantage de publications et de citations par publication dans les domaines des sciences naturelles et notamment de la médecine, en particulier parce que les livres sont très peu couverts par les principales bases de données de citations – WoS et Scopus.

Des tentatives ont été faites pour remédier au biais lié au champ disciplinaire. La normalisation par champ est réalisée en divisant le nombre de citations d'un article par le nombre moyen prévisible de citations dans le même champ disciplinaire et au cours de la même année.

Deux indicateurs d'impact fondés sur les citations ont été élaborés : le nombre de citations normalisé par champ disciplinaire (CPP/FCSm), qui est l'indicateur principal (crown indicator) du classement de Leiden et, plus récemment, le mean normalised citation score (MNCS). Le calcul de ces deux indicateurs, qui s'effectue à partir de données relatives aux citations, est décrit de façon détaillée plus avant dans la partie principale de cette étude. Du point de vue de la méthode de calcul, la principale différence réside dans l'ordre des opérations mathématiques. L'indicateur CPP/FCSm est calculé en additionnant d'abord le nombre de citations par article d'une part, et le nombre moyen de citations (dans le même champ disciplinaire et au cours de la même année) d'autre part, et en divisant ensuite la première somme par la seconde. L'indicateur MNCS, lui, est calculé en sens inverse, en divisant pour chaque article le nombre de citations par le nombre attendu de citations, compte tenu du champ disciplinaire et de l'année de publication, et en additionnant ensuite les résultats.

Il peut être démontré que l'indicateur CPP/FCSm a naturellement tendance à privilégier les publications plus anciennes, qui ont par définition été citées plus souvent. En outre, le fait d'additionner le nombre de citations d'articles dans l'ensemble des champs disciplinaires avant la division rend le résultat moins lisible. Dans le cas de l'indicateur MNCS, le problème

des publications plus anciennes ne se pose pas, et la signification du résultat semble plus claire. Un nouveau problème se pose néanmoins : les publications les plus récentes n'ont été que peu citées. Le nombre moyen de citations à l'échelle mondiale n'est dès lors pas fiable, ce qui rend le résultat de l'indicateur instable. C'est pourquoi un indicateur MNCS2 modifié a été proposé en 2010, qui ne tient pas compte des publications les plus récentes (de l'année précédente).

Si un problème est ainsi résolu, un nouveau problème apparaît. Et en définitive, ces tentatives d'amélioration des méthodes de calcul ne résolvent en aucune façon le problème principal, à savoir que les citations de livres ou de recueils d'articles ne sont toujours pas prises en considération.

17. Concernant le facteur d'impact mesuré à l'échelle des revues scientifiques, il convient de noter que les classements spécialisés ne correspondent pas toujours aux facteurs d'impact, en particulier dans les sciences humaines et sociales (AUBR, 2010). Dans ces domaines, mais aussi dans les sciences de l'ingénieur, d'autres sources, comme les livres et les actes de conférence, sont également importantes. Un avertissement à cet égard a même été publié sur le site Internet de Thomson Reuters, selon lequel « il convient de ne pas utiliser le facteur d'impact sans tenir soigneusement compte des nombreux phénomènes qui influencent les taux de citations »⁸.

18. **Le biais lié à l'évaluation par les pairs** : le terme « évaluation par les pairs » est lui-même ambigu, étant utilisé pour désigner des processus assez différents dans les domaines de l'assurance qualité (AQ) et des classements. Dans l'assurance qualité de la recherche et de l'enseignement, le terme « évaluation par les pairs » désigne généralement une visite d'évaluation effectuée par les pairs, selon des procédures rigoureuses. En revanche, les « évaluations par les pairs » des classements se résument généralement à des enquêtes de réputation. Dans le classement THE-QS, malgré la sollicitation de nombreux personnels d'enseignement et de recherche, seuls quelques 5 % ont effectivement répondu. Par ailleurs, à tout le moins dans le cas du classement THE-QS, les « pairs » ne désignent pas en réalité les universités qu'eux-mêmes jugent excellentes – leur choix est limité par des listes dressées préalablement, sur lesquelles de nombreuses universités, voire des pays entiers ne figurent pas. En outre, l'opinion des « pairs » peut manifestement être influencée par la réputation dont jouissent déjà certaines institutions (AUBR, 2010).

19. **Le biais linguistique et régional** : il a été observé depuis la publication des premiers classements mondiaux que les universités des nations anglophones y sont favorisées parce que les travaux rédigés dans d'autres langues sont à la fois moins publiés et moins cités. Une étude récente réalisée par l'équipe responsable du classement de Leiden révèle que le facteur d'impact d'articles d'universités françaises et allemandes publiés en français et en allemand respectivement, est moindre que le facteur d'impact d'articles publiés en anglais par les mêmes universités (van Raan et al., 2010).

Améliorer la qualité ou améliorer sa position dans les classements ?

20. Dans le souci de renforcer leur position dans les classements, les universités sont fortement **tentées d'améliorer leur performance principalement** dans les domaines mesurés par les indicateurs des classements.

21. On constate dans certains cas une manipulation de données plutôt qu'une amélioration de la performance ; par exemple :

- La fusion d'universités à la seule fin de figurer dans les classements ;
- Le nombre de demandes d'inscription à l'université ;
- Les résultats des candidats aux tests standardisés qui conditionnent l'entrée à l'université dans certains pays ;
- Le nombre d'enseignants-chercheurs ;
- Le taux d'encadrement des étudiants (principalement basé sur le ratio étudiants/enseignants) ; il a été démontré que selon les définitions d'« étudiant » et d'« enseignant », ce coefficient pouvait varier de 6/1 à 39/1 (Baty, 2009) ;
- Les salaires des enseignants-chercheurs ;
- Les enquêtes de réputation auprès des étudiants (en encourageant délibérément les étudiants à mentir) ;
- Même les indicateurs bibliométriques peuvent être corrompus suite à la manipulation de données (AUBR, 2010 : 13).

Comment les classements peuvent-ils être améliorés ?

22. Les producteurs de classements tentent d'**améliorer leur méthodologie**. Toutefois, les améliorations apportées sont plus souvent techniques que conceptuelles. Il est par exemple important d'utiliser des données normalisées par champ disciplinaire. En ce sens, le nouvel indicateur normalisé MNCS2 améliore effectivement les aspects mathématiques par rapport au précédent « crown indicator » CPP/FCSm. Cependant, cette évolution ne contribue en rien à améliorer la situation des sciences humaines, qui restent ignorées par presque tous les indicateurs bibliométriques utilisés dans les classements mondiaux. Améliorer les méthodes de calcul ne suffit pas ; des efforts doivent être consentis pour que les classements s'appliquent de façon égale à tous les domaines de la recherche.

23. Un certain nombre de classements d'université affichent la volonté d'**aider les étudiants** à faire leur choix. Les classements disposent effectivement du potentiel permettant d'aider les étudiants à choisir l'université appropriée dans leur pays ou à l'étranger. Cependant, peu de classements sont actuellement susceptibles de le faire. Le classement du CHE constitue l'un de ces quelques exemples. D'une manière générale, pour répondre aux besoins des étudiants potentiels, la plupart des classements doivent avant tout se fonder sur des indicateurs appropriés et présenter des informations nettement plus précises sur la signification réelle des scores affichés par les indicateurs.

⁸ http://thomsonreuters.com/products_services/science/free/essays/impact_factor/.

⁹ Voir la description du classement de Leiden.

24. Maintenant que les classements ont obtenu l'attention du grand public, y compris celle des responsables politiques, **il existe une véritable demande pour des classements plus « démocratiques »**. Les classements mondiaux n'ont recensé jusqu'à présent que quelques centaines d'universités qui sont les « meilleures » au monde. De sorte que ces classements ont créé des problèmes aux milliers d'universités « normales » qui accomplissent simplement leur mission, qu'elle concerne la formation de spécialistes pour le marché du travail ou la recherche fondamentale ou appliquée. L'obsession actuelle pour les classements semble avoir créé un besoin d'être classé : « qui ne figure pas au palmarès n'existe pas ». Il devrait être possible de modifier considérablement les classements pour permettre à davantage d'EES d'y figurer. Il s'agit là d'un point particulièrement important pour les institutions qui doivent s'acquitter d'une fonction spécifique, comme être au service de la région dans laquelle ils sont situés ou proposer un enseignement supérieur aux personnes qui combinent études et activité professionnelle ou reviennent à l'université dans le cadre de la formation tout au long de la vie. L'inclusion d'un nombre plus important d'universités pourrait être considérée comme un moyen de reconnaître la contribution importante d'institutions performantes qui subissent les « conséquences indésirables » des classements.

25. À l'échelle nationale, les classements encouragent l'acquisition et la publication de données fiables sur l'enseignement supérieur. Dans un contexte international, les classements favorisent l'adoption de définitions communes des aspects sur lesquels des données sont collectées. Les résultats des classements mondiaux peuvent faire l'objet de débats nationaux et d'une analyse approfondie des facteurs en cause, qui peuvent éventuellement déboucher sur un ajustement positif des politiques.

26. À la plupart des classements mondiaux est associée la publication de listes concernant la « **performance** » **des pays**. Ces comparaisons sont établies sur la base de la somme du nombre d'universités de chaque pays figurant sur la liste des universités les mieux notées, un nombre différent de points leur étant généralement attribué selon que l'université figure dans le Top 100, le Top 100-200 ou les centaines suivantes. Les pays les mieux notés dans les listes publiées sont dès lors les États-Unis, le Royaume-Uni, l'Allemagne et la France. Cependant, si les listes publiées étaient « normalisées » en rapportant le nombre d'universités les mieux notées au nombre d'habitants, d'autres pays rejoindraient le haut de ces classements, comme la Suisse, la Suède, la Finlande et le Danemark (Salmi, 2010).

Évolution récente des classements internationaux

27. Les classements existants, à l'exception éventuelle des classements à indicateurs multiples, **sont dans l'impossibilité d'établir un diagnostic de l'ensemble du système d'enseignement supérieur**, car ils ne concernent généralement que les universités de recherche les plus réputées. Les

classements mondiaux actuels ne peuvent par ailleurs fournir que peu d'informations utiles sur des questions comme la qualité de l'enseignement et de l'apprentissage, l'accessibilité, l'engagement régional, la participation à la formation tout au long de la vie, le rapport coût-efficacité et d'autres, tout simplement parce que les indicateurs utilisés ne s'appliquent pas à ces questions. L'évaluation de la recherche universitaire européenne commissionnée par l'UE (AUBR), U-Map, U-Multirank et AHELO sont des initiatives visant à élaborer des instruments de transparence à l'échelle internationale à l'intention de tous les EES, bien que chacune de ces initiatives présente certaines limites.

28. Le groupe de travail AUBR a réalisé une analyse des forces et faiblesses de différents indicateurs de la recherche et examiné leur adéquation à diverses fins d'évaluation, établissant une méthodologie pour **l'évaluation de la recherche universitaire**. Les conclusions du groupe de travail AUBR concernant différents indicateurs sont utiles dans le cadre d'une analyse des classements mondiaux d'universités.

29. U-Map est un projet qui vise à classer l'ensemble des EES en Europe, **indépendamment du type d'établissement, de leur orientation, etc.** Il reflète la diversité des missions et des profils des établissements d'enseignement supérieur en Europe, sans attribuer de score final. U-Map se fonde sur des indicateurs qui caractérisent l'orientation et l'intensité de différentes activités des EES plutôt que la performance, l'impact ou la qualité. Les indicateurs sur lesquels se fonde U-Map portent sur le niveau d'enseignement et le domaine d'étude, la population étudiante, l'intensité de la recherche, le transfert de connaissances, l'orientation internationale et l'engagement régional. U-Map dispose de deux outils de visualisation : l'un pour identifier les établissements d'enseignement supérieur qui correspondent aux caractéristiques établies par l'utilisateur ; l'autre permettant d'établir une comparaison détaillée d'un maximum de trois EES sélectionnés.

Le manque de données comparables à l'échelle internationale constitue un défi pour U-Map. Des définitions communes doivent encore être élaborées pour permettre la comparaison de données de différents pays. Jusqu'à ce que des données comparables aient été collectées dans toute l'UE, U-Map devra se fonder sur des données nationales et institutionnelles et sera par conséquent plus approprié pour comparer des institutions au sein d'un pays plutôt qu'à l'échelle internationale. U-Map est encore en phase de test ; à ce stade, des EES dans les pays volontaires procèdent à la saisie de leurs données.

30. U-Multirank est un projet de classement multidimensionnel qui tient compte de l'ensemble des activités d'un EES – enseignement, recherche, transfert de connaissances et engagement régional. Aucun score final ne sera attribué à un EES, mais il est encore malaisé de déterminer comment éviter que des tiers ne convertissent les résultats du classement en un palmarès (Boulton, 2010). U-Multirank utilisera des données de Thomson Reuters pour établir ses indicateurs bibliométriques. Les autres indicateurs seront basés sur des données autodéclarées par les EES sur les étudiants, les enseignants et sur la recherche (à l'exception des publications

et citations). Un vaste ensemble de données seront tirées d'enquêtes de satisfaction des étudiants.

31. Les problèmes liés au manque de données comparables à l'échelle internationale pourraient concerner U-Multirank plus encore que U-Map. Mais s'agissant d'un classement, d'autres écueils guettent U-Multirank. L'utilisation de données autodéclarées peut être une démarche fiable dans le cas d'un outil de classification comme U-Map. En revanche, dans le cas de U-Multirank, bien que ce classement ne combine pas de scores d'indicateurs afin d'établir un score final, il existe un risque de manipulation des données autodéclarées en vue de renforcer les résultats d'indicateurs particuliers. Il est clairement préférable de se fonder sur des enquêtes de satisfaction des étudiants que de recourir à des sondages de réputation auxquels ont répondu les enseignants-chercheurs (qui ne sont pas utilisés dans U-Multirank). Cependant, les enquêtes de satisfaction des étudiants peuvent s'avérer moins fiables dans les comparaisons internationales, car les étudiants réalisent un classement numérique des EES, encore que les traditions en matière de notation puissent varier selon le pays. U-Multirank en est encore à sa phase pilote, mais semble évoluer vers une méthodologie d'évaluation plutôt que de classement. Il semble que la collecte des données soit l'une des questions les plus délicates.

32. Le projet AHELO de l'OCDE vise à comparer les EES à l'échelle internationale sur la base des acquis de l'apprentissage. Trois instruments d'évaluation seront mis au point dans le cadre d'AHELO : l'un pour évaluer les compétences génériques et deux autres pour évaluer les compétences spécifiques à l'économie et aux sciences de l'ingénieur. Au cours de ces phases initiales du projet, les concepteurs doivent déjà apporter des réponses à un certain nombre de questions, y compris déterminer s'il est possible d'élaborer des instruments pour appréhender des acquis de l'apprentissage qui soient perçus comme valables dans différents contextes nationaux et institutionnels.

L'autorégulation est-elle suffisante ?

33. En automne 2010, le Groupe d'experts internationaux sur les classements (IREG) a annoncé son intention de mettre en place un programme d'audit des classements. Cet audit sera réalisé sur la base de 20 critères inspirés des Principes de Berlin. La procédure d'audit des classements est établie par le Comité exécutif de l'Observatoire de l'IREG, qui nommera par ailleurs les membres de chaque équipe d'audit. Sont membres du Comité exécutif de l'IREG des producteurs de classements et des experts ayant suivi l'évolution des classements.

La procédure d'audit des classements de l'IREG est conçue selon le modèle des procédures d'assurance qualité dans l'enseignement supérieur : un rapport d'auto-évaluation du classement audité est établi sur la base d'un questionnaire ; l'équipe d'audit effectue ensuite une inspection sur place ; après l'inspection, l'équipe d'audit établit un rapport d'audit à soumettre

à l'approbation du Comité exécutif de l'Observatoire de l'IREG. Un label « IREG approved » sera attribué aux classements ayant satisfait aux critères de la procédure d'audit. L'IREG est composé de deux catégories de spécialistes : des experts qui font des recherches sur les classements, mais ne participent pas à leur élaboration, et les producteurs de classements eux-mêmes.

Il n'est pas facile de satisfaire aux exigences posées par les Principes de Berlin, et la plupart, voire la totalité des classements existants auront à apporter certains changements de façon à s'y conformer réellement. Il est bien entendu trop tôt pour évaluer l'audit des classements par l'IREG. Cependant, dans un souci de crédibilité, il conviendrait d'associer des experts extérieurs à la procédure d'audit envisagée, afin de garantir que des voix critiques puissent s'exprimer en toute indépendance.

I. INTRODUCTION

Objet du rapport

Le présent rapport de l'EUA a pour objet d'informer les universités des méthodologies et de l'impact potentiel des classements mondiaux les plus réputés, ainsi que d'autres initiatives prises à l'échelle internationale qui visent à établir un classement, une évaluation ou une classification de la performance des universités.

Il a été décidé de n'examiner que des classements internationaux pour que ce rapport soit utile à l'ensemble des membres de l'EUA. Les classements nationaux présentent des caractéristiques quelque peu différentes des classements internationaux. L'une de ces différences est que les classements nationaux sont le plus souvent établis par les pouvoirs publics, bien que des médias et des agences indépendantes puissent également établir et publier des classements nationaux. Les classements nationaux peuvent être créés pour informer les étudiants des meilleures possibilités d'études dans leur pays. Cependant, en particulier lorsque les classements sont établis par les pouvoirs publics, le but peut être d'utiliser les résultats des classements à des fins d'élaboration des politiques dans le domaine de l'enseignement supérieur, y compris l'allocation de ressources.

Une autre différence réside dans le choix des indicateurs : certains types d'indicateurs, comme différents indicateurs financiers, peuvent être appropriés dans un classement national et ne pas l'être pour autant dans un classement international, en raison d'un manque de données comparables.

Cette étude a été réalisée sur la base de certains principes convenus :

- L'étude porte sur les classements mondiaux d'universités les plus réputés et sur d'autres importantes initiatives prises à l'échelle internationale, y compris des tentatives de mesure

de la performance (classements, évaluations, classifications) qui présentent un intérêt pour les universités européennes. Comme mentionné plus haut, cette étude ne traite pas des classements nationaux ;

- Cette étude ne vise pas à juger ou à classer les classements eux-mêmes, mais à présenter à l'intention des universités une analyse des méthodologies à la base d'une sélection de classements mondiaux, en vue d'aider le lecteur à comprendre ce qui est effectivement mesuré, comment les résultats des classements sont calculés et comment ils devraient être interprétés ;

- Seules des informations publiques et librement accessibles ont été utilisées pour décrire la méthodologie à la base de chaque classement. Il s'agit d'un choix délibéré. En conséquence, aucune donnée n'a été collectée par l'intermédiaire d'enquêtes ou d'entretiens auprès des producteurs de classements. La plupart des classements sont accompagnés de descriptions simplifiées de la méthodologie utilisée ;

- Cette approche imposait un examen approfondi des sites Internet des classements ainsi que des publications des producteurs de classements, ce qui nous a permis de mettre en évidence le degré de transparence de chaque classement ;

- Des efforts ont été déployés pour déterminer ce qui est annoncé comme étant mesuré, ce qui est effectivement mesuré, les méthodes de calcul des scores par indicateur et, le cas échéant, des scores finaux, et définir ce que les résultats signifient réellement ;

- Nous pensons que cette analyse de la méthodologie des classements fournira aux universités des informations utiles à un moment où les classements exercent une influence croissante sur les politiques de l'enseignement supérieur et sur l'opinion publique.

Brève histoire des classements

Le premier classement d'universités à l'échelle nationale a été publié aux États-Unis en 1983 par US News and World Report. Cependant, des classifications et des classements

spécialisés d'universités, de portée plus limitée, avaient déjà été établis aux États-Unis dès 1870 (voir tableau 1 – Salmi et Saroyan, 2007).

Tableau 1 : Chronologie des activités de classement aux États-Unis entre 1870 et 1982 (Salmi et Saroyan, 2007)

1870-1890	La Commission du Bureau américain de l'enseignement entreprend la publication d'un rapport annuel de données statistiques classant les établissements.
1910	L'Association des universités américaines pousse le Bureau américain de l'enseignement à rétablir les classements.
1910-1933	James Cattell, l'un des premiers psychologues américains, professeur à l'Université de Pennsylvanie puis à Columbia, publie « American Men of Science », où il fait un classement des établissements basé sur le nombre de spécialistes éminents associés à un établissement soit comme étudiants, soit comme membres du personnel universitaire ; il prend en compte également le rapport entre les spécialistes dans un établissement donné et le nombre total de professeurs.
1925	Raymond Hugues, Président de l'Université de Miami et plus tard Président du Conseil américain pour l'éducation et de son Comité sur la formation universitaire de cycle supérieur publie « <i>A Study of the Graduate Schools of America</i> », dans lequel il a recours à un classement basé sur la réputation de 26 disciplines dans 36 établissements.
1957	Chesly Manly, du Chicago Tribune, publie six classements différents recensant les dix meilleurs établissements dans les catégories suivantes: universités, « colleges » mixtes, pour hommes, pour femmes, facultés de droit et écoles d'ingénieurs.
1959	Hayward Keniston, de l'Université de Pennsylvanie, publie un classement basé sur la réputation de 15 universités dans différentes disciplines.
1966	Allan Cartter, du Conseil américain pour l'éducation, publie « <i>An Assessment of Quality in Graduate Education</i> », classant 106 établissements.
1973-1975	Blau et Margulies établissent un classement basé sur la réputation des écoles professionnelles.
1982	L'Académie nationale américaine des sciences commande une évaluation de la recherche et des programmes doctoraux aux États-Unis.
1982	Les classements commencent également à couvrir l'enseignement universitaire de premier cycle (ex. Fiske Guide to Colleges, 1982 ; US News and World Report, 1983 ; etc.).

Source : tableau reproduit par Salmi & Saroyan, 2007

L'époque des classements mondiaux a débuté suite à la publication en 2003 des premiers résultats du « classement académique des universités mondiales » de l'Université Jiao Tong de Shanghai, (Academic Ranking of World Universities – ARWU) (ci-après le « classement ARWU »). La publication des résultats du premier classement ARWU en 2003 « a attisé le feu » (van Raan, 2005). Ces résultats ont choqué le monde, et en particulier l'Europe, les universités des États-

Unis et du Royaume-Uni occupant une position dominante dans le Top 20 et le Top 100 du classement ARWU. La première publication du classement mondial d'universités du Times Higher Education Supplement en 2004 (devenu par la suite le classement du Times Higher Education, ci-après le « classement THE ») a constitué dans une certaine mesure une réponse européenne au classement ARWU. D'autres classements mondiaux d'université ont vu le jour depuis lors.

Aperçu des implications des classements

Au cours des dernières années, les classements d'université ont acquis une importance croissante dans le monde entier. Tant que les classements d'universités se limitaient à l'échelle nationale, ils étaient bien accueillis et jouaient un rôle important dans certains pays, mais n'étaient pas de nature à susciter une attention particulière dans d'autres systèmes universitaires. Les classements ont véritablement acquis une dimension internationale en 2003 lors de la publication des résultats du premier classement mondial d'universités par l'Université Jiao Tong de Shanghai. L'importance des classements semble connaître depuis lors une croissance exponentielle.

Les classements ont toujours été de nature à susciter des controverses. Les responsables de l'établissement et de la publication des classements affirment généralement qu'ils sont « objectifs » et que la position d'une université dans un tableau de classement correspond à sa « qualité ». Les détracteurs soutiennent que les résultats d'un classement sont étroitement liés au choix et à la pondération des indicateurs. Ils soulignent par ailleurs qu'étant donné la difficulté, voire l'impossibilité, de mesurer la qualité en tant que telle, les classements se basent sur différents indicateurs plus ou moins indirects dont les valeurs représentent prétendument la qualité et l'excellence. Comme l'indiquent Marginson et van der Wende, « une meilleure façon d'envisager les classements commence par la reconnaissance du fait que tous les classements reposent sur des données de nature partielle, comportent des biais et répondent à des objectifs définis » (Marginson et van der Wende, 2007).

Cependant, le public et les responsables politiques sont souvent favorables à un classement synthétique des universités, établi sur la base de résultats d'indicateurs, et sont convaincus que, à l'image des équipes sportives, chaque université du classement est « meilleure » que la suivante et « moins performante » que la précédente. Les classements font l'objet de nombreuses discussions depuis plusieurs années : dans les médias, entre responsables politiques, dans les milieux universitaires et dans la société en général. L'intérêt croissant que suscitent les résultats des classements a transformé le contexte dans lequel s'inscrit le travail des universités : pour qu'une université soit considérée comme une « bonne » université, il est désormais devenu nécessaire d'améliorer la performance dans les domaines mesurés par les indicateurs des classements.

Les responsables de la politique de l'enseignement supérieur, que ce soit au niveau national ou régional, tendent de plus en plus à consulter les résultats des classements lorsqu'ils prennent des décisions concernant l'allocation de ressources ou la structure des systèmes d'enseignement supérieur.

Les classements d'universités eux-mêmes font l'objet de débats intenses. Les défenseurs des classements mondiaux d'universités affirment que les classements améliorent la transparence et permettent aux étudiants de faire des choix en connaissance de cause, tandis que leurs détracteurs soutiennent que les classements ne tiennent pas compte des différentes fonctions importantes de l'enseignement supé-

rieur, qu'ils reposent sur des indicateurs indirects plutôt que sur une mesure de la qualité en tant que telle et qu'ils comportent des biais considérables (voir p.ex. van Raan, 2005). Le biais le plus important consistant à privilégier la recherche en sciences naturelles et en médecine, au détriment des sciences de l'ingénieur et des sciences sociales, qui sont sous-représentées, et des sciences humaines, qui sont pratiquement voire complètement oubliées. Un autre biais souvent évoqué est celui de l'avantage donné aux publications en langue anglaise (Ibid.).

Qu'on le veuille ou non, les classements mondiaux d'université sont désormais bien établis. Comme l'a indiqué S. Marginson dans un entretien publié dans *The Australian* le 12 janvier 2010, « le nombre de classements d'universités va vraisemblablement augmenter, et non pas diminuer. L'aspect positif est qu'ils deviendront de plus en plus spécialisés » (Hare, 2011).

Il serait naïf d'imaginer que les médias renoncent un jour à utiliser un outil comme les classements mondiaux d'universités, qui attirent des milliers de lecteurs lors de la publication des derniers résultats et entretiennent le suspense pendant une année entière, en publiant des bribes d'information sur des modifications minimales des méthodologies de classement. Le public et les responsables politiques seront toujours intéressés par des classements présentés sous la forme d'un tableau simple et compréhensible, qui indique clairement quelles universités sont les meilleures au monde.

Les universités, quant à elles, sont souvent flattées ou embarrassées, selon leur position dans le classement ou l'évolution de celle-ci par rapport à l'année précédente. Des forces s'exercent tant à l'intérieur qu'à l'extérieur de l'université pour l'inciter à faire toujours plus pour améliorer sa position dans les classements ou simplement pour y figurer. Selon S. Marginson, « les classements sont l'indicateur de la méta-performance et disposent d'un pouvoir qui leur est propre. Ils sont hypnotiques et deviennent une fin en soi, peu importe ce qu'ils mesurent effectivement, s'ils reposent sur des bases solides ou que leur utilisation produise des effets constructifs ou non. L'intérêt pour les classements hiérarchiques l'emporte sur toute autre considération. » (Marginson, 2007).

Cependant, les avantages que présentent les classements sont souvent mis en avant, notamment par les producteurs de classements eux-mêmes. Les classements peuvent par exemple aider un étudiant à choisir un établissement ou promouvoir une culture de la transparence.

Ils renforcent la concurrence entre les universités et entraînent souvent des changements de politique en leur sein, dans le souci d'améliorer leur position dans les classements. Ils présentent des informations simples et aisément lisibles, et commencent à être utilisés pour guider l'allocation de ressources aux universités et plus généralement à des fins d'élaboration des politiques nationales ou régionales de l'enseignement supérieur.

Cependant, il existe une grande variété de classements nationaux et internationaux, qui sont créés à différentes fins, sur la base d'indicateurs différents et souvent d'indicateurs indirects au lieu de paramètres difficiles ou impossibles à mesurer.

Les classements d'universités ont fait l'objet de nombreuses critiques, notamment suite à la publication des premiers classements mondiaux. Les indicateurs utilisés dans les classements sont contestés, en raison de défauts et de biais, qui consistent notamment à privilégier la recherche au détriment de l'enseignement ; à ne tenir aucun compte de la diversité institutionnelle, au bénéfice des universités de recherche ; à négliger, en tout ou partie, les publications sous forme de livre ; auxquels s'ajoutent les biais liés au champ disciplinaire, à la langue, à la géographie, etc.

D'autres critiques plus générales concernent souvent la transparence des classements eux-mêmes. Il n'est par exemple pas toujours facile de déterminer l'identité du producteur de classements, les objectifs d'un classement particulier, ou les publics cibles : le classement est-il établi à l'intention d'étudiants à la recherche de la « meilleure université », est-il destiné au grand public ou s'adresse-t-il aux responsables de la politique de l'enseignement supérieur ? Ou le classement a-t-il été établi pour aider les universités à améliorer leur performance dans des domaines particuliers ?

D'autres griefs concernent la transparence des méthodologies de classement. La signification des indicateurs n'est pas expliquée ou la description en est si générale qu'il est impossible de comprendre :

- Ce qui est effectivement mesuré – s'agit-il d'une mesure directe ou de la valeur d'un indicateur indirect et, le cas échéant, quel est le rapport réel entre la valeur mesurée et la catégorie à laquelle s'applique cette valeur (par exemple, mesurer le taux d'encadrement des étudiants par le ratio étudiants/enseignants et l'intituler « qualité de l'enseignement ») ?
- Comment la valeur de l'indicateur est calculée sur la base de données brutes – par exemple, est-il clair que le nombre 26 à l'indicateur « publications » ne signifie pas que les enseignants-chercheurs publient 26 articles par an, mais que le taux de publications de l'université en question correspond à 26 % du taux de publications de la « meilleure » université du classement au regard de cet indicateur ?
- Comment le score final de l'université est calculé sur la base des résultats des indicateurs – la pondération attribuée aux différents indicateurs n'est pas nécessairement rendue publique.

Principes de Berlin sur le classement des établissements d'enseignement supérieur

Il est devenu manifeste que l'amélioration de la situation supposait la définition de principes communs auxquels pourraient adhérer les producteurs de classements. Un ensemble de lignes directrices a été élaboré par le Groupe d'experts internationaux sur les classements (IREG)¹⁰ en 2006. Il s'agit des « Principes de Berlin sur le classement des établissements d'enseignement supérieur » (IREG, 2006), ci-après les « Principes de Berlin ». Les Principes de Berlin sont, en substance, les suivants (IREG, 2006) :

- **Buts et objectifs** : les classements devraient être explicites quant à leur objectif et à leurs publics cibles ; reconnaître la diversité des établissements et prendre en considération leurs missions et objectifs différents ; faire preuve de clarté quant à leurs sources d'information ; indiquer les contextes linguistiques, culturels, économiques et historiques des systèmes d'enseignement qui font l'objet de classements ;
- **Création et pondération des indicateurs** : les classements devraient être transparents quant à la méthodologie utilisée ; sélectionner les indicateurs en fonction de leur pertinence et de leur validité ; mesurer de préférence, et dans la mesure du possible, les résultats plutôt que les ressources ; mettre en évidence les poids assignés aux différents indicateurs (le cas échéant) et limiter les modifications qui leur sont apportées ;
- **Collecte et traitement des données** : les classements devraient prêter l'attention voulue aux normes éthiques et aux recommandations en matière de bonnes pratiques ; utiliser autant que possible des données auditées et vérifiables ; inclure des données collectées par des procédures appropriées à la collecte scientifique de données ; appliquer des critères d'assurance qualité aux processus de classement eux-mêmes ;
- **Présentation des résultats des classements** : les classements devraient permettre aux utilisateurs d'avoir une compréhension claire de l'ensemble des facteurs déterminant l'établissement d'un classement et leur permettre de choisir le mode de présentation des classements ; être établis de façon à éliminer les erreurs ou à en réduire le nombre dans les données d'origine ; être structurés et publiés de façon à permettre la correction des erreurs et défauts éventuels.

Les Principes de Berlin constituent indéniablement des éléments d'orientation utiles, et les producteurs de classements affirment souvent qu'ils s'y conforment. La réalité semble néanmoins quelque peu différente. À titre d'exemple, une équipe de chercheurs de l'Université du Minnesota a quantifié les Principes de Berlin et classé les producteurs de classements eux-mêmes en fonction de leur conformité aux meilleures pratiques exposées dans les Principes de Berlin (Stoltz et al., 2010). Cette équipe de chercheurs a classé 25 classements d'universités européennes auxquels elle a attribué un score compris entre 1 et 5 : 1 (non-conformité) ; 2 (faible niveau de conformité) ; 3 (niveau de conformité satisfaisant) ; 4 (bon niveau de conformité) ; 5 (excellent niveau de conformité). 13 classements sur 25 ont obtenu un score inférieur à 3 (niveau de conformité satisfaisant), parmi lesquels

¹⁰ L'IREG a été fondé en 2004 par le Centre européen pour l'enseignement supérieur (CEPES) de l'UNESCO et l'Institute for Higher Education Policy de Washington.

le classement mondial du Times Higher Education (THE-QS), qui n'a obtenu qu'un score de 2,25. Les chercheurs ont également constaté que les méthodologies des classements restaient globalement au-dessous des attentes des Principes de Berlin, tandis que ceux-ci étaient mieux respectés dans les domaines de la transparence et de la facilité d'utilisation.

Mais les classements fournissent-ils les informations que recherchent les étudiants ? Il s'agit de l'un des principes de bonnes pratiques en matière de classement énoncés par les Principes de Berlin, selon lesquels les classements doivent être conçus à l'intention d'un public cible. On peut affirmer avec certitude que cet aspect n'a pas été suffisamment pris en compte, au moins jusqu'à une période récente. On a simplement présumé que tous les indicateurs disponibles étaient pertinents et s'appliquaient indifféremment à tous les publics (King, Locke et al., 2008).

À l'automne 2010, l'IREG a annoncé son intention de mettre en place un programme d'audit des classements. Cet audit sera réalisé sur la base de 20 critères inspirés des Principes de Berlin. La procédure d'audit des classements est établie

par le Comité exécutif de l'Observatoire de l'IREG, qui nommera par ailleurs les membres de chaque équipe d'audit. Sont membres du Comité exécutif de l'IREG à la fois des producteurs de classements et des experts ayant suivi l'évolution des classements.

La procédure d'audit des classements de l'IREG est conçue selon le modèle des procédures d'assurance qualité dans l'enseignement supérieur : un rapport d'auto-évaluation du classement audité est établi sur la base d'un questionnaire ; l'équipe d'audit effectue ensuite une inspection sur place ; après l'inspection, l'équipe d'audit établit un rapport d'audit à soumettre à l'approbation du Comité exécutif de l'Observatoire de l'IREG. Un label « IREG approved » sera attribué aux classements ayant satisfait aux critères de la procédure d'audit.

Il n'est pas facile de satisfaire aux exigences des Principes de Berlin. Et comme il a été démontré que les classements existants ne sont le plus souvent pas conformes à certains de ces principes, la participation d'experts indépendants à la procédure d'audit renforcerait considérablement sa crédibilité.

II. MÉTHODOLOGIES DES CLASSEMENTS MONDIAUX LES PLUS RÉPUTÉS

Les classements peuvent être regroupés de différentes façons : en fonction de leur objet, des paramètres mesurés, de la présentation des résultats ou des effets recherchés. Aux fins du présent rapport, les classements internationaux examinés sont regroupés comme suit :

1. Les classements d'universités qui ont pour principal objet d'établir des palmarès

1.1 Le classement académique des universités mondiales (Academic Ranking of World Universities – ARWU) – Shanghai Ranking Consultancy¹¹

1.2 Le classement du Times Higher Education (*Times Higher Education World University Ranking* – THE) – Times Higher Education

1.2.1 en coopération avec Quacquarelli Symonds (jusqu'en 2009)

1.2.2 en coopération avec Thomson Reuters

1.3 Le classement des meilleures universités mondiales (*World's Best Universities Ranking*) – US News & World Report en coopération avec Quacquarelli Symonds

1.4 Le classement mondial des universités – Reitor (Рейтор)

2. Les classements axés uniquement sur la performance de la recherche (avec ou sans palmarès)

2.1 Le classement de Leiden (*Leiden Ranking*) – Université de Leiden

2.2 Le classement du Conseil d'accréditation et d'évaluation de l'enseignement supérieur de Taiwan (*Performance Rankings of Scientific Papers for World Universities*) – Higher Education Evaluation and Accreditation Council of Taiwan

2.3 L'évaluation de la recherche universitaire – Commission européenne

3. Les classements multidimensionnels – classements et classifications d'universités reposant sur un certain nombre d'indicateurs et n'ayant pas pour objet d'établir des palmarès

3.1 Le classement universitaire du CHE – Centre pour le développement de l'enseignement supérieur / Die Zeit

3.1.1 Le classement des universités du CHE

3.1.2 Le classement d'excellence du CHE

3.1.3 Autres classements du CHE

3.2 La classification U-Map – Centre d'études des politiques sur l'enseignement supérieur (Center for Higher Education Policy Studies - CHEPS) de l'Université de Twente

3.3 Le système européen de classement multidimensionnel des universités (U-Multirank) – projet financé par l'UE.

4. Les classements Internet

4.1 Le classement cybermétrique des universités du monde (*Webometrics Ranking of World Universities*)

5. L'étalonnage concurrentiel (benchmarking) sur la base des acquis de l'apprentissage

5.1 Le projet d'évaluation des résultats de l'enseignement supérieur (AHELO) – OCDE

Pour chaque classement, les aspects méthodologiques suivants sont analysés, pour autant qu'ils soient pertinents :

- Les indicateurs utilisés ;
- Ce qui est effectivement mesuré et quels indicateurs indirects sont utilisés ;
- Le calcul de la valeur des indicateurs et du score final le cas échéant ;
- L'évolution de la méthodologie utilisée et son impact ;
- Les forces, les faiblesses et les particularités du classement ;
- Les autres produits associés au classement – analyses complémentaires (par sujet, comparaisons entre pays, informations aux étudiants, classements individuels) ;
- Autres remarques particulières au classement.

Il convient de noter que les classements les plus récents, en particulier U-Map, U-Multirank et AHELO, en étaient encore au stade du développement au moment de la rédaction du présent rapport. Des tentatives ont été faites pour analyser ces développements et fournir des informations qui peuvent présenter un intérêt pour l'EUA. Ces développements seront examinés de façon plus détaillée dans le prochain rapport en 2012.

¹¹ Établi au départ par l'Université Jiao Tong de Shanghai.

1. Classements internationaux publiés sous forme de palmarès

1.1 Academic Ranking of World Universities (ARWU) – Shanghai Ranking Consultancy

L'histoire du classement ARWU commence vraiment en 1998, lorsque l'Université Jiao Tong de Shanghai est choisie par le gouvernement chinois pour faire partie du premier groupe de neuf universités du « projet 985 ». Ce projet a été mis en place en réponse à une déclaration de Jiang Zemin (alors Président de la République populaire de Chine) selon laquelle la Chine devait compter un certain nombre d'universités de réputation mondiale (Liu, 2009). De 1999 à 2001, une équipe de l'Université Jiao Tong a travaillé sur un projet visant à évaluer les universités chinoises les plus réputées et à les comparer aux universités de recherche des États-Unis,

« afin de mesurer l'écart entre les universités chinoises et les universités de réputation mondiale » (Liu et Cheng, 2005). Selon Liu (Liu, 2009), après la soumission du rapport par l'équipe au ministère chinois de l'Éducation, sa publication a suscité des réactions en Chine et à l'étranger favorables à l'établissement d'un véritable classement mondial des universités. Ce classement a été publié pour la première fois en 2003 et a été actualisé chaque année depuis lors. L'Université Jiao Tong a soutenu la publication du classement ARWU jusqu'en 2009, lorsqu'une société indépendante de conseil a été constituée à cette fin.

Une liste des critères et indicateurs utilisés dans le classement ARWU est présentée dans le tableau 3.

Tableau 2 : Liste des 21 champs disciplinaires tels que définis par Thomson Reuters Web of Knowledge ¹²

1. Sciences de l'agriculture	12. Mathématiques
2. Biologie & Biochimie	13. Microbiologie
3. Chimie	14. Biologie moléculaire & Génétique
4. Médecine clinique	15. Neurosciences
5. Informatique	16. Pharmacologie
6. Écologie/Environnement	17. Physique
7. Économie & gestion	18. Botanique & zoologie
8. Ingénierie	19. Psychologie/Psychiatrie
9. Géosciences	20. Sciences sociales – général
10. Immunologie	21. Sciences de l'espace
11. Science des matériaux	

Quelles sont les universités prises en considération dans le classement ARWU ?

Le classement comprend les universités comptant dans leurs rangs des lauréats du Prix Nobel, de la médaille Fields, des chercheurs fréquemment cités ou dont des articles sont publiés dans *Nature* ou *Science* (Liu, 2009). De plus, les universités ayant un nombre important d'articles indexés dans le *Science Citation Index-Expanded* (SCIE) et le *Social Sciences Citation Index* (SSCI) sont aussi comprises (voir la méthodologie ARWU du classement de 2010 – ARWU, 2010).

Par conséquent, le classement ARWU ne vise nullement à établir une comparaison entre toutes les universités du monde – ses critères s'appliquent uniquement aux universités de recherche les plus réputées à l'échelle mondiale. Seules 1 000 des 17 000 universités qui existent au monde sont sélectionnées dans le classement ARWU. Les 500 premières figurent au palmarès.

Critères retenus, indicateurs utilisés et indicateurs indirects
Quatre critères ont été retenus pour le classement des universités (voir le tableau 2 ci-dessous pour les disciplines) :

- *Qualité de l'enseignement* ¹³ ;
- *Qualité des enseignants-chercheurs* ¹⁴ ;
- *Production scientifique* ;
- *Productivité de l'université*.

Quels sont les éléments mesurés et les indicateurs indirects utilisés pour estimer la performance d'une université au regard de chacun de ces critères ?

L'indicateur indirect utilisé pour mesurer la *qualité de l'enseignement* est le nombre de diplômés ayant obtenu le Prix Nobel en sciences ou la médaille Fields en mathématiques.

¹² Ces 21 catégories sont décrites sur le site ISI Highly Cited : http://www.isihighlycited.com/isi_copy/Comm_newse04.htm.

¹³ Voir le tableau 3 pour le critère de la qualité de l'enseignement dans le classement ARWU.

¹⁴ Voir le tableau 3 pour les critères de la qualité du personnel académique dans le classement ARWU.

Deux indicateurs indirects sont utilisés pour estimer la *qualité des enseignants-chercheurs* :

- Le personnel ayant obtenu le Prix Nobel et la médaille Fields ;
- Le personnel figurant sur les listes de chercheurs les plus fréquemment cités dans 21 *champs disciplinaires* (voir tableau 2).

La liste des disciplines à elle seule démontre que les sciences naturelles, la médecine et les sciences de l'ingénieur sont prépondérantes dans l'indicateur de citations, au détriment des sciences sociales. Selon Liu, les arts et les sciences humaines ne sont pas classés à cause des difficultés techniques à trouver des indicateurs internationaux comparables avec des données fiables (Liu, 2009). Certaines disciplines ne sont pas comprises en raison des difficultés liées à leur caractère multidisciplinaire.

La production scientifique est mesurée par les indicateurs suivants :

- Le nombre d'articles publiés dans *Nature* et *Science* au cours des cinq années précédentes ;
- Le nombre d'articles indexés dans le *Science Citation Index-Expanded* et le *Social Sciences Citation Index* de Thomson Reuters au cours de l'année antérieure à celle de l'établissement du classement (il est à noter que seule la publication des « articles » et des « actes » est prise en compte).

La productivité n'est pas mesurée séparément mais est calculée à partir de la valeur des indicateurs précédents, rapportée au nombre des membres des enseignants-chercheurs sur la base de données nationales. Il est à noter que la *productivité* est le seul indicateur ARWU qui tient compte de la taille de l'établissement. C'est-à-dire que le classement ARWU reflète la performance globale d'une université et qu'il est beaucoup plus difficile pour des établissements plus petits, mais excellents, d'y figurer.

Tableau 3 : Critères, indicateurs et pondération utilisés dans le classement ARWU

Critère	Indicateur	Code	Poids
Qualité de l'enseignement	Diplômés d'un établissement ayant obtenu le Prix Nobel et la médaille Fields.	Diplômés	10%
Qualité des enseignants - chercheurs	Personnel d'un établissement ayant obtenu le Prix Nobel et la médaille Fields.	Prix	20%
	Les [200 premiers] chercheurs fréquemment cités dans 21 disciplines.	HiCi	20%
Production scientifique	Articles publiés dans <i>Nature</i> et <i>Science</i> ¹⁵	N&S	20%
	Articles indexés dans le <i>Science Citation Index-expanded</i> et le <i>Social Sciences Citation Index</i> .	PUB	20%
Productivité	Performance académique d'un établissement rapportée au nombre d'enseignants-chercheurs.	PCP	10%
Total			100%

Sources de données

Tous les indicateurs ARWU reposent sur des données recueillies auprès de tiers, tels que le site officiel du Prix Nobel, l'Union mathématique internationale pour les médailles Fields et plusieurs sites Internet de Thomson Reuters pour les citations et les publications. Le nombre d'enseignants-chercheurs pour l'indicateur de productivité est établi sur la base de données nationales. On pourrait certes affirmer que cet ensemble de données est collecté à l'échelle nationale auprès des universités elles-mêmes. C'est pourquoi Usher et Savino (Usher et Savino, 2006) considèrent que les données sur les enseignants-chercheurs sont des données fournies par les EES.

Le calcul des valeurs des indicateurs et leur conversion en scores

Il convient de noter que la *valeur d'un indicateur* et le *score d'un indicateur* sont deux choses différentes.

La *valeur* d'un indicateur est le résultat de la mesure réelle : ex. si l'université X compte 352 publications dans *Nature* ou *Science* (N&S), la valeur de l'indicateur est 352.

Dans le classement ARWU, les scores par indicateur sont calculés en divisant la valeur réelle de l'indicateur par celle de l'université affichant la valeur la plus élevée au regard de ce même indicateur et en multipliant le résultat par 100.

Exemple : l'université X compte 352 publications dans *Nature* ou *Science*, mais l'université Y affiche le meilleur résultat – 398 publications dans *Nature* et *Science* (N&S). Le score de l'université X pour l'indicateur N&S sera le suivant : $N\&SX = (352/398) \times 100 = 88,4$.

¹⁵ Pour les établissements spécialisés dans les sciences humaines et sociales, l'indicateur N&S n'est pas pris en compte et son poids est affecté à d'autres indicateurs.

Ceci implique également que la position dans le classement n'indique pas au lecteur quelles sont les valeurs des indicateurs. Par conséquent, si l'université Z affiche un score de 9,8 pour l'indicateur de productivité (*per capita performance*, PCP), cela ne signifie pas qu'un membre du personnel de l'université Z compte en moyenne 9,8 publications, mais que la performance de l'université Z au regard de l'indicateur PCP représente 9,8 % de la performance de l'université qui affiche le meilleur résultat au regard de ce même indicateur.

Indicateur « Diplômés ». Cette valeur est calculée à partir du nombre de lauréats du Prix Nobel ou de la médaille Fields parmi les diplômés, indépendamment du niveau de diplôme obtenu par le lauréat dans l'université en question. La note attribuée aux lauréats ayant obtenu leur diplôme avant 1991 est dégressive : la note des lauréats ayant obtenu leur diplôme au cours de la période 1981-1990 est de 0,9 ; celle des lauréats ayant obtenu leur diplôme au cours de la période 1971-1980 est de 0,8, et ainsi de suite, une note de 0,1 étant attribuée aux lauréats ayant obtenu leur diplôme entre 1901 et 1910.

Cet indicateur vise à refléter la « qualité de l'enseignement », ce qu'il fait d'une façon très spécifique et assez limitée.

Indicateur « Prix ». Cette valeur est calculée à partir du nombre de lauréats du Prix Nobel ou de la médaille Fields parmi les membres du personnel de l'université. La note attribuée aux lauréats ayant obtenu le prix avant 2001 est dégressive : elle est de 0,9 pour les lauréats de la période 1991-2000 ; de 0,8 pour les lauréats de la période 1981-1990, et ainsi de suite, une note de 0,1 étant attribuée aux lauréats de la période 1911-1920. Si l'un des lauréats est lié à plus d'une université, la valeur est partagée entre les universités, si un prix est partagé par plusieurs personnes, des poids sont attribués aux lauréats suivant la proportion du prix qui leur échoit.

Indicateur HiCi. Cette valeur est calculée à partir du nombre de chercheurs fréquemment cités dans les 21 disciplines listées dans le tableau 2. L'ISI recense les chercheurs fréquemment cités¹⁶ et établit des listes¹⁷ des 250 chercheurs les plus cités dans chacune des 21 disciplines.

Indicateur « Nature et Science ». Cette valeur indique le nombre d'articles publiés dans *Nature* et *Science* au cours des cinq années précédant l'année du classement. Pour les articles rédigés par plusieurs auteurs, la note 1 est attribuée à l'auteur « correspondant » (celui qui a soumis l'article), la note 0,5 est attribuée au premier auteur (ou le deuxième auteur si le premier est aussi l'auteur correspondant), la note 0,25 est attribuée à l'auteur suivant et la note 0,1 aux autres auteurs. Seuls les « articles » ou les « actes » sont pris en compte.

Indicateur « Publications ». Cette valeur indique le nombre d'articles indexés dans le *Science Citation Index-Expanded* et le *Social Sciences Citation Index* au cours de l'année précédente. Seuls les « articles » et les « actes » sont pris en compte. Un article indexé dans le *Social Sciences Citation Index* compte double.

Indicateur de productivité (*per capita performance*). Cette valeur correspond à la somme des scores pondérés des cinq indicateurs précédents divisée par le nombre d'enseignants-chercheurs exprimé en équivalents temps plein. Cependant, si le nombre d'enseignants-chercheurs des établissements dans un pays n'a pas pu être obtenu, le score pondéré des cinq indicateurs précédents est utilisé.

Pondération des indicateurs et critères retenus pour le calcul du score final. Le poids de chaque indicateur est indiqué au tableau 3. Le nombre de diplômés ayant obtenu le Prix Nobel étant un indicateur assez exclusif et le seul indicateur utilisé pour mesurer la qualité de l'enseignement, ce critère n'intervient qu'à hauteur de 10 % dans le calcul du score total.

Le poids total de la qualité des *enseignants-chercheur* est assez important et s'élève à 40 %. Pour obtenir des scores élevés au regard de ce critère, les membres du personnel d'une université doivent être essentiellement actifs dans les domaines des sciences ou de la médecine.

Le poids total de la *production scientifique* est également de 40 %. Les publications dans *Nature* et *Science* représentent la moitié de ce score. Ici encore, les chercheurs dans les domaines des sciences exactes et de la médecine sont favorisés. Une initiative a été prise pour que soient représentées les universités spécialisées dans les sciences humaines ou sociales. Dans ce cas, l'indicateur *Nature* et *Science* n'est pas appliqué et sa contribution au score total est attribuée en parts égales à d'autres indicateurs. Cependant, si cette initiative est utile aux établissements spécialisés dans les sciences humaines ou sociales, elle n'a aucun effet sur les universités pluridisciplinaires, les sciences sociales étant sous-représentées et les sciences humaines n'étant toujours pas prises en compte au regard de l'indicateur N&S.

Le critère de la productivité ne contribue qu'à hauteur de 10 % du score final. Il convient toutefois de noter qu'il s'agit du seul indicateur qui tienne compte de la taille de l'établissement.

Évolution de la méthodologie utilisée

La comparaison des méthodologies utilisées pour les classements ARWU 2004, 2005, 2006, 2007, 2008 et 2009, sur la base des informations disponibles sur le site Internet du classement ARWU, n'a pas permis de mettre en évidence un changement de méthodologie. Il convient toutefois de noter que, dans leur article publié en 2005, Liu et Cheng ont préféré recourir à la formulation « *taille de l'établissement* » plutôt que « *productivité* » (*per capita performance*). Ceci peut donner à penser que la méthodologie initiale a été révisée et que certaines modifications ont été apportées, bien que, si tel est le cas, celles-ci soient probablement mineures.

¹⁶ Voir http://isihighlycited.com/isi_copy/howweidentify.htm.

¹⁷ Les listes par discipline des chercheurs fréquemment cités sont disponibles à l'adresse : <http://hcr3.isiknowledge.com/>.

Présentation du classement et analyses complémentaires

Le classement ARWU peut être consulté sur Internet. Il est présenté sous la forme de listes d'universités : Top 100, Top 101-200, Top 201-300, Top 301-400 et Top 401-500.

Statistiques. Les données du classement ARWU sont également présentées sous la forme de statistiques par région du monde et par pays, qui indiquent le nombre d'universités d'une région du monde ou d'un pays en particulier qui figurent dans le Top 20, le Top 100, 200, 300, 400 et 500.

Des **classements par champ disciplinaire** sont établis pour les sciences naturelles et les mathématiques, les sciences de l'ingénieur, les sciences de la vie, la médecine et les sciences sociales. Les cinq indicateurs utilisés dans les classements par champ disciplinaire sont semblables à ceux du classement mondial des universités, mais leur pondération est différente et leur application est spécifique au domaine considéré¹⁸ (voir les classements ARWU par domaine). L'indicateur de productivité n'est pas utilisé, mais il est remplacé dans les classements par champ disciplinaire par l'indicateur *Top* (voir ci-dessous). Un autre indicateur – *Fonds (funding)* – est uniquement utilisé dans le domaine des sciences de l'ingénieur.

L'indicateur *Diplômés* représente un poids de 15 % et sa valeur correspond au nombre de diplômés ayant obtenu le Prix Nobel ou la médaille Fields (en sciences et en mathématiques) dans leurs disciplines respectives depuis 1951. Cet indicateur ne s'applique pas au domaine des sciences de l'ingénieur.

L'indicateur *Prix* représente un poids de 10% et sa valeur correspond au nombre d'enseignants-chercheurs ayant obtenu le Prix Nobel ou la médaille Fields (en sciences et en mathématiques) dans leurs disciplines respectives depuis 1961. Cet indicateur ne s'applique pas au domaine des sciences de l'ingénieur.

L'indicateur *Fonds* s'applique uniquement au domaine des sciences de l'ingénieur. Son poids est de 25 % (le poids total des indicateurs Diplômés et Prix utilisés pour les autres champs disciplinaires). La valeur de l'indicateur Fonds correspond au total des dépenses liées à la recherche dans le domaine des sciences de l'ingénieur.

L'indicateur *HiCi* a un poids de 25 %. La valeur de cet indicateur correspond au nombre de chercheurs recensés sur les listes des chercheurs fréquemment cités établies par ISI Web of Knowledge. Dans les classements par champ disciplinaire, les 21 disciplines sont classées selon des champs élargis d'étude: le classement par champ disciplinaire des sciences et des mathématiques couvre les mathématiques, la physique, la chimie, les géosciences et les sciences de l'espace ; le classement par champ disciplinaire des sciences de l'ingénieur couvre les sciences de l'ingénieur, la science des matériaux et l'informatique ; le classement par champ disciplinaire des sciences de la vie couvre la biologie et la biochimie, la biologie moléculaire et la génétique, la micro-

biologie, l'immunologie, la neuroscience, les sciences de l'agriculture, la botanique et la zoologie ainsi que l'écologie/environnement ; le classement par champ disciplinaire de la médecine couvre la médecine clinique, la pharmacologie, et, en partie, les sciences sociales (général) ; le classement par champ disciplinaire des sciences sociales couvre, en partie, les sciences sociales (général) l'économie & gestion. Les classements par champ disciplinaire ARWU ne couvrent pas la psychologie/psychiatrie.

L'indicateur *Publications* a un poids de 25 %. Sa valeur correspond au nombre d'articles indexés dans le Science Citation Index-Expanded par champ disciplinaire. Ici encore, seuls les « articles » et les « actes » son pris en compte.

L'indicateur *Top* a un poids de 15%. Il indique le pourcentage d'articles publiés dans le Top 20 % des revues scientifiques dans chaque champ disciplinaire, selon leur facteur d'impact tel que défini par le *Journal Citation Reports* de Thomson Reuters. Le seuil a été fixé à 10 % du nombre moyen d'articles publiés par les trois premiers établissements dans chaque champ disciplinaire. Si le seuil dans un champ spécifique est inférieur à 100 publications, c'est alors la valeur 100 qui est utilisée. Si le nombre d'articles publiés par un établissement est inférieur au seuil minimum, l'indicateur Top n'est pas calculé pour cet établissement et son poids est affecté à d'autres indicateurs.

Classements ARWU par discipline. Outre les classements par champ disciplinaire, qui portent sur plusieurs disciplines, ARWU présente également des classements dans les disciplines suivantes : mathématiques, physique, chimie, informatique et économie & gestion.

L'utilisation des indicateurs est semblable à celle des classements par champ disciplinaire, mais les données sont collectées pour chaque discipline spécifique. Il convient de noter que, dans le cas de l'informatique, le prix Turing est utilisé pour les indicateurs Diplômés et Prix.

Autres produits potentiels. Selon les producteurs du classement ARWU, est à l'étude la possibilité d'établir des classements basés sur différentes catégories d'universités, selon leurs fonctions, les caractéristiques des disciplines étudiées, leur histoire, taille et budget (Liu, 2009). Outre le classement ARWU, Shanghai Ranking Consultancy publiera par ailleurs plusieurs comparaisons internationales et analyses approfondies d'universités de recherche ¹⁹.

¹⁸ Le tableau des indicateurs des classements par domaine peut être consulté à l'adresse : <http://www.arwu.org/ARWUFieldMethodology2009.jsp>.

¹⁹ Voir <http://www.arwu.org/aboutARWU.jsp>.

1.2 THE World University Ranking – Times Higher Education

Le *Times Higher Education World University Rankings* a été publié pour la première fois en 2004. Il s'agissait dans une certaine mesure d'une « réponse » à la première publication du classement ARWU en 2003.

Le classement du *Times Higher Education Supplement*, par après devenu *Times Higher Education*, était établi en coopération avec Quacquarelli-Symonds (QS), la société en charge de la collecte et du traitement des données entre 2004 et 2009. En 2009, le *Times Higher Education* (ci-après le « THE ») a annoncé la fin de sa coopération avec QS et l'établissement d'une nouvelle coopération avec Thomson Reuters (Baty, 2009).

Le THE a depuis lors présenté la nouvelle méthodologie de son classement mondial des universités pour l'année 2010. Entre-temps, QS a conclu un nouveau partenariat avec *US News and World Report* visant à établir un autre classement mondial. La méthodologie du THE au cours de la période 2004-2009 et la méthodologie proposée pour le classement 2010 seront par conséquent présentées séparément.

1.2.1 THE-QS World University Rankings entre 2004 et 2009

L'objectif affiché des classements du *Times Higher Education* (THE World University Rankings - THE) est « de reconnaître les universités comme les organisations polyvalentes qu'elles sont, d'établir une comparaison globale de leur performance par rapport à la mission théorique de rester ou de devenir une référence mondiale » (THE-QS, 2009).

La qualité de la recherche, la qualité de l'enseignement, l'employabilité des diplômés et l'orientation internationale sont décrites comme les quatre piliers d'une université de réputation mondiale et représentent par conséquent les domaines sur lesquels porte le classement THE (Ibid.).

Les classements THE et ARWU diffèrent largement quant à la méthode utilisée. Alors que le classement ARWU examine principalement les résultats de la recherche, le score final dans le classement THE repose dans une large mesure (à l'origine 50%, ensuite 40%) sur une « évaluation par les pairs » des universités, qui en réalité est une enquête de réputation en ligne réalisée auprès des personnels d'enseignement et de recherche. Une autre enquête, réalisée auprès des employeurs, contribue à hauteur de 10 % à l'établissement du score final.

Quelles sont les universités prises en considération ?

Comme mentionné plus haut, le classement tient compte des universités qui ont ou construisent une réputation mondiale. Selon une autre déclaration, les « THE QS World University Rankings™ ont été conçus pour présenter une vue d'ensemble des forces relatives des universités qui jouissent

d'une réputation mondiale », confirmant par là que seule l'élite mondiale des universités est prise en compte dans l'établissement du classement (THE-QS, 2009)²⁰. Il n'est dès lors pas surprenant que la méthodologie utilisée ne retienne qu'environ « 600 universités en tout et 300 dans chacun des cinq principaux domaines académiques » (Ibid.).

Critères retenus, indicateurs utilisés et indicateurs indirects

Les quatre domaines couverts par les indicateurs sont les suivants : deux indicateurs – *l'évaluation par les pairs* et *les citations par membre d'enseignant-chercheur* – sont utilisés pour caractériser la recherche ; *l'évaluation par les employeurs* est utilisée pour estimer l'employabilité des diplômés ; le seul indicateur indirect utilisé pour évaluer la qualité de l'enseignement est le très critiqué *taux d'encadrement des étudiants* (basé sur le nombre d'étudiants par enseignant) ; les deux indicateurs indirects utilisés pour décrire l'orientation internationale des universités sont la proportion du *personnel international* et la proportion d'*étudiants internationaux*.

Indicateur « évaluation par les pairs ». Cette évaluation par les pairs n'est pas une visite d'experts sur site, au sens d'une « évaluation par les pairs » dans les procédures d'assurance qualité. Ici, *l'évaluation par les pairs* est une enquête en ligne par laquelle les « pairs » sont invités à sélectionner jusqu'à 30 universités à partir d'une liste préétablie.

Cet indicateur a le poids le plus important dans les classements THS-QS – 50 % dans le premier classement en 2004, et 40 % entre 2005 et 2009. Selon la méthodologie présentée, « les résultats sont fondés sur les réponses à une enquête diffusée dans le monde entier aux personnes interrogées précédemment et aux abonnés à deux bases de données de premier plan : *World Scientific* (www.worldscientific.com), qui fournit 180 000 adresses de courrier électronique, et *International Book Information Service (IBIS)* »²¹. Le score global pour cet indicateur est établi sur la base de cinq séries de résultats correspondant aux cinq domaines suivants : arts et lettres, ingénierie & technologies de l'information, sciences de la vie & biomédecine, sciences naturelles, sciences sociales (Ibid.).

Les personnes interrogées sont invitées à indiquer leur domaine de compétence et leur appartenance à une région géographique. Les répondants n'ont plus la possibilité de sélectionner leur propre établissement depuis 2007, et les réponses relatives aux universités dans et en dehors de leur pays sont recueillies séparément depuis 2008, afin d'éviter les biais liés aux réponses par pays (THE-QS, 2009).

Cet indicateur soulève un certain nombre de questions importantes. D'une part, les scores sont établis sur la base d'un nombre assez limité de réponses : 9 386 en 2009 et 6 534 en 2008. En réalité, les 3 000 réponses environ à l'enquête de 2009 ont simplement été ajoutées aux chiffres de 2008 (voir le tableau 4 ci-dessous). Le nombre de réponses est lamentablement faible par rapport aux 180 000 adresses de courrier électronique utilisées.

²⁰ <http://www.topuniversities.com/university-rankings/world-university-rankings/methodology/simple-overview>: consulté le 29 décembre 2010

²¹ <http://www.topuniversities.com/university-rankings/world-university-rankings/methodology/academic-peer-review>.

D'autre part, les personnels d'enseignement et de recherche sont invités à « sélectionner jusqu'à 30 universités dans [la] liste internationale qui, selon [eux], réalisent les meilleurs travaux de recherche en (nom du champ disciplinaire approprié) »²². Que sont ces « listes internationales » ? Malgré la publication en ligne de la méthodologie QS 2010 (après la fin de la collaboration avec THE), il n'a pas été possible de trouver des informations publiques sur les auteurs de ces listes dans chaque domaine. Surtout, sur quels critères ceux-ci se fondent-ils pour écarter un nombre important d'universités, voire des pays entiers ? À titre d'exemple, les listes préétablies par le THE-QS dans les cinq domaines ne recensent généralement que des universités de 25 ou 26 pays européens sur les 48 pays que compte l'Espace européen de l'enseignement supérieur. Par ailleurs, les personnels d'enseignement et de recherche interrogés (les « pairs ») ne peuvent pas sélectionner des universités qui ne figurent pas sur la liste, sans parler des universités de pays complètement absents de la liste. La présentation de la méthodologie QS 2010²³ peut contribuer à éclairer ces questions.

Selon l'explication fournie en 2010, la sélection des universités pour le classement s'appuie sur quatre sources :

- Le succès dans les classements nationaux suivis par le département des renseignements de QS (*QS Intelligence Unit*) ;
- Les établissements mentionnés par les répondants aux enquêtes de réputation auprès des personnels d'enseignement et de recherche et des employeurs ;
- L'équilibre géographique (aucune explication complémentaire quant aux critères ou aux principes) ;
- La soumission directe d'une requête – les établissements qui s'adressent directement à QS en vue de leur inclusion dans le classement.

Cependant, ces indications n'expliquent toujours pas pourquoi les listes dans lesquelles les pairs et les employeurs sélectionnent les 30 premières universités sont différentes dans chaque domaine ou qui décide d'inclure une université plutôt qu'une autre dans chacune des listes.

Indicateur « évaluation par les employeurs ». L'évaluation par les employeurs est réalisée selon le modèle de l'évaluation des universités par les pairs, s'agissant aussi d'une enquête en ligne réalisée à l'échelle mondiale. Ici encore, les adresses électroniques des répondants sont obtenues à partir de trois sources principales : l'importante base de données de QS sur les entreprises, un réseau de partenaires avec lesquels QS coopère dans le cadre de ses activités, et les établissements participants qui soumettent une liste des professionnels avec lesquels ils collaborent (THE-QS, 2009)²⁴, induisant ainsi un nouveau biais. Les personnes interrogées sont également invitées à sélectionner jusqu'à 30 universités dont sont issus les meilleurs diplômés du premier cycle (Ibid.).

Les questions qui se posent concernant l'évaluation par les employeurs sont les mêmes que celles concernant l'évaluation par les pairs : qui réalise la présélection des universités figurant sur les listes destinées aux employeurs et selon quels critères ces établissements sont-ils sélectionnés ?

Seules 2336 réponses ont été recueillies à l'échelle mondiale, ce qui n'est pas sans implications. Notamment parce qu'un échantillon aussi réduit d'employeurs dans le monde peut tout simplement ne pas connaître d'excellentes universités dans des pays plus petits, en particulier dans des pays où l'on ne parle ni l'anglais ni une autre langue internationale.

Indicateur « citations par enseignant-chercheur ». Au cours de la période 2004-2006, le classement THE-QS s'est fondé sur les données relatives au nombre de citations fournies par Essential Science Indicators (ESI) du *Thomson Reuters Web of Science*. À partir de 2007, le THE-QS s'est tourné vers Scopus (qui appartient aujourd'hui à Elsevier) « pour un certain nombre de motifs, mais essentiellement en raison d'une plus large couverture des articles permettant d'obtenir des résultats pour un plus grand nombre d'établissements » (THE-QS, 2009, section sur l'indicateur « citations per faculty »²⁵). Le nombre d'enseignants-chercheurs est une valeur totale exprimée en ETP, plutôt que des chiffres distincts sur le nombre d'enseignants et de chercheurs. Cette orientation s'explique essentiellement par des définitions différentes et, par conséquent, par des données incompatibles entre différents pays.

Taux d'encadrement des étudiants (basé sur le nombre d'étudiants par enseignant). Les classements THE-QS au cours de la période 2004-2009 se sont fondés sur le taux d'encadrement des étudiants comme seul indicateur indirect de la *qualité de l'enseignement*. Le poids de cet indicateur y était alors de 20 %. Les nouveaux producteurs du classement THE citent des critiques quant à l'utilisation de ce taux, qu'ils critiquent eux-mêmes (Baty, 2010a). En tout état de cause, peu de professeurs conviendront que leur enseignement peut être jugé selon le nombre d'étudiants dans la classe.

Le nombre total d'étudiants correspond à la somme du nombre d'étudiants des différents cycles (licence, master, doctorat), sur la base des chiffres fournis. Lorsque ces données ne sont pas disponibles ou sont incomplètes, le nombre total d'étudiants est utilisé (THE-QS, 2009, section sur le taux d'encadrement des étudiants)²⁶. Le nombre d'enseignants-chercheurs est une valeur totale exprimée en ETP. Les producteurs du classement admettent néanmoins que « l'idéal serait de distinguer les notions d'enseignement et de recherche et de se fonder sur la première pour le calcul de cet indicateur » (Ibid.). Il a par ailleurs été démontré que, en raison de différentes définitions des catégories de membres du personnel, les universités peuvent facilement manipuler les résultats de cet indicateur. Selon Baty, Marny Scully, Directrice générale Politique et Analyse de l'Université de Toronto, a démontré que selon les définitions, sur la base des mêmes données, ce coefficient pouvait varier de 6/1 à 39/1 (Baty, 2010b). Désireuses d'améliorer leur position dans les classements, les universités peuvent être tentées de présenter leurs données de la façon la plus avantageuse.

Les indicateurs « personnel international » et « étudiants internationaux » ont un poids de 5 % chacun. Il peut sembler simple et direct de calculer les valeurs de ces indicateurs en divisant le nombre de collaborateurs étrangers ou d'étudiants

²² http://research.qsnetwork.com/qs_surveysystem/index.php?viewonly&order=normal&partnerset=0 : consulté le 15 septembre 2010.

²³ <http://www.topuniversities.com/university-rankings/world-university-rankings/methodology/institution-inclusion> : consulté le 29 décembre 2010.

²⁴ Voir <http://www.topuniversities.com/university-rankings/world-university-rankings/methodology/employer-review>.

²⁵ Voir <http://www.topuniversities.com/university-rankings/world-university-rankings/methodology/citations-faculty>.

²⁶ Voir <http://www.topuniversities.com/university-rankings/world-university-rankings/methodology/student-faculty>.

internationaux par le personnel total exprimé en ETP ou par le nombre total d'étudiants. Les informations publiques sur le classement THE-QS ne comportent toutefois aucune définition du personnel international et des étudiants internationaux. La définition d'un étudiant international peut par exemple être liée à son pays d'origine, à son pays de résidence et/ou au pays dans lequel il a fait ses études auparavant. En outre, les étudiants internationaux peuvent être des étudiants qui suivent un cycle d'études complet ou qui n'étudient à l'étranger que pendant une période déterminée. Quant aux collabo-

rateurs étrangers, ils peuvent être employés par l'université ou invités ; leur séjour, permanent ou temporaire, peut être financé par des ressources extra-budgétaires, etc. Comme les chiffres relatifs au nombre d'étudiants internationaux et aux collaborateurs étrangers font partie des informations recueillies auprès des universités elles-mêmes, des nuances dans les définitions peuvent également avoir une incidence sur les résultats.

Tableau 4 : Indicateurs et pondération dans le classement THE-QS en 2005-2009

Indicateur	Explication	Pondération
Évaluation par les pairs	Indice composite établi sur la base d'une enquête auprès de personnels d'enseignement et de recherche (divisée en cinq domaines d'étude). 9 386 réponses en 2009 (6 354 en 2008).	40%
Évaluation par les employeurs	Score basé sur les réponses à une enquête auprès des employeurs. 3 281 réponses en 2009 (2 339 en 2008).	10%
Ratio enseignants/étudiants	Score basé sur le nombre d'étudiants par enseignant.	20%
Citations par enseignant-chercheur	Score basé sur les résultats de la recherche rapportés au nombre d'enseignants-chercheurs.	20%
Personnel international	Score basé sur la proportion de collaborateurs étrangers dans le personnel.	5%
Étudiants internationaux	Score basé sur la proportion d'étudiants internationaux.	5%

Source : QS Quacquarelli Symonds. www.topuniversities.com. Avec autorisation.

Copyright © 2004-2008 QS Quacquarelli Symonds Ltd.

Sources de données

Le classement THE-QS a utilisé plusieurs sources de données au cours de la période 2004-2009.

Des enquêtes ont été réalisées auprès de personnels d'enseignement et de recherche pour élaborer l'indicateur « évaluation par les pairs », au poids important, qui est utilisé pour décrire la réputation de l'université dans le domaine de la recherche.

Des enquêtes ont été réalisées auprès d'employeurs pour déterminer la réputation des universités quant à l'employabilité de leurs diplômés, afin d'élaborer l'indicateur « évaluation par les employeurs ».

Les données utilisées pour élaborer l'indicateur « citations par enseignant-chercheur » ont été recueillies auprès de tiers : Essential Science Indicators (ESI) du Thomson Reuters Web of Science entre 2004 et 2006 et Scopus (Elsevier) à partir de 2007.

Les informations fournies par les universités. Le classement THE-QS a souvent eu recours à des données recueillies auprès des universités elles-mêmes, comme les nombres d'étudiants à temps plein (pour tous les cycles), d'enseignants-chercheurs (ETP), des personnels internationaux et d'étudiants internationaux.

Le calcul des valeurs des indicateurs et leur conversion en scores

Évaluation par les pairs. Les valeurs de cet indicateur pour une université sont d'abord calculées dans chacun des cinq domaines d'étude : arts et lettres, ingénierie & technologies de l'information, sciences de la vie & biomédecine, sciences naturelles, sciences sociales.

Au cours de ce processus :

- Les réponses des pairs relatives aux universités dans et en dehors de leur propre pays sont d'abord recombinaées, « afin d'éviter les biais liés aux réponses par pays » (THE-QS, 2009) ²⁷ ;
- Des coefficients régionaux sont ensuite appliqués à chacun des cinq domaines d'études, « pour garantir une représentation équilibrée des quatre « grandes » régions des Amériques ; Europe ; Moyen-Orient et Afrique ; et la région Asie-Pacifique » (Ibid.) ; Il n'a toutefois pas été possible d'obtenir des informations publiques sur la valeur de cette pondération régionale et sur son application.

Les résultats dans les cinq domaines sont ensuite additionnés.

Évaluation par les employeurs. Comme pour l'indicateur « évaluation par les pairs », les réponses des employeurs re-

²⁷ <http://www.topuniversities.com/university-rankings/world-university-rankings/methodology/academic-peer-review>.

latives aux universités dans et en dehors de leur propre pays sont recombinaées, et la pondération régionale est appliquée.

Citations par enseignant-chercheur. La valeur de cet indicateur est calculée en divisant le nombre de citations par le nombre total d'enseignants-chercheurs (exprimé en ETP), sans distinction entre les enseignants et les chercheurs (voir plus haut).

Taux d'encadrement des étudiants. Bien que souvent appelé le « ratio étudiants/enseignants » (faculty-student), cet indicateur est ici intitulé « ratio enseignants/étudiants », c'est-à-dire le contraire. Il est cependant calculé en divisant le nombre d'étudiants (de tous les cycles) par le nombre total d'enseignants-chercheurs (en ETP), toujours sans distinction entre les enseignants et les chercheurs.

Les valeurs des indicateurs « *personnel international* » et « *étudiants internationaux* » sont établies comme la proportion de collaborateurs étrangers et d'étudiants internationaux par rapport à l'ensemble des enseignants-chercheurs et des étudiants respectivement. Les définitions de membre du personnel international ou d'étudiant international ne sont pas précisées.

Normalisation des résultats

Avant 2007, le score final d'une université dans le classement THE-QS était calculé en divisant le score de l'université par celui de l'université qui affichait la valeur la plus élevée et en multipliant le résultat par 100 (pour plus de précisions sur cette méthode, voir la description du classement ARWU, p.27).

Entre 2007 et 2009, les écarts types et les valeurs moyennes pour tous les indicateurs étaient calculés avant que les scores de chaque indicateur ne soient combinés en un score final. Ensuite, sur la base de méthodes statistiques, on vérifiait si les résultats des différents indicateurs étaient compatibles entre eux et permettaient une combinaison des données (THE-QS, 2009, section sur la normalisation) ²⁸.

La méthode de normalisation utilisée fait intervenir des z-scores. Le logarithme naturel est d'abord appliqué aux résultats de chaque indicateur ²⁹. Sont ensuite calculés la valeur moyenne $\bar{X}$ et l'écart type σ des résultats de chaque indicateur.

Le z-score correspond à la différence entre la mesure x et la valeur moyenne $\bar{X}$, divisée par l'écart type σ :

$$Z = \frac{x - \bar{X}}{\sigma}$$

Comme ils sont sans dimension, les z-scores de différents indicateurs peuvent être combinés en vue d'établir un score global.

Le score final est calculé en multipliant chaque score d'indicateur par son coefficient de pondération, en additionnant les résultats, en arrondissant à une décimale et en mettant la valeur obtenue à l'échelle de l'établissement ayant obtenu le meilleur résultat, auquel est attribué un score de 100 (Ibid.).

Évolution de la méthodologie utilisée

Un certain nombre de modifications ont été apportées au classement THE-QS depuis 2004.

En 2005 :

- L'indicateur « évaluation par les employeurs » a été introduit ;
- Le poids de l'indicateur « *évaluation par les pairs* » a été réduit, de 50 à 40 %.

En 2007 :

- Les paramètres de l'enquête réalisée auprès des personnels d'enseignement et de recherche (indicateur « évaluation par les pairs ») ont été modifiés de façon à ce qu'il ne soit plus possible de sélectionner sa propre université comme faisant partie des 30 universités les plus performantes ;
- Il a été décidé d'appliquer des z-scores pour le calcul des scores de chaque indicateur, ainsi que de combiner les scores des indicateurs en vue d'établir un score global (voir plus haut). C'est-à-dire que le score d'une université n'a plus été calculé en divisant son score par celui de l'université qui affichait la valeur la plus élevée et en multipliant le résultat par 100 ;
- Il a été décidé de vérifier la compatibilité des résultats de différents indicateurs ;
- Les données relatives aux citations n'ont plus été fournies par ESI (Thomson Reuters) mais par Scopus (Elsevier).

En 2008 :

- L'indicateur « évaluation par les pairs » a été ajusté : les réponses relatives aux universités dans et en dehors de leur pays ont été recueillies séparément.

La définition précise des données demandées a par ailleurs évolué progressivement au fil des ans (THE-QS, 2009).

En 2009, suite à la publication du classement THE-QS 2009, THE a annoncé la fin de sa coopération avec QS comme fournisseur de données et a conclu un contrat avec Thomson Reuters (Baty, 2009).

Il va de soi qu'autant de modifications apportées à la méthodologie doivent avoir eu une incidence sur les résultats. Il semble impossible de déterminer exactement comment ces changements multiples peuvent avoir modifié la position des différentes universités figurant dans le classement.

²⁸ <http://www.topuniversities.com/university-rankings/world-university-rankings/methodology/standardization-weightings-and-aggregation>

²⁹ L'application d'une échelle logarithmique réduit considérablement les différences entre les scores. Le logarithme est donc utilisé afin d'éviter l'effet induit par des différences entre les universités, qui semblent importantes mais sont en réalité insignifiantes.

Présentation du classement et analyses complémentaires

Le *Times Higher Education* publie le Top 200 et le Top 50 des universités dans le monde dans cinq domaines d'étude : ingénierie & technologies de l'information, sciences de la vie & biomédecine, sciences naturelles, sciences sociales, arts et lettres.

En 2008, le THE a également présenté un classement des systèmes nationaux d'enseignement supérieur (*System Strength Rankings*), qui recensait les 40 systèmes d'enseignement les plus performants, sans répéter l'exercice par la suite (THE, le 9 octobre 2008). Les scores par pays étaient établis sur la base de quatre séries de données :

- « *System* » vise à examiner la capacité des systèmes d'enseignement à produire des universités de niveau international. Sa valeur correspond au nombre d'universités de chaque pays figurant dans le Top 600, divisé par leur position moyenne ;
- « *Access* » correspond, pour chaque pays, au nombre d'étudiants (ETP) dans le Top 500 des universités, divisé par la racine carrée de la population du pays ;
- « *Flagship* » indique la position du premier établissement de chaque pays dans le classement ;
- « *Economic measure* ». Sa valeur est calculée comme suit : 5 points sont attribués aux universités figurant dans le Top 100, 4 points aux universités dans le Top 101-200, et 3, 2 ou 1 points pour chaque université classée dans le Top 201-300, le Top 301-400 et le Top 401-500 respectivement. La somme des points est ensuite divisée par le PIB par habitant du pays concerné.

Les résultats du THE-QS sont publiés sur le site Internet de QS, et sont disponibles sous forme de listes des Top 100, 101-200, 201-300, 301-400, 401-500 et 500+. Dans la section interactive des classements, les utilisateurs peuvent ajouter de nouveaux domaines au classement, comparer les scores des universités avec ceux des années précédentes, consulter les résultats des universités dans les cinq domaines d'étude ou classer les résultats en fonction des scores par indicateur. Toutefois, le site n'offre pas la possibilité de calculer le score final sur la base d'une pondération différente des indicateurs, comme recommandé par les Principes de Berlin (Principes de Berlin, 2006).

1.2.2 La méthodologie THE-Thomson Reuters du World University Ranking 2010

Le 30 octobre 2009, le *Times Higher Education* a annoncé avoir conclu un accord avec Thomson Reuters pour la fourniture des données de son classement annuel des universités dans le monde (Baty, 2009). Cette décision a été suivie par de considérables modifications apportées à la série d'indicateurs et à la méthodologie générale des classements 2010.

Critères retenus, indicateurs utilisés et indicateurs indirects

La nouvelle méthodologie a été rendue publique en juin 2010. Elle est fondée sur une nouvelle série d'indicateurs « pour rendre les classements plus rigoureux, équilibrés, sophistiqués et transparents » (Baty, 2010c). La description de la méthodologie a été publiée au même moment sous la forme d'une description schématique et peu détaillée des catégories et des indicateurs. D'autres changements ont été apportés à la méthodologie par la suite. La publication du classement THE-Thomson Reuters et de sa méthodologie le 16 septembre 2010 a révélé que les catégories d'indicateurs, leur pondération et certains indicateurs eux-mêmes avaient été considérablement modifiés par rapport à la méthodologie annoncée auparavant (voir Baty, 2010e par rapport à Baty, 2010c).

Le 5 août 2010 (voir Baty, 2010d), le THE a également annoncé l'exclusion de plusieurs catégories d'universités de la nouvelle présentation du classement THE, notamment celles qui n'avaient pas communiqué de données, les écoles doctorales, et les universités ayant publié moins de 50 articles en 2008.

Le classement THE-Thomson Reuters 2010 s'appuie sur 13 indicateurs différents, regroupés en cinq catégories : la recherche – volume, financement, réputation (30 %) ; l'impact de la recherche (32,5 %) ; l'activité économique et l'innovation (2,5 %) ; le degré d'internationalisation – personnel et étudiants (5 %) ; et l'enseignement – l'environnement d'étude (30 %).

Jusqu'à présent, la description de la méthodologie utilisée est moins élaborée que celle des classements THE-QS et ne permet pas de suivre le calcul des scores finaux sur la base des données brutes. La pondération des catégories d'indicateurs et la liste des indicateurs du classement 2010 sont présentées dans le tableau 5.

Tableau 5 : Classement THE-Thomson Reuters 2010
Catégories, indicateurs et pondération

Poids (catégorie)	Catégories	Indicateurs	Poids
2,5%	Activité économique/Innovation		
		Financement de la recherche par l'industrie (par enseignant-chercheur).	2,5%
5%	Taux d'internationalisation – personnel et étudiants		
		Personnel international rapporté au personnel national.	3%
		Nombre d'étudiants internationaux rapporté au nombre d'étudiants nationaux.	2%
30%	Enseignement – l'environnement d'étude		
		Enquête de réputation – enseignement.	1,5%
		Nombre de doctorats décernés rapporté au nombre des enseignants-chercheurs.	6%
		Nombre d'étudiants inscrits (en premier cycle) rapporté au nombre d'enseignants-chercheurs.	4,5%
		Nombre de doctorats décernés rapporté au nombre de licences décernées.	2,25%
		Budget global par enseignant-chercheur.	2,25%
30%	Recherche – volume, financement, réputation		
		Enquête de réputation – recherche.	19,5%
		Budget global de recherche rapporté au nombre d'enseignants-chercheurs.	5,25%
		Articles par enseignant-chercheur.	4,5%
		Financement public de la recherche / rapporté au financement total de la recherche.	0,75%
32,5%	Citations – influence de la recherche		
		Impact en citations (moyenne des citations par article, normalisée par champ disciplinaire et année).	32,5%

Source : Baty, 2010e.

La catégorie « activité économique – innovation » (2,5 %) repose sur un seul indicateur en 2010 :

- Le *financement de la recherche par l'industrie*, rapporté au nombre d'enseignants-chercheurs³⁰. Il a également été affirmé que cette catégorie pourrait comprendre à l'avenir des données sur les articles écrits en collaboration avec des partenaires de l'industrie, ainsi que l'avis des employeurs sur les diplômés (Baty, 2010c).

La catégorie relative à la diversité institutionnelle, annoncée initialement (pondération proposée de 10 %) (Baty, 2010c) est devenue le degré d'internationalisation, dont le poids de 5 % est réparti entre deux indicateurs :

- *Le personnel international rapporté au personnel national* – actuellement de 3 % ;
- *Le nombre d'étudiants internationaux* rapporté au nombre d'étudiants nationaux – actuellement de 2 % (par rapport à 5 % chacun, comme indiqué en juin 2010).

Il avait été annoncé auparavant (Baty, 2010c) que cette catégorie pourrait également comprendre à l'avenir des données sur les articles écrits en collaboration avec des partenaires internationaux.

La catégorie relative aux *indicateurs institutionnels*, d'un poids total de 25 %, comme annoncé en juin 2010 (Baty, 2010c), est devenue la catégorie « Enseignement – environnement d'étude » (poids total de 30 %) en septembre. Cette catégorie repose sur cinq indicateurs :

1. *Enquête de réputation – enseignement* (15 %). Cet indicateur est basé sur une enquête réalisée auprès d'universitaires expérimentés. Cumulée à une enquête de réputation sur la recherche (voir ci-dessous), elle a recueilli 13 388 réponses. Les producteurs du classement fournissent peu d'informations sur les enquêtes de réputation (l'une sur la recherche et l'autre sur l'enseignement) sur le site Internet du THE. Selon d'autres publications, « les répondants estiment la réputation des établissements dans leur propre domaine de compétences » et cette « enquête a été traduite dans de nombreuses langues pour remédier au biais induit par la langue anglaise » (Pratt, 2010). Aucune explication n'est toutefois donnée quant aux procédures d'enquête proprement dites. Il serait utile de connaître les questions posées aux universitaires expérimentés, le nombre d'universités qu'ils peuvent choisir, s'ils peuvent choisir eux-mêmes les universités ou si leur choix est conditionné par une liste préétablie, comme dans les précédents classements du THE.

2. *Nombre de doctorats décernés* (6 %) – le nombre de doctorats décernés par un établissement est rapporté au nombre d'enseignants-chercheurs.

3. *Nombre d'étudiants inscrits* (en premier cycle) rapporté au nombre d'enseignants-chercheurs (4,5 %). Il s'agit en réalité du très critiqué taux d'encadrement des étudiants (ratio étudiants/enseignants), auquel s'applique toutefois un coefficient de pondération nettement inférieur aux 15 % qui lui était attribué dans les classements THE-QS.

4. *Nombre de doctorats décernés rapporté au nombre de licences décernées* (2,25 %). Il n'est pas précisé si seul le premier ou les deux premiers cycles sont pris en compte – ce

qui représente une différence importante dans certains pays.

5. *Budget global par enseignant-chercheur* (2,25 %). Il s'agit de la mesure du budget total d'un établissement rapporté au nombre des enseignants-chercheurs. Il n'est pas précisé si les chiffres utilisés correspondent au nombre total de membres des enseignants-chercheurs, d'enseignants ou de chercheurs. Cette valeur est ajustée à la parité du pouvoir d'achat.

La catégorie « *Citations – influence de la recherche* » (32,5 %) repose sur un seul indicateur, qui est le plus déterminant de la série. Il correspond au nombre de citations par article mentionnées dans 12 000 articles indexés dans la base de données Web of Science de Thomson Reuters sur une période de cinq ans (2004-2008) (Baty, 2010e). Les résultats sont normalisés – le nombre de citations par article est rapporté au nombre moyen de citations par article pour l'ensemble des articles publiés dans le même champ disciplinaire et au cours de la même année (Ibid.), mais la méthodologie de normalisation par champ disciplinaire n'est pas précisée³¹. Les établissements publiant moins de 50 articles par an ne sont pas pris en considération, car l'impact relatif en citations peut être considérablement influencé par un ou deux articles fréquemment cités et, par conséquent, ne reflète pas avec précision leurs résultats habituels.

Recherche – volume, financement, réputation. Cette catégorie (30 %) s'appuie sur les bases de données de Thomson Reuters recensant les articles scientifiques et comprend quatre indicateurs :

- *L'enquête de réputation* (19,5 %) (voir les remarques sur l'enquête de réputation de l'enseignement plus haut) ;
- *Budget global de recherche de l'université* (5,25 %) rapporté au nombre d'enseignants-chercheurs et ajusté à la parité du pouvoir d'achat ;
- *Les articles par enseignant-chercheur* (4,5 %) correspondent au nombre d'articles publiés par les personnels enseignants et de recherche dans les revues scientifiques et indexées par Thomson Reuters ;
- *Le financement public de la recherche rapporté au financement total de la recherche* (0,75 %). Les données utilisées sont recueillies auprès des universités.

Si l'on examine la méthodologie du classement THE-Thomson Reuters 2010 sous un autre angle, la pondération des indicateurs utilisés peut être présentée différemment :

- Les **indicateurs bibliométriques** (citations par article et articles par enseignant-chercheur) représentent la proportion la plus importante de la pondération totale, soit 37 % ;
- Les **indicateurs de réputation** (de la recherche et de l'enseignement cumulés) suivent de près, avec un poids de 34,5 % ;
- Les **indicateurs de financement** (financement de la recherche par l'industrie, budget global de recherche et budget global de l'université – tous par enseignant-chercheur – et le financement public de la recherche rapporté au financement total de la recherche) représentent un poids total de 10,75 % ;
- L'**importance des études doctorales** (le nombre de doctorats décernés rapporté au nombre de licences décernées,

³⁰ D'autres indicateurs mentionnent les « membres du personnel de recherche ». La différence n'a pas encore été précisée.

³¹ Voir l'examen des indicateurs CPP/FCSm, MNCS et MNCS2 dans la section sur le classement de Leiden.

et le nombre de doctorats décernés rapporté au nombre d'enseignants-chercheurs) représente un poids total de 8,25 % ;

- Les **indicateurs d'internationalisation** (nombre d'étudiants internationaux et personnel national rapportés au nombre d'étudiants nationaux et au personnel national respectivement) représentent un poids cumulé de 5 % ;

- Le **nombre d'étudiants rapporté au nombre d'enseignants-chercheurs** représente un poids total de 4,5 %.

Par conséquent, bien que plus d'un tiers de la pondération totale ait été attribué aux indicateurs bibliométriques, les enquêtes de réputation continuent à représenter plus d'un tiers de la pondération totale, en dépit des critiques formulées précédemment. Les enquêtes de réputation ne sont pas suffisamment décrites pour déterminer si la méthodologie sur laquelle elles s'appuient a été améliorée depuis 2009.

Suivent en importance les indicateurs liés aux revenus, qui toutefois n'exercent pas une influence déterminante étant donné le poids de 10 % qui leur a été attribué.

Nul n'a expliqué pourquoi les indicateurs d'internationalisation ont un poids inférieur à celui qui avait été prévu initialement, en particulier quand l'internationalisation est présentée comme l'une des caractéristiques essentielles des « universités de classe mondiale » (Salmi, 2009).

Le poids du très critiqué taux d'encadrement des étudiants est ici relativement faible.

On peut conclure de ce qui précède que le classement THE-Thomson Reuters est fortement axé sur la recherche. Les indicateurs cumulés de volume et d'impact de la recherche représentent un poids de 62,5 %. Et si l'on ajoute l'indicateur sur le financement de la recherche par l'industrie et les indicateurs relatifs aux doctorats (qui caractérisent aussi la recherche), le poids total de la recherche s'élève à 73,2 %.

Il convient également de noter que tous les indicateurs de production dans ce classement sont relatifs (par enseignant-chercheur, par publication, etc.) et que, par conséquent, le score des universités dans le classement n'est pas dépendant de leur taille.

1.3 World's Best Universities Ranking – US News & World Report en coopération avec Quacquarelli Symonds (QS)

Au début de l'année 2010, *US News & World Report* (USNWR) a commencé à coopérer avec QS et a publié en ligne son nouveau classement 2009 le 25 février 2010. Cette publication a pris la forme d'un rapport basé sur les mêmes résultats compilés par QS qui avaient été publiés en 2009 sur le site Internet du THE-QS World Universities Ranking ainsi que sur le site de QS. La différence entre les trois est que le classement USNWR-QS publie une liste des 400 premières universités, tandis que le THE publie un Top 200 et que QS publie une liste de plus de 500 universités. La description de

la méthodologie du *World's Best Universities Ranking* 2009 de USNWR est identique à celle du classement THE-QS 2009, à l'exception de quelques changements mineurs dans le texte, qui consistent essentiellement à remplacer « THE » par « US News World's Best Universities » (voir les phrases « The academic peer review is the centrepiece of the U.S. News World's Best Universities rankings [...] » en 2010 et « The Academic peer Review is the centrepiece of the Times Higher – QS world University Rankings [...] » sur le site Internet de QS l'année précédente).

Il a donc été décidé que, s'agissant de la méthodologie sur laquelle se fonde le classement *US News & World Report / QS World's Best Universities Ranking*, il suffit de se reporter au chapitre sur le classement 2009 du THE-QS.

1.4 Classement mondial des universités - Reitor (Рейтор)

Le classement mondial des universités Reitor est réalisé par une agence de classement³², Reitor (ou Reiter, en russe « Рейтор »), située à Moscou. L'« idéologie » du classement émane toutefois de Reitor et de l'Université d'État de Moscou Lomonossov (STRF, 2008 ; Doneckaja, 2009). Le seul classement mondial des universités réalisé jusqu'à présent par Reitor a été établi en 2008, et les résultats ont été publiés en février 2009. En dépit d'une intention affichée d'en faire un classement périodique, aucune information supplémentaire n'a été communiquée à cet égard.

L'objectif affiché du classement est de répondre aux besoins des milieux universitaires russes, qui manifestent un intérêt croissant pour l'évaluation internationale des universités afin de pouvoir les situer dans l'enseignement supérieur mondial. Il fournit également un instrument d'évaluation de la compétitivité de l'enseignement supérieur professionnel en Russie. Le classement mondial des universités de Reitor a été inclus dans le présent rapport pour deux raisons. D'une part, il a été dit que les méthodologies des classements mondiaux les plus réputés ne reflètent pas la réalité des universités de la Fédération de Russie, ce qui est probablement la première raison de l'établissement du classement Reitor. D'autre part, sur la base d'une première analyse, les indicateurs Reitor (tout au moins ceux qui ont été présentés avant la publication des résultats du classement) présentaient des différences intéressantes par rapport aux indicateurs utilisés dans les autres classements mondiaux.

Quelles sont les universités prises en considération ?

Selon l'objectif affiché, les producteurs du classement ont invité les groupes d'universités suivants à participer au classement :

1. Toutes les universités du monde figurant dans les quatre principaux classements mondiaux (établis par l'Université Jiao Tong de Shanghai (ARWU-500), *The Times* (QS-THES-200),

³² Cette agence est qualifiée d'indépendante sur son site Internet. Comme plusieurs sources indiquent que « l'idéologie » du classement est fortement influencée par l'Université d'État de Moscou Lomonossov, nous nous abstenons d'utiliser le mot « indépendant ».

Webometrics et l'Université nationale de Taïwan ³³ ;

2. Les écoles supérieures les plus réputées de la Fédération de Russie ;
3. Les établissements des pays de l'ex-URSS.

Toutes les universités souhaitant faire partie du classement ont par ailleurs été invitées à remplir et à renvoyer le questionnaire.

Critères retenus, indicateurs utilisés et indicateurs indirects

La **méthodologie Reitor, telle que décrite** dans la section relative à la méthodologie sur le site Internet (voir Reitor – méthodologie), montre que les producteurs du classement Reitor prévoyaient d'utiliser un certain nombre d'indicateurs que la plupart des classements les plus réputés n'utilisent généralement pas, par exemple :

- Le nombre de programmes d'étude par niveau (licence, spécialisation) ;
 - Le nombre d'étudiants lauréats de concours universitaires internationaux ;
 - Le nombre de publications du personnel universitaire autres que des articles dans des revues scientifiques, à savoir les monographies, les manuels et autres ouvrages ;
 - Le nombre de « certificats de découverte » et de brevets d'invention obtenus par l'université et son personnel ;
 - La valeur totale des installations de formation et de laboratoire des universités exprimée en dollars US ;
 - La proportion du personnel enseignant titulaire d'un doctorat ;
 - Le nombre de membres du personnel qui ont obtenu le titre de docteur honoris causa auprès d'universités à l'étranger ;
 - Le nombre de professeurs membres d'académies nationales et internationales des sciences à compter de l'année précédente ;
 - Les caractéristiques des activités « socialement significatives » des diplômés (ex. figures éminentes dans le domaine de la culture, des affaires, de la politique ; cadres supérieurs de l'administration publique ou d'organisations internationales) ;
 - Le nombre de publications par les enseignants-chercheurs (de différents types : articles, manuels, monographies et autres ouvrages) ;
 - Outre les proportions d'étudiants internationaux et de personnel international, d'autres indicateurs du degré d'internationalisation des universités ; en particulier, des indicateurs relatifs à la mobilité à court terme du personnel et des étudiants, en termes d'arrivées et de départs (Reitor, 2008a).
- Outre les indicateurs quantitatifs, l'utilisation des résultats d'une enquête de réputation était également prévue. Cette enquête faisait partie du formulaire général d'enquête envoyé aux universités participant au classement. Ces universités ont donc été invitées à sélectionner les 10 universités étrangères « les plus importantes en termes de qualité de l'éducation et

de la formation des cadres supérieurs » (Reitor, 2008b). Les indicateurs et la méthodologie de l'enquête sont également décrits dans un article publié par les producteurs du classement Reitor (Kruzhalinin et Artjushina, 2008), qui précise que plusieurs experts de la même université peuvent sélectionner les 10 meilleures universités étrangères, mais que leurs avis seront regroupés en un seul point de vue exprimé par l'université. Il est souligné que la sélection d'universités étrangères uniquement « permet d'exclure la composante « patriotique » de l'évaluation et d'élargir la géographie de l'université [sic] » (Reitor, 2008a ; Kruzhalinin et Artjushina, 2008).

Quelle est la méthodologie effectivement utilisée ? Vladimir Moskovkin (2009) attire l'attention sur les fortes divergences entre les sections « Méthodologie » et « À propos du classement » et relève que des informations importantes sur les indicateurs ne sont pas présentées dans les pages « Méthodologie », mais plutôt sur la page « À propos du classement ». Une analyse plus approfondie de la description de la méthodologie débouche sur un résultat étonnant : nombre d'indicateurs mentionnés dans la section « Méthodologie » du site Internet de Reitor et dans l'enquête prévue auprès des experts n'ont en réalité pas été utilisés. Ce sont au contraire des indicateurs bibliométriques qui ont été utilisés. La section « Méthodologie » du site Internet de Reitor demeure toutefois inchangée. Ce changement de méthode peut éventuellement s'expliquer par le nombre d'universités participantes. Bien qu'un large éventail d'universités ait été invité à participer et à remplir le questionnaire du classement, seules 63 universités, dont seulement neuf en dehors de la CEI, y ont répondu. On peut effectivement confirmer que les producteurs du classement n'ont adressé leurs remerciements pour avoir répondu au questionnaire qu'à 63 universités (Reitor, 2009b). Nombre d'indicateurs sélectionnés au départ supposent cependant la contribution de toutes les universités, et non pas d'un nombre limité d'entre elles. Ceci peut expliquer pourquoi les producteurs du classement ont subitement adopté des indicateurs différents, basés sur les données disponibles à l'échelle internationale.

Une analyse approfondie du site Internet de Reitor révèle deux sources intéressantes, qui permettent de mieux comprendre la méthodologie effectivement utilisée. La description suivante de la méthodologie s'appuie sur la section « À propos du classement » du site de Reitor (Reitor, 2009a) et sur la présentation des résultats du classement à l'Université d'État de Moscou Lomonossov (Reitor, 2009c), qui sont également citées par Vladimir Moskovkin (ex. Moskovkin, 2009). Il convient de noter que certaines décisions importantes sur le calcul des résultats du classement ont été laissées à la discrétion du groupe d'experts de Reitor. Selon Reitor, le groupe d'experts détermine :

- « Les échelles d'évaluation pour chaque indicateur de performance ;
- Le poids de chaque indicateur dans une catégorie [d'indicateurs] ;
- La pondération du score de chaque catégorie [dans le score total] » (Reitor, 2008a).

Si la pondération de chaque catégorie d'indicateurs dans le score total peut être consultée dans la section « À pro-

³³ Il s'agit probablement ici du *Performance Ranking of Scientific Papers for World Universities* du Conseil d'accréditation et d'évaluation de l'enseignement supérieur de Taïwan.

pos du classement », le poids des différents indicateurs au sein d'une catégorie ne peut être connu via des sources publiques. Il n'est donc pas possible de suivre le calcul des scores cumulés.

La méthodologie effectivement utilisée. Le tableau 6 (Reitor

2009c) résume cette méthodologie. Elle diffère considérablement de la liste d'indicateurs présentée sur le site Internet de Reitor dans la section « Méthodologie », comme mentionné plus haut.

Tableau 6 : Pondération des catégories d'indicateurs et liste des indicateurs effectivement utilisés

Catégories d'indicateurs	Poids	Indicateurs
I. Activité d'enseignement	20%	1. Nombre de programmes de licence, de spécialisation ³⁴ , de master et de doctorat. 2. Taux d'encadrement des étudiants (ratio étudiants/enseignants).
II. Activité de recherche	20%	3. Nombre de « certificats de découverte » et de brevets depuis 2001. 4. Performance du centre informatique de l'université. 5. Indice h de l'université.
III. Capacité financière (« financial maintenance »)	15%	6. Budget total de l'université par étudiant à temps plein.
IV. Compétence professionnelle des enseignants-chercheurs	20%	7. Nombre de membres du personnel ayant obtenu un prix de niveau international (Prix Nobel, médaille Fields et autres, comme le Prix Descartes, le Prix Abel, la médaille Lomonossov, le prix russe « Global Energy »). 8. Nombre de publications des enseignants-chercheurs. 9. Citations et références aux publications des enseignants-chercheurs.
V. Activité internationale	10%	10. Communautés académiques internationales auxquelles a été associée l'université au cours de l'année universitaire précédente. 11. Proportion d'étudiants étrangers l'année précédente.
VI. VI. Présence sur Internet	15%	12. Volume de produits en ligne. 13. Succès de l'université selon les requêtes. 14. « Page Rank » de la page d'accueil du site de l'université.

Source : informations tirées de la présentation des résultats du classement à l'Université d'État de Moscou Lomonossov (Reitor, 2009c).

Précisions complémentaires à propos des indicateurs :

L'*activité d'enseignement* d'une université est caractérisée par quatre indicateurs directs :

- Le nombre de programmes proposés par l'université dans chacun des trois cycles, qui représente trois indicateurs ;
- Le taux d'encadrement des étudiants (ratio étudiants/enseignants) calculé en ETP.

La *performance de la recherche* est mesurée par :

- Le nombre de brevets et de découvertes enregistrées depuis 2001, sur la base des données fournies par Scopus ;
- La performance du centre informatique de l'université, comme indicateur de la capacité de recherche de l'université. Les données sont tirées du Top 500 des superordinateurs³⁵ ;
- L'indice h de tous les auteurs d'une université, sur la base des données de Scopus.

La *compétence professionnelle des enseignants-chercheurs* est estimée selon :

- Le nombre de lauréats du Prix Nobel, de la médaille Fields et autres récompenses, comme le Prix Descartes³⁶, le Prix Abel³⁷, la médaille Lomonossov de l'Académie russe des sciences et le prix russe « Global Energy » depuis 2001 ;
- Le nombre de publications du personnel universitaire dans les revues référencées depuis 2001 ;
- Les citations et références aux publications des enseignants-chercheurs depuis 2001.

La *capacité financière* (« financial maintenance ») représente un poids de 15 %. La valeur de cet indicateur correspond au budget total de l'université par étudiant à temps plein pour l'année précédente.

³⁴ Les « programmes de spécialisation » dans la Fédération de Russie et d'autres pays de la CEI sont l'équivalent des programmes d'études de 5-6 ans qui existaient en Europe continentale avant la mise en œuvre du processus de Bologne.

³⁵ <http://www.top500.org/>.

³⁶ Le Prix Descartes est une récompense décernée chaque année par l'Union européenne depuis 2000, ainsi nommé en l'honneur du mathématicien et philosophe français René Descartes. Ce prix de la recherche est attribué à des équipes de chercheurs qui ont « atteint des résultats scientifiques ou technologiques exceptionnels dans le cadre d'un projet de recherche collaborative dans toute discipline scientifique, y compris les sciences économiques, humaines et sociales ».

³⁷ Le Prix Abel, décerné pour la première fois en 2003, est doté d'un montant de 6 millions NOK (environ 750 000 euros ou 1 million de dollars US). Ce prix international récompense les travaux scientifiques remarquables dans le domaine des mathématiques, y compris les aspects mathématiques de l'informatique, la physique mathématique, la probabilité, l'analyse numérique, l'informatique scientifique, la statistique.

L'activité internationale d'une université représente un poids de 10 % et est évaluée sur la base de deux indicateurs définis comme suit :

- « L'activité des universités au sein des communautés académiques internationales. En raison du nombre [élevé] de ces communautés dans le monde, cette activité est évaluée par un groupe d'experts » (Reitor, À propos du classement).
- Le nombre total d'étudiants étrangers rapporté au nombre total d'étudiants de l'université.

La présence sur Internet. Trois indicateurs sont utilisés :

- Le volume de produits en ligne est inclus car « il décrit le développement des technologies de l'internet dans les universités quant à la production et à la promotion des publications scientifiques et méthodiques » (Reitor, À propos du classement) ;
- La « popularité » de l'université, en nombre de requêtes reçues sur le site Internet de l'université en 2008 ;
- Le « Page Rank » de Google de la page d'accueil principale du site Internet de chaque université à l'automne 2008.

Sources de données (D'après Reitor 2008b)

Compte tenu du nombre limité d'universités ayant répondu complètement au questionnaire d'enquête – environ 10 % des universités auxquelles il a été envoyé – il est probable que le reste des données ait été tiré des sites Internet des universités ou recueilli auprès d'organismes regroupant des données statistiques sur l'éducation. Les recherches effectuées sur les sites Internet des universités sont par ailleurs mentionnées comme source supplémentaire de données par les producteurs du classement eux-mêmes (Reitor, 2009a) :

« Les responsables de l'évaluation ont réalisé le travail de recherche, d'édition et de mise à l'échelle de données publiques, sur la base des sources d'information suivantes :

- Sites officiels des universités ;
- Rapports annuels des universités ;
- Agences nationales spécialisées dans la collecte et le traitement de données statistiques sur l'éducation ;
- Différentes évaluations susceptibles de comprendre des universités étudiées (ex. le Top 500 mondial des superordinateurs) ;
- La base de données scientométrique Scopus ;
- Les données du moteur de recherche Google ».

Le calcul des valeurs des indicateurs et leur conversion en scores

La description présentée dans la section « Méthodologie » du site Internet de Reitor indique que chaque expert évalue chaque indicateur pour toutes les universités. Le score total pour chaque indicateur particulier correspond à la moyenne de l'ensemble des résultats des évaluations réalisées par les experts. En cas d'écart de plus de 15 % entre les résultats des évaluations réalisées par les experts, ceux-ci examinent la question collégialement afin de d'établir la valeur finale sur la base d'un consensus (Reitor, « Méthodologie »)³⁸. Il est malaisé de déterminer si cette procédure a été appliquée ou non.

Évolution de la méthodologie utilisée

Le classement a été publié pour la première fois en 2009. Aucune suite n'y a été donnée depuis lors. Il convient toutefois de souligner que les descriptions de la méthodologie sur le site Internet de Reitor sont contradictoires. Il existe trois versions différentes – la première est présentée dans la section « Méthodologie », la deuxième sur la page « À propos du classement » et une troisième fournie lors de l'exposé présenté à l'Université d'État de Moscou Lomonossov.

Présentation du classement et analyses complémentaires

Outre le palmarès principal, un classement recense les pays selon le nombre d'universités dans le Top 500 et un autre concerne les universités de la Fédération de Russie et de la Communauté d'États indépendants (CEI).

³⁸ http://www.globaluniversitiesranking.org/index.php?option=com_content&view=article&id=96&Itemid=54

2. Classements axés uniquement sur la performance de la recherche (avec ou sans palmarès)

2.1 Leiden Ranking – Université de Leiden

Le producteur du classement de Leiden est le Centre d'études scientifiques et technologiques (*Centre for Science and Technology Studies* – CWTS) de l'Université de Leiden, qui a élaboré ses propres indicateurs bibliographiques. Les résultats du classement de Leiden ont été publiés pour la première fois en 2008. Des résultats actualisés ont été publiés sur Internet le 2 août 2010.

L'objectif affiché du classement de Leiden est le suivant : « Le classement de Leiden vise à établir une comparaison des institutions de recherche selon des mesures d'impact qui tiennent compte des différences entre les disciplines ».

Il ne présente pas de score global sur la base d'un indice composite, mais des scores selon différents indicateurs (quatre en 2008 ; cinq en 2010). Chacun de ces indicateurs est appelé un « classement » (*ranking*) et présenté séparément. Cependant, si les universités sont classées séparément en fonction de chaque indicateur, leurs résultats au regard des autres indicateurs sont également présentés.

Quelles sont les universités prises en considération ?

Le classement de Leiden porte sur les universités qui comptent le plus grand nombre de publications par an indexées dans *Web of Science*. En 2008, le classement recensait environ 1 000 universités comptant le plus grand nombre de publications (au niveau mondial), et les listes du Top 250 ont été publiées. En 2010, les classements d'universités ont été présentés sous la forme de listes du Top 100, Top 250 et Top 500.

Critères retenus et indicateurs utilisés

Le classement de Leiden est axé uniquement sur la recherche, et plus particulièrement sur la production scientifique. Les cinq indicateurs utilisés dans le classement 2010 sont les suivants :

1. Nombre de publications (P) (indicateur « jaune » selon le CWTS). Cet indicateur mesure le nombre de publications dans les revues scientifiques référencées par des index de citations (*Web of Science*, *Scopus*) au cours d'une certaine période. Le classement de Leiden 2010 porte sur les publications de la période 2003-2008 et prend en compte les citations sur la période 2004-2009.

Cet indicateur mesure la production scientifique de l'ensemble de l'université. Il est par conséquent dépendant de la taille des établissements.

2. Nombre de citations par publication (CPP) – (indicateur « bleu » selon le CWTS). La valeur de l'indicateur CPP est calculée en excluant les auto-citations. Les résultats de la soustraction du nombre d'auto-citations C_s du nombre total de citations pour chaque article C_i sont additionnés. Cette somme est ensuite divisée par le nombre total d'articles P :

$$CPP = \frac{1}{P} \sum_{i=1}^P (C_i - C_s)$$

Sachant que :

CPP est le nombre de citations par article ;
 P est le nombre de publications (articles) ;
 C_i est le nombre total de citations pour l'article ;
 C_s est le nombre d'auto-citations pour l'article.

Le CPP permet d'estimer l'impact scientifique moyen de l'université. L'inconvénient du CPP utilisé comme indicateur est qu'il ne tient pas compte du fait que les articles plus anciens recueillent généralement un plus grand nombre de citations, et ignore les différences entre taux de citations selon le type de publication (articles, actes, etc) et le champ disciplinaire (Rehn et al., 2007).

L'indicateur CPP est par ailleurs exprimé par un nombre relatif, qui par conséquent ne reflète pas à lui seul la performance globale d'une université.

3. Nombre de citations par publication normalisé par champ disciplinaire (CPP/FCSm) – (indicateur « vert clair » selon le CWTS). L'indicateur CPP/FCSm a été introduit en 1993, bien que le mécanisme de normalisation remonte aux années 1980 (Waltman et al., 2010). Les producteurs du classement de Leiden qualifient eux-mêmes cet indicateur de « crown indicator » dans le classement 2008, s'agissant d'un indicateur indépendant de la taille de l'établissement et normalisé par champ disciplinaire.

La valeur de cet indicateur est établie en divisant le résultat de l'indicateur CPP (bleu) précédent (voir plus haut) par le nombre moyen de citations par champ disciplinaire (FCSm).

$$\frac{CPP}{FCSm} = \frac{\sum_{i=1}^P \frac{C_i}{P}}{\sum_{i=1}^P \frac{e_i}{P}} = \frac{\sum_{i=1}^P C_i}{\sum_{i=1}^P e_i} = \frac{C_1 + C_2 + C_3 \dots}{e_1 + e_2 + e_3 \dots}$$

Sachant que :

C_i est le nombre de citations de l'article i ;
 e_i est le nombre attendu de citations de l'article i selon le champ disciplinaire et l'année de publication de l'article i ;
 P est le nombre de publications (articles).

La normalisation des valeurs des citations s'opère à partir de la somme des citations et des citations par champ disciplinaire. Cette méthode présente toutefois certains inconvénients (voir Leydesdorf et Opthof, 2010). Il a déjà été noté que les taux de citation ne sont pas ici normalisés au niveau de chaque publication, mais à un niveau d'agrégation supérieur. Cette méthode de calcul renforce le poids des publications plus anciennes, en particulier des critiques (« reviews »), et les publications dans des champs disciplinaires où les taux de citations sont généralement plus élevés (Rehn et al., 2007).

4. « Mean normalised citation score » (MNCS2) – (indicateur « vert foncé » selon le CWTS). En 2010, le CWTS a élaboré un nouvel indicateur pour remplacer le « crown indicator » utilisé précédemment. Le « mean normalised citation score » (MNCS) avait déjà été introduit en 1990. Il est calculé comme suit :

$$MNCS = \frac{1}{P} \sum_{i=1}^P \frac{C_i}{e_i} = \frac{1}{P} * \left(\frac{C_1}{e_1} + \frac{C_2}{e_2} + \frac{C_3}{e_3} + \dots \right)$$

La comparaison du calcul du CPP/FCSm (voir la description de l'indicateur précédent) et du MNCS révèle que l'indicateur CPP/FCSm est normalisé selon le calcul d'un rapport de moyennes, tandis que l'indicateur MNCS est normalisé selon le calcul d'une moyenne de rapports (Waltman et al., 2010). La différence entre les deux est que, dans le cas du CPP/FCSm, les citations (au numérateur) sont d'abord additionnées, comme l'est le nombre attendu de citations de la publication e_i au dénominateur. Ensuite, la première somme est divisée par la seconde. Les publications plus récentes étant moins citées, elles ont peu d'influence sur la valeur des sommes, tant au numérateur qu'au dénominateur. L'impact des nouvelles publications est par conséquent très faible.

Dans le cas de l'indicateur MNCS, le nombre de citations d'une publication est divisé individuellement par le nombre attendu de futures citations. Par conséquent, le nombre de citations pour les nouvelles publications est faible, comme l'est également le nombre attendu de citations. C'est pourquoi l'impact des publications récentes ne se distingue plus de celui des publications plus anciennes. Le calcul du MNCS présente néanmoins un nouvel inconvénient (Leydesdorf et Opthof, 2010). En effet, il est difficile de prévoir le nombre de citations futures de publications très récentes, ce qui rend les résultats moins stables.

Enfin, l'indicateur MNCS (ou MNCS1) a été modifié. Il est devenu l'indicateur MNCS2, qui, à la différence du MNCS1, exclue les publications récentes, c'est-à-dire celles qui ont eu moins d'un an pour être citées. L'analyse empirique de Waltman et al. (Waltman et al., 2010) démontre que l'indicateur MNCS2 peut effectivement remplacer le précédent « crown indicator » CPP/FCSm.

L'indicateur CPP/FCSm et les deux indicateurs MNCS1 et MNCS2 reflètent « l'impact » des publications. Par exemple, si la valeur de l'un de ces indicateurs pour une université X est de 1,57, cela signifie que les publications de l'université X recueillent en moyenne 1,57 fois plus de citations que n'en

recueillent habituellement les publications dans ce domaine particulier au cours de la même année. Ou encore, si la valeur de l'un de ces indicateurs pour une université Y est de 0,4, cela signifie que « l'impact » de ses publications est moindre et que, par conséquent, pour recueillir le même nombre de citations, l'université Y doit publier 2,5 fois plus d'articles que les autres.

5. L'indicateur de « force brute » P*CPP/FCSm – (indicateur « orange »). Afin de refléter la performance réelle d'une équipe de recherche ou d'une université dans le monde, le nombre de citations par publication normalisé par champ disciplinaire est multiplié par le nombre total de publications.

$$P * \frac{CPP}{FCSm} = P * \frac{\sum_{i=1}^P C_i}{\sum_{i=1}^P e_i}$$

Par conséquent, l'indicateur de « force brute » dans le classement de Leiden représente dans un sens le nombre total de publications, corrigé de façon à tenir compte de leur « impact ».

Sources de données

Seules sont utilisées des données bibliométriques tirées de bases de données de citations.

Le calcul des valeurs des indicateurs et leur conversion en scores

Les résultats des différents indicateurs ne sont pas convertis en une « valeur finale », et aucun classement global n'est établi.

Évolution de la méthodologie utilisée

Par rapport à 2008, un nouvel indicateur a été ajouté en 2010 : le mean normalised citation score.

Présentation du classement et analyses complémentaires

Les listes du Top 100 et du Top 250 (celle-ci ajoutée en 2010) sont présentées tant à l'échelle mondiale qu'européenne. Une liste du Top 500 à l'échelle mondiale a également été établie en 2010.

L'équipe du classement de Leiden a participé à l'élaboration d'indicateurs et de méthodes de normalisation. Elle a également calculé les scores de certains indicateurs pour d'autres classements, y compris des projets de classement soutenus par l'UE, comme le classement d'excellence du CHE et U-Multirank.

L'équipe du classement de Leiden a récemment comparé l'impact des publications (articles et autres formes de publications) en anglais à l'impact de l'ensemble des publications des 69 universités allemandes et françaises les plus fréquemment citées, et leur incidence sur la position de ces universités dans le classement.

Les résultats montrent que la fréquence des citations pour les publications en anglais est systématiquement plus élevée que la fréquence des citations pour l'ensemble de leurs publications³⁹. Par conséquent, la position de ces universités dans le classement, sur la base de l'impact en citations de leurs publications en anglais, est systématiquement meilleure que leur position basée sur les résultats de l'impact en citations de l'ensemble de leurs publications (van Raan et al., 2010).

Il convient d'ajouter que cet effet est probablement encore plus prononcé pour des langues autres que le français et l'allemand, comme par exemple le russe, l'espagnol, le chinois ou le japonais.

2.2 Performance Rankings of Scientific Papers for World Universities – Classement du Higher Education Evaluation & Accreditation Council of Taiwan

Le classement du Conseil d'accréditation et d'évaluation de l'enseignement supérieur de Taïwan (*Higher Education Evaluation & Accreditation Council of Taiwan*) (ci-après « le classement HEEACT ») est un classement mondial établi chaque année depuis 2007. Le HEEACT évalue la performance des universités selon le critère de la publication d'articles scientifiques et établit un classement mondial des 500 premières universités, sur la base des données indexées dans le SCI et le SSCI.

Les producteurs du classement soulignent qu'ils mettent l'accent sur les résultats de la recherche et que c'est en cela que le classement HEEACT se distingue du classement THE « qui vise essentiellement à classer les établissements, et [du] classement académique produit par ARWU » (Huang, 2010).

Le classement repose sur huit indicateurs regroupés dans trois catégories principales : la productivité de la recherche (20 % du poids total), l'impact de la recherche (30 %) et l'excellence de la recherche (50 %).

Le classement HEEACT tient également compte des fusions d'universités et des instituts universitaires spécialisés ou encore de différents campus dans un système universitaire. Il inclut par conséquent les publications d'organismes liés à une université, comme les centres de recherche et les hôpitaux universitaires.

Quelles sont les universités prises en considération ?

La sélection des universités est basée sur le nombre d'articles et de citations dans des revues scientifiques. Pour établir une liste du Top 500, 700 établissements sont d'abord sélectionnés sur les 4 000 institutions de recherche recensées par Essential Science Indicators (ESI). Les institutions qui ne sont pas des universités sont retirées et les établissements restants sont comparés aux listes des classements THE, ARWU et US News and World Report. Cela aboutit à une liste de 725 universités.

Critères retenus et indicateurs utilisés

Quatre des huit indicateurs sont calculés par membre du personnel (ETP), dans un souci des producteurs du classement d'atténuer le biais induit par la taille. L'accent est mis sur l'impact plutôt que sur la performance. L'ensemble des indicateurs d'impact représente 80 % du score total d'une université. Des tentatives ont été faites pour refléter les succès récents des universités. Le score total repose notamment à hauteur de 55 % sur des indicateurs basés sur des données de l'année ou des deux années précédentes (Huang, 2009).

La catégorie relative à la **productivité de la recherche** comporte deux indicateurs :

- Le *nombre d'articles* publiés au cours des 11 dernières années dans des revues scientifiques évaluées par les pairs [par membre du personnel – ETP]⁴⁰ ;
- Le *nombre d'articles* publiés au cours de l'année précédente [par membre du personnel – ETP].

Le poids total de la catégorie relative à la productivité de la recherche est relativement faible et s'élève à 20 %.

La catégorie relative à l'**impact de la recherche** comporte trois indicateurs et son poids total est de 30 % :

- Le *nombre de citations au cours des 11 dernières années* correspond au nombre total de citations des articles de l'université en question au cours des 11 années précédentes rapporté au nombre de membres du personnel exprimé en ETP ;
- Le *nombre de citations au cours des 2 dernières années* correspond au nombre total de citations recensées par SCI et SSCI par membre du personnel en ETP ;
- Le *nombre moyen de citations au cours des 11 dernières années* correspond au nombre total de citations d'une université au cours des 11 années précédentes rapporté au nombre total de publications de l'université au cours de la même période.

La catégorie relative à l'**excellence de la recherche** comporte trois indicateurs, qui représentent 50 % du score final :

- L'*indice h des deux dernières années*, où la valeur h est le nombre d'articles publiés par une université au cours des deux années précédentes, cités pas moins de h fois ; Cet indicateur représente 20 % du score total.

³⁹ C'est-à-dire les publications des universités françaises et allemandes dans toutes les langues y compris en français et en allemand, respectivement.

⁴⁰ Le tableau des indicateurs sur le site Internet du classement HEEACT prête à confusion, en l'absence de précision quant aux indicateurs calculés en ETP ou non.

- Le *nombre d'articles fréquemment cités* correspond au nombre absolu d'articles de l'université en question qui figurent dans le top 1 % des articles les plus souvent cités recensés par ESI et publiés au cours des 11 années précédentes ;

- Le *nombre d'articles dans des journaux à fort facteur d'impact au cours de l'année précédente* correspond au nombre absolu de publications de l'université en question au cours de l'année précédente dans le top 5 % des revues classées selon leur facteur d'impact.

Tableau 7 : Critères et indicateurs du classement du Higher Education Evaluation and Accreditation Council of Taiwan

Critères / Poids	Indicateurs de performance globale	Poids
Productivité de la recherche (20 %)	Nombre d'articles publiés au cours des 11 dernières années ⁴¹ [par membre du personnel – ETP].	10%
	Nombre d'articles publiés au cours de l'année précédente ⁴² [par membre du personnel – ETP].	10%
Impact de la recherche (30%)	Nombre de citations au cours des 11 dernières années [par membre du personnel – ETP].	10%
	Nombre de citations au cours des 2 dernières années ⁴³ [par membre du personnel – ETP].	10%
	Nombre moyen de citations [par publication] des 11 dernières années.	10%
Excellence de la recherche (50%)	Indice-h des 2 dernières années.	20%
	Nombre d'articles fréquemment cités au cours des 11 dernières années.	15%
	Nombre d'articles publiés dans des revues à grand impact au cours de l'année précédente	15%

Les données utilisées pour évaluer la performance des universités ont été fournies par *Thomson Reuters Web of Science*, y compris *Essential Science Indicators* (ESI), le *Science Citation Index* (SCI), le *Social Sciences Citation Index* (SSCI) et le *Journal Citation Reports* (JCR).

Bien que le *Arts & Humanities Citation Index* (A&HCI) soit également disponible, le classement HEEACT ne se fonde pas sur cette base de données, car « elle pourrait ne pas représenter objectivement et précisément la performance des chercheurs dans le domaine des arts et des sciences humaines ». La base de données indexe essentiellement des articles publiés en anglais, tandis que les chercheurs dans les domaines des arts et des sciences humaines publient souvent dans leur langue maternelle et sous diverses formes (ex. livres). Selon les producteurs du classement, privilégier les données indexées dans le SCI et le SSCI permet d'établir des comparaisons plus objectives entre les universités à l'échelle mondiale.

Le *nombre d'enseignants-chercheurs* (ETP) est obtenu auprès de QS, sur les sites Internet des universités elles-mêmes ou par l'intermédiaire de l'administration de l'enseignement supérieur des différents pays.

Le *nombre de citations au cours des 11 dernières années* et le *nombre d'articles fréquemment cités* sont fournis par ESI, tandis que le *nombre de citations au cours des 2 dernières années* et l'*indice h des 2 dernières années* sont établis sur la base des données indexées par le SCI et le SSCI (WoS), qui disposent de statistiques récentes. Le *nombre d'articles publiés dans des revues à fort facteur d'impact au cours de l'année précédente* est basé sur les données du JCR.

Calcul des scores finaux

Scores par indicateur. À l'instar d'autres classements, le score de chaque indicateur est calculé en divisant la valeur de l'indicateur de l'université en question par celle de l'université ayant obtenu le meilleur résultat. Le résultat est ensuite multiplié par 100 ⁴⁴. Calculés de la sorte, les scores des différents indicateurs sont des valeurs relatives sans dimension, qui peuvent par conséquent être combinées.

Score global. Le score global est calculé en multipliant le score de chaque indicateur par son poids et en additionnant l'ensemble des résultats.

Évolution de la méthodologie utilisée

Le premier classement HEEACT de 2007 reposait sur une série de neuf indicateurs. L'indice h représentait un poids de 20 %, tandis que chacun des huit autres indicateurs contribuait à hauteur de 10 % au score final. À partir de 2008, l'indicateur relatif au *nombre de domaines dans lesquels l'université fait preuve d'excellence* a été supprimé, et son poids a été attribué en parts égales aux indicateurs relatifs au *nombre d'articles fréquemment cités* et au *nombre d'articles publiés dans des revues à fort facteur d'impact*.

Le score de l'indicateur (aujourd'hui supprimé) relatif au *nombre de domaines dans lesquels l'université fait preuve d'excellence* correspond au nombre de disciplines (parmi les 21 disciplines listées par Thomson Reuters Web of Knowledge) pour lesquelles l'université en question est recensée

⁴¹ Soit 1999-2009 pour le classement 2010 et en conséquence pour les classements des années antérieures.

⁴² Soit 2009 pour le classement 2010 et en conséquence pour les classements des années antérieures.

⁴³ Soit 2008-2009 pour le classement 2010 et en conséquence pour les classements des années antérieures.

⁴⁴ Pour plus de précisions, le lecteur se reportera utilement au chapitre « Le calcul des valeurs des indicateurs et leur conversion en scores » relatif au classement ARWU.

dans la base de données de citations de l'ESI (seul le top 1 % des établissements le plus souvent cités est inclus dans ces listes). Cet indicateur favorisait des universités plus petites ou spécialisées faisant preuve d'excellence dans des domaines spécifiques et augmentait leur visibilité. Sa suppression a, dans une certaine mesure, affaibli la position de ces établissements. En réalité, les autres indicateurs de la catégorie relative à l'excellence de la recherche, qui « aggrègent » l'excellence dans différents domaines, ne tiennent généralement pas compte des établissements petits ou spécialisés.

Présentation du classement et analyses complémentaires

Les résultats du classement HEEACT sont présentés sous la forme d'une liste d'universités du Top 500, déclinée en tableaux 1-100, 101-200, 201-300, 301-400 et 401-500. Les utilisateurs peuvent consulter ces listes par score total ou selon le score pour chaque indicateur. Des classements par continent et par pays sont également présentés.

Le HEEACT publie également les résultats des classements dans les six champs disciplinaires suivants :

1. Agriculture et sciences de l'environnement ;
2. Ingénierie, informatique et technologie ;
3. Médecine clinique ;
4. Sciences de la vie ;
5. Sciences naturelles ;
6. Sciences sociales.

Les résultats des classements par champ disciplinaire sont présentés sous la forme de listes du Top 100 par champ, de listes par continent et par pays. Les classements par champ reposent sur les mêmes indicateurs que le classement mondial des universités. Les classements par champ s'en distinguent toutefois par des échantillons d'universités différents. Dans chaque champ disciplinaire, 400 universités sont sélectionnées dans la base de données de l'ESI selon le nombre d'articles publiés et le nombre de citations recensées au cours des 11 dernières années. Seuls sont pris en compte les établissements qui délivrent des diplômes dans les trois cycles d'étude. Les listes initiales sont complétées par des comparaisons avec les listes des classements ARWU et THE. Elles comprennent en définitive entre 460 et un peu plus de 500 universités, selon le champ spécifique.

2.3 Assessment of University-Based research – Commission européenne

En 2008, la Direction générale de la Recherche de la Commission européenne a constitué un groupe de travail sur l'évaluation de la recherche universitaire (Assessment of University-Based Research) (ci-après le « groupe de travail AUBR »). Le rapport final du groupe (AUBR, 2010) a été publié au début de l'année 2010.

La mission de ce groupe de travail comprenait les objectifs suivants :

1. « Examiner les besoins des différents types d'utilisateurs de la mesure de la qualité de la recherche dans les universités ;
2. Examiner les principales méthodologies d'évaluation/de classement de la qualité de la recherche des universités, notamment les évaluations/classements internationaux existants et d'autres méthodologies en cours de développement ;
3. Proposer dans la mesure du possible une approche méthodologique multidimensionnelle consolidée répondant aux besoins des différents types d'utilisateurs et définir les données et indicateurs nécessaires (le cas échéant proposer différentes approches pour différents types d'utilisateurs » (Ibid.).

Par conséquent, la mission du groupe de travail AUBR ne consistait pas à créer un classement, mais à analyser les indicateurs relatifs à la recherche utilisés dans les classements d'universités existants et à proposer une méthodologie pour une évaluation multidimensionnelle de la recherche universitaire.

Analyse des indicateurs utilisés pour l'évaluation de la recherche

Le groupe de travail a notamment établi une liste d'indicateurs pouvant être utilisés pour évaluer la recherche universitaire et tenté d'identifier les points forts et les faiblesses de ces différents indicateurs (AUBR, 2010, pp. 44-48 et 69-82). Le groupe de travail AUBR a par ailleurs identifié un certain nombre de mesures qui devraient être prises préalablement à l'utilisation de chaque indicateur. Par exemple, avant d'appliquer l'indicateur relatif à la production scientifique par chercheur, il convient de s'accorder sur la définition d'un « chercheur ». De même, avant d'appliquer l'indicateur relatif au *financement de la recherche rapporté au nombre d'enseignants-chercheurs ou ETP*, les données doivent être ajustées à l'échelle et à la mission de l'université en question (Ibid., p.45).

Quelques réflexions du groupe de travail AUBR sur les indicateurs de la recherche utilisés, directement ou indirectement, dans les classements mondiaux d'universités sont présentées ci-dessous :

- *Nombre de publications et autres types de production scientifique.* Différentes disciplines génèrent différents types de production scientifique. Cet indicateur met l'accent sur la quantité de publications ;
- *Nombre/pourcentage de publications dans les revues à fort facteur d'impact les mieux classées.* Les classements spécialisés ne correspondent pas toujours aux facteurs d'impact, en particulier dans les sciences humaines et sociales. Dans ces domaines, ainsi que dans les sciences de l'ingénieur, d'autres sources sont également importantes (livres, actes) ;
- *Citations.* Les citations présentent peu d'intérêt dans les disciplines qui ne sont pas bien couvertes par les index de citations, en particulier dans certaines branches des sciences humaines et sociales et des sciences de l'ingénieur ;
- *Nombre de récompenses et de prix nationaux/internationaux prestigieux.* Il n'existe pas d'équivalent reconnu et appliqué à l'échelle internationale ;
- *Postes de chercheurs invités.* Il n'existe pas d'équivalence qui soit reconnue et appliquée à l'échelle internationale et facilite la comparaison entre les disciplines.
- *Participation aux comités de lecture de revues scientifiques nationales et internationales prestigieuses.* Il n'existe pas d'équivalence reconnue et appliquée à l'échelle internationale ;
- *Valorisation de la propriété intellectuelle générée par les brevets, licences ou la création d'entreprises.* Les brevets sont un très faible indicateur de la commercialisation. Ils sont étroitement liés au contexte national et à la discipline ;
- *Nombre de collaborations et de partenariats.* Il peut être difficile d'obtenir et de vérifier les données en raison d'un manque de clarté ;
- *Pourcentage des enseignants-chercheurs actifs en recherche par rapport au nombre total d'enseignants-chercheurs.* Il n'existe pas de définition claire de « actif en recherche » (« research active ») ;

- *Niveau de financement mobilisé par les chercheurs et les universités auprès de sources extérieures.* Les niveaux de financement externe varient largement selon les disciplines ;
- *Financement de la recherche par enseignant-chercheur ou ETP.* La comparabilité dépend de la mission institutionnelle, du contexte et de la discipline ;
- *Investissement total en R&D.* Il est difficile d'obtenir des données institutionnelles valables et comparables ;
- *Infrastructures et installations de recherche.* Il est difficile d'obtenir des données valables et comparables. Les universités plus anciennes et mieux dotées sont favorisées. (ibid., p. 43-47).

Grille d'évaluation multidimensionnelle de la recherche

Un résultat important présenté dans le rapport final du groupe de travail AUBR est la grille d'évaluation multidimensionnelle de la recherche (Tableau 8 – AUBR, 2010).

Cette grille permet l'identification d'indicateurs appropriés dans cinq catégories – productivité de la recherche ; qualité et rayonnement académique ; innovation et contribution sociale ; viabilité et échelle ; infrastructure de recherche – en fonction de l'objectif de l'évaluation : allouer des ressources ; favoriser la différenciation des missions de recherche ; renforcer l'engagement régional/local ; améliorer la performance de la recherche ; évaluer l'optimisation des ressources ou le rapport coûts-avantages de la recherche ; promouvoir la coopération internationale ; soutenir la recherche multidisciplinaire.

Tableau 8 : Grille d'évaluation multidimensionnelle de la recherche (AUBR, 2010)

OBJECTIF	PRODUCTIVITÉ DE LA RECHERCHE	QUALITÉ ET RAYONNEMENT ACADEMIQUE	INNOVATION ET CONTRIBUTION SOCIALE	VIABILITÉ ET ÉCHELLE	INFRASTRUCTURES DE RECHERCHE
Allouer des ressources	- Production scientifique/données bibliométriques	- Données de citations - Évaluation par les pairs - Discours liminaires, prix et récompenses, etc.	- Financement de la recherche		- Proportion d'enseignants-chercheurs actifs en recherche par rapport au nombre total d'enseignants-chercheurs. - Bibliothèques, équipement, etc.
Favoriser la différenciation des missions de recherche	- Production scientifique/données bibliométriques - Production par chercheur	- Évaluation par les pairs - Auto-évaluation	- Financement de la recherche rapporté au financement de l'enseignement - Financement externe de la recherche	- Étudiants du 1er cycle rapportés aux étudiants des 2e et 3e cycles	
Renforcer l'engagement régional/local	- Publications, rapports stratégiques, etc.	- Évaluations par l'utilisateur final - Discours liminaires, récompenses médiatiques, etc.	- Pourcentage du financement provenant des utilisateurs finaux - Brevets, licences, création d'entreprises	- Nombre de collaborations et de partenariats	
Améliorer la performance de la recherche	- Production scientifique/données bibliométriques	- Données de citations - Nombre et pourcentage de publications dans les revues à fort facteur d'impact les mieux classées - Évaluation par les pairs			
Évaluer l'optimisation des ressources ou le rapport coûts-avantages de la recherche	- Production scientifique/données bibliométriques - Production par chercheur	- Évaluation par les pairs et/ou données de citations - Données liées à la valorisation et au transfert - Évaluations par l'utilisateur final	- Indicateurs d'impact social, économique, culturel et environnemental - Financement externe de la recherche - Employabilité des titulaires d'un doctorat	- Nombre de collaborations et de partenariats	
Promouvoir la coopération internationale	- Production scientifique/données bibliométriques axées sur les collaborations aux plans européen et international		- Pourcentage du financement de la recherche provenant de sources internationales	- Nombre de collaborations et de partenariats	
Soutenir la recherche multi-disciplinaire	- Production scientifique/données bibliométriques axées sur les domaines interdisciplinaires	- Évaluation par les pairs - Auto-évaluation	- Nouveaux domaines de recherche, programmes d'enseignement interdisciplinaires, etc.		- Travaux de recherche réalisés par des personnes de disciplines différentes

Source : WG AUBR, 2010

3. Classements multidimensionnels

Les classements multidimensionnels sont des classements d'université qui reposent sur un plus grand nombre d'indicateurs

et ne présentent généralement pas de palmarès, mais des résultats par indicateur ou groupe d'indicateurs.

3.1 CHE Ranking – Centre for Higher Education Development / Die Zeit

3.1.1 CHE University Ranking

Le Centre pour le développement de l'enseignement supérieur ⁴⁵ (CHE – Allemagne) réalise un classement sur la base d'indicateurs multiples qui a été publié pour la première fois en 1998. Le classement universitaire du CHE vise à présenter des informations objectives, utiles et pertinentes à l'intention des jeunes qui se préparent à choisir un EES pour leurs études, des étudiants déjà inscrits, et des EES eux-mêmes.

Les résultats du classement universitaire du CHE peuvent être visualisés de différentes façons (voir ci-dessous). Cette initiative peut être considérée à la fois comme une évaluation et un classement. Dans certaines présentations, les universités sont répertoriées par ordre alphabétique. Les résultats pour plusieurs indicateurs sont présentés sous la forme d'un tableau indiquant seulement si l'université figure dans le groupe supérieur, moyen ou inférieur. Dans d'autres présentations, les universités sont classées en fonction d'un seul indicateur. Ici encore, les résultats n'indiquent que l'une des trois catégories dans laquelle figure l'université. Cependant, les résultats pour les autres indicateurs sont également présentés simultanément.

Un « score global » n'est établi que dans la section « My Ranking », dans laquelle les utilisateurs ont la possibilité de sélectionner jusqu'à 10 indicateurs qu'ils jugent importants pour choisir un établissement. Le système détermine ensuite les universités qui répondent le mieux aux besoins de l'utilisateur.

Les classements d'universités du CHE sont très interactifs. En utilisant les sites Internet de Die Zeit ⁴⁶ ou du DAAD ⁴⁷, un utilisateur individuel peut choisir différents indicateurs (ou groupes d'indicateurs) et obtenir un classement personnalisé des universités selon les critères sélectionnés. Les classements du CHE sont essentiellement conçus pour aider les étudiants à choisir l'université la mieux adaptée à leurs besoins. Comme mentionné plus haut, ces classements sont disponibles sur les sites Internet de Die Zeit et du DAAD.

Quelles sont les universités prises en considération ?

Le classement universitaire du CHE a initialement été conçu pour les établissements allemands d'enseignement supérieur

et ne présentent généralement pas de palmarès, mais des résultats par indicateur ou groupe d'indicateurs. rieur, y compris les universités et les Fachhochschulen ou universités de sciences appliquées. L'internationalisation du classement universitaire du CHE a débuté dans le cadre de projets pilotes auxquels ont été associées des universités autrichiennes en 2004 et des universités suisses en 2005, bien qu'aucun de ces projets n'ait été développé par la suite.

À partir de 2007, les classements du CHE ont inclus des EES des Pays-Bas et de la Communauté flamande de Belgique ⁴⁸. Cette initiative a débouché sur l'inclusion de presque toutes les universités et autres EES néerlandais en 2009.

Cette internationalisation du classement CHE se poursuit par son extension à des universités dans des pays où la langue allemande est utilisée, comme langue unique d'enseignement ou parallèlement à la langue nationale, comme en Italie et en Hongrie.

Critères retenus, indicateurs utilisés et indicateurs indirects

L'un des principaux objectifs du classement CHE est de fournir aux étudiants potentiels des informations sur lesquelles ils peuvent se fonder pour choisir un EES, en tenant compte de leur intérêt pour un domaine d'étude particulier. C'est pourquoi le classement CHE porte sur tous les champs disciplinaires et est structuré par domaine.

La description du classement universitaire du CHE recense plus de 100 indicateurs regroupés dans neuf catégories. Toutefois, le nombre d'indicateurs réellement utilisés dans les différentes versions du classement présentées ci-dessous n'est généralement pas supérieur à 30. Si les études dans le domaine sélectionné sont proposées à la fois par les universités et les universités de sciences appliquées (*Fachhochschulen*), l'utilisateur doit d'abord choisir l'une de ces deux options. La série d'indicateurs proposée à l'utilisateur diffère selon qu'on choisit l'un ou l'autre de ces types d'établissement.

Le CHE traite les données et a accumulé des résultats pour un vaste éventail d'indicateurs. Cependant, les indicateurs réellement utilisés dans les différentes versions du classement universitaire du CHE reposent essentiellement sur l'évaluation par les étudiants d'un certain nombre d'aspects particulièrement pertinents, comme le soutien offert par les

⁴⁵ Le CHE est une société à but non lucratif constituée en 1994 par la Fondation Bertelsmann et la Conférence des recteurs allemands.

⁴⁶ www.zeit.de.

⁴⁷ www.daad.de.

⁴⁸ Dans le cadre d'un projet soutenu par l'UE.

enseignants, l'évaluation globale de la qualité de l'enseignement, la qualité des bibliothèques et des ressources informatiques, les salles de cours et de séminaires et les sports universitaires. Cette évaluation porte également sur l'orientation de recherche de l'université et la préparation au marché du travail. Les évaluations réalisées auprès des enseignants-chercheurs sont utilisées moins souvent. Elles portent essentiellement sur la réputation de la recherche et/ou de l'enseignement dans d'autres universités. Certains indicateurs sont basés sur des « faits » (ex. statistiques ou données bibliométriques).

Une comparaison des indicateurs utilisés dans les classements du CHE dans 20 domaines d'étude choisis au hasard met en évidence la prépondérance de l'avis des étudiants, sur lequel reposent 70 % des indicateurs utilisés dans certains domaines, cette proportion étant ramenée à un peu plus de 40 % dans les domaines pour lesquels la méthodologie a été récemment actualisée. La proportion d'indicateurs basés sur des « faits » se situe entre 10 et 30 %, tandis que le poids cumulé des indicateurs basés sur l'évaluation par le personnel est en recul et se situe entre 5 et 10 %. Ceux-ci représentent jusqu'à 20 % des indicateurs lorsqu'ils comprennent également l'avis des diplômés.

Pour l'ensemble des indicateurs basés sur des évaluations, les répondants attribuent un score selon une échelle de 1 (très bon) à 6 (très mauvais).

Évolution de la méthodologie utilisée

La méthodologie est progressivement actualisée par domaine. Les séries d'indicateurs sont par conséquent plus à jour dans certains domaines d'étude que dans d'autres.

Dans les domaines pour lesquels la méthodologie a récemment été actualisée, les opinions des diplômés sont prises en compte et remplaceront à terme les opinions des étudiants. On observe également une tendance à accroître l'importance des données statistiques et bibliométriques par rapport aux indicateurs basés sur la réputation.

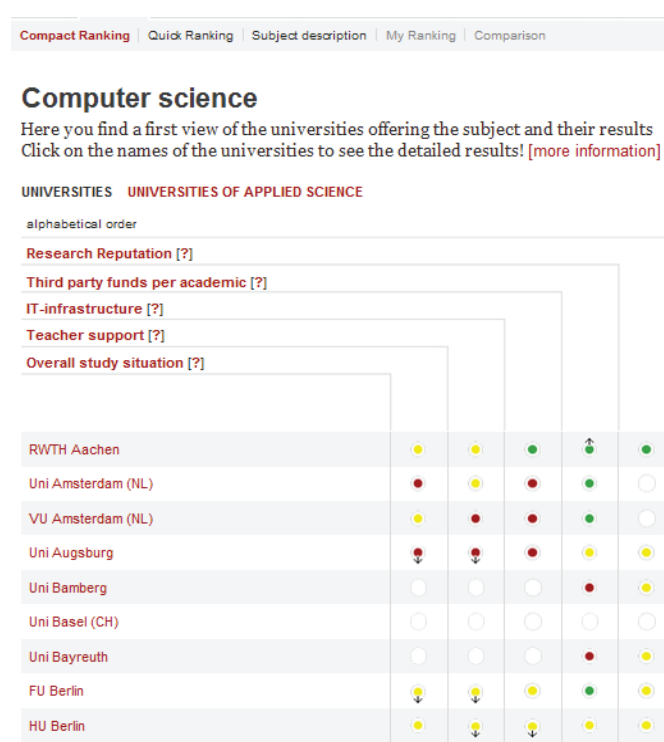
Présentation du classement et analyses complémentaires

Les différentes versions du classement universitaire du CHE sont les suivantes : le « *Compact Ranking* » (également dénommé « *Ranking Overview* »), le « *Quick Ranking* » (site Internet de Die Zeit), « *My Ranking* » et « *University Comparison* » (site Internet du DAAD).

Le « *Compact Ranking* » du CHE (appelé « *Ranking Overview* » sur le site Internet du DAAD) permet la comparaison de programmes d'études dans un domaine particulier entre différentes universités ou Fachhochschulen. Après avoir opté pour les universités ou les Fachhochschulen, l'utilisateur accède à une liste d'établissements proposant des programmes dans le domaine choisi. La liste est présentée par ordre alpha-

bétique. Elle indique dans quel groupe – supérieur, moyen ou inférieur – figure l'EES au regard de chacun des cinq indicateurs utilisés (voir l'exemple à la figure 1). Il est possible de modifier la présentation des résultats finaux selon chacun des cinq indicateurs. Toutefois, les résultats ne sont jamais présentés sous forme d'un palmarès, les universités étant classées par ordre alphabétique, même au sein des groupes supérieur, moyen ou inférieur. Les cinq indicateurs peuvent varier selon le domaine d'étude : les sciences, les sciences de l'ingénieur, la médecine et les sciences humaines et sociales utilisent des combinaisons d'indicateurs quelque peu différentes, basées sur une sélection par les producteurs du classement des indicateurs les plus appropriés pour comparer les EES proposant le type de programmes en question.

Figure 1 : Exemple du « Compact Ranking » du CHE – capture d'écran du site Internet de Die Zeit



Les principes sur lesquels repose la sélection des cinq indicateurs utilisés pour le « Compact Ranking » semblent être les suivants :

1. Un indicateur basé sur l'évaluation par les étudiants des conditions générales des études est utilisé pour tous les programmes ;
2. Un indicateur basé sur les avis des professeurs est également appliqué à tous les programmes. Dans la plupart des cas, il s'agit de l'indicateur de *réputation de la recherche*, mais dans le cas des programmes proposés par les *Fachhochschulen*, il peut également s'agir des *avis des professeurs sur la qualité des études supérieures et de l'enseignement* ;
3. L'indicateur de *soutien des enseignants*, qui est aussi basé sur l'évaluation par les étudiants, est souvent utilisé comme deuxième indicateur de la qualité des études et de l'enseignement ;

4. Un deuxième indicateur de la recherche est souvent utilisé, qui porte sur les *citations par publication* (dans les sciences, les sciences de l'ingénieur, la médecine) ou sur le *financement externe de la recherche par enseignant-chercheur* (dans d'autres domaines) ;
5. Au moins un indicateur lié à l'infrastructure de l'établissement est utilisé. Dans la plupart des cas, il s'agit d'une

évaluation par les étudiants des bibliothèques ou des laboratoires. Pour certains types de programmes, cet indicateur peut être très spécifique, comme le nombre de lits d'hôpitaux par centaine d'étudiants en médecine, les activités de terrain dans les géosciences ou les stages dans les sciences de l'ingénieur.

Tableau 9 : Exemples d'indicateurs généralement sélectionnés pour former les combinaisons de cinq indicateurs du « Compact Ranking » du CHE

Indicateur	Signification de l'indicateur	Application
Conditions générales des études	Contexte général de l'enseignement/des études évalué par les étudiants selon une échelle graduée de 1 (très bon) à 6 (très mauvais).	Tous les domaines.
Soutien des enseignants	Évaluation par les étudiants de l'accessibilité, des horaires de permanence, des activités de conseil/orientation, de l'examen des travaux, etc. selon une échelle graduée de 1 (très bon) à 6 (très mauvais).	Utilisé souvent comme deuxième indicateur de la qualité des études.
Réputation des études supérieures et de l'enseignement	Établissements d'enseignement supérieur recommandés par les professeurs pour la qualité de leur enseignement.	Utilisé pour les Fachhochschulen.
Réputation de la recherche	La réputation de la recherche est basée sur les avis des professeurs auxquels il est demandé d'identifier les EES les plus performants en termes de recherche (à l'exception de leur propre établissement).	Utilisé pour les universités.
Citations par publication	Nombre moyen de citations par publication.	Utilisé comme deuxième indicateur de la recherche (ex. dans les sciences, la médecine, la pharmacie).
Financement de la recherche	Financements externes (industrie, fondations, etc.) par enseignant-chercheur (à l'exception des établissements financés par des tiers).	Utilisé comme deuxième indicateur de la recherche dans les domaines autres que les domaines cités ci-dessus.
Laboratoires	Disponibilité et état des salles de laboratoire, etc. selon une échelle graduée de 1 (très bon) à 6 (très mauvais).	Utilisé par exemple dans les sciences, les sciences de l'ingénieur.
Bibliothèque	Disponibilité des publications, aide à l'utilisateur, possibilités de recherche de documents et disponibilité des postes de travail, etc. selon une échelle graduée de 1 (très bon) à 6 (très mauvais).	Utilisé en remplacement de l'indicateur relatif aux laboratoires.

Le « **Quick Ranking** » du **CHE** ([site Internet du DAAD](#)) permet aux étudiants d'obtenir rapidement des informations pour choisir les universités qui correspondent le mieux à leurs souhaits particuliers. Les étudiants sélectionnent d'abord le domaine et optent ensuite pour les universités ou les *Fachhochschulen* (FH) ou, le cas échéant, pour des études sanctionnées par un diplôme de licence ou un *Diplomstudium*. Une liste d'indicateurs regroupés dans six ou sept catégories leur est ensuite proposée. Chaque catégorie comporte généralement plusieurs indicateurs. Les utilisateurs peuvent choisir les indicateurs qu'ils jugent importants (voir le tableau 10 ci-après).

Tableau 10 : Le « Quick Ranking » du CHE – indicateurs proposés aux utilisateurs

Catégories	Indicateurs	Remarques
Études supérieures et enseignement		
	Contact entre les étudiants et le personnel.	Proposé dans tous les cas.
	Conseil / orientation	Proposé dans tous les cas.
	Cours proposés	Proposé dans tous les cas.
	Apprentissage en ligne.	Proposé dans tous les cas.
	Organisation des études.	Proposé dans tous les cas.
	Évaluation de l'enseignement.	Proposé dans tous les cas.
	Orientation de recherche.	À l'exception des FH et de certaines études universitaires.
	Autres indicateurs spécifiques au domaine possibles	
Équipement		
	Infrastructure informatique	
	Bibliothèque.	
	Salles de cours	
	Installations, comme les laboratoires, les installations sportives, les lits d'hôpitaux, etc	Le cas échéant.
Orientation internationale		Dans la plupart des programmes, mais pas tous.
	Orientation internationale ou soutien aux séjours à l'étranger.	L'un ou l'autre indicateur dans la plupart des programmes.
Marché de l'emploi et orientations de carrière		
	Préparation au marché de l'emploi.	Proposé dans tous les cas.
	Soutien pendant le semestre d'expérience pratique	Proposé dans tous les cas.
	Plusieurs autres indicateurs possibles	Dans certains programmes
Avis général		
	Conditions générales des études	Proposé dans tous les cas.
	Réputation de la recherche.	Pour les études universitaires.
	Réputation des études supérieures et de l'enseignement	Pour les études proposées par les FH.
Recherche		
	Proportion de financement externes	Proposé dans tous les cas
	Nombreux doctorats.	À l'exception des études proposées par les FH.
	Nombreuses citations.	Certains programmes universitaires.
	Nombreuses publications.	Certains programmes universitaires.
Ville et université		
	Sports de l'enseignement supérieur.	Proposé dans tous les cas
	Loyer modéré	Proposé dans tous les cas
	Site universitaire de taille réduite.	Proposé dans tous les cas

Après avoir sélectionné ou supprimé des indicateurs, l'utilisateur accède à une page qui illustre quels sont les EES les plus proches du centre de la « cible ».

Figure 2 : Exemple du « Quick Ranking » du CHE – capture d'écran du site Internet de Die Zeit

CLOSE WINDOW X
REFINE YOUR SELECTION WITH "MY RANKING" »

Quick Ranking: Chemistry

Universities

You can only choose up to 10 indicators at once.

view rank groups in greyscale »

MY INDICATORS:
(S)= Student's judgements (F)= Facts (P)= Professor's judgements

Academic studies and teaching ☒ Contact between students (S) ☐ Counselling (S) ☒ Courses offered (S) ☐ E-Learning (S) ☐ Research orientation (S) ☐ Study organisation (S) ☐ Teaching evaluation (S) Equipment ☐ IT-infrastructure (S) ☐ Laboratories (S) ☒ Library (S) ☐ Library - computer workstations ☒ Rooms (S)	International orientation ☒ Support for stays abroad (S) Job market and career-orientation ☐ Job market preparation (S) ☐ Practice Support (S) Overall opinions ☒ Overall study situation (S) ☐ Reputation for academic studies and teaching (P) ☒ Research Reputation Research ☐ Inventions (F) ☒ many citations (F) ☐ many doctorates (F) ☒ many publications (F) ☐ much third party funding (F) Town and University ☒ Higher education sport (S) ☐ low rent (F) ☐ small university location (F)
--	---

La présentation « **My Ranking** » du CHE est une version modifiée du « Quick Ranking » disponible sur les sites Internet du DAAD et de Die Zeit. La différence entre le « Quick Ranking » et « *My Ranking* » réside dans le choix des indicateurs,

« *My Ranking* » ne permettant aux utilisateurs de choisir que cinq indicateurs sur la liste. Dans le cas du « Quick Ranking », la liste des indicateurs proposés varie selon le domaine et le choix entre les universités et les Fachhochschulen.

Figure 3 : Exemple de « My Ranking » du CHE – capture d'écran du site Internet du DAAD

My Ranking: Law | [view course-description »](#)
for Universities

Compact Ranking: [show a short overview »](#)

University comparison: [choose universities for yourself »](#)

What are the five most important criteria for me?

Select up to five of the criteria which are most important for you. You also have the opportunity to decide whether only the top group, the top and middle group or all universities are to be displayed.

only top-group
★★★

top- and middle-group
★★★★
★★★☆☆

top-, middle- and bottom-group
★★★★★
★★★★☆
★★★☆☆

1. Indicator:

2. Indicator:

3. Indicator:

4. Indicator:

5. Indicator:

● ● ● ● ●

● ● ● ● ●

● ● ● ● ●

● ● ● ● ●

● ● ● ● ●

(F) stands for Facts, (S) for Students opinions and (P) for Professor's opinion.

[Submit](#)

Universities (sorted in alphabetical Order)	1. choosen indicator	2. choosen indicator	3. choosen indicator	4. choosen indicator	5. [%] choosen indicator
Europ. Uni Frankfurt/O.	★★★★ 2.2	★★★★ 2.1	★★★★ 2.1	★★★☆☆ 0.8	★★★★ 0.9
Bucerius LS Hamburg (priv.)	★★★★ 1.2	★★★★ 1.4	★★★★ 1.2	★★★☆☆ 0.5	★★★★ 1.2
Uni Düsseldorf	★★★★ 2.3	★★★★ 2	★★★★ 2.4	★★★★ 2.6	★★★★ 1.5

Figure 4 : Exemple de la présentation « University Comparison » - capture d'écran du site Internet du DAAD

University Comparison: Chemistry | [view course-description »](#)
for Universities

Compact Ranking: [show a short overview »](#)

University comparison: [choose universities for yourself »](#)

Here you can find all important information about studying at the universities of your choice.

★★★★ top group

★★★☆☆ middle group

★★★☆☆ bottom group

○ not specified

Department	Uni Tübingen » Fakultät für Chemie und Pharmazie »	Uni Ulm » Fakultät für Naturwissenschaften »	Uni Würzburg » Fakultät für Chemie und Pharmazie »
Total main subject students ^(?)	334	394	374
Teaching Professors ^(?)	11	13	16
New appointments ^(?)	2	3	3
Publications per academic ^(?)	★★★★ 16.8	★★★☆☆ 5.1	★★★★ 13
Citations per publication ^(?)	★★★☆☆ 6	★★★★ 6.7	★★★☆☆ 8.1
Third party funds per academic ^(?)	★★★★ 73.8 T€	★★★★ 68.1 T€	★★★☆☆ 46.4 T€
Doctorates per professor ^(?)	★★★★ 2	★★★☆☆ 1.4	★★★☆☆ 1.7
Inventions per 10 Scientists ^(?)	★★★★ 1.8	○	★★★☆☆ 0.5
Teacher support ^(?)	★★★☆☆ 2.3	★★★☆☆ 2.2	★★★☆☆ 2.6
Contact to students ^(?)	★★★☆☆ 1.9	★★★☆☆ 1.8	★★★☆☆ 1.9
Courses offered ^(?)	★★★☆☆ 3.1	★★★★ 2.8	★★★☆☆ 2.9
Research orientation ^(?)	★★★☆☆ 3.5	★★★★ 2.6	★★★☆☆ 3.1
Study organisation ^(?)	★★★☆☆ 2.2	★★★★ 2.5	★★★☆☆ 2.5

La présentation « **Comparison of Universities** » du CHE permet de comparer deux à trois universités dans un domaine déterminé. Elle utilise une série d'indicateurs similaires à ceux sur lesquels repose le « *Quick Ranking* », à l'exclusion de la catégorie « ville et université » (voir la figure 4).

3.1.2 CHE Excellence Ranking

Le classement d'excellence du CHE recense les universités, ou plus précisément les facultés de ces universités, qui font preuve d'excellence en biologie, chimie, mathématiques, physique, sciences politiques, économie et psychologie.

Le classement d'excellence du CHE ne convertit pas les résultats des classements individuels en un score final et n'utilise pas ces résultats pour établir un classement unique. Aucune pondération n'est, par conséquent, appliquée aux différents indicateurs.

Quelles sont les universités prises en considération et quels sont les critères retenus ?

Le classement d'excellence du CHE est une opération réalisée en deux étapes. Au cours de la première étape, les universités qui font preuve d'excellence dans les domaines spécifiques sont sélectionnées. Cette sélection se fonde sur le nombre d'« étoiles » attribuées sur la base d'indicateurs de présélection (voir ci-dessous). Les facultés présélectionnées sont ensuite soumises à une analyse approfondie.

Le CHE a publié son classement d'excellence pour la première fois en 2007. Il portait alors exclusivement sur les sciences naturelles. En 2009, les sciences politiques, l'économie et la psychologie ont également été incluses dans l'analyse. En 2010, le CHE a réalisé une nouvelle évaluation dans le domaine des sciences naturelles, dont les résultats ont été combinés aux résultats obtenus en 2009 en sciences politiques, en économie et en psychologie (Berghoff et al., 2010).

Critères de présélection

La présélection est basée sur le nombre d'« étoiles » obtenues par une université dans le domaine considéré. Pour l'économie, les sciences politiques et la psychologie, l'établissement est présélectionné s'il compte un total de deux étoiles, dont une au moins au regard des publications ou des citations. Pour les sciences naturelles et les mathématiques, l'établissement est présélectionné s'il compte deux étoiles au regard des indicateurs publications ou citations, ou trois étoiles en tout.

Une étoile est attribuée aux établissements qui représentent au moins 50 % de l'ensemble des réalisations dans le domaine. Cependant, pour être prise en considération, une université doit compter au moins 3 000 publications indexées dans le Web of Science au cours de la période 1997-2007 dans les domaines des sciences naturelles et des mathématiques, ainsi que des publications dans tous les autres domaines au cours de la période 1999-2006.

Indicateurs de présélection

1. Nombre de publications indexées dans le Web of Science (appliqué à tous les domaines). Une étoile est attribuée aux établissements qui figurent dans le groupe d'institutions qui affichent le nombre le plus élevé de publications et qui, comme groupe, représentent 50 % du nombre total de publications. En d'autres mots, il s'agit là d'une méthode de classement des universités selon le nombre de publications. Un seuil a été établi à 50 % du nombre total de publications, c'est-à-dire que les universités qui se situent au-dessus de ce seuil reçoivent une étoile.

2. Citations (appliqué à tous les domaines). Les citations par publication normalisées par champ disciplinaire (CPP/FCSm), ou le « *crown indicator* » du classement de Leiden, sont utilisées (voir le chapitre sur le classement de Leiden pour plus de précisions sur cet indicateur). La valeur de cet indicateur correspond au nombre de citations par publication⁴⁹ (CPP) rapporté au nombre moyen de citations par publication dans le même champ disciplinaire au cours de la même année (FCSm). Cela signifie que, si les publications d'une université sont citées avec une fréquence « habituelle » dans ce champ disciplinaire, la valeur de l'indicateur CPP/FCSm sera égale à 1. Une valeur supérieure à 1 indique que les publications de l'université dans ce champ particulier recueillent un nombre de citations supérieur à la fréquence « habituelle ». Dans le classement d'excellence du CHE, une étoile est attribuée aux universités dont la valeur de l'indicateur CPP/FCSm est au moins égale à 1,1.

3. Chercheurs éminents (appliqué aux sciences naturelles et aux mathématiques). Une étoile est attribuée à une université dans un domaine particulier si elle compte actuellement parmi son personnel enseignant des lauréats du Prix Nobel, du Prix Körber pour la science européenne ou de la médaille Fields en mathématiques. Un chercheur éminent suffit.

4. Nombre de projets dans le cadre du programme Marie Curie (appliqué aux sciences naturelles et aux mathématiques). Sur la base des informations tirées de la base de données Cordis, sont prises en compte dans le cadre du programme Marie Curie les activités suivantes :

- IEF – *Intra-European Fellowships for Career Development* (bourses de recherche intra-européennes pour le développement de carrière) ;
- IRG – *International Reintegration Grants* (bourses de réintégration internationale) ;
- ITN – *Initial Training Networks* (réseaux de formation initiale) ;
- ERG – *European Reintegration Grants* (bourses de réintégration européenne) ;
- IAPP – *Industry-Academia Partnerships and Pathways* (partenariats et passerelles industrie-académie) ;
- IOF – *International Outgoing Fellowships for Career Development* (bourses internationales « sortantes ») ;
- IIF – *International Incoming Fellowships* (bourses internationales « entrantes ») ;

⁴⁹ À l'exclusion des auto-citations.

- IRSES – *International Research Staff Exchange Scheme* (système international d'échange de personnel de recherche).

Le mode d'attribution des étoiles dans ce contexte est semblable à celui qui est utilisé pour l'indicateur de publications. Cependant, pour obtenir une étoile, trois projets sont nécessaires en biologie, deux en physique et en chimie, et un seul en mathématiques.

5. Mobilité des étudiants (appliqué à tous les domaines). Une étoile est attribuée aux universités qui affichent la mobilité la plus élevée et font partie du groupe représentant globalement 80 % des étudiants mobiles (50 % dans les sciences). Bien que ceci ne soit pas explicitement mentionné, le contexte donne à penser que le groupe de référence est celui des étudiants entrants de deuxième et troisième cycles. Le nombre total d'étudiants mobiles nécessaire à l'obtention d'une étoile varie de 35 en économie à 20 en physique.

6. Mobilité du personnel enseignant (appliqué à tous les domaines). Les membres du personnel qui ont effectué un court séjour d'enseignement à l'étranger dans le cadre du programme Erasmus sont comptabilisés pour cet indicateur, des points étant attribués tant à l'établissement d'origine qu'à l'établissement d'accueil. Une « étoile » est attribuée aux universités qui affichent la mobilité la plus élevée et font partie du groupe représentant 80 % des enseignants mobiles (50 % dans les sciences). En pratique, 3-4 enseignants mobiles suffisent pour recevoir une « étoile ».

7. Master Erasmus Mundus (appliqué à tous les domaines). Une « étoile » est attribuée pour un programme de Master conjoint dans le cadre du programme Erasmus Mundus. Une « étoile » est parfois attribuée à une seule faculté, plutôt qu'à tous les partenaires car, comme le fait observer le CHE, d'autres partenaires peuvent éventuellement participer à des parties du programme qui ne relèvent pas des domaines d'étude pris en compte par le classement d'excellence. Cet aspect est soigneusement vérifié.

8. Bourses du Conseil européen de la recherche (appliqué aux sciences naturelles et aux mathématiques). Une « étoile » est attribuée à l'établissement d'origine et à l'établissement d'accueil pour chaque bourse allouée en 2007 et en 2008.

9. Citations de livres (appliqué à l'économie, aux sciences politiques et à la psychologie). Le CHE affirme que, si l'indicateur de citations de livres ne peut pas offrir une analyse correspondant précisément aux citations d'articles, il constitue néanmoins un moyen d'éviter la discrimination à l'égard des champs disciplinaires dans lesquels la publication de livres représente le moyen privilégié de diffuser les résultats de la recherche.

Bien qu'ils ne soient pas pondérés, certains indicateurs sont plus importants que d'autres pour le processus de « présélection » ou, en d'autres mots, pour être inclus dans les résultats finaux du classement d'excellence. Pour y figurer, une université doit compter au moins 3 000 publications indexées dans le *Web of Science* et se voir attribuer un certain nombre

d'« étoiles » au regard des indicateurs de publications ou de citations.

Jusqu'à présent, le classement d'excellence a recensé et répertorié des centres d'excellence en Europe dans les sciences naturelles, l'économie, les sciences politiques et la psychologie. Les rapports ont été publiés en 2009.

Analyse approfondie

Les universités présélectionnées sont ensuite soumises à une analyse approfondie. Les données sont collectées sur la base d'enquêtes institutionnelles et d'enquêtes réalisées auprès d'étudiants des 2^{ème} et 3^{ème} cycles. Les étudiants répondent à des questions sur le contexte général des études, la disponibilité des conseillers et la qualité des centres de conseil et d'orientation professionnelle, les examens, les laboratoires, la bibliothèque, la formation, l'organisation des études, l'infrastructure informatique, le conseil, les sites Internet, les salles de cours, les relations sociales, la participation aux conférences, la communauté scientifique, le temps nécessaire pour mener à bien un projet de doctorat, les salles de travail, les ateliers, etc. L'enquête institutionnelle comporte différents types d'informations sur le personnel, les étudiants et l'université.

Présentation des résultats. Les résultats du classement d'excellence du CHE sont publiés sur le site Internet de *Die Zeit*. La sélection du domaine d'étude pris en compte dans ce classement permet d'accéder à la « liste d'excellence » des établissements présélectionnés dans le domaine choisi. Les listes sont présentées par ordre alphabétique. Les utilisateurs peuvent sélectionner deux ou trois universités à des fins de comparaison. Sur cette base, des informations plus détaillées peuvent être consultées, y compris des données recueillies dans le cadre des procédures de présélection ainsi que par l'intermédiaire des enquêtes réalisées auprès des étudiants et des établissements. Il est également possible de consulter les informations sur les programmes d'études dans ces universités et de les comparer.

Quant à la sélection des listes d'équipes de recherche, il est également possible de consulter les résultats sous une présentation différente. Après avoir choisi une branche du domaine d'étude (ex. chimie analytique), l'utilisateur accède à une liste des équipes de recherche. Y sont indiqués le nom du directeur de laboratoire, les membres de l'équipe, leur domaine de recherche spécifique, ainsi qu'une liste succincte des publications les plus importantes de l'équipe.

Dans l'ensemble, le classement d'excellence du CHE est un outil d'information utile, qui permet aux futurs étudiants de deuxième et troisième cycles de déterminer quelles universités font preuve d'excellence dans leur domaine d'intérêt particulier.

3.1.3 Autres classements du CHE

Le classement de recherche du CHE repose sur des indicateurs multiples. Il vise à comparer la performance des universités allemandes selon une série d'indicateurs, sans toutefois présenter les résultats de ce classement sous forme de palmarès.

L'« évaluation de l'employabilité » du CHE vise à évaluer les programmes de licence sur la base des moyens mis en œuvre pour promouvoir les compétences et les aptitudes qui contribuent à renforcer les capacités professionnelles des diplômés. Le premier rapport a été publié en 2008.

3.2 Classification U-Map des EES - CHEPS

L'objectif premier du projet U-Map est l'élaboration d'un outil de classification de l'enseignement supérieur en Europe qui reflète la diversité des missions et des profils des établissements d'enseignement supérieur (U-Map, 2010, p.11). Cet instrument examinera les différences entre les établissements (diversité institutionnelle) quant à leurs missions et profils (diversité horizontale). Le projet U-Map est financé par la Commission européenne et dirigé par le Centre d'études des politiques sur l'enseignement supérieur (*Center for Higher Education Policy Studies* - CHEPS) de l'Université de Twente, aux Pays-Bas.

U-Map est un outil à indicateurs multiples qui ne calcule pas le score final d'un établissement d'enseignement supérieur et par conséquent n'établit pas de palmarès. Des efforts ont été déployés pour présenter les données de façon à ce que des tiers ne puissent se baser sur ces données pour établir un classement. Les valeurs absolues des indicateurs ne peuvent être consultées que sur la base d'une sélection de trois EES comparés par indicateur. Des efforts prodigieux devraient ainsi être déployés pour recueillir toutes les informations nécessaires à l'établissement d'un classement sur la base des données U-Map – bien que cela ne soit pas complètement impossible. Cet instrument de classification s'adresse à tous les EES européens, sans discrimination quant au type d'établissement ou au domaine d'étude.

Le projet U-Map a débuté au deuxième semestre 2005 et le rapport final a été publié en janvier 2010. Au moment de la rédaction du présent rapport, le projet en est à la phase pilote de « l'insertion préalable » de données, auquel sont associés des EES de Norvège et, depuis juin 2010, des Pays-Bas et de la Communauté flamande de Belgique.

Indicateurs et classification des résultats

Les indicateurs sont regroupés selon six « profils » : profil d'enseignement, profil étudiant, l'intensité d'activité de recherche, le transfert de connaissances (« 3e mission »), l'orientation internationale et l'engagement régional (voir tableau 11). Les valeurs des indicateurs étant présentées visuellement (voir la section sur les outils de visualisation ci-dessous), les résultats de tous les indicateurs, à l'exception de trois d'entre eux, sont classés selon les catégories « *major* », « *substantial* », « *some* », « *none* ». L'outil de visualisation « *Profile viewer* » présente les quatre catégories de résultats sous la forme de quatre dimensions différentes du secteur attribué à chaque indicateur. Bien que la plupart des indicateurs de la liste soient explicites, des remarques complémentaires sont présentées au tableau 11, concernant essentiellement les indicateurs auxquels ne s'appliquent pas exactement les catégories *major*, *substantial*, *some* et *none* (ces indicateurs relèvent tous du profil étudiant).

Tableau 11 : Profils et indicateurs (U-Map, 2009a)

Indicateurs	Calcul	Seuils
Profil d'enseignement et d'apprentissage		
Niveau de diplôme prioritaire	% de diplômes de doctorat dans le nombre total de diplômes délivrés. % de diplômes de master dans le nombre total de diplômes délivrés. % de diplômes de licence dans le nombre total de diplômes délivrés. % de diplômes de niveau inférieur dans le nombre total de diplômes délivrés.	≥5% de diplômes de doctorat – priorité aux doctorats ; ≥25% de diplômes de master – priorité aux masters ; ≥40% de diplômes de licence – priorité aux licences ; ≥5% de certificats de cycle court – priorité aux certificats de cycle court
Offre académique	Nombre de champs disciplinaires offerts tels que définis par les 8 catégories de l'UNESCO/OCDE.	>6 champs – offre générale ; 3-6 champs – offre large ; ≤3 champs – offre spécialisée.
Orientation des diplômes	Proportion des diplômes délivrés : - Professions réglementées ; - Autres types de formations professionnalisantes ; - Formation générale.	Orientation ⁵⁰ : >1/3 des diplômés dans des programmes débouchant sur des professions réglementées – orientation vers les professions réglementées ; >1/3 des diplômés dans d'autres types de formations professionnalisantes ; >1/3 des diplômés dans des programmes de formation générale – orientation générale.
Dépenses d'enseignement	% des dépenses d'enseignement	>40% – major ; 10% - 40% – substantial ; 1% - 10% – some ; <1% – none.
Profil de l'étudiant		
Étudiants adultes / plus âgés	% d'étudiants adultes / plus âgés.	>20% – major ; 10% - 20% – substantial ; 5% - 10% – some ; <5% – none.
Étudiants à temps partiel	% d'étudiants à temps partiel.	>20% – major ; 10% - 20% – substantial ; 5% - 10% – some ; <5% – none.
Étudiants à distance	% d'étudiants à distance.	>20% – major ; 10% - 20% – substantial ; 5% - 10% – some ; <5% – none.
Nombre d'étudiants	Nombre total d'inscriptions.	30 000 – très grand ; 15 000 – 30 000 – grand ; 5 000 – 15 000 – moyen ; < 5 000 – petit
Intensité de l'activité de recherche		
Publications évaluées par les pairs ⁵¹	Publications par enseignant-chercheur.	>2 – major ; 1 - 2 – substantial ; 0,1 - 0,75 – some.
Diplômes de doctorat décernés	Diplômes de doctorat décernés rapporté au nombre d'enseignants-chercheurs.	1,5 – major ; 0,75 – 1,5 – substantial ; 0,1 - 0,75 – some.
Dépenses de recherche	% de dépenses de recherche.	>40% – major ; 10% - 40% – substantial ; 1% - 10% – some ; <1% – none.

⁵⁰ Une orientation mixte est possible si deux ou trois catégories représentent > 1/3 des diplômés.

⁵¹ Les livres et monographies sont pris en compte. Il est malaisé de déterminer si les revues sont comparables entre les pays.

Transfert de connaissances		
Création d'entreprises	Nombre d'entreprises créées (en ‰ du nombre d'enseignants-chercheurs).	>10 – major ; 5 - 10 – substantial ; 1 - 5 – some ; <1 – none.
Demandes de brevets	Nombre de demandes de brevet par (en ‰ du nombre des enseignants-chercheurs).	>10 – major ; 5 - 10 – substantial ; 1 - 5 – some ; <1 – none.
Activités culturelles	Nombre d'expositions, de concerts, de représentations.	>100 – major ; 50 - 100 – substantial ; 0 - 50 – some.
Revenus des activités de transfert des connaissances	% du revenu des accords de licence, des contrats, des droits d'auteur et des donations.	>40% – major ; 11% - 40% – substantial ; 1% - 10% – some ; <1% – none.
Orientation internationale		
Étudiants de pays tiers inscrits en vue de l'obtention d'un diplôme	% d'étudiants de pays tiers cherchant à obtenir un diplôme par rapport au nombre total d'étudiants	>7% – major ; 2,5% - 7,5% – substantial ; 0,5% - 2,5% – some ; <0,5% – none.
Étudiants entrants dans le cadre de programmes internationaux d'échange	% d'étudiants entrants dans le cadre de programmes internationaux d'échange par rapport au total des étudiants inscrits	>2% – major ; 1% - 2% – substantial ; 0% - 1% – some ; <0,5% – none.
Étudiants sortants dans le cadre de programmes internationaux d'échange	% d'étudiants sortants dans le cadre de programmes internationaux.	>2% – major ; 1% - 2% – substantial ; 0% - 1% – some ; <0,5% – none.
Enseignants-chercheurs internationaux	Personnel international en % des effectifs.	15% – major ; 5% - 15% – substantial ; 1% - 5% – some ; <1% – none.
Importance des sources internationales de financement	% de revenus issus de sources non nationales, à l'exclusion des droits d'inscription.	>10% – major ; 5% - 10% – substantial ; 1% - 5% – some ; <1% – none.
Engagement régional		
Diplômés travaillant dans la région	% de diplômés travaillant dans la région.	>10% – major ; 5% - 10% – substantial ; 1% - 5% – some ; <1% – none.
Étudiants de première année de licence originaires de la région	% d'étudiants de première année de licence originaires de la région.	>10% – major ; 5% - 10% – substantial ; 1% - 5% – some ; <1% – none.
Importance des sources de financement locales/régionales	% de sources de financement locales/régionales.	>10% – major ; 5% - 10% – substantial ; 1% - 5% – some ; <1% – none.

Source : information from (U-map, 2009a)

Niveau de diplôme prioritaire. Les indicateurs du « niveau de diplôme prioritaire » se fondent sur quatre éléments d'information. Les valeurs de ces quatre indicateurs correspondent simplement au pourcentage de diplômes de doctorat, de master, de licence et les certificats de cycle court⁵² décernés par rapport au nombre total de diplômes délivrés. L'EES est considéré comme accordant une attention particulière au niveau du doctorat si la proportion de diplômes de doctorat par rapport au nombre de diplômes décernés est de 5 %. Une attention particulière au niveau du master implique une proportion de 25 % de diplômes de master, une attention prioritaire au niveau de la licence impose une proportion de 40 % de diplômes de master, la priorité accordée à un diplôme de niveau inférieur suppose une proportion de 5 % de diplômes de niveau inférieur. Comme la somme des pourcentages mentionnés ci-dessus ne s'élève qu'à 75 %, un EES a la possibilité de satisfaire à plusieurs, voire à tous les critères. Un EES peut par conséquent avoir des priorités diverses. Aucune explication n'est fournie (tout au moins dans l'espace public du site Internet de U-Map) sur la manière dont les pourcentages plus faibles de ces diplômes sont associés aux trois tailles inférieures du segment correspondant dans le graphique. À titre d'exemple, si un EES compte 15 % d'étudiants en licence, quelle sera la taille du segment représentant l'indicateur du niveau de diplôme prioritaire appliqué aux étudiants de licence?

L'indicateur **Offre académique** dépend du nombre de champs disciplinaires proposés par l'EES tels que définis par l'UNESCO/OCDE (huit domaines principaux) : l'offre des universités qui proposent plus de six de ces champs est qualifiée de « générale », celle des établissements qui proposent de trois à six de ces champs est qualifiée de « large » et celle des EES qui en proposent moins de trois est qualifiée de « spécialisée ». Une analyse du nombre réel de champs disciplinaires proposés par les EES dont l'offre est qualifiée de « spécialisée » selon U-MAP révèle cependant que la situation sur le terrain est quelque peu différente. Dans certains cas, les établissements « spécialisés » proposent en réalité des programmes dans les huit champs disciplinaires tels que définis par l'UNESCO/OCDE. Ceci semble indiquer que sont également utilisés d'autres critères qui n'ont pas été présentés ou mentionnés sur l'espace public du site Internet.

L'indicateur **Orientation des diplômes/qualifications** établit une distinction entre les programmes d'études débouchant sur des professions réglementées/soumises à une autorisation d'exercice, les programmes proposant d'autres orientations de carrière et les programmes proposant une formation générale. La désignation correspondante est attribuée si la proportion des programmes respectifs s'élève à 1/3 de l'offre globale de l'établissement. Ici encore, un EES peut par conséquent afficher deux, voire théoriquement trois orientations.

L'indicateur **Dépenses d'enseignement** est décrit mais n'est en réalité pas utilisé. Tout au moins, il ne figure pas parmi les présentations publiques des résultats.

Un certain nombre d'indicateurs, comme la *création d'entreprises*, les *demandes de brevets*, les *activités culturelles*, les *publications évaluées par les pairs* et les *diplômes de doctorat décernés par enseignant-chercheur*, devraient être mesurés au cours d'une période déterminée. Cet aspect est mentionné, mais aucune information n'est disponible quant à la période au cours de laquelle chacun de ces indicateurs est mesuré.

Dans l'ensemble, on peut affirmer – ce qui est pertinent pour un système de classification – que les indicateurs U-Map sont plus étroitement liés à l'orientation et à l'intensité, plutôt qu'à la qualité des différentes activités des EES. À titre d'exemple, le nombre de publications est pris en considération, à la différence des citations ou de l'impact ; de même, les demandes de brevets sont prises en compte, mais pas les brevets délivrés, etc. Il ne s'agit pas d'une critique, mais d'une mise en garde concernant les limites d'un système conçu comme une classification. Cela peut également expliquer pourquoi cette initiative a été amenée à une nouvelle étape : U-Multirank.

Présentation des résultats U-Map

Les résultats d'un EES peuvent être visualisés à l'aide de deux outils en ligne : le Profile Finder et le Profile Viewer.

À l'aide du *Profile Finder*, il est possible de trouver des EES selon une combinaison de critères définis par l'utilisateur. La grille du Profile Finder permet à l'utilisateur de choisir une combinaison d'indicateurs et de catégories pour chaque indicateur. Le système affiche les universités qui répondent aux critères définis par l'utilisateur. Par exemple, si l'utilisateur choisit une université de *taille moyenne avec diverses priorités*, quelques (*some*) *activités culturelles*, un *financement régional important (substantial)* et comptant quelques (*some*) collaborateurs étrangers, le système extrait 2 des 67 EES enregistrés⁵³. L'utilisateur peut ensuite comparer les EES affichées par le *Profile Finder* de façon plus détaillée en utilisant l'outil *Profile Viewer*.

Dans le *Profile Viewer*, chaque EES est présenté sous la forme d'un graphique éclaté (« sunburst chart », voir figure 5). Chaque indicateur y est représenté par un secteur circulaire. Ce secteur peut avoir quatre tailles ou longueurs différentes. À l'exception des trois indicateurs présentés plus haut, les quatre tailles différentes des secteurs sont utilisées pour indiquer les valeurs *major*, *substantial*, *some* et *none*, comme décrit plus haut. Bien que les seuils pour chaque indicateur aient été établis de manière quelque peu subjective, il est possible de consulter les valeurs réelles des indicateurs en pointant le curseur sur le segment relatif à l'indicateur particulier (voir figure 6). L'outil *Profile Viewer* permet la sélection d'un maximum de trois EES figurant dans la base de données à des fins de comparaison par indicateur.

⁵² Le rapport n'apporte aucune précision complémentaire. Cependant, l'éventail de données recueillies donne à penser que ces « sub-degrees » sont délivrés au terme d'une formation professionnelle de type court.

⁵³ Testé le 1er novembre 2010.

Figure 5 : Le « sunburst chart » d'un EES (U-Map, 2010)

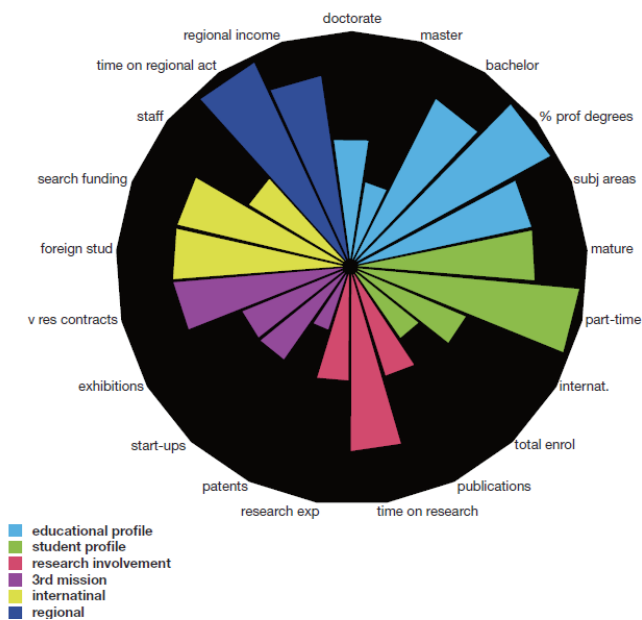
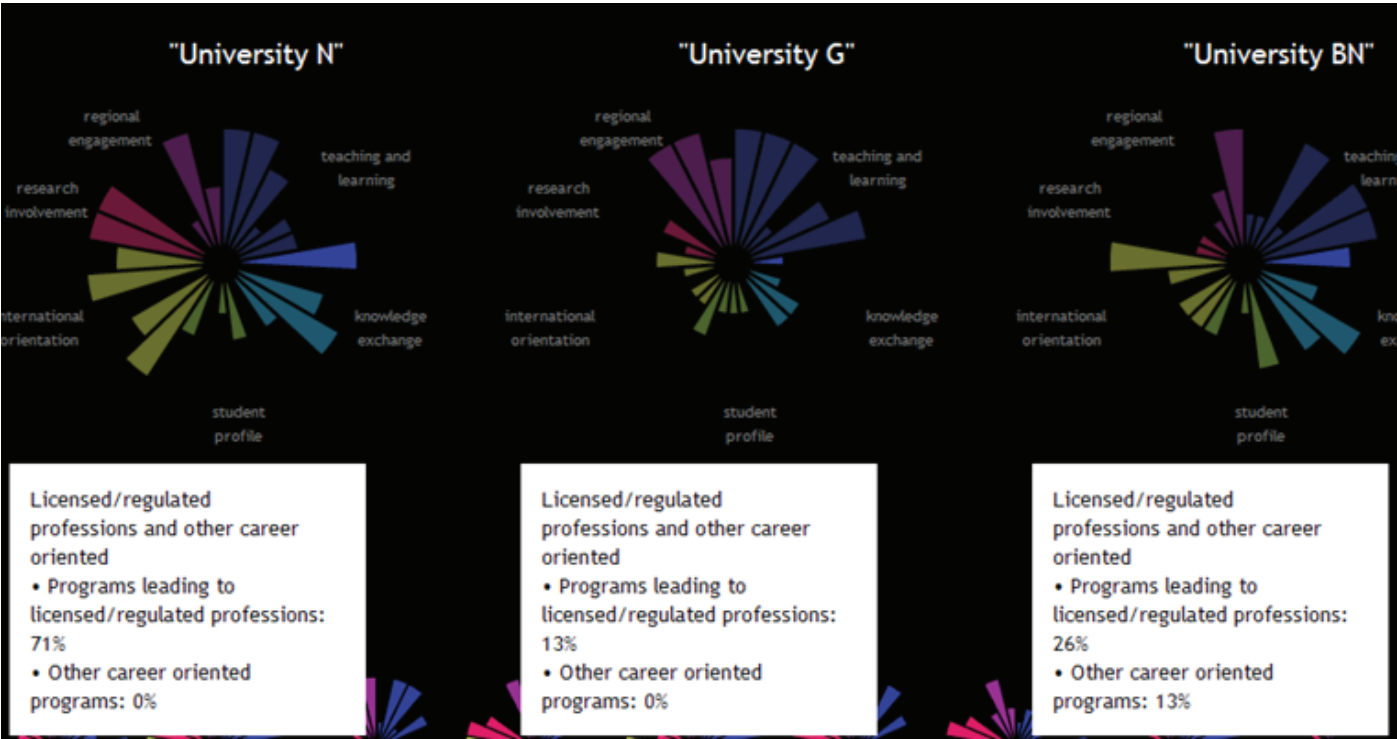


Figure 6 : Comparaison de trois EES dans le Profile Viewer (U-map.eu)



Quelles sont les universités prises en considération ?

Cet outil de classification est destiné à tous les établissements d'enseignement supérieur en Europe.

Sources de données

L'accès à des sources de données appropriées semble être le principal défi à relever pour utiliser U-Map comme outil de comparaison des EES au niveau international. L'équipe du projet U-Map a examiné les possibilités d'obtenir dans différents pays des données comparables. Leur conclusion est qu'il n'existe pas (encore) de système de données à l'échelle européenne et qu'il est en effet difficile de créer un tel système sur la base des séries de données nationales (exis-

tantes) (U-Map, 2010, p.18). La Commission européenne et EUROSTAT ont lancé une initiative visant à soutenir l'élaboration d'un « recensement » de l'enseignement supérieur et de la recherche, mais cette activité en est encore aux premiers stades de développement. C'est pourquoi U-Map repose essentiellement sur des données nationales ou institutionnelles. Dans le cadre de la phase pilote d'insertion préalable de données, U-Map a utilisé des données tirées de bases de données nationales sur l'enseignement supérieur et d'enquêtes auxquelles ont participé les EES eux-mêmes. Le manque de données comparables à l'échelle européenne s'explique par les énormes différences entre les pays quant à la définition, à la collecte, à l'interprétation et à l'utilisation des données.

Il semble par conséquent que les comparaisons fonctionnent beaucoup mieux à l'échelle nationale qu'au plan européen, tout au moins jusqu'à ce que ces nouvelles initiatives européennes commencent à produire des résultats.

3.3 European Multidimensional University Ranking System (U-Multirank) – projet financé par l'UE.

Le système européen de classement multidimensionnel des universités (U-Multirank) est un projet financé par l'UE visant à établir un classement mondial d'universités qui devrait, idéalement, éviter les principaux inconvénients des classements mondiaux d'universités actuels. Selon les objectifs fixés par l'UE (Commission européenne, 2009), le classement devrait être **multidimensionnel**, soit refléter les différentes missions des établissements, comme l'enseignement, la recherche, l'innovation, l'internationalisation, l'ancrage territorial et l'employabilité des étudiants; **indépendant**, vis-à-vis des pouvoirs publics et des universités; **mondial**, c'est-à-dire ne pas se limiter aux établissements européens. Ce projet est mené par le réseau CHERPA, sous la direction du Centre d'études des politiques sur l'enseignement supérieur (*Center for Higher Education Policy Studies* - CHEPS) de l'Université de Twente, aux Pays-Bas, et du Centre pour le développement de l'enseignement supérieur (*Zentrum für Hochschulentwicklung* – CHE), en Allemagne. Ce projet est en cours de développement, et la seule source d'information disponible est le premier rapport intérimaire (CHERPA, 2009) sur la phase de conception du classement. Une description complète du projet U-Multirank sera par conséquent présentée au terme du projet et après la publication des résultats.

Deux classements sont réalisés dans le cadre du projet U-Multirank : un classement spécifique des établissements (*Focused institutional ranking*) et un classement par champ disciplinaire (*Field-based ranking*).

Le classement spécifique des établissements permet des comparaisons d'établissements selon une seule dimension de l'activité institutionnelle, comme l'enseignement, la recherche, l'internationalisation ou le transfert de connaissances (CHERPA, 2010, p.77). Selon le réseau CHERPA (2010), les scores obtenus au regard de ces différentes di-

mensions ne seront pas combinés en vue d'établir un score global. Cependant, il reste à savoir s'il sera possible de présenter les résultats du classement U-Multirank de façon à éviter que des tiers ne combinent les résultats en vue d'établir un palmarès. Il existe des craintes sérieuses que cela ne soit inévitable (voir p.ex. Boulton, 2010, parag. 28).

Le classement par champ disciplinaire sera conçu comme un classement multidimensionnel d'un ensemble de programmes d'études dans un domaine ou discipline spécifique, proposés par des établissements au profil comparable. Selon les informations présentées sur le site Internet U-Multirank⁵⁴, l'établissement de classements par champ disciplinaire sur la base d'une classification objective des établissements est un objectif principal du projet ; en effet, les classements de programmes d'études ne peuvent être interprétés de façon satisfaisante que dans le cadre plus large de la classification multidimensionnelle de l'établissement dans son ensemble.

Quelles sont les universités prises en considération ?

Le classement U-Multirank portera sur des établissements en Europe et au-delà, en particulier aux États-Unis, en Asie et en Australie (Commission européenne, 2009).

Critères retenus

U-Multirank repose sur une approche multidimensionnelle. Il rend compte des différentes missions des établissements (*dimensions*), comme décrit plus haut. La Commission européenne rappelle que les classements existants tendent à se focaliser sur la recherche dans les sciences exactes et ne tiennent pas compte de la performance des universités dans des domaines comme les sciences humaines et sociales, et d'aspects comme la qualité de l'enseignement et l'ancrage territorial. C'est pourquoi la Commission affirme que le projet U-Multirank doit viser à couvrir au mieux l'ensemble des « dimensions » et des domaines d'études (Commission européenne, 2009).

Indicateurs et indicateurs indirects

La liste initiale d'indicateurs du classement U-Multirank s'inspire largement des indicateurs utilisés dans le classement universitaire du CHE en Allemagne. Ces indicateurs sont regroupés dans les catégories suivantes : « *Enabling* » (à son tour subdivisée en *Input* et *Process*) et « *Performance* » (subdivisée en *Output* et *Impact*).

Certains indicateurs sont destinés à être utilisés à la fois dans le *Focused institutional ranking* et le *Field-based ranking*, tandis que d'autres ne seront utilisés qu'à une seule fin. La liste d'indicateurs n'a pas été arrêtée définitivement au moment de la rédaction du présent rapport. Elle n'est par conséquent pas présentée.

⁵⁴ <http://u-multirank.eu/project/>.

Sources de données

Le rapport intérimaire du projet U-Multirank (CHERPA, 2010) présente une excellente analyse des avantages et inconvénients des différentes sources de données disponibles. Cependant, au moment de la rédaction du présent rapport, il restait malaisé de déterminer quelles mesures seront prises pour améliorer la collecte de données dans le cadre du projet U-Multirank par rapport aux classements existants.

On ne peut que souscrire à l'analyse de l'équipe du projet U-Multirank selon laquelle les bases de données WoS et Scopus ne reflètent pas de façon satisfaisante la recherche dans les arts et les sciences humaines, ne prennent pas en compte les livres autant que les revues scientifiques, et induisent un biais linguistique. Toutefois, il reste à voir comment l'équipe du projet U-Multirank prévoit d'éviter ces écueils.

Le rapport intérimaire précise que les indicateurs bibliométriques utiliseront les bases de données de Thomson Reuters et Scopus, et que des bases de données sur les brevets seront également utilisées. En outre, les données autodéclarées par les universités sont appelées à jouer un rôle notable tant dans le classement des établissements que dans le classement par champ disciplinaire (CHERPA, 2010, p.88). Ces données autodéclarées porteront sur le personnel, les étudiants, les ressources et les installations, la recherche (à l'exception des publications et des citations), le transfert de connaissances (à l'exception des brevets) et l'engagement régional, ainsi que sur l'enseignement et l'apprentissage.

Les enquêtes réalisées auprès des étudiants constitueront la troisième catégorie de sources de données utilisée. U-Multirank n'exploitera pas les enquêtes de réputation auprès des universitaires, s'agissant d'un indicateur peu fiable dans les classements internationaux (Federkeil, 2009).

Conversion des valeurs des indicateurs en scores

L'intention n'est pas de calculer un score global ou d'appliquer une pondération aux différents indicateurs.

Évolution de la méthodologie utilisée

La méthodologie est en cours de développement dans le cadre de la première phase du projet.

Présentation du classement et analyses complémentaires

Au moment de la rédaction du présent rapport, l'équipe du projet U-Multirank n'a pas encore précisé comment les résultats seraient présentés. Les informations disponibles indiquent que le *Focused institutional ranking* et le *Field-based ranking* établiront une comparaison entre programmes d'études semblables au sein de groupes d'EES au profil similaire.

4. Classements Internet

4.1 Webometrics Ranking of World Universities

Le classement cybermétrique des universités du monde (ci-après le « Webometrics ») est une initiative du Laboratoire de cybermétrie, un groupe de recherche du *Centro de Ciencias Humanas y Sociales* (CCHS), qui fait partie du Conseil national de la recherche espagnol.

L'objectif affiché du projet est « de convaincre les milieux universitaires et politiques de l'importance des [...] publications en ligne, non seulement à des fins de diffusion des [...] connaissances académiques, mais aussi afin de mesurer les activités, l'importance et l'impact scientifiques » (<http://www.webometrics.info/>).

Le classement Webometrics mesure le « volume » et la « visibilité » des sites Internet des universités (Aguillo et al., 2008). Le volume est défini par le nombre de pages sur le site Internet de l'université, ainsi que par le nombre de publications et de fichiers disponibles (.pdf, .ppt, .doc et .ps). La « visibilité » de l'université est mesurée par le nombre de liens entrants vers le site Internet de l'université.

L'équipe du Webometrics utilise des moteurs de recherche commerciaux pour la collecte de données parce que « les sites Internet peuvent être directement analysés par des logiciels de recherche qui collectent des informations de base par navigation hypertextuelle ou [parce que] des statistiques peuvent être extraites de bases de données analysées antérieurement et obtenues grâce à des moteurs de recherche commerciaux. Malgré les biais induits par des contenus incomplets et d'autres défauts, si une page Web n'est pas indexée par ces moteurs de recherche, elle ne peut être utilisée à aucune fin » (Aguillo et al., 2009, p.542). Le classement Webometrics est actualisé tous les six mois. Les données sont recueillies en janvier et en juillet et publiées un mois plus tard. La collecte de données est automatique, mais la position finale des universités dans le classement est calculée manuellement. Des comparaisons par rapport à l'année précédente sont également établies.

Quelles sont les universités prises en considération ?

Sont prises en considération toutes les universités qui ont un domaine indépendant, c'est-à-dire que les universités dont le domaine relève d'un autre établissement ne sont pas prises en compte. Plus de 20 000 établissements d'enseignement supérieur ont été analysés, et 12 000 ont été inclus dans la liste. Les résultats de tous les autres EES ne sont pas suffisamment significatifs. Lorsqu'une université compte plusieurs domaines, tous sont analysés, bien que seul soit inclus le domaine le mieux classé.

Critères retenus, indicateurs utilisés et indicateurs indirects

Le classement Webometrics mesure le volume, la visibilité et l'impact des pages Web des universités, avec une attention particulière accordée à la production scientifique.

Les créateurs du classement Webometrics considèrent qu'une forte présence sur Internet fournit des informations sur un vaste ensemble de facteurs qui peuvent clairement être mises en corrélation avec la qualité globale de l'université en question, mais également sur : « la disponibilité généralisée des ressources informatiques, le degré de formation à l'utilisation de l'internet dans le monde, les politiques en faveur de la démocratie et de la liberté d'expression, la concurrence pour une visibilité à l'échelle internationale ou encore le soutien aux initiatives de libre accès, etc. » (Aguillo et al., 2008, p.235). Bien que davantage de paramètres soient mesurés, comme le nombre de liens externes, le nombre de sous-domaines et le nombre de visites du site Internet, seuls quatre indicateurs sont inclus dans le classement. Ces indicateurs sont liés à deux critères : le « volume » et la « visibilité ».

Le volume global d'information publiée est mesuré par trois indicateurs :

1. Le nombre de pages sur le site Internet de l'université ;
2. Le nombre de fichiers (.pdf, .ppt, .doc et .ps) publiés ;
3. Le nombre d'articles publiés extraits par Google Scholar.

Les indicateurs de *volume* sont une mesure indirecte de l'intensité de la production académique de l'université, encore que d'une manière très particulière, qui regroupe les publications de recherche, les présentations, le matériel didactique, des données brutes, des versions préliminaires et d'autres documents relatifs à la recherche et à l'enseignement, ainsi que des informations de nature administrative publiées par l'université (Ibid.).

Cette démarche suppose une analyse de différents types de fichiers quant à leur pertinence en termes de production académique et scientifique. Les quatre types de fichiers ont été sélectionnés comme étant les plus pertinents.

Le nombre d'articles publiés et le nombre de citations de ces articles est tiré de Google Scholar. Il convient de noter que les résultats issus de la base de données Scholar concernent à la fois les articles, les rapports et d'autres documents de caractère universitaire.

La *visibilité* de l'université sur Internet est définie par le nombre de liens externes entrants vers le site Internet de l'université, qui permettent de jauger l'intérêt suscité par l'université.

L'établissement d'un lien vers le site Internet d'une université peut dans une certaine mesure être comparé à une citation. Ce lien, établi comme référence à une information présentée sur le site Internet, permet à des tiers d'accéder à l'information initiale ou de consulter les documents originaux.

La description et la pondération des quatre indicateurs sont illustrées au tableau 12.

Tableau 12 : Indicateurs utilisés dans le Webometrics Ranking of World Universities

Indicateur	Définition	Description	Poids
Visibilité (liens externes)	Nombre total de liens externes uniques reçus (liens entrants) par un site : il ne peut être obtenu avec assurance que sur la base de Yahoo Search.	Les résultats sont normalisés et ensuite combinés pour générer le classement.	50%
Volume du site Internet de l'université	Nombre de pages extraites de : Google, Yahoo, Live Search et Exalead.	Pour chaque moteur de recherche, les résultats sont normalisés. Ensuite, pour chaque domaine, les valeurs situées aux extrêmes sont exclues, et une position dans le classement est attribuée à chaque université selon la somme cumulée.	20%
Fichiers disponibles	Nombre de fichiers Adobe Acrobat (pdf), Adobe PostScript (ps), Microsoft Word (doc) et Microsoft PowerPoint (ppt).	Le nombre de ces types de fichiers est extrait à partir de Google. Les résultats pour chaque type de fichier sont normalisés et ensuite combinés.	15%
Scholar	Nombre d'articles et de citations, tiré de Google Scholar.	Ces résultats issus de la base de données Google Scholar représentent le nombre d'articles, de rapports et d'autres publications universitaires.	15%

Sources de données

Les données relatives au volume du site Internet de l'université sont tirées de Google, Yahoo, Live Search et Exalead. Les valeurs situées aux extrêmes sont exclues.

Le nombre de fichiers disponibles est obtenu à partir de Google, qui offre la possibilité technique d'extraire et de comptabiliser séparément les différents types de fichiers disponibles.

Les moteurs de recherche commerciaux sont utilisés parce qu'ils reposent sur des logiciels de recherche performants. Leurs bases de données sont fréquemment actualisées. Ils sont par ailleurs équipés d'outils automatiques qui peuvent être adaptés aux besoins de l'utilisateur et de calculateurs puissants pour l'extraction de données (Aguillo et al., 2008, p.235). Les moteurs de recherche commerciaux présentent toutefois des limites et inconvénients, tels que des résultats incohérents et arrondis, des biais régionaux et linguistiques ou encore des modifications fréquentes et opaques de leurs procédures de fonctionnement (Ibid.).

Le calcul des valeurs des indicateurs et leur conversion en scores

La valeur d'indicateur pour une université correspond à sa position dans le classement au regard de cet indicateur. La position finale est déterminée sur la base de la pondération présentée au tableau 12.

Évolution de la méthodologie utilisée

Webometrics n'a pas modifié sa méthodologie, mais les moteurs de recherche commerciaux utilisés effectuent des changements fréquents et opaques de leurs procédures de fonctionnement (Aguillo et al., 2008, p.235).

Présentation du classement et analyses complémentaires

Le classement global reflète la position globale. La position de chaque université au regard de chacun des quatre indicateurs est également présentée. Outre le classement principal, des classements par continent et par pays sont également présentés.

Webometrics réalise également des classements de centres de recherche non universitaires, d'écoles de commerce, d'hôpitaux et de dépôts de données.

5. L'étalonnage concurrentiel sur la base des acquis de l'apprentissage

5.1 Projet d'évaluation des résultats de l'enseignement supérieur (AHELO) – OCDE

En 2008, l'OCDE a entamé des discussions sur les comparaisons internationales de la qualité de l'enseignement supérieur. Dans le prolongement de ces discussions, l'OCDE a lancé une étude internationale sur les connaissances et les compétences des étudiants au moment de l'obtention de leur diplôme : l'évaluation des résultats de l'enseignement supérieur (ci-après « AHELO »).

Le projet AHELO vise à définir des critères permettant d'évaluer la pertinence et la qualité de ce qui est enseigné aux étudiants à travers le monde (OCDE, 2010). Si les résultats sont concluants, le projet AHELO pourra fournir des informations sur les stratégies d'enseignement les plus efficaces pour garantir que les acquis de l'apprentissage correspondent aux objectifs de l'enseignement. Les ambitions sont grandes : les

responsables du projet AHELO considèrent que les résultats aideront les EES à réduire les taux d'abandon en cours d'études, permettront à davantage d'étudiants de terminer leurs études avec succès, encourageant ainsi l'équité, et mettront par ailleurs à la disposition des pouvoirs publics un outil leur permettant de mesurer l'efficacité des systèmes d'enseignement supérieur (Ibid.).

Trois instruments d'évaluation sont en cours de développement : l'un pour évaluer les compétences génériques et deux autres pour évaluer les compétences spécifiques à l'économie et aux sciences de l'ingénieur. Les premiers tests concerneront une population ciblée d'étudiants sur le point d'achever des programmes de trois ou quatre ans (Lancette, 2010).

Le projet AHELO compte deux phases⁵⁵. La première phase, de janvier 2010 à juin 2011 (OCDE, 2010), porte sur le développement des cadres d'évaluation et des dispositifs de test adaptés à une mise en œuvre dans un contexte international. La deuxième phase, de janvier 2011 à décembre 2012 (Ibid.)

⁵⁵ Les différentes publications et présentations des responsables du projet sont quelque peu contradictoires quant au calendrier des deux phases. Les informations présentées ici sont issues du rapport publié en décembre 2010 (OCDE, 2010), s'agissant du rapport le plus récent au moment de la rédaction du présent rapport.

portera sur la mise en œuvre des trois instruments d'évaluation dans un groupe restreint de différents établissements d'enseignement supérieur. Des questionnaires contextuels à l'intention des étudiants, des enseignants-chercheurs et des établissements seront établis dans le cadre de chacune des trois évaluations (compétences génériques, économie et ingénierie), afin de définir un certain nombre d'hypothèses de départ quant à la relation entre contexte et acquis de l'apprentissage. Il est prévu de mettre en œuvre ces trois instruments (compétences génériques, économie et ingénierie) à la fin 2011/début 2012 dans une dizaine d'établissements d'enseignement supérieur par pays. C'est aussi à ce moment que les études contextuelles seront réalisées (Ibid.).

Cependant, les responsables doivent encore répondre à un certain nombre de questions importantes, notamment en ce qui concerne la phase initiale du projet (Lalancette, 2010) :

- Est-il possible de développer des instruments pour saisir les résultats de l'enseignement qui soient jugés valables dans des contextes nationaux et institutionnels différents ?
- Les éléments des questionnaires fonctionnent-ils comme prévu et les résultats des tests sont-ils conformes aux

normes psychométriques prédéfinies en termes de validité et de fiabilité ?

- Est-il possible de noter les éléments complexes de manière systématique pour tous les pays ?
- Est-il possible de saisir des informations sur les contextes d'enseignement et d'apprentissage qui contribuent à expliquer les différences de performance des étudiants ?

En outre, l'utilisation des acquis de l'apprentissage à des fins de comparaison de la performance des processus d'enseignement et d'apprentissage ne peut s'avérer concluante que si les établissements et les pays participants ont effectivement adopté une approche fondée sur les acquis de l'apprentissage tant en ce qui concerne le processus d'enseignement que l'évaluation des étudiants. Si la mise en rapport des programmes et des cours spécifiques d'une part et des acquis de l'apprentissage d'autre part est une démarche admise au sein de l'Espace européen de l'enseignement supérieur, des rapports d'étape récents sur le processus de Bologne indiquent toutefois que la généralisation de cette pratique prendra un certain temps (Vesterhijden et al., 2010 ; Ravhargers et al., 2009 : 25).

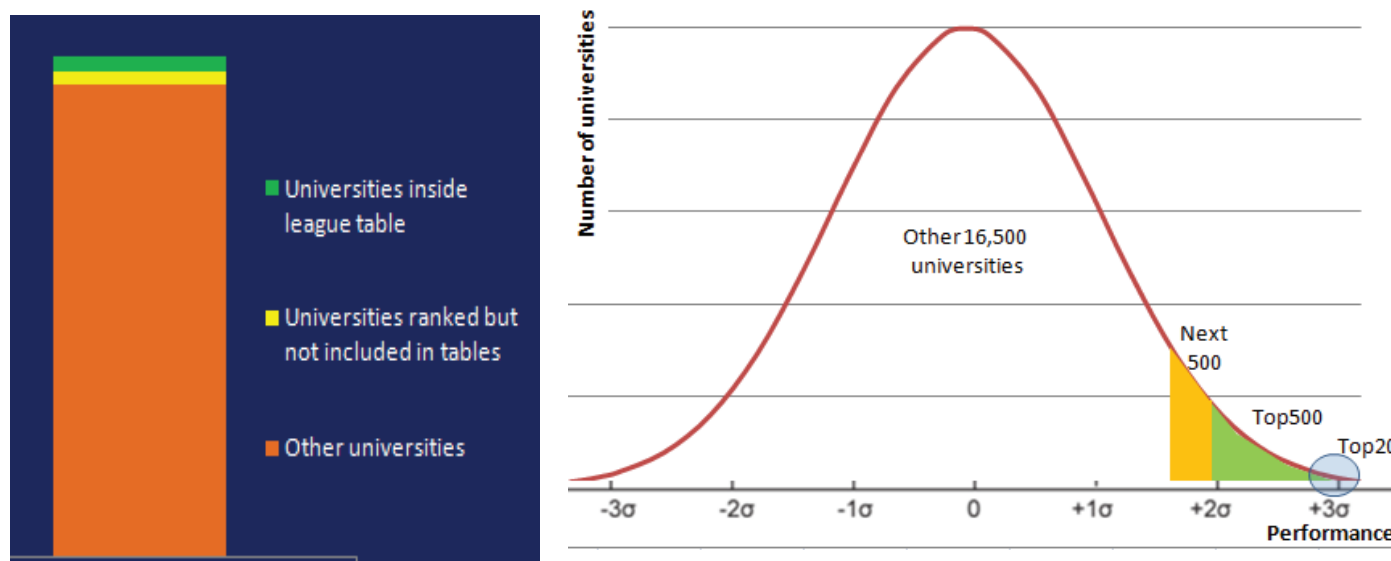
III. ANALYSE DES RÉSULTATS (IMPLICATIONS POUR LES UNIVERSITÉS)

Quelles sont les universités prises en considération dans les classements mondiaux d'université publiés sous forme de palmarès ?

Les classements mondiaux les plus populaires (ARWU, THE-QS et THE-Thomson Reuters, USNWR, HEEACT, Reitor et d'autres) prennent uniquement en considération les universités de recherche les plus réputées. Les critères sur lesquels se fondent ces classements sont tels que leur application ne permet de sélectionner qu'environ 1 200 universités dans le cas du classement ARWU et un total de quelques 600 universités et environ 300 par champ disciplinaire dans le cas du classement THE-QS. En d'autres mots, les méthodologies de

ces classements sont conçues de telle sorte que leurs résultats sont dépourvus de signification pour les universités qui ne font pas partie de ce groupe restreint. Bien que davantage d'universités souhaitent éventuellement se situer dans l'un de ces classements, quand bien même elles ne figureraient pas dans les listes du Top 100 ou du Top 500, elles n'en ont pas la possibilité : les méthodologies de ces classements mondiaux ne le leur permettent tout simplement pas.

Proportion d'universités prises en considération dans les classements mondiaux par rapport au nombre total d'universités dans le monde.



Par conséquent, les classements mondiaux ne tiennent compte que d'une très faible proportion des quelques 17 000 universités qui existent dans le monde. Ces universités de réputation mondiale sont certes très importantes, mais leur classement cause des problèmes ou entraîne des « conséquences indésirables », selon la formule des producteurs de classements : en effet, les responsables politiques et la société en général tendent de plus en plus à juger l'ensemble de l'enseignement supérieur dans le monde à l'aune des critères utilisés pour recenser un nombre limité d'universités de recherche de premier plan. Il convient de souligner que le seul fait de figurer dans ces classements mondiaux

révèle que ces universités font partie de l'élite des universités de recherche. Des pays, responsables politiques et même des universités affichent régulièrement leur ambition de voir leur institution ou groupes d'universités figurer au Top 20, au Top 100, ou simplement d'entrer dans les listes de classement à l'avenir. Ces déclarations sont fréquentes, mais rarement réalistes. Les termes de Jamil Salmi sont sans équivoque : « Combien d'universités peuvent figurer parmi les 500 premières ? Cinq cent » (Salmi, 2010).

La plupart des classements mondiaux s'accompagnent de listes concernant la « performance » de différents pays. Ces

comparaisons sont établies sur la base du nombre d'universités de chaque pays qui figurent sur les listes des universités les plus renommées. Un nombre de points différent est généralement attribué aux universités, selon qu'elles figurent au Top 100, au Top 101-200, etc. Selon cette méthode de calcul, les pays les mieux notés dans les listes publiées sont les États-Unis, le Royaume-Uni, l'Allemagne et la France. Cependant, si les listes publiées étaient « normalisées » en rapportant le nombre d'universités les mieux notées au nombre d'habitants, d'autres pays rejoindraient le haut de ces classements, comme la Suisse, la Suède, la Finlande et le Danemark (Salmi, 2010).

Ce que les classements sont et ce qu'ils ne sont pas.

Les classements font l'objet de débats passionnés et sont fortement critiqués. Cependant, malgré leurs nombreux défauts, biais et lacunes, « les parties prenantes et le grand public les acceptent facilement car ils proposent des informations simples, prêtes à consommer » (Groupe d'experts sur l'évaluation de la recherche universitaire – AUBR, 2009). D'aucuns affirment que les classements ne sont pas près de disparaître. Ils ont probablement raison. Leur nombre devrait d'ailleurs augmenter (Marginson, 2011), c'est pourquoi il est utile d'examiner les effets positifs et négatifs éventuels des classements.

Parmi les effets positifs des classements, il convient de mentionner les points suivants. Les classements d'universités peuvent éventuellement aider les étudiants à choisir l'université appropriée dans leur pays ou à l'étranger. À cette fin, les classements devraient toutefois fournir de meilleures explications quant à la signification réelle des scores des indicateurs. L'utilisation d'indicateurs de sélection plus « démocratiques » serait également utile, car cela signifierait que les

classements ne sont plus limités à l'élite des universités de recherche. Au plan national, les classements encouragent la collecte et la publication de données fiables sur l'enseignement supérieur. Dans un contexte international, les classements favorisent l'adoption de définitions communes des aspects pour lesquels les données sont collectées. Les résultats des classements mondiaux peuvent par ailleurs susciter un débat à l'échelle nationale et une analyse précise des facteurs déterminants du succès dans les classements, qui, à leur tour, peuvent provoquer d'utiles changements de politique.

Les classements, tout au moins ceux qui sont publiés sous forme de palmarès, sont toutefois dans l'impossibilité d'établir un diagnostic de l'ensemble du système d'enseignement supérieur, car ils ne concernent généralement que les universités de recherche les plus réputées. Les classements mondiaux actuels ne peuvent, par ailleurs, fournir que peu d'informations utiles sur des questions comme la qualité de l'enseignement et de l'apprentissage, l'accessibilité, l'engagement régional, la participation à la formation tout au long de la vie, le rapport coût-efficacité et d'autres, tout simplement parce que les indicateurs utilisés ne prennent pas ces éléments en compte. Dans un livre publié récemment, Ellen Hazelkorn affirme que « les classements ne tiennent en outre pas compte, par exemple, de la contribution des industries créatives/culturelles à l'innovation, ou de la mesure dans laquelle l'innovation sociale est à l'origine d'une évolution fondamentale de l'économie sociale, via de nouvelles formes d'action mutuelle, de nouveaux modes de gestion des économies, de nouvelles formes de consommation, et de l'organisation et du financement des pouvoirs publics » (Hazelkorn, 2011).

Les classements et la mission de recherche des universités

Les classements portent beaucoup plus sur la performance de la recherche des universités que sur la qualité de l'enseignement. Les principales catégories d'indicateurs et leur utilisation dans les classements mondiaux les plus renommés sont résumées comme suit :

Le **nombre de publications**, qui est appliqué de diverses façons :

- Le classement ARWU comporte deux indicateurs dont le poids total représente 40 % du score final : les publications dans *Nature* et *Science* et les publications indexées dans le *Science Citation Index-expanded* et le *Social Sciences Citation Index* ;
- Le classement HEEACT compte quatre indicateurs qui représentent un poids total de 50 % : le nombre d'articles par enseignant-chercheur publiés au cours des 11 années précédentes (10 %), le nombre d'articles par enseignant-chercheur publiés au cours de l'année précédente (10 %), le

nombre d'articles fréquemment cités au cours des 11 années précédentes, y compris le nombre d'articles figurant dans le top 1 % des articles les plus souvent cités recensés par ESI (15 %) et le nombre d'articles publiés dans le top 5 % des revues à fort facteur impact au cours de l'année précédente (15 %) ;

- Le classement THE-QS (jusqu'en 2009) n'utilise pas le nombre de publications ;
- Le THE-TR (depuis 2010) comporte un indicateur, dont le poids représente 4,5 % : le nombre de publications par enseignant-chercheur dans les revues scientifiques indexées par Thomson Reuters ;
- Le classement de Leiden comporte un indicateur (sans score global) : le nombre de publications dans les revues référencées dans les index de citations ;
- Le classement Reitor compte un indicateur : le nombre de publications des enseignants-chercheurs extraites de la base

de données Scopus (cet indicateur, dont le poids n'est pas précisé, fait partie d'une catégorie de trois indicateurs, qui représente un poids total de 20 %) ;

- Le CHE comprend un indicateur (sans score global) : le nombre de publications par enseignant-chercheur ;
- U-Map compte un indicateur (sans score global) : le nombre de publications évaluées par les pairs⁵⁶ par enseignant-chercheur ;
- U-Multirank comportera les indicateurs suivants sur le nombre de publications (sans score global) : le nombre de publications dans des revues scientifiques internationales à comité de lecture rapporté au personnel ETP, le nombre de publications les plus souvent citées, ainsi que des indicateurs sur plusieurs types de publications conjointes (aux plans national, régional et international).

Le nombre de publications est par conséquent utilisé de diverses façons. Le classement ARWU et le classement de Leiden s'y réfèrent en tant que mesure du nombre effectif de publications. Il convient toutefois de noter que les données extraites des bases de données ESI ou Scopus le sont à l'exclusion des publications sous forme de livres, c'est-à-dire au détriment des champs disciplinaires dans lesquels les traditions en matière de publications sont différentes⁵⁷. Le nombre de publications peut également être utilisé en relation avec le nombre d'enseignants-chercheurs (THE-TR, CHE, U-Map), afin d'illustrer l'intensité de publication, ou encore en tant que termes d'impact, afin de refléter l'excellence de la recherche (HEEACT, U-Multirank).

Le **nombre de citations**, qui est également appliqué de différentes façons :

- Le classement ARWU utilise le nombre de citations indirectement, plutôt que directement (pondération de 20%) : soit le nombre d'enseignants-chercheurs qui figurent parmi les 200 chercheurs les plus fréquemment cités dans 21 disciplines ;
- Le classement HEEACT utilise trois indicateurs (poids total de 30 %) : le nombre de citations par enseignant-chercheur au cours des 11 années précédentes (10%) ; le nombre de citations par enseignant-chercheur au cours des 2 années précédentes (10 %) ; le nombre moyen de citations par publication au cours des 11 années précédentes (10 %) ;
- Le THE-QS (jusqu'en 2009) utilise un indicateur (poids de 20 %) : le nombre de citations par membre du personnel ETP sur la base des données extraites de Scopus (Thomson Reuters au cours de la période 2004-2006) ;
- Le THE-TR (en 2010) utilise un indicateur (poids de 32,4 %) : la moyenne des citations par article normalisée par champ disciplinaire⁵⁸ au cours de la période 2004-2008 ;

- Le classement de Leiden utilise quatre indicateurs (sans score global) : un indicateur « simple » – le nombre de citations par article ; deux indicateurs normalisés – les citations par publication normalisées par champ disciplinaire (CPP/FCSm) le mean normalised citation score (MNCS2)⁵⁹ ; l'indicateur de « force brute », qui correspond au nombre normalisé de citations par publication multiplié par le nombre total de publications ;
- Reitor utilise un indicateur : les citations et références aux publications des enseignants-chercheurs (cet indicateur, dont la pondération n'est pas précisée, fait partie d'une catégorie de trois indicateurs dont le poids total représente 20 %) ;
- Le CHE utilise un indicateur (sans score global) : le nombre de citations par publication ;
- U-Multirank comprendra un indicateur relatif au taux de citations normalisé par champ disciplinaire (sans score global).

À l'instar du nombre de publications, le nombre de citations peut également être utilisé de différentes façons. Il peut être utilisé en tant que nombre absolu de citations (Reitor⁶⁰). Dans d'autres classements, le nombre de citations est utilisé en tant que citations par membre du personnel ETP (HEEACT, THE-QS), en tant que citations par publication (THE-TR, CHE, U-Multirank), ou les deux (HEEACT, classement de Leiden). Le nombre de citations est normalisé par champ disciplinaire dans plusieurs classements (Leiden, THE-TR, U-Multirank).

L'indice h. L'indice h de l'université est utilisé dans les classements Reitor et HEEACT. Dans ce dernier, l'indice h de l'université est établi sur la base des publications de l'université au cours des deux années précédentes (poids de 20 %). Dans le classement Reitor, l'indice h, dont la pondération n'est pas précisée, fait partie d'une catégorie d'indicateurs qui représente un poids total de 20 %.

Les enquêtes de réputation de la recherche. L'« évaluation par les pairs » dans le THE-QS (poids de 40 %) est une enquête de réputation plutôt qu'une évaluation par les pairs au sens de la terminologie de l'assurance qualité. La réputation de la recherche est utilisée dans le classement THE-Thomson Reuters (poids de 19,5 %). Outre les deux classements du THE, les enquêtes de réputation sont également largement utilisées dans le classement universitaire du CHE.

Dans le classement THE-QS, les « pairs » ont été invités à sélectionner jusqu'à 30 universités qu'ils estimaient excellentes dans leur domaine de recherche, sur la base de listes préétablies. Le classement THE-Thomson Reuters n'apporte pas plus de précisions sur l'enquête de réputation de la recherche.

⁵⁶ Les livres et monographies sont inclus. Il est malaisé de déterminer si les revues sont comparables entre les pays (U-Map, 2010).

⁵⁷ Il convient de noter ici que, dans le classement ARWU, les publications indexées dans le Social Sciences Citation Index comptent double, dans le souci d'atténuer le biais lié aux sciences sociales.

⁵⁸ La méthodologie de normalisation par domaine utilisée n'est pas précisée.

⁵⁹ Voir la section sur le classement de Leiden pour les différences entre les deux indicateurs.

⁶⁰ Selon la description de la méthodologie, cet indicateur peut correspondre au nombre de citations par publication des membres du personnel ou au nombre de publications par membre du personnel.

Le financement de la recherche. Lorsque cet indicateur est utilisé, il peut être appliqué de plusieurs façons :

- Le THE-TR compte trois indicateurs : le budget de recherche rapporté au personnel (poids de 5,25 %), le financement de la recherche issu de sources publiques par rapport au financement total de la recherche (0,75 %), et le financement de la recherche par l'industrie rapporté aux enseignants-chercheurs (2,5 %) ;
- Le classement universitaire du CHE, qui n'établit aucun score final, utilise l'indicateur de financement de la recherche par des tiers rapporté aux enseignants-chercheurs (autres domaines que les sciences, la médecine et la pharmacie) ;
- U-Map utilise le pourcentage des dépenses alloué à la recherche (sans score global) ;
- U-Multirank se base sur le financement externe de la recherche rapporté au personnel ETP (sans score global).

Le financement de la recherche est par conséquent rapporté au personnel ETP (financement global de la recherche dans le THE-TR et U-Multirank ; financement de la recherche par des tiers dans le THE-TR et le CHE). Ces valeurs peuvent être ajustées à la parité du pouvoir d'achat (THE-TR). Le pourcentage de ressources alloué à la recherche comme partie du revenu total (THE-TR) ou des dépenses totales (U-Map) peut également être utilisé.

L'intensité de l'activité doctorale

Le nombre de diplômes de doctorat décernés rapporté au nombre d'enseignants-chercheurs est un indicateur utilisé dans le THE-TR (poids de 6 %), U-Map et U-Multirank. Le THE-TR utilise également le nombre de diplômes de doctorat décernés rapporté au nombre de diplômes de licence décernés (2,25 %).

Les classements et la mission d'enseignement des universités

Certains producteurs de classements mondiaux soulignent l'importance de la question de la qualité de l'enseignement dans les classements. Il s'agit effectivement d'un problème sérieux. Les points suivants décrivent comment l'enseignement est actuellement inclus dans les différents classements examinés dans le présent rapport :

Dans le **classement ARWU**, la *qualité de l'enseignement* est mesurée par le nombre de diplômés d'un établissement ayant obtenu le Prix Nobel. On peut se demander si cet indicateur reflète l'enseignement ou, plutôt, la performance de la recherche de l'université dans laquelle le lauréat du Prix Nobel a étudié. Et si cet indicateur vise à refléter la qualité de l'enseignement, il le fait d'une façon très spécifique et limitée, ne s'appliquant d'ailleurs pas à un certain nombre d'universités qui figurent pourtant parmi les 500 premières du classement.

Dans les **classements THE-QS**, l'évaluation par les pairs porte exclusivement sur la réputation de la recherche. On peut se faire une idée de la qualité de l'enseignement sur la base des résultats de l'enquête réalisée auprès des employeurs, qui étaient invités à sélectionner les 30 meilleures universités de leur point de vue. Cette enquête était essentiellement destinée à de grandes entreprises (la plupart d'entre elles clientes de QS), et seules 3 281 réponses ont été recueillies à l'échelle mondiale (2 339 en 2008). Par conséquent, la seule « mesure de la qualité de l'enseignement » est le très critiqué taux d'encadrement des étudiants (ratio étudiants/enseignants), qui est une donnée d'entrée. Il peut donner une idée de la qualité s'il est appliqué au sein de la même discipline et dans un contexte national plutôt que dans le cadre d'une comparaison internationale.

Dans le **classement universitaire du CHE**, un certain nombre d'indicateurs concernent l'enseignement et l'apprentissage : le contexte général d'enseignement/des études, l'accessibilité, les horaires de permanence, les cours proposés, l'organisation des études, l'évaluation de l'enseignement, les possibilités d'apprentissage en ligne, la disponibilité et l'état des laboratoires et bibliothèques, la préparation au marché du travail, le soutien et l'encadrement (tous évalués par les étudiants), ainsi que la réputation de l'enseignement (évaluée par les professeurs). Ces informations étant basées sur l'expérience d'autres étudiants, elles peuvent aider les jeunes dans leur choix d'une université et d'un programme d'études particulier. Toutefois, ces indicateurs sont tous des indicateurs indirects, plutôt que des mesures objectives de la qualité de l'enseignement et de l'apprentissage.

Le **classement THE-Thomson Reuters 2010** comporte cinq indicateurs dans la catégorie *Enseignement – environnement d'étude* : une enquête de réputation sur l'enseignement, le nombre de diplômes de doctorat décernés rapporté au nombre d'enseignants-chercheurs, la proportion des diplômes de doctorat et de licence décernés, le nombre d'étudiants de licence inscrits rapporté au nombre d'enseignants-chercheurs (c.-à-d. le fameux taux d'encadrement des étudiants) et le financement rapporté au nombre d'enseignants-chercheurs. Ces cinq indicateurs peuvent décrire l'*environnement* d'étude, mais ici encore, ces indicateurs ne permettent d'évaluer l'enseignement que de manière indirecte. L'enquête de réputation fournit des informations sur les universités dans lesquelles l'enseignement est considéré comme étant le « meilleur » dans un ou plusieurs domaines, tandis que les universités offrant un « bon » enseignement ne sont pas notées. Un nombre élevé de diplômes de doc-

⁶¹ Les chiffres sont ajustés à la parité du pouvoir d'achat.

torat décernés est certes un bon signe, mais il n'est pas directement lié à la qualité de l'enseignement au niveau de la licence ou du master. Le taux d'encadrement des étudiants (ratio étudiants/enseignants) et le financement par enseignant-chercheur sont des indicateurs encore plus indirects.

En toute logique, les indicateurs utilisés par **U-Map** sont conçus à des fins de classification, plutôt que de classement ou d'assurance qualité, ce qui explique pourquoi les aspects liés à l'enseignement portent sur le niveau des diplômes, l'offre académique, l'orientation des diplômes et les dépenses d'enseignement. Ces indicateurs ne permettent pas d'évaluer la performance ou la qualité de l'enseignement.

U-Multirank comprendra plusieurs indicateurs relatifs à l'enseignement⁶². S'ils sont appelés à décrire le contexte de l'enseignement et de l'apprentissage, aucun de ces indicateurs n'en mesurera directement la qualité. Certains d'entre eux, comme les dépenses d'enseignement, pourraient s'avérer difficiles à utiliser dans des comparaisons internationales. La prise en considération du temps nécessaire à l'obtention d'un diplôme et des taux de réussite pose également des problèmes, car des établissements pourraient être tentés d'améliorer leurs taux de réussite en abaissant simplement leurs exigences. Des indicateurs comme les taux relatifs de chômage des diplômés et de revenus des diplômés peuvent aussi être utilisés pour estimer la qualité de l'enseignement, mais les résultats de ces indicateurs dépendent également de la situation et de l'évolution de l'économie nationale. La

proportion d'enseignants-chercheurs titulaires d'un doctorat peut constituer un indicateur pertinent pour autant qu'il soit tenu compte des différents types d'établissements d'enseignement supérieur et des traditions nationales. Si les indicateurs relatifs à la satisfaction des étudiants concernant le matériel informatique, les bibliothèques, les salles de cours et les stages professionnels dans le cadre des programmes d'études, sont utiles, ils relèvent néanmoins de critères de mise en capacité (*enabling*) plutôt que de performance.

Les classements basés sur la performance en ligne des universités sont encore plus loin de représenter une mesure directe de la performance ou de la qualité de l'enseignement et de l'apprentissage.

En conclusion, il semble que pratiquement tous les indicateurs utilisés pour décrire l'enseignement et l'apprentissage fournissent en réalité des informations sur *l'environnement* de l'enseignement et de l'apprentissage⁶³, plutôt qu'une mesure de la qualité de l'enseignement et de l'apprentissage en tant que telle.

Le projet AHELO est une initiative visant à évaluer la qualité de l'enseignement et de l'apprentissage par une mesure effective des acquis de l'apprentissage. Ce projet en est encore aux premiers stades de son développement, il est par conséquent trop tôt pour déterminer dans quelle mesure il atteindra ses objectifs.

Biais et défauts

Le biais en faveur des sciences naturelles et de la médecine au détriment des sciences sociales. De nombreuses publications ont démontré que la médecine et les sciences naturelles sont nettement mieux représentées dans les indicateurs bibliométriques que les sciences de l'ingénieur, les sciences sociales, et surtout les sciences humaines, dont il n'est souvent tenu aucun compte (voir p.ex. Leuwen et al., 2001 ; Moed, 2005 ; van Raan, 2005 ; AUBR, 2010 ; CHERPA, 2010). Ce biais est dans une large mesure induit par le fait que les indicateurs bibliométriques portent essentiellement sur les publications. Comme le souligne le groupe d'experts de l'UE sur l'évaluation de la recherche universitaire, « les spécialistes des sciences naturelles publient des livres, mais la publication d'articles dans les revues à comité de lecture constitue le mode de diffusion privilégié des résultats de leurs travaux de recherche. En ingénierie, les scientifiques publient principalement dans des actes de conférences, bien qu'ils publient aussi des articles dans des revues et conçoivent des prototypes. Les spécialistes des sciences humaines et sociales utilisent de nombreux modes de diffusion, dont les livres, qui représentent un moyen de communication important, tandis que la production artistique est diffusée sous la forme d'œuvres d'art, de compositions et de productions médiatiques » (AUBR, 2010, p.26).

Les producteurs du classement ARWU ont récemment entrepris une étude (Cheng, 2010) qui examine les traditions en matière de publication et de citations dans plusieurs champs disciplinaires. Cheng démontre que, sur la base de données recueillies aux États-Unis, le nombre de publications par enseignant-chercheur s'élève à 7,62 dans les sciences biologiques, mais atteint 6,04 dans les mathématiques, et est de 2,14 dans les sciences sociales et du comportement, ce qui reflète des pratiques de publication très différentes selon les champs disciplinaires. Quant au nombre de citations, les biologistes peuvent revendiquer 7,82 citations par article, contre 2,95 pour les ingénieurs et à peine 2,56 pour les spécialistes des sciences sociales et du comportement. Selon des données recueillies en Chine, le financement par enseignant-chercheur dans les sciences naturelles, les sciences de l'ingénieur et la médecine est 3,2 fois plus important que dans les sciences sociales. Ces données et celles de nombreux autres indicateurs mettent en évidence la nécessité, à tout le moins, de normaliser les données par champ disciplinaire, dans un souci d'objectivité.

La normalisation par champ disciplinaire – solutions et autres défauts. Des tentatives ont été faites pour remédier au biais lié au champ disciplinaire, parmi lesquelles le « *crown indicator* » du classement de Leiden – CPP/FCSm –,

⁶² Voir les tableaux U-Multirank 1 et U-Multirank 2.

⁶³ Ce terme est utilisé dans le classement THE-Thomson Reuters 2010.

dont la valeur correspond au nombre de citations par publication (CPP) rapporté au nombre moyen de citations par publication dans le même champ disciplinaire au cours de la même année. Cependant, les publications plus récentes étant moins citées, elles ont peu d'influence sur la valeur des sommes, tant au numérateur qu'au dénominateur. L'impact des nouvelles publications est par conséquent très limité⁶⁴. Il en va de même pour les champs disciplinaires dans lesquels les citations sont moins fréquentes : leur influence sur la valeur des sommes au numérateur et au dénominateur est également moindre. Cet indicateur, bien que permettant une certaine normalisation, comporte par conséquent un nouveau biais, favorable aux publications plus anciennes et aux champs disciplinaires caractérisés par un nombre élevé de citations. Par ailleurs, faire la somme de l'ensemble des citations et du nombre prévisible de citations ne contribue pas à présenter une image cohérente.

Une autre initiative est le *mean normalised citation score* (MNCS), calculée en divisant d'abord le nombre de citations de chaque publication par le nombre moyen attendu de citations dans le même champ disciplinaire et au cours de la même année. Le biais lié au champ est en l'occurrence supprimé, mais les résultats deviennent instables, parce que les nouvelles publications recueillent moins de citations. Le MNCS2 a été créé pour corriger ce défaut, mais en excluant les publications de l'année précédente, cet indicateur risque d'induire un nouveau biais. En outre, aucune solution n'est apportée pour remédier au biais lié à la publication de livres et d'actes de conférences, qui ne sont pas pris en compte, indépendamment de la méthode de calcul utilisée.

Le facteur d'impact – à utiliser avec précaution. Il convient de rappeler que « les classements spécialisés ne correspondent pas toujours aux facteurs d'impact, en particulier dans les sciences humaines et sociales. Dans ces domaines, et les sciences de l'ingénieur, d'autres sources sont également importantes (livres, actes) » (AUBR, 2010, p.43). Par ailleurs, « le champ d'application des deux bases de données [WoS et Scopus] reste insatisfaisant au regard des champs disciplinaires dans lesquels la publication d'articles dans des revues ou des actes de conférences ne constitue pas le mode principal de diffusion des connaissances : les arts et les sciences humaines en particulier » (CHERPA, 2010). Un avertissement est également affiché sur le site Internet de Thomson Reuters, publié initialement en 1994 : « il convient de ne pas utiliser le facteur d'impact sans tenir soigneusement compte des nombreux phénomènes qui influencent les taux de citations, comme par exemple le nombre moyen de références citées dans l'article moyen. Le facteur d'impact doit être utilisé sur la base d'une évaluation par les pairs réalisée en connaissance de cause » (Garfield, 1994).

Dans un livre publié récemment, Ellen Hazelkorn en conclut que « les classements proposent une définition étroite de « l'impact », qui se limite aux « pairs » de la communauté académique, définition qui résulte d'une évaluation de l'activité et de l'impact de la recherche limitée aux seules publications et citations » (Hazelkorn, 2011).

Le biais lié à l'évaluation par les pairs. Le terme « évaluation par les pairs » est lui-même ambigu, étant utilisé pour désigner des processus assez différents dans les domaines de l'assurance qualité et des classements. Dans l'assurance qualité de la recherche et de l'enseignement, le terme « évaluation par les pairs » désigne généralement une évaluation réalisée par les pairs, selon des procédures rigoureuses. Traditionnellement, des collègues compétents examinent les publications de l'entité évaluée, que ce soit une université, une faculté, un institut de recherche, un programme ou un département (voir p.ex. U-Multirank, 2010). Dernièrement, cette procédure a également inclus l'analyse d'un rapport d'auto-évaluation et d'autres documents fournis par l'entité en question, suivie généralement d'une visite d'experts sur place. Des études récentes ont examiné certaines préoccupations liées à l'impact des classements sur les pairs dans le détail (AUBR, 2010 ; U-Multirank, 2010). Le groupe de travail AUBR les résume de la façon suivante : « les examinateurs [les pairs] peuvent être influencés par des pressions concurrentielles, notamment au regard d'implications éventuelles pour leur propre travail ou celui de leurs collègues. Ils peuvent évaluer la recherche en fonction de leurs connaissances [...]. Ainsi, des idées originales et audacieuses sont susceptibles d'être marginalisées parce qu'elles remettent en cause des conceptions établies. Enfin, les pairs [...] peuvent être influencés par la réputation d'un chercheur plutôt que par sa contribution effective à l'avancement de la connaissance scientifique » (AUBR, 2010), p.40).

En outre, dans le contexte des classements, le terme « évaluation par les pairs » désigne généralement une simple enquête de réputation, à laquelle peuvent certes participer d'éminents chercheurs et professeurs. Cette « évaluation par les pairs » résumée à une enquête de réputation dans le cadre des classements pose problème. Premièrement, bien que de nombreux enseignants-chercheurs soient sollicités, peu y participent. Le THE-QS a reçu 9 386 réponses (dont 6 354 tirées de l'enquête 2008) après avoir contacté environ 180 000 personnes. L'équipe du classement Reitor a supprimé l'indicateur de réputation en raison du nombre extrêmement limité de répondants. Deuxièmement, les « pairs » n'ont pas la possibilité de sélectionner librement les universités qu'ils jugent excellentes. Leur choix est déterminé par des listes préétablies dont sont exclues de nombreuses universités, voire des pays entiers, tout au moins dans le cas du THE-QS. Troisièmement, l'avis des « pairs » est influencé par la réputation dont jouissent déjà certains établissements. Autrement dit, une université qui figure dans le haut d'un classement a de fortes chances d'occuper une position dominante dans un autre classement. De même, un établissement peut être jugé globalement excellent parce qu'il est réputé dans un domaine particulier (AUBR, 2010, p.20). Dans certains cas, des pairs ont jugé qu'une université faisait preuve d'excellence... dans un domaine où elle n'exerçait aucune activité.

Ellen Hazelkorn fait observer que les classements reposent sur une conception quelque peu dépassée de la recherche (Mode 1) au lieu d'en saisir les nouveaux modèles organisationnels en développement (Mode 2) : « Dans le Mode 1, la qualité des résultats des travaux de recherche est assurée par l'évaluation par les pairs, tandis que le Mode 2 privilège

⁶⁴ Voir l'analyse des expressions mathématiques et références dans la description du classement de Leiden.

gie la responsabilité sociale et la réflexivité. Alors que le premier repose sur un modèle élitiste traditionnel de création du savoir, le second démocratise la production, l'application et l'échange des connaissances [...]. » (Hazelkorn, 2011). Elle en conclut que « les classements [...] réaffirment la hiérarchie de la production traditionnelle du savoir » (Ibid.).

Le biais linguistique. Depuis la publication des premiers classements mondiaux, il a été observé que ces classements favorisent les universités dans les pays de langue anglaise, parce que les travaux dans d'autres langues sont à la fois moins publiés et moins cités (Marginson et van der Wende, 2007). Selon Altbach, les scientifiques américains citent essentiellement d'autres Américains et tendent à se désintéresser de les travaux de recherche menés dans d'autres pays, ce qui peut doper artificiellement le classement des universités des États-Unis (Altbach, 2006). En outre, la base de données de l'ISI (aujourd'hui WoS) recense essentiellement des articles publiés en anglais, sélectionnés selon les pratiques en vigueur dans les systèmes universitaires des États-Unis et du Royaume-Uni. Ceci est également susceptible de renforcer le biais lié au champ disciplinaire, l'anglais n'étant pas nécessairement la langue véhiculaire dans les sciences humaines, le droit et un certain nombre d'autres champs (Ibid.).

Le biais régional. Depuis l'apparition des classements mondiaux, certaines régions du monde, notamment l'Amérique

du Nord et l'Europe – et en particulier l'Europe occidentale – sont mieux représentées dans les classements que d'autres. D'ailleurs, les classements mondiaux se réfèrent implicitement au modèle anglo-saxon de l'organisation de la recherche (CHERPA, 2010, p.24).

Ce biais régional peut dans une large mesure s'expliquer par les mêmes raisons qui expliquent le biais linguistique. Cependant, les traditions en matière de publication et de citations peuvent varier selon les régions du monde. Même dans un champ disciplinaire particulier, ces traditions peuvent varier selon les régions. Les enquêtes de réputation en particulier posent des problèmes : en effet, le risque existe que les universités de certains pays ou régions ne soient sélectionnées dans les enquêtes de réputation que si elles figurent déjà dans les classements existants. C'est pourquoi le THE-QS applique une pondération régionale à chaque champ disciplinaire en vue de garantir une représentation équitable des quatre « grandes » régions : les Amériques ; l'Europe ; le Moyen-Orient et l'Afrique ; la région Asie-Pacifique (THE-QS, 2009). Le classement THE-Thomson Reuters 2010 applique un mécanisme de « prise en considération de facteurs internationaux » (Pratt, 2010). Cependant, dans les deux cas, la pondération régionale et les facteurs ou critères sur lesquels elle se base ne sont pas présentés sur les sites Internet publics.

Le risque d'accorder trop d'importance aux positions dans les classements

L'existence de classements incite sans aucun doute les universités à améliorer leur performance. Il reste néanmoins à déterminer les types de décisions qu'ils entraînent. Dans le souci de renforcer leur position dans les classements, les universités sont fortement tentées d'améliorer leur performance en priorité dans les domaines mesurés par les indicateurs sur lesquels repose l'élaboration des classements.

Les universités les mieux classées doivent déployer des efforts considérables pour maintenir leur position, car leurs concurrents évoluent également (CHERPA, 2010). Salmi assimile ce phénomène à l'« effet de la reine rouge » (Salmi, 2010). « Ici, nous devons continuer à courir pour rester au même endroit », dit la Reine Rouge dans le roman « De l'autre côté du miroir » de Lewis Carroll. Selon ce principe, qui a également été formulé en biologie évolutive (Van Valen, 1973), « le développement continu est nécessaire au maintien du rapport de forces entre un système évolutif et les autres systèmes avec lesquels il évolue » (van Heylingen, 1993).

Ce phénomène peut être à l'origine d'initiatives incohérentes ou irrationnelles à la seule fin d'obtenir de meilleurs résultats

dans les classements, par exemple :

- Davantage de lauréats du Prix Nobel parmi le personnel – la plupart des universités n'en comptent aucun, pourquoi ne pas tenter d'en engager un ?
- Davantage de publications et davantage de publications par ETP – il pourrait s'agir d'une évolution positive, mais en pratique cela signifie simplement plus d'articles dans les revues référencées par Thomson Reuters ou Scopus, plutôt que des livres et d'autres types de publications. Déployer des efforts pour améliorer la performance de la recherche dans le seul but d'améliorer les scores dans les classements équivaut à un soutien unilatéral de la recherche en médecine et dans les sciences naturelles, au détriment des sciences sociales et à l'exclusion des sciences humaines.
- Davantage de citations : par ETP et par article – Bien qu'il soit difficile de manipuler ces éléments si le classement en question procède à une normalisation par champ disciplinaire, un moindre soutien aux sciences humaines et sociales et une attention prioritaire aux sciences exactes et à la médecine sont susceptibles d'améliorer les scores dans les classements.
- La performance de tous les indicateurs auxquels sont

associées des valeurs exprimées en ETP (personnel de recherche, personnel enseignant, nombre total d'enseignants-chercheurs) est susceptible d'être « améliorée » en jouant sur les définitions des catégories de personnel. Dans les classements qui utilisent le taux d'encadrement des étudiants comme seul indicateur indirect de la qualité de l'enseignement, quelques modifications suffisent à entraîner des améliorations considérables.

- Comme plusieurs classements publiés sous forme de palmarès utilisent des indicateurs qui reflètent la production totale de l'université (nombre total de lauréats du Prix Nobel, d'articles, de citations, etc.), les fusions d'universités peuvent améliorer la position dans ces classements ou aider des établissements à y faire leur entrée.
- Les indicateurs relatifs à la proportion de membres du personnel international et d'étudiants internationaux dépendent de la définition de ces catégories. Ces indicateurs sont sujets à manipulation, tout au moins tant que l'on ne disposera pas de définitions exactes. Par exemple, les étudiants (ou membres du personnel) de nationalité étrangère résidant sur le territoire national peuvent-ils être considérés comme « internationaux » ?

Des cas de manipulation motivée par l'obtention de meilleurs résultats dans les classements ont effectivement été recensés. Plusieurs exemples de manipulation de différents indicateurs dans le classement USNWR sont cités dans le rapport intérimaire du projet U-Multirank (CHERPA, 2010, p.60) :

- Certains établissements ont déployé des efforts considérables pour encourager les étudiants à déposer des demandes d'inscription, bien qu'ils n'aient aucune chance d'être sélectionnés (Schreiterer, 2008). Ce faisant, ces institutions s'assuraient un score plus élevé au regard du critère de sélectivité ;
- Afin d'accroître la moyenne des résultats aux tests standardisés des candidats qui s'inscrivent, certains établissements ont décidé de rendre facultative la soumission de ces résultats lors de l'inscription, sachant que seuls les candidats ayant obtenu de bons résultats seraient enclins à le faire ;
- Comme le classement USNWR recense le nombre de membres du personnel ETP pour établir son ratio étudiants/enseignants en automne, des facultés et départements ont encouragé leurs enseignants-chercheurs à prendre leur congé sabbatique au printemps plutôt qu'à l'automne (Espeiland et Sauder, 2007) ;
- Dans le but d'illustrer à quel point le taux d'encadrement des étudiants, mesuré par le ratio étudiants/enseignants, est peu fiable, Marny Scully, de l'Université de Toronto, a démontré que, sur la base des mêmes données, sa valeur pouvait varier de 6/1 à 39/1 (Baty, 2010b) ;
- Les salaires des enseignants-chercheurs sont également pris en considération dans les classements ; certains exemples montrent que des établissements ont augmenté les salaires pour cette raison ;
- Il existe un cas bien connu au Royaume-Uni où, en 2008, il a été demandé à plus de 100 étudiants de faire l'éloge de leur université pour améliorer sa position dans les classements ;

- Lorsqu'un classement repose notamment sur une enquête de réputation concernant d'autres établissements, il est manifestement dans l'intérêt de l'EES de manipuler ces résultats (CHERPA, 2010). C'est pourquoi le CHE supprime progressivement de ses classements les enquêtes réalisées auprès des personnels d'enseignement et de recherche ;
- Le groupe de travail AUBR a découvert que même les indicateurs bibliométriques peuvent être corrompus à cause de manipulations de données (AUBR, 2010, p.13).

Amélioration des scores et accomplissement des différentes missions des EES.

Les efforts déployés par les universités européennes pour améliorer leurs positions dans les classements peuvent les détourner de certaines missions qui n'influencent pas directement leurs scores. Les auteurs du rapport Trends 2010 de l'EUA mettent en garde contre le développement d'indicateurs et de statistiques à des fins de classement, compte tenu des effets mitigés que cela pourrait avoir sur les établissements et les étudiants de l'enseignement supérieur. Cette évolution pourrait affaiblir l'importance des partenariats dans le cadre du processus de Bologne ainsi que le développement et l'amélioration de la qualité, et par conséquent vider d'une bonne partie de sa substance la philosophie à la base du processus de Bologne (Sursock & Smidt, 2010, p.27). En d'autres mots, on peut affirmer qu'il est peu probable que la mise en œuvre du processus de Bologne contribue à améliorer le score des universités européennes dans les classements, et que le « besoin » d'améliorer leurs résultats dans les classements risque de conduire les universités à affecter des ressources à des actions dont elles attendent qu'elles aident à remplir cet objectif.

Accorder trop d'attention à l'amélioration des scores dans les classements est, par conséquent, susceptible de nuire à l'accomplissement d'autres missions importantes des établissements d'enseignement supérieur.

IV. PRINCIPALES CONCLUSIONS

- 1.** Il ne fait aucun doute que l'apparition des classifications et classements internationaux d'universités a galvanisé le monde de l'enseignement supérieur. Depuis l'apparition des classements mondiaux, les universités n'ont pas pu éviter les comparaisons nationales et internationales, ce qui a entraîné des changements dans le mode de fonctionnement des universités.
- 2.** Les classements, et en particulier les classements mondiaux publiés sous forme de palmarès, ont adopté des méthodologies qui s'appliquent uniquement aux universités de recherche les plus réputées. De fait, ces méthodologies ne produisent des résultats stables que pour 700 à 1000 universités, qui ne représentent qu'une faible proportion des 17 000 universités dans le monde. La majorité des universités du monde n'entrent pas dans l'équation. Si cette approche peut être utile pour établir des listes d'universités les plus réputées, le tourbillon d'activité que suscitent les classements, souvent à l'initiative des producteurs de classements eux-mêmes, se répercute sur l'ensemble des milieux universitaires. Cette agitation qui accompagne la publication de ces classements donne à penser que l'ensemble des établissements d'enseignement supérieur sont concernés, alors que les critères utilisés ne sont appropriés que pour les universités de recherche les plus renommées.
- 3.** Jusqu'à présent, les classements ne portent que sur certaines missions de l'université. Peu d'entre eux en examinent la diversité, que ce soit au niveau des types d'établissement ou des missions qui leurs sont confiées.
- 4.** D'aucuns affirment que les classements contribuent à la « transparence » des universités. Toutefois, les méthodologies des classements existants, en particulier des palmarès les plus réputés, continuent à manquer elles-mêmes de transparence. Il est difficile, voire impossible, sur la seule base d'informations publiques, de suivre les calculs selon lesquels les données brutes sont converties en valeurs d'indicateurs, à leur tour combinées en vue d'établir un score global.
- 5.** « Il n'existe pas d'indicateur objectif » (AUBR, 2010). Le manque d'indicateurs appropriés devient évident quand il s'agit de mesurer la qualité de l'enseignement universitaire, pour laquelle il n'existe pas d'indicateur adapté. L'évaluation de la performance de la recherche pose moins de problèmes. Cependant, même les indicateurs bibliométriques utilisés pour mesurer la performance de la recherche induisent des biais et présentent des défauts. Les producteurs de classements s'efforcent d'améliorer les méthodologies utilisées, mais ces améliorations portent généralement sur la méthode de calcul, alors que le vrai problème réside dans l'utilisation d'indicateurs indirects inadéquats ou dans l'omission d'une partie des informations imputable aux contraintes méthodologiques. Les indicateurs indirects peuvent être améliorés, mais ils restent des indicateurs indirects.
- 6.** Actuellement, il serait difficile d'affirmer que les avantages des informations présentées par les classements, ainsi que la plus grande « transparence », surpassent les effets négatifs de ce qu'il est convenu d'appeler les « conséquences indésirables » des classements.
- 7.** Les récentes initiatives visant à développer des classifications, des classements et des évaluations de l'ensemble des établissements d'enseignement supérieur et de leurs différentes missions, comme l'évaluation de la recherche universitaire européenne commissionnée par l'UE (AUBR), les projets U-Map, U-Multirank et AHELO, ont toutes pour intention d'améliorer la situation. Il est toutefois trop tôt pour déterminer dans quelle mesure ils atteindront leurs objectifs. Ils en sont encore à différents stades de développement ou à la phase pilote de mise en œuvre et doivent relever de multiples défis, liés notamment à la collecte de données et à l'élaboration de nouveaux indicateurs.
- 8.** Les décisions politiques relatives à l'enseignement supérieur ne devraient pas se fonder uniquement sur les données présentées par les classements.

V. INTERPRÉTER LES RÉSULTATS DES CLASSEMENTS

Indications générales

1. Identifier l'objet et les publics cibles du classement. Pour certains classements, ces informations permettent de comprendre la sélection des indicateurs et d'autres caractéristiques. Dans d'autres cas, elles peuvent éventuellement révéler que le classement ne répond pas aux objectifs affichés.

2. Déterminer quels éléments sont effectivement mesurés et les comparer aux informations présentées sur les indicateurs ou groupes d'indicateurs utilisés. Ceci permet de déterminer si une valeur correspond à la mesure effective d'un aspect particulier ou si elle est le résultat d'un indicateur indirect.

3. Vérifier s'il est possible de suivre la méthode de calcul utilisée pour établir le résultat de chaque indicateur à partir de données brutes. Ceci permet de comprendre ce que la valeur finale représente réellement.

- Les indicateurs ont des dimensions différentes. C'est pourquoi les résultats doivent être recalculés afin de supprimer leur caractère dimensionnel et de les convertir en un score final unique. Pour ce faire, le résultat de l'université en question est souvent divisé par le résultat de l'université qui affiche la valeur la plus élevée au regard de cet indicateur, et le résultat obtenu est ensuite multiplié par 100. Par conséquent, le chiffre qui figure dans le tableau ne correspond généralement pas au nombre de publications, de citations, d'étudiants, etc. (comme indiqué en titre de colonne), mais au pourcentage du meilleur résultat.

- Vérifier si les résultats sont normalisés selon un aspect particulier. Les résultats des indicateurs de recherche sont normalisés par champ disciplinaire dans certains classements. C'est-à-dire que les résultats sont rapportés, par exemple, au nombre moyen de publications par membre du personnel ETP dans ce champ particulier, au nombre de citations par article dans ce champ particulier, etc. Cette méthode permet d'établir des comparaisons, mais elle implique également que le score de l'indicateur ne correspond pas à la mesure effective, mais représente une valeur composée qui fait intervenir une multiplication ou une division par certains facteurs – qui résultent d'un choix fait par les producteurs du classement. À cet égard, il serait utile de connaître les valeurs de ces facteurs, mais les présentations simplifiées des méthodologies utilisées, publiées avec les classements, n'apportent généralement pas ces précisions.

La normalisation par champ disciplinaire n'est pas le seul processus de normalisation appliqué. Le classement THE-Thomson Reuters 2010 applique par exemple un mécanisme de « prise en considération de facteurs internationaux » (Pratt, 2010). Cependant, le site Internet public n'apporte aucune précision sur ces facteurs régionaux.

4. Accorder une attention particulière à la pondération des indicateurs et des groupes d'indicateurs – les différences peuvent être importantes. À titre d'exemple, dans le classement THE-Thomson Reuters 2010, le poids des citations représente 32,5 %, contre 2,5 % pour le poids attribué au financement de la recherche par l'industrie. L'indicateur de citations est par conséquent 13 fois plus important que l'impact du financement de la recherche par l'industrie.

5. Examiner le mode de collecte de données et identifier les sources de données. La compréhension du classement est facilitée lorsqu'il est possible de connaître la proportion de données objectives qui peuvent être vérifiées, de données basées sur des avis exprimés par des pairs ou des employeurs, de données basées sur des opinions formulées par des étudiants ou des membres du personnel de l'université, de données extraites de statistiques nationales, ou de données autodéclarées par des universités.

Spécificités des différents classements d'universités

Academic Ranking of World Universities (ARWU) - Shanghai Ranking Consultancy

- Le classement ARWU ne concerne pas toutes les universités. Seules répondent aux critères de sélection environ 1 200 universités qui comptent parmi les plus réputées au monde ;
- Le classement ARWU est parfaitement adapté à son objectif initial : comparer les universités chinoises aux meilleures universités de recherche des États-Unis. S'il est utilisé comme un classement mondial, ses résultats peuvent avoir des conséquences négatives. En particulier lorsque cet instrument de comparaison de la capacité de recherche de l'élite mondiale des universités de *recherche* est utilisé afin d'évaluer la performance d'établissements divers ;
- Les 21 disciplines sur lesquelles porte le classement ARWU illustrent la prépondérance des sciences naturelles, de la médecine et, dans une certaine mesure, des sciences de l'ingénieur, tandis que les sciences humaines et les domaines interdisciplinaires sont délibérément écartés ;
- Le classement ARWU compare la performance totale des universités : tous les indicateurs, à l'exception d'un seul, sont basés sur des valeurs absolues (nombre de lauréats du Prix Nobel, de chercheurs fréquemment cités, d'articles, etc.). Bien que le dernier indicateur soit exprimé comme valeur rapportée au nombre d'enseignants-chercheurs, il ne représente que 10 % du score total. En d'autres mots, la méthodologie ARWU favorise les grandes universités ;
- Il s'ensuit que les fusions d'universités permettent une amélioration des résultats dans le classement ARWU, tandis que la scission des facultés, en particulier dans les domaines de la médecine, des sciences naturelles ou des sciences de l'ingénieur, affaiblit la position d'un établissement ;
- L'amélioration de la qualité de l'enseignement ou le renforcement de l'engagement régional n'améliore pas les scores d'un établissement dans le classement ARWU ;
- Les succès enregistrés dans les domaines des arts et des sciences humaines sont sans effet sur la position dans le classement ARWU. On ne peut qu'espérer que des universités ne décideront pas d'économiser dans ces domaines afin de se renforcer dans les sciences naturelles et la médecine, qui exercent une forte influence sur les scores dans le classement ARWU ;
- La publication de livres est sans effet sur la valeur de l'indicateur de publications ;
- Le classement ARWU est-il utile aux étudiants dans leur choix d'un établissement ? Pour les étudiants excellents qui souhaitent étudier les sciences naturelles, la médecine et les sciences de l'ingénieur : oui. Savoir s'il est utile à de jeunes gens talentueux qui recherchent « simplement » une formation universitaire de qualité peut prêter à discussion ;

- Le classement ARWU favorise les universités très performantes en sciences, principalement dans les pays de langue anglaise. L'anglais est généralement la langue de la recherche. Les travaux dans d'autres langues sont moins publiés et moins cités. Par ailleurs, en ce qui concerne les universités des États-Unis, il semble que les Américains tendent surtout à citer d'autres Américains (Altbach, 2006).

THE-QS World University Ranking (2004-2009)

- Les classements du THE examinent seulement les universités de réputation mondiale ;
- Les critères retenus par la méthodologie du THE correspondent aux « quatre piliers » des universités de classe mondiale. Cette méthodologie recense un total de quelque 600 universités et environ 300 universités dans chaque domaine ;
- L'« évaluation par les pairs » se résume à un sondage d'opinion réalisé auprès de personnels d'enseignement et de recherche et ne comprend pas de visite d'experts sur site, à la différence des évaluations réalisées dans le cadre des procédures d'assurance qualité ;
- Les pairs ou les employeurs sélectionnent les universités qu'ils considèrent comme les meilleures, sur la base de listes préétablies, desquelles sont exclues de nombreuses universités, voire des pays entiers ;
- Deux enquêtes contribuent à hauteur de la moitié du score global : l'une auprès des pairs, l'autre auprès des employeurs. Les taux de réponses aux deux enquêtes sont faibles. Les autres données, à l'exception du nombre de citations, sont fournies par les universités elles-mêmes ;
- Le seul indicateur indirect utilisé pour mesurer la qualité de l'enseignement est le taux d'encadrement des étudiants, mesuré par le ratio étudiants/enseignants ;
- Une pondération régionale est appliquée dans un souci d'équilibre entre les différentes régions du monde ;
- La normalisation des données et des z-scores sont utilisés pour calculer le score final.

THE-Thomson Reuters World University Ranking (2010)

- Le Times Higher Education a débuté sa coopération avec Thomson Reuters à l'automne 2009, après la publication du dernier classement THE-QS ;
- Le classement THE-Thomson Reuters reste axé sur les universités de réputation mondiale. Il exclut explicitement plusieurs catégories d'universités, en particulier les « graduate schools » et les universités ayant publié moins de 50 articles.

- La description de la méthodologie telle qu'elle est présentée actuellement ne permet pas de reconstituer le calcul des scores ;
- Le classement 2010 du THE se réfère aux publications de la période 2004-2008, excluant ainsi celles de 2009 ;
- Les indicateurs bibliométriques représentent la proportion la plus importante de la pondération totale (37 %) ;
- Les résultats de l'indicateur de citations sont normalisés. Le nombre de citations par article est rapporté au nombre moyen de citations recueillies par l'ensemble des articles publiés dans le même champ disciplinaire au cours de la même année ;
- Les enquêtes de réputation de la recherche et de l'enseignement conservent un impact énorme : leur poids cumulé représente plus d'un tiers du score global (34,5 %) ;
- Les valeurs des indicateurs dans ce classement sont relatives (par enseignant-chercheur, par publication, etc.). Par conséquent, le score dans le classement ne dépend pas de la taille de l'établissement.

World's Best Universities Ranking - US News & World Report en coopération avec QS

- Jusqu'à présent, la méthodologie est identique à la méthodologie utilisée pour le classement THE-QS 2009.

Classement mondial des universités – Reitor

- L'objet de ce classement est de déterminer la position des universités russes et d'élaborer une méthodologie plus adaptée aux caractéristiques spécifiques de l'enseignement supérieur en Russie ;
- Les universités invitées à participer à l'établissement de ce classement sont celles qui affichent les meilleurs résultats dans les classements ARWU, THE et HEEACT, en outre des universités situées dans les pays de l'ex-URSS. Toute université peut néanmoins y participer ;
- Ce classement n'a été établi qu'une fois, en 2008. Des intentions d'en faire un classement mondial publié régulièrement ont été formulées, mais aucune activité n'a été recensée jusqu'à présent ;
- Le nombre d'universités russes dans le Top 500 dans le classement Reitor s'élève à 69, contre 7 dans le classement ARWU 2009 et 4 dans le classement 2009 du THE-QS ;
- La structure du classement manque de transparence :
 - Les sections « méthodologie » et « À propos du classement » du site Internet de Reitor présentent des informations contradictoires.
 - Un certain nombre de décisions importantes ont été laissées à la discrétion de l'équipe d'experts de Reitor ; ces décisions ne sont pas publiées.
 - La pondération des différents indicateurs n'a pas été publiée, seuls sont présentés les poids par catégorie.

- Aucune précision n'est apportée quant à la méthode de calcul des scores.

- Le système d'indicateurs présenté initialement dans la section « Méthodologie » du site Internet de Reitor n'a en réalité pas été utilisé. C'est en revanche une série d'indicateurs différente qui a été utilisée. Elle comprend davantage d'indicateurs bibliographiques et des indicateurs pour lesquels des données peuvent être recueillies sur les sites Internet des universités ou extraites des statistiques nationales sur l'enseignement supérieur.

Leiden Ranking – Université de Leiden

- Le classement de Leiden mesure uniquement la production scientifique ;
- Seuls des indicateurs bibliométriques sont utilisés ;
- Le classement de Leiden n'établit aucun score global ; les universités sont classées séparément, selon chaque indicateur ;
- Les indicateurs reflètent le nombre total de publications (P) et le nombre de citations par publication – en tant que tel (indicateur CPP), normalisé par champ disciplinaire (indicateur CPP/FCSm) et le *mean normalised citation score* (indicateur MNCS2). Un indicateur reflétant la performance totale d'une université dans le monde (son rayonnement académique) est également utilisé ($P \cdot CPP/FCSm$) ;
- Les deux « *crown indicators* » ne tiennent pas compte des publications les plus récentes, le CPP/FCSm parce que les articles plus récents ont recueilli moins de citations que les articles plus anciens, et le MNCS2 parce que les publications de l'année précédente sont exclues délibérément ;
- Le classement de Leiden ne tient pas compte de la diversité ou de la variété des missions de l'université ;
- Pour figurer dans le classement, une université doit compter un nombre élevé de publications recensées par les index de citations ;
- Deux indicateurs sont dépendants de la taille (c.-à-d. qu'ils favorisent les grandes universités) : (P) et l'indicateur de « force brute » ($P \cdot CPP/FCSm$).

Performance Rankings of Scientific Papers for World Universities (HEEACT) – Taiwan Higher Education Accreditation and Evaluation Council

- Le HEEACT classe les universités exclusivement sur la base d'indicateurs bibliométriques. Il est axé sur la productivité, l'impact et l'excellence de la recherche. Le classement HEEACT répond par conséquent à son objectif de classer les universités purement sur la base de la performance de la recherche ;
- Le classement HEEACT met l'accent sur les performances actuelles de la recherche des établissements, ce qui le distingue des classements THE ou ARWU ;

- Le classement HEEACT ne classe pas toutes les universités du monde. Il sélectionne 700 universités parmi les plus réputées pour son classement global et environ 500 universités de premier plan dans chaque domaine ;
- À la différence du classement de Leiden, qui tient compte à la fois des universités et d'autres organismes de recherche, le classement HEEACT examine uniquement la recherche universitaire. Pour figurer dans les classements par champ disciplinaire, l'université en question doit proposer des programmes de licence dans le domaine concerné ;
- Le classement HEEACT se fonde sur les index de citations SCI et SSCI, à l'exclusion de l'index des citations des sciences humaines, qui sont par conséquent exclues du classement ;
- À l'instar du classement ARWU, le HEEACT couvre les 21 disciplines de l'ESI, avec pour conséquence, ici encore, la prépondérance des sciences naturelles, de la médecine et, dans une certaine mesure, des sciences de l'ingénieur ;
- Le classement HEEACT tente de pallier le biais lié à la taille de l'université (à la différence, par exemple, des classements ARWU ou de Leiden) : 50 % des indicateurs sont calculés par enseignant-chercheur (ETP) ;
- Conformément à son objectif affiché, le HEEACT se désintéresse de l'enseignement et de l'apprentissage et ne tient pas compte des autres missions de l'université et de la diversité des établissements d'enseignement supérieur ;
- Le HEEACT tente d'éviter les cas dans lesquels plusieurs universités font partie d'un système universitaire global (ex. les systèmes universitaires aux États-Unis). Ces systèmes sont considérés comme une seule université. De même, les départements dépendants d'une université sont pris en considération pour mesurer la performance de cette université le cas échéant ;
- Comme la position dans le classement HEEACT n'est pas influencée par la réussite dans les arts et les sciences humaines, le risque existe de voir des universités réaffecter des ressources aux domaines qui influencent les scores dans le classement HEEACT, afin d'améliorer leur position dans ce classement ;
- La publication de livres est sans effet sur la valeur de la plupart des indicateurs ;
- Le classement HEEACT est-il utile aux étudiants dans leur choix d'une université ? Pour les étudiants qui veulent faire carrière dans la recherche dans les sciences naturelles, la médecine, les sciences de l'ingénieur et, en partie, dans les sciences sociales : dans une certaine mesure, oui. Pour les autres : non.

Assesment of University-Based research – Commission européenne

- AUBR n'est pas un classement d'universités. Il s'agit d'une méthodologie d'évaluation de la recherche universitaire ;
- La méthodologie AUBR propose une approche fondée sur des indicateurs multiples, sans calcul d'un score global et, par conséquent, ne vise pas à établir de palmarès ;

- Le groupe de travail AUBR a analysé les points forts et les faiblesses de différents indicateurs de recherche ;
- Des combinaisons d'indicateurs adaptées à l'objet de l'évaluation sont présentées dans la grille d'évaluation multidimensionnelle de la recherche ;
- Les conclusions du groupe de travail AUBR concernant différents indicateurs sont utiles dans le cadre d'une analyse des classements mondiaux d'universités.

CHE University Ranking

- Le CHE est un classement d'universités à indicateurs multiples qui a pour principal objet d'aider les étudiants à choisir un établissement d'enseignement supérieur approprié, mais aussi de fournir des informations utiles aux établissements d'enseignement supérieur eux-mêmes ;
- Le classement n'a pas pour objet de calculer le score final des universités considérées. En revanche, les universités sont incluses dans des groupes supérieur, moyen et inférieur en fonction de leurs réalisations au regard d'aspects particuliers ;
- Le classement du CHE porte sur les universités dans les pays de langue allemande et les universités dans lesquelles l'enseignement est proposé en allemand ;
- L'utilisateur peut choisir entre différents indicateurs dans les groupes présentés et réaliser un classement personnalisé qui affiche les universités les mieux adaptées aux conditions définies ;
- Les indicateurs couvrent l'enseignement et la recherche, mais aussi d'autres aspects, comme les bâtiments universitaires, la taille du campus, l'offre sportive, etc ;
- La liste d'indicateurs à sélectionner présentée à l'utilisateur diffère selon le choix préalable entre universités ou universités de sciences appliquées (*Fachhochschulen*) ;
- La plupart des indicateurs actuellement utilisés dans le classement universitaire du CHE sont surtout basés sur les évaluations faites par les étudiants au regard de différents aspects des universités, et nettement moins sur les évaluations réalisées par les enseignants-chercheurs. Des indicateurs reposent également sur certains « faits » (statistiques, données bibliométriques). La méthodologie semble évoluer dans le sens d'une plus grande importance donnée aux évaluations réalisées par les diplômés, aux statistiques et aux données bibliométriques.

Classification U-Map – Center for Higher Education Policy Studies (CHEPS) de l'Université de Twente – Pays-Bas

- U-MAP est un outil de classification à indicateurs multiples. Aucun score final n'est calculé. À l'exception de trois indicateurs, chaque aspect est qualifié de *major*, *substantial*, *some* et *none* ;
- U-Map a été conçu pour classer l'ensemble des EES en Europe, indépendamment du type d'établissement, de leur orientation ou spécialisation, etc ;

- Classification plutôt que classement, U-Map se fonde sur des indicateurs qui caractérisent l'orientation et l'intensité de différentes activités des EES plutôt que la performance, l'impact ou la qualité ;
- U-Map dispose de deux outils de visualisation : le *Profile Finder*, permettant d'identifier les établissements d'enseignement supérieur qui correspondent aux caractéristiques établies par l'utilisateur, et le *Profile Viewer*, permettant de comparer dans le détail jusqu'à trois EES sélectionnés ;
- Les indicateurs portent sur le niveau d'enseignement et le domaine d'étude, la population étudiante, l'intensité des activités de recherche, le transfert de connaissances, l'orientation internationale et l'engagement régional ;
- Le manque de données comparables à l'échelle internationale constitue un défi pour U-Map. La Commission européenne et Eurostat ont lancé des initiatives visant à recueillir des données comparables à travers l'Europe. Cependant, avant la mise en œuvre complète de ces plans, U-Map devra se fonder sur des données nationales et institutionnelles et sera par conséquent plus approprié pour comparer des institutions au sein d'un même pays plutôt qu'à l'échelle internationale ;
- Une autre condition préalable du succès de cet outil européen de classification est l'adoption de définitions communes à l'échelle européenne de termes comme « enseignant-chercheur », « enseignant », « chercheur », « publications évaluées par les pairs » (notamment en ce qui concerne les livres et les monographies) ;
- U-Map est un nouvel instrument, mis au point au début 2010. Il est encore en phase de test ; des EES des pays volontaires procèdent à la saisie de leurs données ;
- Il est évident que les responsables du projet ont tout fait pour éviter que des tiers n'utilisent les données des indicateurs de U-Map pour établir un palmarès. En effet, rassembler toutes les informations nécessaires pour établir un palmarès sur la base des données de U-Map supposerait un travail colossal. Toutefois, ce n'est pas absolument impossible.

European Multidimensional University Ranking System (U-Multirank) – projet financé par l'UE

- U-Multirank est un classement multidimensionnel actuellement développé avec le soutien de l'UE. Ce projet s'inscrit dans le prolongement de U-Map, la classification européenne des EES ;
- U-Multirank tiendra compte de l'ensemble des activités d'un EES, y compris l'enseignement, la recherche, le transfert de connaissances et l'engagement régional ;
- Deux types de classements sont en cours de développement :

1. Les *classements spécifiques des établissements*, qui permettent des comparaisons d'établissements selon une seule dimension de l'activité institutionnelle, comme l'enseignement, la recherche, l'engagement régional, etc.

2. Les *classements par champ disciplinaire*, qui permettent des comparaisons des programmes d'études dans le même domaine dans un groupe d'établissements au profil similaire.

- Aucun score final ne sera calculé. Cependant, il reste à savoir s'il sera possible de protéger les résultats de façon à éviter que des tiers ne convertissent les résultats du classement en un palmarès ;
- Ce classement sera basé sur des indicateurs multiples. Certains indicateurs seront utilisés à la fois dans le classement des établissements et le classement par champ disciplinaire, tandis que d'autres ne le seront que dans l'un des deux ;
- Les indicateurs portent sur le niveau d'enseignement et le domaine d'étude, la population étudiante, l'intensité de la recherche, le transfert de connaissances, l'orientation internationale et l'engagement régional ;
- En dépit des critiques et des inconvénients mis en évidence dans le rapport intérimaire, U-Multirank reposera essentiellement sur les informations fournies par Thomson Reuters et Scopus en ce qui concerne les éléments bibliométriques et sur les données communiquées par les EES eux-mêmes sur les étudiants, les personnels enseignants et la recherche (à l'exception des publications et des citations). Un vaste ensemble de données sera tiré d'enquêtes de satisfaction des étudiants. Les enquêtes auprès des personnels enseignants ne seront pas utilisées ;
- Les résultats du classement sont destinés aux étudiants, aux enseignants-chercheurs mais aussi à d'autres parties prenantes au sein de la société ;
- Des analyses complémentaires seront présentées au terme de la phase de développement de U-Multirank ;

Webometrics Ranking of World Universities

- L'objet initial du classement Webometrics était d'inciter les milieux académiques à publier davantage en ligne. Webometrics vise à promouvoir les initiatives de libre accès et l'accès électronique aux publications scientifiques et à d'autres contenus académiques. Le classement Webometrics tient compte de tous les établissements d'enseignement supérieur qui disposent de leur propre domaine, tandis que la plupart des classements mondiaux portent exclusivement sur les universités de recherche qui font partie de l'élite scientifique ;
- Webometrics mesure deux aspects principaux : le *volume du site Internet de l'université*, représenté par trois indicateurs, soit le nombre de pages sur le site Internet de l'université, le nombre de publications et le nombre de fichiers disponibles jugés pertinents dans un contexte académique (.pdf, .ppt, .doc et .ps) ; la *visibilité* de l'université sur Internet, représentée par un indicateur, à savoir le nombre de liens entrants externes (c'est-à-dire les liens établis à partir d'autres sites vers le site Internet de l'université) ;
- L'indicateur de *volume* est utilisé comme indicateur indirect de la production académique de l'université, encore que d'une manière particulière, en regroupant les publications

de recherche, les présentations, le matériel didactique, des données brutes, des versions préliminaires et d'autres documents relatifs à la recherche et à l'enseignement, ainsi que des informations de nature administrative publiées par l'université ;

- L'établissement d'un lien vers le site Internet de l'université peut être comparé à une citation – ce lien est établi comme référence à une information présentée sur le site Internet et pour permettre à des tiers d'accéder rapidement à l'information initiale. Le nombre de liens entrants vers le site Internet de l'université est par conséquent utilisé comme indicateur indirect des citations ;
- Bien que les indicateurs indirects semblent assez lointains, on observe une bonne corrélation entre les résultats de Webometrics et ceux des autres classements mondiaux ;
- 20 000 EES sont en cours d'analyse, et les listes du classement Webometrics recensent 12 000 universités ;
- Webometrics utilise des données fournies par quatre moteurs de recherche : Google, Yahoo, Live Search et Exalead ;
- Les producteurs du classement Webometrics ont rendu publics tous les biais pertinents dans le cadre de leur analyse (ex. les biais géographiques et linguistiques induits par les moteurs de recherche).

- L'équipe du projet est consciente des difficultés potentielles liées à l'utilisation d'outils d'évaluation à l'échelle internationale, c'est pourquoi elle rassemble des informations de référence susceptibles de contribuer à la réalisation de comparaisons internationales ;

- Des analyses complémentaires seront présentées au terme de la phase actuelle du projet AHELO.

Évaluation des résultats de l'enseignement supérieur (AHELO) – OCDE

- Le projet AHELO de l'OCDE vise à comparer les EES à l'échelle internationale sur la base des acquis de l'apprentissage ;
- Le projet AHELO a été lancé pour déterminer les possibilités de mesure et de comparaison des acquis de l'apprentissage dans l'enseignement supérieur à l'échelle internationale ;
- Le projet ambitionne non seulement de mesurer les acquis de l'apprentissage en tant que tels, mais aussi de fournir des informations sur les stratégies d'enseignement efficaces pour garantir l'acquisition des résultats de l'apprentissage. L'objectif est de permettre à davantage d'étudiants de terminer leurs études avec succès, encourageant ainsi l'équité dans les différents systèmes d'enseignement supérieur. AHELO vise par ailleurs à mettre à la disposition des pouvoirs publics un outil leur permettant de mesurer l'efficacité des systèmes d'enseignement supérieur ;
- L'outil d'évaluation des compétences génériques est basé sur le *US Collegiate Learning Assessment*, qui impose aux étudiants d'utiliser un ensemble intégré de compétences axé sur la réflexion critique, le raisonnement analytique, la résolution des problèmes et la communication écrite ;
- En ce qui concerne les outils d'évaluation des compétences spécifiques à certaines disciplines, l'approche du projet Tuning est utilisée comme base sur laquelle seront définis l'ensemble des résultats de l'apprentissage attendus et souhaités ;
- Chaque volet du projet sera évalué dans plusieurs pays dans différentes régions du monde. Des étudiants de 10 EES dans chaque pays participant seront évalués à l'approche du terme de leur premier cycle d'études (licence) ;

BIBLIOGRAPHIE

- Altbach. (2006). The Dilemmas of Ranking. International Higher Education, No. 42, winter 2006.
- Annapolis group (2007). Annapolis group statement on rankings and ratings, 19 June 2007. Retrieved on 12 Jan. 2011 from: <http://www.collegenews.org/x7131.xml>
- Aquillo I., Ortega J., Fernandez, M. (2008). Webometric Ranking of World Universities: Introduction, Methodology, and Future Developments. Higher Education in Europe, Vol. 33, No. 2/3, July–October 2008, pp. 233-244.
- Aquillio, I. (2009). Measuring the institution's footprint in the web. Library Hi Tech, Vol. 27 No. 4, pp. 540-556.
- Aguillo I., Bar-Ilan J., Levene, M., Ortega J. (2010). Comparing university rankings. Scientometrics, 85, pp. 243-256.
- ARWU. (2009). Methodology. Retrieved on 12 Jan. 2011 from: <http://www.arwu.org/ARWUMethodology2009.jsp>
- Assessment of University-Based Research Expert Group (AUBR). (2010). Assessing Europe's University-Based Research, K1-NA-24187-EN-N, European Commission, Brussels, pp. 151. Retrieved on 12 Jan. 2011 from: <http://ec.europa.eu/research/era/docs/en/areas-of-actions-universities-assessing-europe-university-based-research-2010-en.pdf>
- Baty, P. (2009). New data partner for World University Rankings. Times Higher Education, 30 Oct. 2009. Retrieved on 1 Nov. 2010 from: <http://www.timeshighereducation.co.uk/story.asp?storycode=408881>
- Baty, P. (2010a). THE University rankings, 25 March 2010. <http://www.timeshighereducation.co.uk/story.asp?storycode=410974>
- Baty, P. (2010b). THE World University Rankings 27 May 2010. Retrieved on 11 Jan. 2011 from: <http://www.timeshighereducation.co.uk/story.asp?storyCode=411739§ioncode=26>
- Baty, P. (2010c). THE unveils broad, rigorous new rankings methodology, Times Higher Education, 3 June 2010. Retrieved on 12 Jan. 2011 from: <http://www.timeshighereducation.co.uk/story.asp?sectioncode=26&storycode=411907&c=1>
- Baty, P. (2010d). We have had to make some extraordinarily difficult decisions this week - on which institutions to exclude from our 2010 World University Rankings. THE World University Rankings 5 August 2010. Retrieved on 12 Jan. 2011 from: <http://www.timeshighereducation.co.uk/story.asp?sectioncode=26&storycode=412832&c=1>
- Baty, P. (2010e). Robust, transparent and sophisticated. Times Higher Education, 16 Sep 2010. Retrieved on 30 Dec 2010 from: <http://www.timeshighereducation.co.uk/world-university-rankings/2010-2011/analysis-methodology.html>
- Berghoff, S., Federkeil G., Giebisch, P., Hachmeister, C., Hennings, M., Roessler I., and Ziegele F. (2009). CHE Hochschulranking Vorgehensweise und Indikatoren, CHE Arbeitspapier 19 May 2009, pp. 90. Retrieved on 12 Jan. 2011, from: http://www.che.de/downloads/CHE_AP119_Methode_Hochschulranking_2009.pdf
- Berghoff, S., Brandenburg, U., Carr-Boulay, D., Hachmeister, C., Hennings, M., Leichsenring, H., and Ziegele F. (2010). Identifying the best: The CHEE Excellence Ranking 2010. Centre for Higher Education, Working paper, No. 137, pp. 49. Retrieved on 25 Feb 2011 from: http://www.che.de/downloads/CHE_AP137_ExcellenceRanking_2010.pdf
- Boulton, G. (2010). University rankings: diversity, excellence and the European initiative. Advice paper, No. 3, League of European Research Universities (LERU). Retrieved 31 Dec. 2010 from: http://www.leru.org/files/publications/LERU_AP3_2010_Ranking.pdf
- Buela-Casal, G., Gutiérrez-Martínez, O., Bermúdez-Sánchez, M., Vadillo-Muño, O. (2007). Comparative study of international academic rankings of universities, Scientometrics, Vol. 71, No. 3, pp. 349-365.
- Cattau, D. (2010). 'Fight the college rankings madness', blog Opinion, Dallas News, 27 Aug 2010. Retrieved 30 Oct. 2010 from: http://www.dallasnews.com/sharedcontent/dws/dn/opinion/viewpoints/stories/DN-cattau_29edi.State.Edition1.13c4754.html
- CHE (2010). Methodology. Die Zeit website retrieved at Aug 17 2010 <http://ranking.Zeit.de/che2010/en/ueberdasranking/methodik>
- Cheng, Y. (2010). 'To better Measure social sciences performance. A review of existing ranking indicators'. Presentation at IREG 2010 conference, Berlin, 6-8 Oct. 2010. Retrieved 30 Oct. 2010 from: http://www.ireg-observatory.org/index.php?option=com_content&task=view&id=92&Itemid=86

- CHERPA. (2010). U-Multirank: Design Phase of the Project 'Design and Testing the Feasibility of a Multi-dimensional Global University Ranking', CHERPA Network, pp. 217. Retrieved 12 Jan. 2011 from: http://www.u-multirank.eu/UMR_IR_0110.pdf
- Doneckaja, S. (2009). Российский подход к ранжированию ведущих университетов мира. ЭКО, No. 8, pp. 137-150. Retrieved on 12 Aug. 2010 from: http://www.sibupk.su/stat/publicat_nir/smk_4.pdf
- EU Commission. (2009). Commission launches feasibility study to develop a multi-dimensional university ranking, EU commission press release IP/09/859, Brussels, 2 June 2009. Retrieved 12 Jan. 2011 from: http://europa.eu/rapid/pressReleasesAction.do?reference=IP/08/1942&format=HTML&age_d=0&language=EN&guiLanguage=en
- Garfield, E., (1994). The Thomson Reuters impact factor, published in ISI Current Contents print editions on 20 Jun 1994. Retrieved 11 Jan. 2011 from: http://thomsonreuters.com/products_services/science/free/essays/impact_factor/
- Federkeil, G. Ying C. (2010). IREG Ranking audit - purpose, criteria and procedure. Presentation at IREG-5 Conference, Berlin, 7 October 2010.
- Florian, R.V. (2007). Irreproducibility of the results of the Shanghai academic ranking of world universities. Scientometrics, Vol. 72, No. 1, pp. 25-32.
- Hare, J. Coming up, body to rank rankers. The Australian, 12 Jan 2011. Retrieved 12 Jan 2011 from: <http://www.theaustralian.com.au/higher-education/coming-up-body-to-rank-rankers/story-e6frgcjx-1225985832449>
- Hazelkorn, E. (2011). Rankings and the Reshaping of Higher Education. The Battle for World-Class Excellence. Plagrove.
- Hicks, D., Tomizawa, H., Saitoh, Y. and Kobayashi, S. (2004). Bibliometric techniques in the evaluation of federally funded research in the United States. Research evaluation, 13, pp. 78-96.
- Heylighen F. (1993): «Red Queen Principle», in: F. Heylighen, C. Joslyn and V. Turchin (editors): Principia Cybernetica Web (Principia Cybernetica, Brussels).
- Huang, M. (2010). Performance Ranking of Scientific Papers for World Universities. Retrieved 12 Jan 2011 from: <http://ranking.heeact.edu.tw/en-us/2010/Page/Background>
- IREG. (2006). Berlin Principles on Ranking of Higher Education Institutions adopted at IREG's 2nd meeting in Berlin, 18-20 May, 2006-
- IREG. (2010). IREG-Ranking Audit: Purpose, Criteria and Procedure, Draft version November 2010. Retrieved 10 Jan. 2011 from: http://www.ireg-observatory.org/index.php?option=com_content&task=view&id=110&Itemid=101
- King, R., Locke, W., et al. (2008). Counting what is measured or measuring what counts? League tables and their impact on higher education institutions in England. Hefce issues paper, 14 April 2008. Retrieved from: http://www.hefce.ac.uk/pubs/hefce/2008/08_14/
- Kruzhalinin V., and Artjushina, I. (2008). Rankings and quality of education. Retrieved 5 Aug. 2010 from: <http://reitor.ru/ru/observatory/global/index.php?id19=474>
- Lalancette D. (2010). Assessment of Higher Education Learning Outcomes: a ground-breaking initiative to assess quality in higher education on an international scale. Presentation at IREG-5 Conference, Berlin, 7 Oct. 2010. Retrieved 17 Dec. 2010 from: http://www.ireg-observatory.org/presentation/2010/S2_2_Lalancette.pdf
- Leiden ranking. (2008). <http://www.cwts.nl/ranking/Leiden-RankingWebSite.html>
- Leydesdorff L., and Opthof, T. (2010). Normalisation, CWTS indicators, and the Leiden Rankings: Differences in citation behavior at the level of fields. Retrieved on 17 Dec. 2010 from: <http://www.leydesdorff.net/reply2cwts/reply2cwts.pdf>
- Liu, N.C., and Cheng, Y. (2005). The Academic Ranking of World Universities. Higher Education in Europe, Vol. 30, No. 2, pp. 127-135.
- Liu, N. C. (2009). The Story of Academic Rankings. International Higher Education, No. 54, 2-3. winter 2009.
- Marginson, S. (2007). Rankings: Marketing Mana or Menace? The Big Picture. 16th Annual New Zealand International Education Conference, 8-10 August 2007, Christchurch, New Zealand. Retrieved 11 Jan 2011 from: http://www.cshe.unimelb.edu.au/people/staff_pages/Marginson/KeyMar8-10Aug07.pdf
- Marginson, S., and Wan der Wende, M. (2007). To Rank or To Be Ranked: The Impact of Global Rankings in Higher Education. Journal of Studies in International Education, Vol. 11, No. 3/4, Fall/Winter 2007, pp. 306-329. Retrieved 10 Jan. 2011 from: http://www.cshe.unimelb.edu.au/people/staff_pages/Marginson/Marginson&VanDeWendeJSIEFallWinter07.pdf
- Marmolejo, F. (2010). The madness of rankings. Chronicle of Higher education, 22 Sept. 2010. Retrieved 12 Jan. 2011 from: <http://chronicle.com/blogPost/The-Madness-of-Rankings-/27100/>

- Moskovkin, V. (2009). From shifting places... What is the use of global ranking for Russian universities? Poisk, newspaper of scientific society, 28 April 2009. Retrieved from: http://www.poisknews.ru/2009/04/28/ot_peremeny_mest.html
- Moed, H. (2005). Citation Analysis in Research Evaluation. Dordrecht, Springer, pp. 353.
- OECD. (2008). Measuring Improvements in Learning Outcomes: Best Practices to Assess the Value-Added of Schools. ISBN: 9789264050228.
- OECD. (2010). AHELO (assessment of higher education learning outcomes). Project update - December 2010. Retrieved 17 Dec. 2010 from: <http://www.oecd.org/dataoecd/8/26/45919053.pdf>
- Olds, K. (2010). Governing world university rankers: an agenda for much needed reform, 21 Sept. 2010 at global-highered. Retrieved from: <http://globalhighered.wordpress.com/2010/09/21/governing-world-university-rankers-an-agenda-for-much-needed-reform/>
- Pratt, S. (2010). Times Higher Education World University Rankings Global Institutional Profiles Project. Presentation at IREG-5 Conference, Berlin, 7 October 2010. Retrieved 12 Jan. 2011 from www.ireg-observatory.org/prezentacje/2010/S4_1_Pratt.pdf
- Rauhvargers, et al. (2009). Bologna Stocktaking report 2009, pp. 25, Retrieved 28 Dec. 2010 from: http://www.ond.vlaanderen.be/hogeronderwijs/bologna/conference/documents/Stocktaking_report_2009_FINAL.pdf
- Rehn C., Kronman, U., and Wadskog, D. (2007). Bibliometric indicators – definitions and usage at Karolinska Institutet. http://ki.se/content/1/c6/01/77/43/Bibliometric%20indicators%20-%20definitions_1.0.pdf
- U-Multirank. (2010). Interim Progress Report, CHERPA Network, January 2010, internet: www.u-multirank.eu/UMR_IR_0110.pdf
- Reitor. (2008a). Methodology. Retrieved 31 Dec. 2010 from: http://www.globaluniversitiesranking.org/index.php?option=com_content&view=article&id=68&Itemid=128
- Reitor. (2008b). Questionnaire. Retrieved 31 Dec. 2010 from: http://www.globaluniversitiesranking.org/images/banners/Questionnaire_1.doc
- Reitor. (2009a). About the ranking. <http://www.globaluniversitiesranking.org/index.php?view=article&id=96%3A2009-02-11-13-58-17&tmp> retrieved on 09 aug 2010
- Reitor. (2009b). From the ranking compilers (list of universities that actually answered). Retrieved 31 Dec. 2010 from: http://www.globaluniversitiesranking.org/index.php?option=com_content&view=article&id=95&Itemid=1
- Reitor. (2008c). Presentation at Moscow University. Retrieved on 6 Aug. 2010 from: http://www.reitor.ru/common/img/uploaded/files/msu_present_19-20.11.pdf
- Sadlak, J. and Liu N.C., Eds. (2007). The world-class university and ranking: Aiming beyond status. Bucharest UNESCO-CEPES, pp. 378. ISBN: 978-92-9069184-0.
- Saisana, M., and D'Hombres, B. (2008). Higher Education Rankings: Robustness Issues and Critical Assessment. How much confidence can we have in Higher Education Rankings? Luxembourg, Office for Official Publications of the European Communities. Retrieved 12 Jan. 2011 from: <http://global-highered.files.wordpress.com/2008/11/eur23487.pdf>
- Science and technology in the Russian Federation. (2008). Мировой рейтинг вузов: Российская версия, 11 Apr. 2008. Retrieved 12 Aug. 2010 from: http://strf.ru/organisation.aspx?CatalogId=221&d_no=14080
- Salmi, J., and Saroyan, A. (2007). League tables as policy instruments: Uses and misuses. Higher. Education Management and Policy, Vol. 19, No. 2, pp. 31-68.
- Salmi, J. (2009). The challenge of World class universities, World bank, Washington, pp. 136. Retrieved 12 Jan. 2011 from: <http://siteresources.worldbank.org/EDUCATION/Resoues/278200-1099079877269/547664-1099079956815/547670-1237305262556/WCU.pdf>
- Salmi, J. (2010). If ranking is the disease, is benchmarking the cure? Presentation at IREG 2010 conference, Berlin, 7-8 October 2010.
- Samarasekera, I. (2007). Rising Up Against Rankings. Inside Higher Education, April 2. Retrieved 12 Jan 2011 from: <http://www.insidehighered.com/views/2007/04/02/samarasekera>
- Stolz, I., Hendel D., and Horn, A. (2010). Ranking of rankings: benchmarking twenty-five higher education ranking systems in Europe. Higher Education, Online, pp. 22. ISSN 0018-1560. Springer, Netherlands.
- Sursock, A., Smidt H., et al. (2010). Trends 2010: A decade of change in European Higher Education. European University Association, Brussels, pp. 123.
- THE (2008). World University Rankings 2008. Page retrieved 29 Dec. 2010 from: <http://www.timeshighereducation.co.uk/hybrid.asp?typeCode=416&pubCode=1&navcode=137>

THE (2008b). World University Rankings 2008: Strong measures. Retrieved 29 Dec. 2010 from: <http://www.timeshighereducation.co.uk/story.asp?storycode=407247>

THE (2009). World University Rankings 2009. Retrieved 29 Dec 2010 from: <http://www.timeshighereducation.co.uk/hybrid.asp?typeCode=431&pubCode=1&navcode=148>

THE-QS (2009). Methodology: purpose and approach. Retrieved 28 Dec. 2010 from: <http://www.topuniversities.com/university-rankings/world-university-rankings/methodology/purpose-approach>

THE (2009b). Rankings 09: Talking points. Retrieved 29 Dec. 2010 from: <http://www.timeshighereducation.co.uk/story.asp?storycode=408562>

Tuning (2004). Tuning project methodology. Tuning Project, 2004. Retrieved 8 Jan. 2011 from: <http://tuning.unideusto.org/tuningeu/index.php?option=content&task=view&id=172&Itemid=205>

U-Map. (2009a). Overview of indicators and data-elements, by dimension. Retrieved 22 Sept. 2010 from: <http://www.u-map.eu/U-Map%20dimensions%20and%20indicators%20overview.pdf>

U-Map. (2009b). Overview of dimensions and indicators. Retrieved 12 Jan. 2011 from: <http://www.u-map.eu/U-Map%20dimensions%20and%20indicators%20overview.pdf>

van Vught, F., Kaiser, F., File, J., Gaethgens, C., Peter, R., and Westerheijden, D. (2010). The European Classification of Higher Education Institutions, CHEPS, Enschede, pp. 46. Retrieved 22 Sep 2010 from: http://www.u-map.org/U-MAP_report.pdf

van Vught, F., Bartelse, J., Huisman J., and van der Wende, M. (2005). Institutional profiles, Towards a typology of higher education institutions in Europe. CHEPS, Enschede. Retrieved 22 Sept. 2010 from: <http://doc.utwente.nl/53776/1/engreport05institutionalprofiles.pdf>

van Vught, F., Kaiser F., Bohmert D., File J., and van der Wende M. (2008). Mapping diversity; Developing a European classification of higher education institutions. Enschede, CHEPS.

van Vught, F., Ed. (2009). Mapping the Higher Education Landscape: Towards a European Classification of Higher Education. Dordrecht etc., Springer., pp. 170.

Vestrehijden D. et al. (2010). The first decade of working on the European Higher Education Area. The Bologna Process Independent Assessment, Vol. 1, pp. 39. Retrieved 28 Dec. 2010 from: http://www.ond.vlaanderen.be/hogeronderwijs/bologna/2010_conference/documents/IndependentAssessment_1_DetailedRept.pdf

Usher, A., and Savino, M. (2006). A World of Difference: A Global Survey of University League. Tables. Toronto, ON: Educational Policy Institute. Retrieved 12 Jan. 2011 from: <http://www.educationalpolicy.org/pdf/world-of-difference-200602162.pdf>

van Raan, A. (2005). Fatal Attraction: Ranking of Universities by Bibliometric Methods Fatal Attraction: Conceptual and methodological problems in the ranking of universities by bibliometric methods. Scientometrics Vol. 62, No. 1, pp. 133-143. Retrieved 10 Jan. 2011 from: <http://www.cwts.nl/TvR/documents/AvR-RankingScientom.pdf>

van Raan, A. (2010). Presentation at school of Scientometrics, Humboldt university 16 June. Retrieved 12 Jan 2011 from: http://www.scientometrics-school.eu/images/esss1_vanRaan.pdf

van Raan et al. (2010). Anthony F.J. van Raan, Thed N. van Leeuwen, and Martijn S. Visser Van. Germany and France are wronged in citation-based rankings. Retrieved 11 Jan. 2010 from: <http://www.cwts.nl/pdf/LanguageRanking22122010.pdf>

Waltman, L., van Eck, N., van Leeuwen, T., Visser, M., and van Raan, A. (2010). Towards a new crown indicator: An empirical analysis. Centre for Science and Technology Studies, Leiden University, The Netherlands. Retrieved 3 Jan. 2011 from: Centre for Science and Technology Studies, Leiden University, The Netherlands http://arxiv.org/PS_cache/arxiv/pdf/1004/1004.1632v2.pdf

van Noorden, R. (2010). A profusion of measures. Vol. 465, 17 June 2010, pp 864-866. Retrieved 12 Jan 2011 from: http://planetsci.uwo.ca/PS%20Seminar/PS%20Seminar/Papers/A_Profusion_of_Measures.pdf

