

# The Unbearable Lightness of Ranking

Nadine Rons  
Karen Vandevelde

Nadine Rons is coördinator Kwaliteitsevaluatie en Beleidsvoorbereidende Studies , Vrije Universiteit Brussel (nrns@vub.ac.be). Karen Vandevelde is Beleidsmedewerker Kwaliteitszorg, Universiteit Gent (Karen.Vandevelde@UGent.be).

**Het rangschikken van universiteiten op basis van output- en kwaliteitscriteria bestaat al geruime tijd. Universiteiten volgen de resultaten met argusogen, hoewel ze weten dat er sterke bedenkingen zijn bij die rankings.**

De methodologie, de volledigheid, de betekenis en de nauwkeurigheid van rankings laten wel eens te wensen over. De periode is evenwel voorbij dat universiteiten de impact van dergelijke rankings konden negeren. Stakeholders gaande van subsidiërende overheden tot professoren en studiekeizers volgen rankings tegenwoordig met heel wat meer aandacht dan pakweg vijf jaar geleden.

## **Interesse voor rankings**

De grote belangstelling voor rankings van universiteiten is sterk gerelateerd aan twee recente fenomenen in het hoger onderwijs. Een eerste fenomeen is de toenemende globalisering van onderwijs en onderzoek. Dit is merkbaar zowel vanuit het standpunt van studenten die zoeken naar de instelling die het beste aansluit bij hun verwachtingen, als vanuit het perspectief van docenten en onderzoekers die willen werken aan de instelling met de beste uitstraling en loopbaanperspectieven. Het tweede fenomeen is de toenemende verantwoording die moet worden afgelegd. Zowel overheden als externe geldschieters willen indicatoren om retrospectief te kunnen beoordelen of hun investering heeft geleid tot kwalitatief hoogstaande prestaties, en willen zo mogelijk ook vooraf kunnen voorspellen of hun beschikbare middelen in een rendabele omgeving zullen terechtkomen.

In elk van deze situaties is het kernwoord 'kwaliteit'. Diverse groepen stakeholders hebben nood aan objectieve, transparante en internationaal bruikbare maatstaven voor kwaliteit. Maar een nauwkeurige en genuanceerde invulling hiervan op internationale schaal vraagt enerzijds

een grote inspanning, en is anderzijds niet altijd mogelijk door de aard van de beschikbare gegevens. Rankings kunnen – bij gebrek aan betere alternatieven – slechts in zekere mate tegemoetkomen aan deze verwachtingen.

Hoe dan ook heeft het initiële wantrouwen en zelfs verzet vanwege de academische wereld en beleidswerkers tegenover de makers van dergelijke Top-100's ruimte gemaakt voor toenadering. Diverse belanghebbenden, zowel uit de groep van 'rankers' als vanuit de groep die 'gerankt' wordt, hebben zich sinds 2005 verenigd in de International Rankings Expert Group (IREG). Deze organisatie heeft tot doel de methodologie van rankings en dus hun bruikbaarheid te verbeteren. De *Berlin principles* die deze groep in 2006 opstelde zijn een eerste stap hierin: ze creëren een referentiekader dat geldt als kwaliteitsnorm voor de kwaliteitsmeting in het hoger onderwijs. Zo worden commerciële rankings onder druk gezet door academische experts, universiteiten en beleidsorganen om hun werkwijze te professionaliseren en hun geloofwaardigheid te verhogen. Zowel voor een transparant en objectief beleid op universitair, nationaal en internationaal niveau, als voor een transparante en objectieve profilering van de universiteiten is er immers nood aan kwalitatieve maatstaven.

### **Diversiteit van doelstellingen**

Rankings hebben uiteenlopende initiële doelstellingen, zoals het wereldwijd in kaart brengen van de sterkste kennisproducenten en van 'stijgers' en 'dalers' in het hoger onderwijs (World University Rankings van Times Higher Education Supplement), het bepalen van de internationale positie en vooruitgang van hogeronderwijsinstellingen in een bepaald land (Academic Ranking of World Universities van de Shanghai Jiao Tong University) of het informeren van potentiële studenten bij hun studiekeuze (CHE-Ranking van het Centrum für Hochschulentwicklung). De brede bekendheid die aan rankings wordt gegeven leidt evenwel vaak tot andere toepassingen dan oorspronkelijk beoogd. Frustrerend voor de gerankte instellingen is het als rankings worden opgesteld uitgaande van methodes en doelstellingen die ver verwijderd liggen van die welke de onderwijs- en onderzoeksinstituten zelf vooropstellen.

Publiek gemaakte rankings zouden volgens de Berlin Principles gebaseerd moeten zijn op een wetenschappelijk onderbouwde methodologie, gestoeld op betrouwbare data en vergezeld van een duidelijke uitleg over de betekenis en de eventuele beperkingen. Jammer genoeg zijn de

bestaande rankings van variabele kwaliteit. Een reeks technische en methodologische problemen werd reeds aangekaart in de wetenschappelijke literatuur. Zo wordt voor het onderzoeksluik doorgaans voor een groot deel gebruik gemaakt van bibliometrische gegevens, wat gepaard gaat met problemen aangaande het linken van citaties aan geciteerde publicaties, de toewijzing van publicaties aan een instelling, de beperkte dekking van een aantal domeinen door de ISI-citatie indexen, bias naar land of taal ... (Van Raan, 2005).

Het systeem dat momenteel volgens experts het beste tegemoetkomt aan de vereisten voor een goede ranking, is dat ontwikkeld door het Duitse Centrum für Hochschulentwicklung<sup>1</sup>. Het betreft geen 'klassieke' ranking in één overzichtslijst van 1 tot 500. CHE is veeleer een buitenbeentje met een meerdimensionale, opleidingsgewijze voorstelling, specifiek gericht op studiekeuze. De filosofie daarachter is dat de kwaliteit van een instelling immers niet in één 'gemiddeld' of 'samengesteld' cijfer te vatten is. Accenten inzake onderzoek liggen naargelang de instelling in andere domeinen, al dan niet ten gevolge van een expliciet speerpuntenbeleid. Een gemiddeld beeld over alle disciplines heen heeft bijgevolg weinig inhoudelijke betekenis. Rankings gebeuren dus beter per discipline en niet voor een universiteit als geheel. Ook Vlaams minister van Onderwijs Vandenbroucke sprak zich reeds uit voor een meerdimensionale ranking, met het oog op meer transparantie en mobiliteit van studenten en docenten.

Het bovenstaande maakt duidelijk dat het woord 'ranking' sterk verschillende ladingen kan dekken. Het beeld van de eendimensionale ranglijst dat het woord nog bij velen oproept, lokt vooral negatieve reacties uit. Verschillende criteria of gewichten van criteria kunnen immers leiden tot verschillende rangordes (Van Leeuwen et al., 2003). Een dergelijk opzet biedt dan ook ruimte voor manipulatie (Grupp & Mogege, 2004). Het CHE-systeem is een heel ander soort ranking. Ze biedt voor een zeer ruime waaier van criteria een indeling van universiteiten in drie grote groepen: 'top', 'intermediate' en 'bottom'. Potentiële studenten kunnen *à la carte* criteria selecteren volgens hun eigen wensen en verwachtingen. Geïnspireerd door het CHE-systeem werd in Nederland een gelijkaardige interactieve portaalsite ontwikkeld op basis van reeds centraal beschikbare gegevens<sup>2</sup>. Momenteel wordt via een pilootproject onderzocht of Vlaanderen en Nederland aan het CHE-systeem kunnen deelnemen (CHE et al., 2006). Een beperkt aantal disciplines en universiteiten werkt hieraan mee op vrijwillige basis. Op termijn wordt een Europees systeem beoogd, hoewel het niet evident is om voor verschillende landen vergelijkbare cijfergegevens in beeld te brengen.

## Diversiteit van instellingen

Er bestaat ook een grote diversiteit onder instellingen voor hoger onderwijs inzake omvang, wetenschappelijke domeinen, onderzoeksbeleid, onderwijsbeleid, internationalisering en maatschappelijke rol. Een professionele bacheloropleiding heeft bijvoorbeeld een andere doelstelling dan een academische bacheloropleiding. Een universiteit die een cruciale rol speelt in regionale ontwikkeling heeft andere prioriteiten dan een onderzoeksuniversiteit zoals MIT in Boston. Onderwijsinstellingen kunnen bovendien sterk veranderen in de loop der tijd. Een objectieve vergelijking ligt dus allerminst voor de hand, maar binnen eenzelfde type instellingen (zie bv. de classificatie van instellingen in de VS door de Carnegie Foundation for the Advancement of Teaching<sup>3</sup>) kunnen toch aannemelijke vergelijkingen worden gemaakt door verschillende aspecten van kwaliteit in kaart te brengen.

In wereldwijde rankings van instellingen, zoals deze uitgebracht door *Times Higher Education Supplement* (THES) en *Shanghai Jiao Tong University* (SJTU)<sup>4</sup>, wordt met de diversiteit aan instellingen vaak geen of nauwelijks rekening gehouden. Nochtans kunnen naargelang de situatie of het domein goede prestaties gekenmerkt zijn door verschillende parameterwaarden. Zo varieert het gemiddelde aantal ISI-publicaties per vorser<sup>5</sup> van één per jaar voor sociale wetenschappen tot meer dan drie per jaar voor natuurkunde en sterrenkunde (Moed, 2005). Zo zijn er ook typische aantallen citaties per publicatie, PhD's per senior staff, ... De globale rankings van THES en SJTU tellen respectievelijk citaties per vorser en publicaties over hele instellingen. Universiteiten die hun zwaartepunten vooral hebben in domeinen met hoge typische outputaantallen zullen dus daardoor, los van de kwaliteit, al beter scoren in dergelijke rankings. Zo betekent bijvoorbeeld het beschikken over een medische faculteit een groot pluspunt in rankings waarbij citaties worden geteld.

Inmiddels brengen zowel THES als SJTU ook deelrankings uit per groot wetenschapsdomein<sup>6</sup>. Deze onderverdelingen zijn zeker een stap vooruit, doch ook daarbinnen bestaat nog aanzienlijke variatie in typisch publicatiegedrag van de subdomeinen.

De vraag is tot welk detail een ranking moet gaan om informatie aan te bieden die nuttig en relevant is. Het behoort tot de verantwoordelijkheid van de 'rankers' om dit zo goed mogelijk af te stemmen op de beoogde doelstelling. Voor de gebruikers betekent dit dat de verschaft

informatie niet automatisch bruikbaar is voor andere doeleinden dan die welke door de ranking zelf wordt vooropgesteld.

De CHE-ranking beoogt in de eerste plaats het informeren van potentiële studenten over de prestaties van een reeks instellingen in het door hen gekozen studiegebied. Ook dit momenteel meest gewaardeerde systeem is niet zonder zwakke punten, want binnen een bepaald studiegebied kunnen eveneens grote verschillen bestaan in typische publicatie-omvang of financieringsmogelijkheden. De vraag is of de student goed wordt geïnformeerd via naast elkaar geplaatste cijfers van instellingen die gespecialiseerd zijn in verschillende subdomeinen, zonder afweging ten opzichte van wat in die verschillende subdomeinen mag worden verwacht. Anderzijds is een meer gedetailleerde domeinkeuze wellicht nog niet aan de orde voor de potentiële student. Als tussenoplossing zouden, naast de cijfers, de voornaamste deeldomeinen kunnen worden vernoemd<sup>7</sup>. Dit biedt de student niet alleen nuttige bijkomende informatie, maar vestigt ook zijn aandacht op de relativiteit van de cijfers.

Een ranking die in de eerste plaats de ambitie heeft om de kwaliteit van een instelling (en dus niet van een opleiding of studiegebied) in te schatten, moet zo nauwkeurig mogelijk rekening houden met de bestaande diversiteit. Kwaliteit in al haar verscheidenheid zichtbaar maken via een spectrum van indicatoren is echter niet evident<sup>8</sup>. Zolang met de verschillende vormen van diversiteit niet afdoende rekening wordt gehouden bestaat het gevaar dat, louter ten gevolge van 'technische beperkingen', een verschraling van het universitair en wetenschappelijk landschap in de hand wordt gewerkt. Bepaalde instellingstypes of domeinen zullen immers, los van hun kwaliteit, beter 'scoren' op de gekozen indicatoren dan andere.

Ook wanneer hiervoor door de auteurs wordt gewaarschuwd, kunnen de 'gemakkelijke' vergelijkingen die rankings aanreiken nog steeds leiden tot verkeerde indrukken of beslissingen. Dit gevaar mag niet worden onderschat, gezien de ruime bekendheid die doorgaans aan rankings wordt gegeven via de pers en het gehoor dat zij vinden bij overheden. Er bestaan wel degelijk rankings die werken met genormaliseerde indicatoren, die de prestaties per onderzoeksdomein afwegen ten opzichte van de verwachtingen daarbinnen<sup>9</sup>. Ook dan is bij interpretatie echter nog de nodige voorzichtigheid aangewezen, gezien de moeilijkheden die worden ondervonden bij de gegevensverzameling.

Feedback vanuit de wetenschappelijke instellingen die het onderwerp van de rankings uitmaken is dan ook geen overbodige luxe, naast de bespreking van technische en methodologische problemen in de wetenschappelijke literatuur. Dit zal de kwaliteit van de rankings ten goede komen en de gebruikers helpen bij een correcte interpretatie. Zo formuleerden verschillende universiteiten reeds opmerkingen over de correctheid van de verzamelde basisgegevens<sup>10</sup> bij de rankings van THES en SJTU. Recent werd een 'alternatieve' THES-ranking, rekening houdend met de diversiteit aan instellingsprofielen, uitgewerkt door professor Tulkens van de UCL (2007). Of er ooit een betrouwbare wereldwijde, voldoende genuanceerde en multidimensionale weergave van de kwaliteit van instellingen voorhanden zal zijn, is nog zeer de vraag.

### **The importance of being ranked**

Ondanks de moeilijkheden die er nog mee gepaard gaan, spelen rankings van onderwijs- en onderzoeksinstellingen nu reeds een belangrijke rol in de competitieve hogeronderwijsmarkt. Het huidige hoger onderwijs vertoont immers veel vergelijkingspunten met een vrijmarkteconomie, waarin stimulerende competitie, marketing en consumentenbinding een specifieke onderwijs- of onderzoekstoepassing hebben gevonden. Elke instelling heeft er belang bij om in de bestaande rankings voor te komen en zichtbaar te zijn voor potentiële studenten of andere stakeholders. Verder kunnen rankings, hoe onperfect ook, interessante informatie leveren voor het beleid van een instelling. Dit is een belangrijk aspect aangezien rankings op zich geen instrument zijn voor de bevordering van de kwaliteit. Dit gebeurt via het beleid dat een instelling voert, zich positionerend tegenover haar 'concurrenten'.

Rankings zijn van nature dan ook een dynamisch gegeven, afhankelijk van het actuele beleid en aanwezige potentieel. In die evolutieve context brengen rankings de toestand in beeld zoals die was in de laatste jaren. Eventuele probleempunten kunnen door een instelling worden aangepakt, als dit nog niet gebeurde op basis van eigen beleidsstudies. Doorgaans vraagt dit enige tijd. Hier houdt de potentiële *impact* van een ranking een gevaar in, indien het effect op de leefbaarheid van een domein of instelling te snel en ingrijpend zou zijn om een bijsturing kans op succes te bieden. Dit zou leiden tot een paradox, waarbij rankings een kwaliteitsverbetering verhinderen, terwijl zij het tegendeel voor ogen hebben.

De nood aan objectievere en meer betrouwbare rankings is daarom eens zo reëel, want de effecten van rankings gaan vaak heel wat verder dan hun oorspronkelijke doelstellingen. Een hoge plaats in een ranking levert geen ‘bewijs’ van kwalitatieve prestaties, maar slechts een beeld ervan ‘bij benadering’. Hoe dan ook is de druk om goed te presteren in rankings voor alle hogeronderwijsinstellingen groot. Een strategisch beleid kan erop gericht zijn hoog te scoren op bepaalde indicatoren, wat een omwisseling van doel en middel betekent. Een degelijk onderzoeks- en onderwijsbeleid zal een positie in rankings niet als doel op zich vooropstellen, maar ze gebruiken als een van de middelen om kwaliteit te meten. De aantrekkingskracht van een instelling bij studenten en vorsers vloeit voor een belangrijk deel voort uit dit dagelijkse beleid, waarvan de essentie niet steeds zichtbaar is in de cijfers van de rankings.

Een onderzoeksproject van de OESO ging recent via surveys bij diverse universiteitsbesturen na wat de impact is van rankings op de interne werking van de universiteiten (Hazelkorn, 2007). De eerste resultaten wijzen er al op dat 71% van de respondenten de ambitie hebben om tot de internationale top-25 te behoren. Aangezien er wereldwijd zo’n 40.000 instellingen van hoger onderwijs zijn, is deze collectieve ambitie verre van realistisch. Uit de survey blijkt ook dat het er vaak niet toe doet of het beeld dat rankings creëren authentiek is of niet. De perceptie alleen al is in vele gevallen voldoende om universiteiten aan te zetten een strategisch beleid te voeren. Op zich is daar niets mis mee, zolang het middel (kwaliteitsmeting) niet het doel zelf wordt.

In plaats van met spanning de eerstvolgende wereldwijde ranking af te wachten en stijgers en dalers op te lijsten, is het belangrijker en interessanter na te gaan op welke manier universiteiten hun rankingspositie situeren binnen een algemeen kwaliteitsbeleid. Heeft een ontgoochelende positie het universitair beleid er bijvoorbeeld toe aangespoord bepaalde incentivemaatregelen in te voeren? Heeft een rankingspositie aangezet tot nadenken over nationale of internationale profilering? Is een verrassende stijging het resultaat van een strategisch onderzoeksbeleid, van een nieuwe Nobelprijswinnaar of van een internationale studentenrekrutering? Als rankings erin slagen het kwaliteitsconcept in hoger onderwijs te verfijnen, aan diverse instellingen een medium bieden om hun eigenheid te profileren en pas in laatste instantie worden gebruikt voor marketingdoeleinden, dan zullen zij een waardevol beleidsinstrument leveren. Zowel overheden als universiteiten hebben er daarom alle belang bij de recente initiatieven tot verbetering te ondersteunen.

## Bronnen

- Moed, H.F. (2005), 'Citation Analysis in Research Evaluation', *Information Science and Knowledge Management*, Volume 9, Springer (Chapter 21: 'Did Scientific Publication Productivity Increase?').
- Tulkens, H. (2007), 'Ranking universities: How to take better account of diversity', CORE, Université catholique de Louvain, 17 pages (<http://www.core.ucl.ac.be/staff/tulkens.html>).
- Centrum für Hochschulentwicklung (CHE), Universiteit Twente, Universiteit Leiden, Ministerie van Onderwijs en Vorming van de Vlaamse Gemeenschap van België (2006), 'The CHE Ranking of European Universities', Socrates Programme, Proposals on Higher Education Reform (Lisbon Strategy and Bologna Process), Accompanying Measures, Call EAC 01/06.
- Hazelkorn, E. (2007) 'The Impact of League Tables and Ranking Systems on Higher Education Decision Making' *Higher Education Management and Policy* 19 (2) (te verschijnen)
- Grupp, H. and Moguee, M.E. (2004) 'Indicators for national science and technology policy: how robust are composite indicators?', *Research Policy*, 33, pages 1373-1384.
- IREG (2006) Berlin Principles on Ranking of Higher Education Institutions.  
[http://www.ihep.org/Organization/Press/Berlin\\_Principles\\_Release.pdf](http://www.ihep.org/Organization/Press/Berlin_Principles_Release.pdf)
- Van Leeuwen, T.N., Visser, M.S., Moed, Nederhof, T.J. en Van Raan, A.F.J. (2003), 'The Holy Grail of science policy: Exploring and combining bibliometric tools in search of scientific excellence', *Scientometrics*, 57 (2), pages 257-280.
- Van Raan, A.F.J. (2005), 'Fatal attraction: Conceptual and methodological problems in the ranking of universities by bibliometric methods', *Scientometrics*, 62 (1), pages 133-143.

---

## Eindnoten

- <sup>1</sup> CHE-website: <http://www.che.de>
- <sup>2</sup> <http://www.studiekeuze123.nl>
- <sup>3</sup> <http://www.carnegiefoundation.org/classifications>
- <sup>4</sup> Respectieve websites: <http://www.thes.co.uk> en <http://www.sjtu.edu.cn>
- <sup>5</sup> Publicaties geïndexeerd in de citatie-indexen van Thomson Scientific (<http://www.thomsonscientific.com/>).
- <sup>6</sup> THES- deelrankings: 'Science', 'Biomedicine', 'Social Science' en 'Arts and Humanities'; SJTU- deelrankings: 'Natural Sciences and Mathematics', 'Engineering / Technology and Computer Sciences', 'Life and Agriculture Sciences', 'Clinical Medicine and Pharmacy', 'Social Sciences'.
- <sup>7</sup> Afleidbaar uit ISI-publicaties.
- <sup>8</sup> Zie bijvoorbeeld het in rekening brengen van publicaties in de humane wetenschappen, de variërende instellingstypes en context van hogeronderwijsstructuren waarbinnen de prestaties worden geleverd ...
- <sup>9</sup> Zie het bibliometrische project 'Mapping Excellence in Science and Technology across Europe' (<http://www.cwts.nl/ec-coe/>).
- <sup>10</sup> Aangaande naamsverwarringen, niet-uniformiteit van personeelscategorieën ...