

NETZWERK EXZELLENZ
AN DEUTSCHEN
HOCHSCHULEN



Rankings im Wissenschaftssystem

Zwischen Wunsch und Wirklichkeit

Angela Borgwardt

FRIEDRICH
EBERT 
STIFTUNG

Rankings im Wissenschaftssystem

Zwischen Wunsch und Wirklichkeit

Publikation zur Konferenz der
Friedrich-Ebert-Stiftung
am 17. März 2011

Angela Borgwardt

Schriftenreihe des
Netzwerk Exzellenz
an Deutschen Hochschulen

ISBN: 978-3-86872-805-7

1. Auflage

Copyright by Friedrich-Ebert-Stiftung

Hiroshimastraße 17, 10785 Berlin

Abt. Studienförderung

Redaktion: Marei John-Ohnesorg, Marion Stichler, Stefanie Mensching

Satz & Umschlag: minus Design, Berlin

Druck: Bonner Universitäts Buchdruckerei

Printed in Germany 2011

INHALT

ZEHN THESEN	7
DEUTSCHE UND INTERNATIONALE RANKINGS IM ÜBERBLICK	13
Simone Burkhart	
Wie misst man Qualität? Zur Methodik von Rankings	13
Rankings in Deutschland	15
Überregionale und weltweite Rankings	18
Auswirkungen von Rankings	20
KRITERIEN, AUSSAGEKRAFT UND WIRKUNGEN VON RANKINGS	23
Angela Borgwardt	
Forschungsrating des Wissenschaftsrats	23
Hochschulranking des CHE Centrum für Hochschulentwicklung	27
Förder-Ranking der Deutschen Forschungsgemeinschaft (DFG)	31
Kontrovers diskutiert: Bedeutung des Indikators Drittmittel?	36
Nutzen von Rankings	38
Kritik an Rankings	39
U-MULTIRANK – EIN ALTERNATIVES KONZEPT FÜR EIN MULTI-DIMENSIONALES INTERNATIONALES RANKING	43
Gero Federkeil	
Hintergrund	43
Die Kritik an den bestehenden globalen Rankings	44
Das U-Multirank-Projekt	47
INTERNATIONALE RANKINGS – WIE KANN MAN SIE BESSER MACHEN?	51
Angela Borgwardt	
Diversität und Profilbildung der Hochschulen berücksichtigen	51
Fachbezogene Rankings entwickeln	52
Politische Folgewirkungen beachten	53
Brauchen wir ein europäisches Ranking?	54
Herausforderungen im Umgang mit Rankings	57

ZEHN THESEN

Auf Grundlage der Konferenz der Friedrich-Ebert-Stiftung zum Thema: „Rankings im Wissenschaftssystem – Zwischen Wunsch und Wirklichkeit“ am 17. März 2011 hat die Planungsgruppe des ‚Netzwerk Exzellenz an deutschen Hochschulen‘ folgende Thesen entwickelt:

1. Rankings bilden Realität vereinfacht ab und schaffen eine eigene Realität.

Rankings haben das Ziel, mehr Transparenz zu schaffen und einen Beitrag zur Orientierung zu leisten. Rankings im Hochschul- und Wissenschaftssystem stellen allerdings Realität nur ausschnittshaft dar und vereinfachen Komplexität. Zum einen geht ein großer Teil an differenzierter Aussagekraft verloren, wenn aus heterogenen Daten Durchschnittswerte gebildet werden. Zum anderen müssen Qualitäten – um Vergleiche überhaupt zu ermöglichen – auf Zahlenwerte reduziert werden (wenn sie denn überhaupt erfasst werden können). Für das Rankingergebnis ist zudem die Definition der zugrundeliegenden Kriterien und Indikatoren entscheidend, die immer eine mehr oder weniger angemessene Setzung bedeuten. Rankings können zwar Orientierung und Transparenz ermöglichen, Wirklichkeit aber auch stark verzerrt und verkürzt abbilden. Sie schaffen letztlich eine eigene Realität, indem sie in der Öffentlichkeit auf Resonanz stoßen und reale Konsequenzen auslösen. Umgekehrt schafft auch die Unkenntnis von Vergleichsdaten eine eigene Realität, die von den innerhalb von Einrichtungen oder Wissenschaftssystemen etablierten Kräften vermittelt wird.

2. Rankings entfalten erhebliche Steuerungswirkungen.

Rankings beeinflussen nicht nur die Entscheidungen von Studierenden, Hochschulleitungen, Professor/innen und Politiker/innen, sondern auch die gesamte öffentliche Debatte über das Hochschul- und Wissenschaftssystem. Selbst wenn Rankings gravierende methodische Fehler aufweisen, können sie wissenschaftspolitisch enorme Steuerungswirkungen entfalten:

So beeinflussten die sehr selektiven Daten des weltweiten Shanghai-Ranking durch das schlechte Abschneiden deutscher Hochschulen die politische Diskussion und verstärkten den Wettbewerb im Hochschulsystem. Die enorme Wirkungsmacht veröffentlichter Rankings – unabhängig von ihrer Qualität – birgt die Gefahr von Fehlsteuerungen und dysfunktionalen Effekten. So sind – zum Nachteil deutscher Hochschulen – in manchen Ländern Stipendiat/innen gehalten, sich nur an ausländischen Institutionen mit einem bestimmten Rankingergebnis zu bewerben.

3. Mit Rankingergebnissen muss sehr sorgfältig und verantwortungsvoll umgegangen werden.

Die Auswirkungen von Rankings sind kaum zu überschätzen und erfordern einen sehr sorgsam Umgang mit ihren Ergebnissen. Das betrifft nicht nur die Berücksichtigung methodischer Eigenheiten und Mängel von Rankings, sondern auch die Problematik der interessenbestimmten Interpretation von Daten und weitere Gefahren, wie die Steuerung von Forschungsschwerpunkten und mögliche normierende Wirkungen im Wissenschaftssystem. Hochschulleitungen und politische Entscheidungsträger/innen haben eine besonders große Verantwortung, wenn die Ergebnisse von Rankings in ihre Entscheidungen einfließen, da ihre Schlussfolgerungen mit weitreichenden Konsequenzen für das deutsche Hochschul- und Wissenschaftssystem verbunden sein können.

4. Der Gefahr der Instrumentalisierung von Rankings muss begegnet werden.

Rankings und Ratings können langfristige, zum Teil auch ungeplante und unvorhergesehene Auswirkungen auf Forschungstätigkeit, Infrastruktur, Ausbildung etc. im Hochschulsystem haben. Das bringt große Einflussmöglichkeiten der Ranking-Einrichtungen mit sich. Es muss verhindert werden, dass diese Einrichtungen ihre Kategorien gezielt an eigenen Interessen oder den Interessen ihrer Auftraggeber ausrichten und damit unter Umständen den „Wissenschaftsmarkt“ langfristig manipulieren. Es müssen unbedingt die unbeabsichtigten Nebenwirkungen im Auge behalten werden, besonders dann, wenn man bestimmte Adressatengruppen befriedigen will. Deshalb sind Bemühungen der EU, der Europäischen Hochschulkonferenz, der HRK, der DFG und des Wissenschaftsrates auf diesem Gebiet zu begrüßen.

5. An der Verbesserung von Rankingverfahren muss intensiv gearbeitet werden.

Die Ergebnisse von Rankings werden durch die Methodik sowie durch die Auswahl und Gewichtung der Indikatoren bestimmt: Über die Definition des Verfahrens und der Kategorien kann das Ergebnis entscheidend beeinflusst werden. Zudem weisen viele bestehende Rankings methodische Defizite auf, z. B. problematische Erhebungsmethoden, willkürliche Auswahl der Indikatoren, fehlende Offenlegung der Datensätze, zu kleine Stichprobengrößen. Besonders problematisch sind die bekannten internationalen Rankings, die eine begrenzte Aussagekraft und methodische Schwächen aufweisen und dabei versuchen, Hochschulen als Ganzes in Ranglisten zu erfassen.

Deshalb ist es wichtig, bessere Rankingverfahren zu entwickeln. Inzwischen gibt es in Europa und Deutschland verschiedene Initiativen, um geeignete Methoden, valide Indikatoren und Kriterien für Rankings zu erarbeiten, die arbeitsfähig und transparent sind und den Ansprüchen an differenzierte Rankings gerechter werden. Es könnte sinnvoll sein, wenn HRK, Wissenschaftsrat, DFG und DAAD und andere Institutionen eine Initiative ergreifen, um „Regeln guter Rankings“ festzulegen (siehe auch die ‚Empfehlungen zu Rankings im Wissenschaftssystem‘ des Wissenschaftsrats, 2004).

Bisher entwickelt jede Hochschule ihre internen Datenbanken nach anderen Parametern, so dass kein einheitliches Datenset existiert. Deshalb sollte, z.B. durch den Akkreditierungsrat und das iFQ (Institut für Forschungsinformation und Qualitätssicherung) ein Datenkonzept und Qualitätssicherungssystem für Rankings entwickelt werden: Welche Daten sollen einbezogen werden, wie sollen sie erhoben und vergleichbar gemacht werden? Dieses Datenkonzept könnte dann zu einer Grundlage für deutsche (und später möglicherweise auch europäische) Wissenschaftsrankings werden. Auf europäischer Ebene sollten Institutionen wie das ENQA (European Association for Quality Assurance in Higher Education) einbezogen werden, auf internationaler Ebene ist IREG (International Observatory on Academic Ranking and Excellence) führend. IREG hat 2006 die ‚Berlin Principles on Ranking of Higher Education Institutions‘ verabschiedet, die als Grundlage dienen können und alle relevanten Punkte aufgreifen.

6. Unterschiedliche Bedarfe, Ziele und Adressaten erfordern unterschiedliche Rankings.

Die große Aufmerksamkeit für Rankings spricht dafür, dass dadurch manifeste Informationsbedürfnisse befriedigt werden; zugleich sind Rankings Ausdruck einer internationalen Entwicklung. Ein wesentliches Ziel von Rankings besteht darin, durch Leistungsmessungen und -vergleich mehr Transparenz im nationalen und internationalen Wettbewerb der Hochschulen zu schaffen (Benchmarking). Rankings können als Entscheidungshilfen fungieren und Entscheidungen besser begründen und legitimieren. Die meisten Rankings zielen darauf, Studieninteressierten und Studierenden wichtige Orientierung bei der Hochschulwahl zu bieten und/oder Akteure mit Führungsverantwortung in Hochschulen und Politik dabei zu unterstützen, Entscheidungen auf der Basis empirischer Daten zu treffen. Um den Informationsbedarf dieser Adressatengruppen differenziert zu befriedigen, muss nach den besten Kriterien, Indikatoren und Methoden für verschiedene Rankings gesucht werden. Nur dann ist der größtmögliche Nutzen für unterschiedliche Bedarfe, Ziele und Adressaten zu erreichen. Gleichzeitig obliegt Entscheidungsträger/innen selbst die Verantwortung, für den jeweiligen Zweck passende Rankings auszuwählen. Selbst noch so gute und differenzierte Rankings werden immer erfordern, dass die Nutzer/innen dieser Rankings sich sorgfältig über Indikatoren und Methoden informieren, bevor sie Entscheidungen darauf basieren.

7. Rankings sollten auf mehreren Indikatoren und fachspezifischen Verfahren beruhen.

Rankings, die den Anspruch haben, Einrichtungen umfassend zu vergleichen, müssen eine Vielzahl von Indikatoren berücksichtigen. Um aussagekräftige Ergebnisse zu erhalten, sollten quantitative und qualitative Faktoren (wie Begutachtung durch Expert/innen, Studierende) kombiniert werden. Nur dann ist es möglich, die Leistungen verschiedener Hochschulen, einzelner Disziplinen, Fakultäten und auch Forscher/innen bezogen auf Forschung, Lehre und Ausstattung differenziert zu erfassen und Strategien abzuleiten. Aufgrund der Ausdifferenzierung der deutschen Hochschullandschaft und der Profilbildung der Hochschulen sind Gesamtbewertungen von Hochschulen (mittels einer Durchschnittsbildung der verschiedenen Fächer) ungeeignet und höchstens in der Bewertung von Management und Strategiefähigkeit aussagekräftig. Dadurch entstehen

Fehler, die sich durch Ranglisten noch weiter verstärken. Eine angemessene Bewertung von wissenschaftlichen Leistungen setzt ein fachspezifisches Verfahren voraus, bei dessen Entwicklung Fachleute einbezogen werden müssen, um wissenschaftsadäquate Methoden und Indikatoren zu erhalten (Definition der Kriterien, Auswahl und Gewicht der Indikatoren im jeweiligen Fach etc.). Das schließt nicht aus, dass es in Einzelfällen auch Rankings gibt, die nur auf einem Indikator (z.B. Zufriedenheit von Absolvent/innen) beruhen – sofern dies transparent ist und nicht zu falschen Schlüssen einlädt.

8. Der Indikator Drittmittel gibt wichtige Hinweise, darf aber nicht überbewertet werden.

Die erfolgreiche Einwerbung von Drittmitteln ist durchaus Ausdruck einer hohen Qualität von Forschungsprojekten, doch wäre es gefährlich, sie als Kriterium zu stark zu gewichten. Es müssen immer mehrere Indikatoren herangezogen werden, um ein differenziertes Gesamtbild über die gesamte Hochschullandschaft bzw. Forschungsleistungen in allen Fächern zu erhalten. Dabei sollten auch die Ergebnisse der Forschung (Publikationen, Promotionen, Patente etc.) berücksichtigt und die unterschiedlichen Arten von Forschung abgebildet werden. Drittmittelinwerbungen sind in den verschiedenen Fächern unterschiedlich relevant: in den Naturwissenschaften eher wichtig, in den Geisteswissenschaften eher untergeordnet. Deshalb sind Rankings, die sich ausschließlich auf diesen Indikator stützen, je nach Wissenschaftsbereich unterschiedlich aussagekräftig.

9. Drittmittel dürfen die Grundfinanzierung der Hochschulen ergänzen, nicht ersetzen.

Grundsätzlich wird eine problematische Folge von Drittmitteln deutlich: Aufgrund der sinkenden Grundfinanzierung sind die Hochschulen dazu gezwungen, Forschung zunehmend über Drittmittel zu finanzieren. Wenn darüber hinaus Drittmittel zu einem entscheidenden Indikator für jegliche Mittelverteilung werden und die Verteilung der Normalbudgets immer leistungsabhängiger gestaltet wird, droht eine Abwärtsspirale für Hochschulen und Fachbereiche mit geringer Drittmittelinwerbung. Diese Entwicklung gefährdet die Freiheit von Forschung und Lehre und die

Grundausrüstung der Hochschulen – und damit die Basis dessen, was lange Zeit die hohe Qualität der deutschen Hochschulen ermöglichte. Ein Beispiel für eine eklatante Fehlnutzung von Rankings wäre die Kopplung der Grundfinanzierung an die Drittmittelquote oder die Verknüpfung der leistungsorientierten Mittelvergabe allein mit den eingeworbenen Drittmitteln.

10. Ein europäisches Ranking muss der Diversität und Profilbildung der Hochschulen gerecht werden.

Bisherige internationale Rankings sind nicht geeignet, die besonderen Qualitäten des Wissenschafts- und Hochschulstandorts Deutschland angemessen zu erfassen. So bleiben die herausragenden Forschungsleistungen an außeruniversitären Forschungsinstituten unberücksichtigt. Notwendig ist deshalb ein Ranking, das sich nicht – wie die meisten internationalen Rankings – primär am angloamerikanischen Hochschulmodell orientiert, sondern der Vielfalt europäischer Hochschulsysteme gerecht wird. Da der europäische Hochschulraum stark diversifizierte Systeme vereint, ist es bei einem europäischen Ranking sehr wichtig, Unterscheidungen in Bezug auf die unterschiedlichen Profile der Hochschulen zu treffen. Im deutschen Hochschulsystem kann jede Hochschule ihre Mission zunehmend selbst definieren, ihr eigenes Profil entwickeln und ihre Ziele auf hohem Niveau umsetzen. Das ist für ein europäisches Ranking eine Herausforderung, der es aber gerecht werden muss. Deshalb setzen die momentan diskutierten europäischen Ansätze auf multiperspektivische und fachgebundene Rankings.

DEUTSCHE UND INTERNATIONALE RANKINGS IM ÜBERBLICK

Dr. Simone Burkhart Deutscher Akademischer Austauschdienst (DAAD)

Rankings sind aus dem Hochschulwesen nicht mehr wegzudenken. Während sie im angelsächsischen Raum eine lange Tradition haben, erfreuen sie sich im deutschen Hochschulraum erst seit etwa zehn Jahren wachsender Aufmerksamkeit. Jeder dritte Studienanfänger bezieht bei der Wahl des Studienortes Ergebnisse von Rankings in die Entscheidung ein und für Hochschulleitungen, aber auch für politische Entscheidungsträger liefern Rankingergebnisse wichtige Hinweise auf die Verortung im nationalen und internationalen Wettbewerb.

Gerade aufgrund der rasant gestiegenen Bedeutung von Rankings ist eine kritische Reflexion über die Aussagekraft und die Vielfalt von Rankings unerlässlich. In diesem Zusammenhang soll dieser Artikel einen kurzen Überblick über die Methodik und die Vielfalt von nationalen und internationalen Rankings geben.

Wie misst man Qualität? Zur Methodik von Rankings

Eine zentrale Frage von Rankings ist, was gerankt werden soll: Universitäten, Fachbereiche oder einzelne Fächer? Etliche, vor allem bekannte internationale Rankings, versuchen Hochschulen als Ganzes zu ranken. Die Probleme sind evident: Viele Hochschulen sind in einer Disziplin sehr gut, in anderen weniger. Eine gute Experimentalphysik bedeutet nicht, dass die Universität auch eine gute Pädagogik beheimatet. Schlimmer noch: Viele Indikatoren, mit denen Qualität gemessen wird, sind abhängig von einer Fächerkultur.¹ In den Lebenswissenschaften wird der Publikation in Fachzeitschriften hohes Gewicht beigemessen, in den Natur- und Ingenieurwissenschaften sind Patente ein wichtiger wissenschaftlicher Output und in den Geisteswissenschaften kommt es auf Monografien an.

¹ Vgl. hierzu und im Folgenden: Hornbostel 2007.

Eine gute Lehre in der Chemie setzt unter anderem gute Labore voraus, in den Geisteswissenschaften ist vor allem eine gut ausgestattete Bibliothek wichtig. Will man also ganze Hochschulen ranken und vergleichen, müsste man fachbezogene Indikatoren verwenden, die dann zu einem Gesamtergebnis aggregiert werden – ein beträchtlicher Aufwand, der bei keinem Ranking ganzer Hochschulen konsequent betrieben wird.

Rankings unterscheiden sich weiterhin in der Frage, ob sie Ranglisten (1. Platz, 2. Platz etc.) oder Ranggruppen (z. B. Spitzengruppe, Mittelgruppe, Schlussgruppe) erstellen. Das Problem der Ranglisten ist, dass sie eine Genauigkeit suggerieren, die nicht vorhanden ist. Kleine Unterschiede in den Werten der Indikatoren können dazu führen, dass eine Hochschule nicht auf Platz 30, sondern auf Platz 45 landet. Dieses Problem wird bei der Bildung von Ranggruppen abgeschwächt, doch bleibt es letztlich auch existent: Wo zieht man die Grenze zwischen den „schlechten“ der Spitzengruppe und den „guten“ der Mittelgruppe, und wie geht man mit Dimensionen um, in denen alle Hochschulen gute Werte erreichen, also auch die „Schlussgruppe“ an sich gute Werte erzielt?

Welche Indikatoren für die Messung von Qualität herangezogen werden und wie diese einzelnen Indikatoren in ein Endergebnis überführt werden sollen, ist ein drittes – und sicher das schwierigste – Problem, dem Rankings gegenüberstehen. Zunächst muss klar sein, welche Dimensionen von Qualität eigentlich gemessen werden sollen: Qualität in der Forschung, Qualität in der Lehre, Ausstattung der zentralen Einrichtungen (z. B. Bibliothek), Lebensqualität am Hochschulort – die Liste ließe sich beliebig fortsetzen. Häufig wird suggeriert, dass es eine allumfassende Qualitätsangabe geben kann, das heißt, es wird nicht klar zwischen einzelnen Dimensionen getrennt.

Da Qualität nicht objektiv messbar, sondern eine Konstruktion ist, wird sie auf verschiedenen Wegen erhoben. Erstens über Befragungen, die subjektive Beurteilungen ermitteln. Beliebte sind in diesem Zusammenhang Reputationsbefragungen bei kompetenten Peers (in der Regel Professor/innen). Viele Rankings berücksichtigen zudem Befragungen von Studierenden oder Arbeitgebern. Zweitens sind quantitative Indikatoren zu nennen, die auf Indikatoren beruhen, die Gutachtermeinungen quantifizieren (z.B. Höhe der Drittmittel, die in Gutachterverfahren vergeben werden, Anzahl der Publikationen in Zeitschriften mit *Peer Review*). Drittens sind quantitative Input- und Output-Indikatoren von Bedeutung, die über diverse Statistiken oder direkte Abfragen an den Hochschulen erhoben werden. Alle diese Indikatoren sind problematisch, sie haben Mess-

und Interpretationsprobleme, über die jedoch selten reflektiert wird.

Problematisch ist auch die Verfügbarkeit der Daten. Häufig werden Indikatoren nicht davon bestimmt, was gebraucht wird, sondern davon, was einfach zu erheben ist. Aber auch auf den ersten Blick objektive Datengrundlagen wie der *Science Citation Index* (SCI) haben ihre Tücken: Je nachdem, ob man z. B. nach Personen, die zu einem Institut gehören, oder nach institutionellen Adressen sucht, bekommt man verschiedene Ergebnisse. Für diese und andere Datenquellen muss also vorab die Konstruktion der Indikatoren gut überlegt sein, die Daten müssen dann sorgsam bereinigt werden. Wie für vieles gilt damit auch für Rankings: Anspruchsvolle und „gute“ Rankings sind häufig kostenintensiv.

Das CHE-Hochschulranking

Für den Hochschulstandort Deutschland gibt es zahlreiche Rankings. Sie unterscheiden sich hinsichtlich ihrer Methode, der adressierten Zielgruppe, der Auswahl von Indikatoren und der Offenlegung der Ergebnisse und Datenerhebungsmethoden.

Das CHE-Ranking ist das wohl populärste Ranking für Studienanfänger/innen in Deutschland. Ob ein Fachbereich einer Hochschule am CHE-Ranking teilnimmt, entscheidet dieser selbst. Beim CHE-Ranking wird eine Vielzahl von Indikatoren erhoben, die aus unterschiedlichen Datengrundlagen stammen. Verwendet werden Umfragen, bibliografische Daten, Daten zu Patenten und weitere Statistiken. Je nach Fachbereich variiert die Zahl der Indikatoren, maximal sind es 34 Vergleichsindikatoren.

Das CHE-Ranking ist außerordentlich differenziert. Es wird sowohl nach Fachbereichen, als auch nach einzelnen Qualitätsmerkmalen (Forschungsreputation, Betreuung, Studiensituation ...) unterschieden. Es handelt sich um ein fachbezogenes Ranking, d. h. es erfolgt kein Vergleich von ganzen Hochschulen, sondern von Studienfächern. Das CHE-Ranking verwendet keine Rangreihenfolge sondern eine Einteilung in Gruppen.

Da die Ergebnisse nach einzelnen Fächern differenziert ausgewiesen werden, eignet sich das CHE-Ranking nicht für die einfache Darstellung von Ergebnissen. Die differenzierte Darstellung ermöglicht es den Nutzern, sich ein eigenes Ranking nach individuellen Bedürfnissen zusammenzustellen (z. B. im Internet).

Das CHE-Ranking ist eines der elaboriertesten Rankings weltweit. Positiv hervorzuheben sind die verwendeten Methoden sowie die im Vergleich zu anderen Rankings hohe Transparenz. Das CHE-Ranking bietet damit für seine primäre Zielgruppe (Studienanfänger/innen) hilfreiche Informationen. Einige Hochschulen und Fakultäten haben sich in den letzten Jahren vom CHE-Ranking zurückgezogen. Die Gründe sind zum Teil generelle Kritik an Rankings, zum Teil wurde der Ausstieg aber auch mit dem zunehmenden personellen und zeitlichen Mehraufwand, der mit der Beantwortung von Befragungen verbunden ist, begründet.

DFG-Förderranking

Das DFG-Förderranking wird seit 2005 erhoben. Es informiert über die Beteiligung deutscher Wissenschaftseinrichtungen an den Förderprogrammen der DFG sowie nationaler und internationaler Förderinstitutionen. Ein wesentliches Ziel dieses Rankings ist es, die sich aus drittmittelgeförderten Forschungsvorhaben ableitenden Schwerpunktsetzungen der Forschungseinrichtungen auszuweisen. Das Ranking richtet sich an Wissenschaftler/innen sowie an Hochschulleitungen und politische Entscheidungsträger.

Das DFG-Ranking erstellt Ranglisten, was in diesem Kontext sinnvoll ist, da alle verwendeten Indikatoren eine interpretierbare Einheit (Förderbeträge in Euro) aufweisen. Die Rangliste wird auf der Grundlage folgender Daten erstellt: DFG-Bewilligungen, Daten der FuE-Projektförderungen des Bundes und die Anzahl deutscher Beteiligungen am EU-Forschungsrahmenprogramm. Die Daten werden von den öffentlichen Förderinstitutionen geliefert. Die Ergebnisse werden sehr differenziert in einer jährlichen Publikation dargestellt. Dabei werden fachliche Forschungsprofile, regionale Förderverteilungen und Netzwerke sichtbar.

Das Ranking versucht eine Brücke zu schlagen zwischen Rankings ganzer Hochschulen und ausschließlich fachbezogenen Rankings. Die Datenqualität ist durch die direkte Abfrage bei den Mittelgebern sehr hoch, Verzerrungen durch fehlerhafte Daten werden dadurch minimiert. Das Ranking vermittelt einen guten Einblick hinsichtlich der Forschungsschwerpunkte einzelner Hochschulen und forschungsbezogenen Vernetzungen zwischen Hochschulen. Allerdings deckt das Ranking nur einen bestimmten Teil der Forschungsleistungen ab, nämlich Input-Indikatoren. Nicht untersucht wird, inwieweit dadurch wissenschaftlicher Output (Publikationen, Patente etc.) erzeugt wird.

Das Rating des Wissenschaftsrats

Ziel des seit 2004 existierenden Konzepts des Forschungsratings ist der Vergleich der Forschungsleistungen aller deutschen Universitäten und außeruniversitären Forschungseinrichtungen. In einer Pilotstudie wurde das Verfahren für die Fachbereiche Soziologie und Chemie erprobt. Das Verfahren wurde anschließend weiterentwickelt und wird gegenwärtig für ein geisteswissenschaftliches Fach (Anglistik/Amerikanistik) und ein technikkundenschaftliches Fach (Elektro- und Informationstechnik) erprobt. Nach Abschluss dieses zweiten Probelaufs wird über eine Versteigerung entschieden. Das forschungsbezogene Rating richtet sich primär an Wissenschaftler/innen, die betreffenden Fachbereiche, Hochschulen, außeruniversitäre Forschungseinrichtungen und die Politik.

Neben Universitäten werden auch außeruniversitäre Forschungseinrichtungen bewertet. Die Forschungsleistung wird differenziert betrachtet. Es werden keine Ranglisten erstellt, sondern die Bewertung wird in einer Notenskala vorgenommen. Die zentrale Neuheit dieses Rankings liegt in seiner Methode des sogenannten *Informed Peer Review*. Dabei wird die eigentliche Bewertung von Fachgutachtern vorgenommen, denen für jede Hochschule oder Forschungseinrichtung umfangreiche quantitative und qualitative Daten (inklusive Kontextinformationen) vorgelegt werden. Weiterhin werden die Bewertungsmaßstäbe und die Indikatoren in einem engen Dialog mit den Fachbereichen erarbeitet. Die Ergebnisse der Pilotstudie wurden im Internet veröffentlicht, zudem liegen ausführliche Dokumentationen zur Methodik vor.

Das entwickelte Verfahren des Forschungsratings ist sehr differenziert, aufwendig und methodisch anspruchsvoll. Durch den starken Einbezug der Fachbereiche bei der Entwicklung der Kriterien ist von einer relativ hohen Akzeptanz unter den Begutachteten auszugehen. Eine positive Besonderheit ist die Einbeziehung der außeruniversitären Forschungseinrichtungen. Das Rating ist abhängig von der Akzeptanz des Fachbereichs. Als geisteswissenschaftliche Disziplin sollte ursprünglich die Geschichte evaluiert werden. Dies lehnte der Historikerverband 2009 ab. Insgesamt wurde generell die Sinnhaftigkeit von Rankings bezweifelt und der damit verbundene Steuerungsgedanke abgelehnt. Weitere Kritikpunkte waren, dass das Verfahren in regelmäßigem zeitlichen Abstand wiederholt werden müsse, wobei bezweifelt wurde, dass Aufwand und Ertrag in einem sinnvollen Verhältnis stehen. Diskutiert werden muss weiterhin, ob das aufwendige Verfahren in einem angemessenen Kosten-Nutzen-Verhältnis steht.

Shanghai-Ranking (ARWU)

Seit mehreren Jahren gibt es globale Rankings, die versuchen, Hochschulen im internationalen Vergleich zu bewerten. Zu den einflussreichsten und bekanntesten gehört das Ranking der Shanghai Jiao Tong University (ARWU) sowie das *Times Higher Education World University Ranking* (THEWUR).

Das Shanghai-Ranking (Academic Ranking of World Universities, ARWU) ist ein internationales Ranking, das seit 2004 existiert und von der Universität Shanghai jährlich aktualisiert und veröffentlicht wird. Seit 2009 wird das Ranking von der ausgegründeten *Shanghai Ranking Consultancy* publiziert. Da der Fokus fast ausschließlich auf der Forschung liegt und neben Publikationen im Wesentlichen Nobelpreise berücksichtigt werden, richtet sich dieses Ranking primär an Wissenschaftler/innen. Wegen dem hohen Bekanntheitsgrad wird es aber auch von der internationalen Öffentlichkeit, z. B. international mobilen Studierenden, als Datenquelle genutzt.

Das Ranking betrachtet die Universitäten als Ganzes; seit 2007 wird auch nach Fächergruppen differenziert. Folgende Indikatoren werden für die Berechnung herangezogen: die Anzahl der Alumni (rückwirkend bis 1901) sowie der derzeitigen Wissenschaftler/innen mit einem Nobelpreis oder wichtigen Forscherpreis, die Zitations- sowie die Publikationshäufigkeit. Die Ergebnisse werden anhand der Größe der Hochschulen gewichtet. Das Ranking veröffentlicht für die ersten 50 Universitäten genaue Rangplätze, danach folgen nur noch Ranggruppen. Die Ergebnisse der Einzelindikatoren werden zudem separat ausgewiesen.

Unter den 20 besten Universitäten befanden sich 2010 nur drei nicht US-amerikanische: neben Cambridge (Platz 4) und Oxford (Platz 10) noch die Universität Tokio (Platz 20). Die ersten deutschen Universitäten belegen die Plätze 52 (LMU München), 56 (TU München), 63 (Universität Heidelberg), 93 (Universität Bonn und Universität Göttingen). Unter den Top 500-Universitäten befinden sich insgesamt 39 deutsche Hochschulen.

Obwohl häufig als ganzheitliches Hochschulranking interpretiert, fokussiert das ARWU ausschließlich auf die Forschungsperformanz. Die Transparenz der verwendeten Indikatoren und Daten ist gut, bei der Sinnhaftigkeit der Indikatoren sind jedoch Zweifel angebracht: Aufgrund der langen Beobachtungszeiträume, bis 1901 bei den Nobelpreisträgern, werden aktuelle Bewertungen mit historischen vermischt. Außerdem ist die Zuordnung der Nobelpreise fraglich (abgestellt wird auf die Institution, in der der Preis-

träger zum Zeitpunkt der Verleihung angestellt war und nicht auf jene, an der die ausgezeichnete Leistung erbracht wurde). Für Deutschland stellt ein Problem dar, dass der Anteil der an außeruniversitären Einrichtungen stattfindenden Forschung nicht berücksichtigt wird.

Times Higher Education World University Ranking (THEWUR)

Das Ranking der *Times Higher Education* hat das Ziel, die besten Universitäten der Welt zu küren. Im Gegensatz zum Shanghai-Ranking erhebt es den Anspruch, eine Hochschule ganzheitlich (d. h. nicht nur nach wissenschaftlicher Exzellenz) zu bewerten. Es wurde von 2004 bis 2009 mit dem Partner Quacquarelli Symonds (QS) unter dem Namen THES vermarktet. Nach heftiger Kritik wurde 2010 die Zusammenarbeit beendet und ein neuer Partner, Thomson Reuters, verpflichtet. Seither erscheint das Ranking unter dem Label *THE World University Ranking*. Wegen dem hohen Bekanntheitsgrad und dem Anspruch des Rankings richtet sich das Ranking sowohl an Studierende und Wissenschaftler/innen, aber auch an alle Akteure im Hochschulbereich.

Das Ranking betrachtet die Universitäten als Ganzes, doch werden die Ergebnisse auch für sechs Fachbereiche getrennt ausgewertet. Das Ranking verwendet seit der Zusammenarbeit mit Thomson Reuters 13 Indikatoren, wobei Reputationsumfragen sowie die Zitationshäufigkeit das größte Gewicht besitzen. Veröffentlicht werden Rangplätze für die gesamte Hochschule (insgesamt die TOP 200) sowie für einzelne Fachbereiche (50 pro Fachbereich).

Unter den 20 besten Universitäten finden sich fünf nicht US-amerikanische: neben Cambridge (Platz 6) und Oxford (ebenfalls Platz 6) das Imperial College London, die ETH Zürich als einzige nicht-englischsprachige Hochschule und die University of Toronto. Die ersten deutschen Universitäten befinden sich auf den Plätzen 43 (Universität Göttingen), 61 (LMU München) und 83 (Universität Heidelberg). Unter den Top 200-Universitäten befinden sich insgesamt 14 deutsche Hochschulen, Deutschland nimmt damit die dritte Position in der „Länderwertung“ ein (im Vergleich: 71 aus den USA, 29 aus Großbritannien, 11 aus den Niederlanden, 10 aus China und Hongkong, 9 aus Kanada, 4 aus Frankreich).

Insgesamt hat sich das Ranking mit der Überarbeitung leicht positiv ver-

ändert (etwas verbesserte Indikatoren, Ausweis nach Fachbereichen, ansatzweise höhere Transparenz). Gleichwohl bleiben viele Probleme wie z.B. das hohe Gewicht von Reputationsurteilen (ein Drittel der Bewertung): Obwohl mit knapp 13.400 Rückmeldungen mehr Urteile als in dem Vorgängerranking vorliegen, fehlt doch jeder Hinweis auf Rücklaufquoten und regionale oder fachliche Verteilung. Zudem stellen auch 13.400 Urteile bei einem globalen Ranking nur einen geringen Anteil an allen Forscher/innen weltweit dar. Ebenso scheint der Indikator „Zitationen“ fragwürdig berechnet zu sein: Das gute Ergebnis der Universität Göttingen (Platz 43 und beste deutsche Universität) geht zum großen Teil auf diesen Indikator zurück, der damit höher ist als für die Yale-University. Hier ist zu vermuten, dass wenige, sehr häufig zitierte Artikel stark durchschlagen (wahrscheinlich aus den *physical science*, wo die Universität Göttingen das beste Ergebnis aller Hochschulen weltweit bei diesem Indikator erreicht). Solche guten Einzelperformanzen dürften sich nicht in dieser Weise in einem Ranking auswirken, das den Anspruch erhebt, die gesamte Hochschule abzubilden.

U-Multirank

Nicht zuletzt die aus kontinentaleuropäischer Sicht ärgerliche Dominanz von angelsächsischen Hochschulen in globalen Rankings sowie deren methodische Defizite haben die Europäische Kommission dazu veranlasst, die Entwicklung von methodisch anspruchsvollen weltweiten und europaweiten Rankings zu fördern. Dazu gehört das Projekt „U-Multirank“, das von einem Konsortium mit mehreren europäischen Partnern (Federführung CHE) durchgeführt wird. Das Ziel ist zu prüfen, ob eine Methodik für ein belastbares, multidimensionales globales Hochschulranking entwickelt werden kann, es geht also zunächst noch nicht um die Erstellung eines konkreten Rankings.² Die ersten Ergebnisse dieser Machbarkeitsstudie werden im Juni 2011 veröffentlicht.

Auswirkungen von Rankings

Rankings liefern eine wichtige Orientierung für Studierende, Wissenschaftler/innen und politische Entscheidungsträger/innen. Zudem verstärken sie den Wettbewerb unter den Hochschulen. Globale wie nationale Rankings

² Einen Überblick über das Projekt liefert der Artikel von Gero Federkeil in dieser Publikation.

haben zu institutionellen und hochschulinternen Rückwirkungen geführt, die jedoch kaum abgeschätzt werden können. Es ist anzunehmen, dass die häufig forschungsorientierten Rankings Bestrebungen unterstützt haben, prestigeträchtige Forschungsuniversitäten zu etablieren und zu fördern.

Unabhängig von Problemen der Rankings im Detail oder dem kritisierten, häufig feststellbaren Mangel an Transparenz gibt es auch eine grundlegende Kritik, die letztlich Rankings die Existenzberechtigung insgesamt abzuschreiben versucht. Ein wichtiges Argument in diesem Zusammenhang ist, dass sich Rankings zu sich selbst erfüllenden Prophezeiungen entwickeln können: Hochschulen, die bei Rankings gut abschneiden, würden automatisch gute Studierende und exzellente Wissenschaftler/innen anziehen und könnten dann einfacher an (finanzielle) Ressourcen gelangen. Genau entgegengesetzt funktioniert die Spirale bei vermeintlich „schlechten“ Hochschulen. Rankings verstärken so den Trend, Gewinner und Verlierer zu generieren, und verfestigen diese Einteilung weiter. Rankings, so die Kritik, liefern nicht nur fragwürdige Ergebnisse, sondern tragen durch ihre Existenz auch noch dazu bei, dass problematische Steuerungswirkungen entfaltet werden.

Eine weitere Gefahr besteht in Fehlsteuerungen aufgrund von falschen Anreizen. Insbesondere die Verwendung von quantitativen Indikatoren, die im qualitativen Sinn Qualität messen sollen, könnten zu Fehlsteuerungen führen (im schlimmsten Fall zu weniger Qualität im quantitativen Sinn). Ein Beispiel: Wenn klar ist, dass die Zahl ausländischer Studierender ein wichtiger Indikator für Rankings ist, dann verbessert man die Position im Ranking nicht durch sorgsame Auswahl, gute Propädeutik und hohe Abschlussquoten ausländischer Studierender, sondern vor allem durch eine erleichterte Zulassung. Mittel können also fehlgeleitet werden: weg von der (schwer zu messenden) Verbesserung der Qualität hin zu einer Steigerung der (einfach zu messenden) Quantität.

Eine dritte Befürchtung im Kontext von Rankings ist, dass sie „Einheitsbrei“ statt „Diversität“ fördern. Da Rankings darauf abgestellt sind, nach gleichen Kriterien zu bewerten, werden Aspekte gewünschter und erfolgreicher Spezialisierung häufig nicht abgebildet. Wenn sich alle Hochschulen an denselben Indikatoren orientieren, besteht die Gefahr, dass die sinnvolle Profilbildung vernachlässigt wird.

Es gibt zweifellos viel berechtigte Kritik an Rankings. Trotzdem kann man davon ausgehen, dass eine Hochschule, die eine positive Entwicklung in Lehre und Forschung aufweist, langfristig auch in den meisten Rankings besser abschneiden wird.

Literatur

Hornbostel, Stefan (2007): Theorie und Praxis von Hochschulrankings. In: Statistisches Bundesamt (Hrsg.): Statistik und Wissenschaft. Amtliche Statistik und Hochschulrankings. Band 11. Wiesbaden.

KRITERIEN, AUSSAGEKRAFT UND WIRKUNGEN VON RANKINGS

Dr. Angela Borgwardt wissenschaftliche Publizistin, Berlin

Forschungsrating des Wissenschaftsrats

Entstehungskontext. Im Zuge der zunehmenden Autonomie der Hochschulen ist auch der Informationsbedarf der Hochschulleitungen gestiegen, da sie eine bessere Basis für hochschulpolitische Entscheidungen brauchten. So entstand laut Dr. Rainer Lange, Leiter des Referats Forschung im Wissenschaftsrat, das Konzept des Forschungsratings. Mangels Alternativen hätten die Hochschulleitungen auf verfügbare Statistiken und Rankings zurückgegriffen, deren Eignung sich jedoch als fraglich darstellte, z. B. aufgrund methodischer Probleme und Vereinfachungen. Der Wissenschaftsrat diskutierte über mögliche Folgen dieser Entwicklung und habe aufgrund des zweifellos bestehenden Informationsbedarfs nur einen Weg gesehen: Wenn man Fehlsteuerungen durch eine falsche oder missbräuchliche Verwendung von einfachen Statistiken verhindern wollte, musste man etwas Besseres anbieten. Man wollte ein wissenschaftsadäquates Verfahren entwickeln, das Fehler bisheriger Rankings vermeidet. Von Anfang an sei klar gewesen, dass ein solches Instrument nicht pauschal für ganze Hochschulen konzipiert werden kann, sondern dass eine sinnvolle Bewertung fachspezifische Verfahren voraussetzt.

Ziel des Forschungsratings ist der Vergleich und die Bewertung der Forschungsleistungen der deutschen Hochschulen und der (von Bund und Ländern geförderten) außeruniversitären Forschungseinrichtungen. Der Wissenschaftsrat entschied sich zunächst für die Durchführung eines Pilotprojekts, um festzustellen, ob ein solches fachspezifisches Rating überhaupt realisierbar ist und die Ergebnisse die Erwartungen erfüllen.

Adressaten. Das Rating richtet sich nicht – wie die populärsten Rankings – primär an Studienanfänger/innen, die vor der Aufgabe stehen, eine Hochschule bzw. einen Studienort auszuwählen. Vielmehr soll es eine breite Gruppe von Adressaten erreichen und unterschiedlichen Zwecken dienen:

- An erster Stelle stehen die *Leitungspersonen der Hochschulen* (Instituts-, Fachbereichs-, Fakultätsdirektoren, Präsident/innen, Rektor/innen etc.). Sie sollen dabei unterstützt werden, die Selbsteinschätzung ihrer Institution oder Organisationseinheit zu überprüfen, durch Vergleich im Wissenschaftssystem eine Benchmark zu erstellen und bei Bedarf eine individuelle Evaluation durchzuführen.
- Ein weiterer wichtiger Adressat sind die *Fachgesellschaften/-gemeinschaften*. Sie werden durch das Verfahren herausgefordert, sich innerhalb ihres Fachs über Leistungskriterien zu verständigen, die auch gut nach außen kommunizierbar sind.
- Erreicht werden sollen auch Entscheidungsträger im Bereich *Politik (Ministerien und Landtage)*. Sie sollen in die Lage versetzt werden, sich über die Leistungsfähigkeit der Hochschulen in den jeweiligen Fächern besser zu informieren und damit eine Ergänzung zu den Indikatoren der leistungsorientierten Mittelvergabe erhalten.
- Die vierte Gruppe sind *mögliche Kooperationspartner* (insbesondere aus Wirtschaft/Industrie).

Prinzipien des Verfahrens. Folgende Punkte stellen die wesentlichen Grundzüge des Ratings dar:

- Die Bewertung eines Fachs erfolgt über einen *Informed Peer Review* durch fachspezifische internationale Gutachter/innen – im Gegensatz zu den üblichen Rankings, wo Bewertungen direkt auf der Basis von Daten vergeben werden. Solche gutachtergestützten Verfahren (z. B. auch *Research Assessment Exercise* in Großbritannien) werden meist als Ratings und nicht als Rankings bezeichnet.
- Von zentraler Bedeutung ist *der fachspezifische Ansatz*: Gutachtergruppen, die mit Unterstützung der jeweiligen Fachgesellschaft sorgfältig ausgewählt werden, bewerten einzelne Fächer. Zudem sind sie an der Operationalisierung des Verfahrens beteiligt, d. h. an der Auswahl der Indikatoren und der Definition von fachspezifischen Kriterien.
- Charakteristisch ist die *mehrdimensionale Bewertung*. Von Anfang an werden verschiedene Teilkriterien definiert und bewertet, die bei der Veröffentlichung nebeneinander gestellt, also nicht zu einer einzigen Gesamtbewertung verrechnet werden. Dadurch sind in den Ergebnis-

sen auch die unterschiedlichen Aufgabenprofile der Einrichtungen ablesbar (z.B. ob eher wissenschaftsinterne Qualitäten oder die Transferleistungen positiv zu bewerten sind).

- Für die Veröffentlichung der Ergebnisse ist eine *abgeschichtete Kommunikation* kennzeichnend: Die Bewertungen in den verschiedenen (Teil-)Kriterien werden publiziert, eine Reihe weiterer differenzierter Informationen wird jedoch nicht veröffentlicht. Dies hängt zum einen damit zusammen, dass eine stark differenzierte Bewertung teilweise Rückschlüsse auf einzelne Personen zulässt, die aus Gründen des Persönlichkeitsschutzes nicht veröffentlicht werden können. Zum andern soll verhindert werden, dass diese Daten von Dritten für (vereinfachende) Rankings verwendet werden. Die Daten können den Hochschulleitungen aber für den internen Gebrauch zur Verfügung gestellt werden und dann z. B. hinsichtlich der Frage ausgewertet werden, ob sich die jeweilige Hochschule bei den einzelnen Kriterien (Publikationen, Drittmittel etc.) im innerdeutschen Vergleich im oberen, mittleren oder unteren Drittel befindet. Die Ergebnisse können somit auch Zwecken des internen Benchmarking dienen.

Beim Forschungsrating wird in Bezug auf Indikatoren ein komplexes Set an Informationen herangezogen, das verschiedene quantitative und qualitative Informationen umfasst. Alle Gutachtergruppen berücksichtigen bei ihrer Bewertung grundsätzlich die Einwerbung von Drittmitteln, doch werde dieser Indikator fachspezifisch gewichtet. Zudem werden die Ergebnisse der Forschung, also Publikationen, Promotionen, Zitationen etc. einbezogen. Wichtig sei, so Lange, das Gesamtbild, wobei sich in jedem Fach unterscheide, welche Bedeutung die verschiedenen Indikatoren haben und was zur üblichen Praxis gehört. Zur Begutachtung gehören auch Befragungen, die dem Einzelnen die Gelegenheit geben, die individuelle Forschungsstrategie zu beschreiben, etwa: Welche Forschungsschwerpunkte wurden gewählt und weshalb? Warum wird beispielsweise nicht kurzfristig an einem Drittmittelprojekt, sondern langfristig an einer Monografie gearbeitet? Die Gutachter/innen können sich auf die jeweilige Argumentation einlassen und in ihrer Bewertung verschiedene Informationen des Einzelfalls berücksichtigen, wie es nur in einem Begutachtungsverfahren möglich ist.

Vor- und Nachteile. Den größten Vorteil dieses Verfahrens sieht Lange darin, dass sehr differenzierte und aussagekräftige Ergebnisse gewonnen werden. Der größte Nachteil bestehe in dem großen Aufwand, der mit einer fächerspezifischen Begutachtung verbunden ist. Da es keinen sche-

matischen Weg von den Indikatoren zur Bewertung gibt, lebe das Verfahren zudem von einer „gewissen Intransparenz“. Man könne nicht vorher sagen, welches Gewicht die einzelnen Indikatoren (Drittmittel, Publikationen etc.) haben, sondern es werde fallweise entschieden. Deshalb müsse man nach außen deutlich vermitteln, dass die Unabhängigkeit der Gutachter/innen gewahrt ist und dafür sorgen, dass Vertrauen in dieses Verfahren erzeugt wird.

Die Entwicklung des Forschungsratings wurde allerdings auch von kritischen Stimmen begleitet: So hatte z. B. der Historikerverband am Pilotprojekt nicht teilnehmen wollen und sich dagegen ausgesprochen, die Geschichtswissenschaft in das Rating einzubeziehen. Als ein Hauptargument war die Vermutung geäußert worden, man kenne die Ergebnisse ohnehin schon im Voraus und wisse, wer im eigenen Fach „die guten Leute“ sind. Dies habe sich während der Erprobungsphase aber nicht bewahrheitet, so Lange. Die „zehn Besten“ eines Faches seien zwar bekannt, nicht aber die zahlreichen anderen guten Wissenschaftler/innen. Zudem müsse das Wissen um besondere Stärken auch begründet an Hochschulen und Politik vermittelt werden – und nicht nur fachinterner *common sense* bleiben.

Das zweite Hauptargument des Historikerverbandes gründete auf der Befürchtung unabsehbarer Folgen: Man könne nicht wissen, welche Auswirkungen die Ergebnisse auf die Forschungspolitik haben werden. Diese Ergebnisoffenheit gehört für Lange jedoch zum Konzept des Verfahrens: Der Wissenschaftsrat wolle ja keinen Automatismus errichten, sondern strategische Entwicklungsprozesse und einen Dialog innerhalb der Hochschulen unterstützen. Außerdem gehe es bei einem Forschungsrating dieser Art nicht unbedingt darum, überraschende Ergebnisse zu erzielen. Es könne seinen Zweck auch dann erfüllen, wenn sich gar nicht viel ändern muss, sondern durch eine Bestätigung von außen eine gewisse Beruhigung in einer Forschungsinstitution erreicht werden kann.

Perspektive. Um den Aufwand des Verfahrens zu reduzieren, setzte der Wissenschaftsrat zur Weiterentwicklung die Unterarbeitsgruppe „Standardisierung der Datenerhebung“ ein, die seit Juli 2010 Empfehlungen zur Optimierung der Datenerhebung erarbeitet (Vorsitz: Prof. Dr. Peter Lichter, Deutsches Krebsforschungszentrum Heidelberg). Die Entscheidung über die Zukunft des Ratings ist gegenwärtig noch offen; sie wird nach Abschluss der Pilotstudie (Ende 2012) gefällt werden.

Forschungsrating des Wissenschaftsrates

Die Pilotstudie des Forschungsratings läuft seit 2005. In der 1. Phase wurden die Fachgebiete Chemie und Soziologie exemplarisch untersucht, die Ergebnisse wurden im Winter 2007/2008 veröffentlicht. Aktuell befindet sich das Projekt in der 2. Phase, in der das Verfahren wiederum an zwei Fächern (Elektro- und Informationstechnik, Anglistik/Amerikanistik) erprobt wird. Die Veröffentlichung der Ergebnisse ist im Sommer 2011 bzw. im Herbst/Winter 2012 zu erwarten.

Quelle: Website des Wissenschaftsrates,
<http://www.wissenschaftsrat.de/arbeitsbereiche-arbeitsprogramm/forschungsrating/>

Hochschulranking des CHE Centrum für Hochschulentwicklung

Entstehungskontext. Wie entstanden Rankings in Deutschland? Prof. Dr. Detlef Müller-Böling von CHE Consult, der ehemalige Leiter des Centrums für Hochschulentwicklung, erläutert ihre Entstehungsgeschichte: Der Anlass zur Entwicklung der CHE-Rankings (wie auch anderer Rankings) sei das Hochschulranking des Nachrichtenmagazins „Der Spiegel“ 1989 gewesen, das die Universität-Gesamthochschule-Siegen zum Spitzenreiter gekürt hatte. Die Ergebnisse des Rankings riefen nach Müller-Böling eine regelrechte Eruption in Wissenschaft und Öffentlichkeit hervor: Einerseits hätten zahlreiche Hochschulen das Ranking mit dem Hinweis auf methodische Mängel grundsätzlich abgelehnt und sich dadurch gegen jede Kritik immunisiert, andererseits sei es in der Öffentlichkeit auf sehr große Resonanz gestoßen. Der enorme Absatz dieses Spiegel-Titels hatte das „unglaubliche Interesse an solchen Informationen über Hochschulen“ deutlich gemacht, so Müller-Böling. In der Folge brachten weitere große Magazine Hochschulrankings heraus. Während der „Spiegel“ Studierende befragt hatte, führte der „Stern“ eine Professorenbefragung durch und „Bild der Wissenschaft“ publizierte bibliometrische Analysen. Jedes Mal seien vollkommen andere Ergebnisse herausgekommen, was in den Hochschulen die Immunisierung und Abwehrhaltung nur noch weiter verstärkt habe.

In der Hochschulrektorenkonferenz (HRK) habe sich aber zunehmend die Einsicht durchgesetzt, dass man „mit dem Einmauern und nur Geld-

verlangen im Rahmen einer massifizierten Hochschullandschaft“ nicht mehr weitermachen könne, und dass man sowohl gegenüber der Öffentlichkeit als auch nach innen Rechenschaft nach vergleichbaren Kriterien ablegen müsse. Lange Zeit hätten sich die Hochschulleitungen zur Steuerung lediglich auf einen quantitativen Indikator (Kapazitätsverordnung) gestützt, darüber hinaus habe es allenfalls verbreitete Meinungen über die wissenschaftliche Qualität von Kolleg/innen gegeben. Die Qualität von Wissenschaft sei nicht zugänglich oder transparent gewesen. Zudem habe das Kollegialitätsprinzip häufig einen Einblick in die tatsächlichen Forschungsleistungen verhindert.

Vor diesem Hintergrund sei immer deutlicher geworden, dass auch im Wissenschaftssystem mehr Transparenz und Wettbewerb notwendig waren. Die HRK hatte daraufhin das Projekt „Profilbildung“ aufgelegt, in dessen Rahmen einige Hochschulen Informationen zusammentrugen mit dem Ziel, zu wissenschaftsadäquateren Ergebnissen als die großen Magazine zu kommen. Letztlich sei bei diesem Projekt aber nichts Substantielles herausgekommen. Ein wichtiger Gründungsauftrag des Centrums für Hochschulentwicklung³ war deshalb nach Müller-Böling, die Transparenz über das sich wandelnde Hochschulsystem zu erhöhen, um die Steuerungs- und Wettbewerbsfähigkeit der Hochschulen zu verbessern. Man sei zu der Erkenntnis gelangt, dass das Hochschulsystem reformiert werden musste, und ein Ranking sollte Bestandteil davon sein.

Adressaten. Ein wesentlicher Grund für das Scheitern des HRK-Projekts Profilbildung wurde darin gesehen, dass die angestrebten Zielgruppen zu breit gestreut waren (Hochschulleitungen, Landesregierungen, Studierende, Eltern etc.). Dadurch sei die erhobene Datenmenge für die jeweilige Zielgruppe nicht mehr zu verarbeiten gewesen. Das CHE habe deshalb aus diesem Projekt zwei getrennte Projekte entwickelt: zum einen den Benchmarking Club, der auf eine verbesserte Entscheidungsfähigkeit innerhalb der Hochschulen zielt und auf internem Datenaustausch beruht, und zum anderen das HochschulRanking, das vor allem der Unterstützung von Studieninteressierten bei der Studienwahlentscheidung, aber auch der Information von Studierenden und breiterer Öffentlichkeit dienen soll.

³ Das Centrum für Hochschulentwicklung gGmbH wurde 1994 von der HRK und der Bertelsmann Stiftung als gemeinnützige GmbH gegründet und versteht sich als „Reformwerkstatt für das deutsche Hochschulwesen“. Es wird wesentlich von der Bertelsmann Stiftung finanziert. 2001 entstand als Ausgründung die kommerziell arbeitende CHE Consult GmbH, die Beratungsdienste für Hochschulen anbietet und Projekte für Ministerien und Stiftungen durchführt. Vgl. <http://www.che.de/>.

CHE-Ranking

Das CHE-Ranking umfasst verschiedene Rankings: Das CHE-HochschulRanking, das 1998 zum ersten Mal vorgestellt wurde, ist ein Ranking deutschsprachiger Universitäten, Fachhochschulen und Berufsakademien und richtet sich vor allem an Studienanfänger/innen und Studierende, soll aber auch den Hochschulen selbst und der breiteren Öffentlichkeit als Informationsquelle dienen. Das CHE-ForschungsRanking bereitet Daten unter dem Gesichtspunkt der universitären Forschungsleistungen auf und weist besonders leistungsstarke Fakultäten aus. Das CHE-ExcellenceRanking identifiziert herausragende Fachbereiche in Europa (zunächst nur für naturwissenschaftliche Fächer, VWL, Politologie und Psychologie) und das CHE-/dapm-Employability-Rating bewertet Bachelorstudiengänge im Hinblick auf die Beschäftigungsbefähigung der Absolvent/innen.

Quelle: <http://www.che-ranking.de>.

Grundzüge des Verfahrens. Das CHE-HochschulRanking hat folgende wesentliche Merkmale:

- *fachbezogener Vergleich:* Es werden Fächer und nicht ganze Hochschulen miteinander verglichen, da die Hochschulen über verschiedene Profile mit eigenen Stärken und Schwächen in unterschiedlichen Fächern verfügen.
- *multidimensionales Verfahren:* Um wissenschaftsadäquate Indikatoren zu erarbeiten, wurden von Anfang an externe Fachleute (Dekane, Fachgesellschaften) einbezogen. Es werden so viele Einzelindikatoren wie möglich (und angemessen) analysiert, die nebeneinander stehen bleiben. Innerhalb eines Faches werden die Einzelindikatoren also nicht gewichtet und zu einem Gesamtwert aufaddiert. Auf diese Weise sollen die unterschiedlichen Profile der jeweiligen Hochschulen und Institute sichtbar werden.
- *Perspektivenvielfalt:* Die Hochschulen werden aus unterschiedlichen Perspektiven betrachtet, um ein differenziertes Bild zu erhalten. Einbezogen werden Daten zu den einzelnen Fachbereichen und Studiengängen, aber auch die Perspektiven der Hochschullehrer/innen und Studierenden.

- *Ranggruppen statt Rangplätze:* Die Hochschulen erhalten keine einzelnen Rangplätze, sondern sie werden Ranggruppen zugeordnet (Spitzen-, Mittel- und Schlussgruppe).

Vor- und Nachteile. Der entscheidende Vorteil dieses Rankings ist für Müller-Böling, dass die Gewichtung der Einzelwertungen nicht vorgeschrieben wird, d. h. jeder könne die Indikatoren nach seinen individuellen Informationsbedürfnissen herausgreifen. Studieninteressierte können zum Beispiel nach ihren persönlichen Präferenzen die Hochschule oder den Studiengang auswählen. Bei vielen anderen Rankings, die Indikatoren zu einer Gesamtwertung zusammenfassen, sei dagegen die Rangliste automatisch durch die Gewichtung von bestimmten Indikatoren festgelegt.

Als weitere positive Auswirkung nannte Müller-Böling, dass durch das CHE-Ranking der Wettbewerb und die Profilbildung der Hochschulen befördert worden seien: Die Ergebnisse hätten die Strukturentwicklung an den Hochschulen außerordentlich angeregt und beflügelt. Wie erwünscht seien hochschulinterne Diskussionen auf der Basis dieser Ergebnisse geführt worden, die auch Folgen für strategische Entscheidungen auf Fakultäts- oder Hochschulebene hatten. Allerdings habe sich auch ein nicht intendierter, dysfunktionaler Effekt eingestellt: Einige Fächer hatten kritisiert, dass sie in der hochschulpolitischen Diskussion nicht mehr vorkommen, weil sie im CHE-Ranking nicht erfasst worden waren. Daraufhin habe man weitere Fächer wie Romanistik oder Sportwissenschaft mit aufgenommen. Nach Müller-Böling zeigte sich noch eine weitere unbeabsichtigte, allerdings sehr positive Auswirkung: Die im CHE-Ranking erhobenen Daten seien eine wichtige Informationsquelle gerade für Studieninteressierte aus Nichtakademikerhaushalten, die bei der Information über Studienmöglichkeiten und -bedingungen per se benachteiligt seien.

Bezogen auf die Differenzierung nach Fächern, die Multidimensionalität und die Perspektivenvielfalt stelle sich ein grundsätzliches Dilemma: Studienanfänger/innen und deren Eltern wünschen sich im Entscheidungsfindungsprozess über die geeignete Hochschule möglichst übersichtliche, klare bis hin zu einfachen Informationen („die beste Hochschule“), während Wissenschaftler/innen und hochschulpolitische Entscheider/innen ganz andere und wesentlich detailliertere Ergebnisse benötigen. Dazwischen einen ausgewogenen Weg zu finden, sei eine große, bleibende Herausforderung, für die immer im konkreten Einzelfall eine Lösung gesucht werden müsse.

Perspektive. Das CHE-Ranking will auch auf die beginnende Entstehung eines europäischen Hochschulraums reagieren. Deshalb wurden auch

Hochschulen in den Niederlanden sowie in der Schweiz und in Österreich beteiligt. Bei den Pilotprojekten in Österreich und der Schweiz habe man jedoch sehr viel Lehrgeld bezahlen müssen, so Müller-Böling, auch im Umgang mit den Institutionen.⁴ 2009 wurde das Ranking noch einmal erweitert, indem einzelne Hochschulen im Ausland aufgenommen wurden, die für Studierende aus Deutschland interessant sein könnten, da sie z. B. ein deutschsprachiges Studienangebot bereitstellen; so sind mittlerweile auch deutschsprachige medizinische Hochschulen in Ungarn dabei.

Förder-Ranking der Deutschen Forschungsgemeinschaft (DFG)

Entstehungskontext. Die DFG hat sich als mitgliedergetragene Einrichtung schon früh der Diskussion um den grundlegenden Wandel des Hochschulsystems gestellt und die fachliche und strukturelle Ausdifferenzierung der Hochschullandschaft in einem nationalen und internationalen Wettbewerb in ihre Überlegungen einbezogen. Dr. Jürgen Güdler, Leiter der Gruppe Informationsmanagement der Deutschen Forschungsgemeinschaft, erläuterte, dass bereits 1996 einige große Mitgliedshochschulen auf die DFG zugekommen seien mit der Frage, wie viele DFG-Drittmittel ihnen im Vergleich zu anderen Hochschulen bewilligt wurden. Die DFG veröffentlichte daraufhin ein Ranking der Top Ten. Danach hätten auch andere Hochschulen daran teilnehmen wollen. Das erste Ranking war noch ein reines „Tonnagen-Ranking“, bei dem nur die Höhe der DFG-Drittmittel dargestellt wurde. In den Top-Hochschulen sei diese Information zwar für die Öffentlichkeitsarbeit wichtig, aber für die Hochschulleitung nicht planungsrelevant gewesen. Es sei deutlich geworden, dass für einen aussagekräftigeren und regelmäßigen Service für die Hochschulen wesentlich mehr Aufwand betrieben werden musste.

Drittmittel seien eine „harte Währung“, so Güdler, insbesondere DFG-Drittmittel, die im *Peer Review*-Verfahren vergeben werden. Das alle drei Jahre erscheinende DFG-Ranking bezieht nicht nur die Gesamtsumme der Drittmittelinwerbungen, sondern auch die Anteile der jeweiligen Fächer

4 Bei den Pilotprojekten waren alle Hochschulen in der Schweiz und in Österreich einbezogen. 2007 kam die Zusammenarbeit mit den Rektorenkonferenzen und den Datenerhebungsagenturen der beiden Länder zu einem Ende. Damit einher ging eine öffentliche Auseinandersetzung (Kritik an „methodischen Mängeln“ versus „veränderte Internationalisierungsstrategie“). Heute nehmen noch einzelne Hochschulen in der Schweiz und in Österreich am CHE-Ranking teil.

ein. Hier stand die DFG von Anfang an im Austausch mit der Hochschulrektorenkonferenz (HRK) mit dem Ziel, den Hochschulen in ihrer fachlichen Profilbildung gemeinsam als Servicepartner zur Seite zu stehen. Die ersten DFG-Rankings 1997 und 2000 berücksichtigten nur die DFG-Bewilligungen. Dann wurde das Spektrum Schritt für Schritt ausgeweitet: 2003 wurden Daten, die über die von DAAD und Alexander von Humboldt-Stiftung (AvH) geförderten Aufenthalte internationaler Gastwissenschaftler/innen Auskunft geben, aufgenommen. Damit war es möglich, den Drittmittelerfolg bei der DFG der „Abstimmung mit den Füßen“ renommierter Spitzenwissenschaftler/innen gegenüberzustellen – mit dem Befund einer hohen Übereinstimmung.

Förder-Ranking der DFG

Der Bericht der DFG erscheint seit 1997 alle drei Jahre. Die letzte und fünfte Ausgabe wurde 2009 publiziert und weiterentwickelt, indem fachbezogene Analysen ausgeweitet wurden: So wird z. B. mit Blick auf die Forschungsförderung der DFG zwischen 48 verschiedenen Forschungsfeldern unterschieden. Zudem wurde die Darstellung der fachlichen Profilbildung anhand der DFG-Bewilligungen durch neue Förderinformationen ergänzt (direkte Projektförderung des Bundes, 6. EU-Forschungsrahmenprogramm). Als weitere Neuerung wurde bei dem eingeworbenen DFG-Bewilligungsvolumen nach dem Geschlecht der Antragstellenden differenziert. Hinzu kommt eine Verstärkung der „europäischen Dimension“, indem zum ersten Mal Daten zur Beteiligung an den Förderlinien des im Jahr 2007 neu etablierten *European Research Councils* (ERC) herangezogen wurden.

Quelle: Deutsche Forschungsgemeinschaft: Förder-Ranking 2009. Institutionen – Regionen – Netzwerke. Fachliche Profile von Hochschulen und außeruniversitären Forschungseinrichtungen im Licht öffentlich geförderter Forschung. Weinheim 2009, S. 10, http://www.dfg.de/download/pdf/dfg_im_profil/evaluation_statistik/ranking/ranking_2009/gesamtbericht.pdf.

Mit der seit 2006 erfolgten Integration von Daten zur Förderung durch die Ministerien des Bundes sowie der Europäischen Union (seit 2009 einschließlich ERC) gelang es schließlich, das monetäre „Förderer-Profil“ auf breiterer Basis zu erfassen und abzubilden. Inzwischen deckt das DFG-Förder-Ranking etwa 90 Prozent aller Drittmittel ab, die öffentliche Institutionen an die deutsche Wissenschaft vergeben. Das Statistische Bundesamt stellt ergänzend Daten bereit, die Auskunft über

das Engagement der Wirtschaft für die Hochschulforschung geben. Damit verfügt das Förder-Ranking der DFG im internationalen Vergleich nach Ansicht von Güdler über eine einzigartige empirische Grundlage für Aussagen zu den Forschungs- und Förderprofilen von Hochschulen sowie – auch dies eine Besonderheit – von außeruniversitären Forschungseinrichtungen.

Adressaten. Die DFG-Rankings sind vorrangig ein Service für die DFG-Mitgliedshochschulen. Mit der vergleichenden Darstellung ihrer Forschungsprofile will die DFG den Hochschulen ein wichtiges Monitoring-Instrument ihrer (öffentlich geförderten) Forschungsaktivitäten an die Hand geben. Die aufbereiteten Daten können darüber hinaus auch als Instrument für internationales Forschungsmarketing dienen. Studienanfänger/innen sind keine wesentliche Zielgruppe, wohl aber fortgeschrittene Studierende, die sich für eine wissenschaftliche Laufbahn entscheiden.

Seit 2009 werden die Daten nicht nur in einer Publikation veröffentlicht, sondern die DFG berät ihre Mitglieder auch auf Anfrage, wie sie das Ranking interpretieren und welche Kennzahlen zu strategischen Planungszwecken der Hochschule relevant sein können. Seit 2009 werden den Mitgliedshochschulen auch jährlich individuelle Berichte zur Verfügung gestellt, die das spezielle Profil der Hochschule noch detaillierter und zeitnäher abbilden. Der Profilbildungs-/Benchmarking-Charakter des Rankings spielt auch im Kontext der Exzellenzinitiative und der dazugehörigen Antragstellung eine wichtige Rolle.

Prinzipien des Verfahrens. Die wesentlichen Grundzüge des DFG-Rankings beschreiben folgende Punkte:

- *Fokussierung auf Drittmittelwerbung* der Hochschulen bei der DFG und anderen Institutionen der öffentlichen Hand
- *Fächer- und themengebietebezogene Perspektive:* Aussagen zum fachlichen Profil einer Hochschule beziehen sich bislang nicht auf deren Fachbereiche und Fakultäten, sondern greifen auf die Informationen zurück, die Förderer bezüglich der fachlichen Ausrichtung der von ihnen bewilligten Projekte erfassen. Im Falle der DFG wird hierzu eine 48 Fächer unterscheidende Klassifikation genutzt. Im Falle von Bund und EU sind es die von diesen ausgeschriebenen Förderlinien (z.B. Energieforschung und -technologie), die das Profil einer Forschungseinrichtung prägen. Dargestellt wird, wie viele Drittmittel ein bestimmtes Fach eingeworben hat, und zum anderen wird die ganze Einrichtung

betrachtet und verdeutlicht, wie sich das Profil dieser Einrichtung in ihrer Breite und Differenzierung zusammensetzt.

qualitative und quantitative Dimension: Die DFG-Gelder werden auf Basis eines Begutachtungsverfahrens durch Expert/innen vergeben. Bei Drittmitteln spielt nicht nur die Höhe der Finanzierung eine Rolle, sondern auch die Größe der Hochschule und in welchen Fächern und mit welchen Kooperationspartnern das Geld eingeworben wurde.

Nach Güdler ermöglicht der Indikator Drittmittel eine recht hohe Differenzierung, wenn verschiedene Drittmittelquellen und -verwendungen einbezogen werden, da diese jeweils anderes aussagen. Auf diese Weise könne man verschiedene Facetten der Profilbildung betrachten, z. B. Anwendungs- oder Grundlagenorientierung, aber auch Kooperationen mit außeruniversitären Einrichtungen. Zudem werden kartografische Darstellungen erstellt, die z. B. Aussagen über Investitionen des Bundes in verschiedenen Regionen oder über die Art von geförderter Forschung ermöglichen. Auf diese Weise kann auch das regionale Profil verdeutlicht und damit das wissenschaftliche Umfeld der an einem Ort angesiedelten Hochschule bzw. Hochschulen dargestellt werden. Die DFG habe intensiv daran gearbeitet, um aus den vermeintlich eindimensionalen Drittmitteldaten ein möglichst facettenreiches Bild der Forschungslandschaft in Deutschland zu zeichnen.

Zur Bedeutung des Indikators „Drittmittel“ gibt auch eine Befragung von Hochschulprofessor/innen Auskunft, die das Institut für Forschungsinformation und Qualitätssicherung (iFQ) im Auftrag der DFG durchgeführt hat. Darin erklären die Professor/innen, wie oft und in welchem Umfang sie in den letzten fünf Jahren einen Antrag auf Drittmittel bei der DFG gestellt haben. Immerhin 73 Prozent haben nach eigener Aussage in diesem Zeitraum einen Förderantrag gestellt. Zudem zeigte sich, dass 89 Prozent aller Wissenschaftler/innen überhaupt schon einmal Drittmittel beantragt haben.⁵ Um forschen zu können, sei die Einwerbung von Drittmitteln heute in allen Fächern Normalität, auch wenn sie bei Lebens- und Naturwissenschaftler/innen eine höhere Bedeutung habe als bei Geistes- und Sozialwissenschaftler/innen. Nach Güdlers Auffassung ist der Drittmittelindikator sehr aussagekräftig, während die meisten anderen Indikatoren, wie z. B.

⁵ Böhmer, Susan/Neufeld, Jörg/Hinze, Sybille/Klode, Christian/Hornbostel, Stefan: Wissenschaftler-Befragung 2010: Forschungsbedingungen von Professorinnen und Professoren an deutschen Universitäten. iFQ-Working Paper No. 8, März 2011, S. 36 ff.

Publikationen⁶, Patente oder Stimmungsbilder von Professor/innen nicht im Ansatz so belastbar seien, weil sie immer nur kleine Ausschnitte abbilden und zudem in der Regel ein deutlich weniger aktuelles Bild zeichnen.

Vor- und Nachteile. Einer der größten Vorteile des DFG-Rankings ist laut Güdler, dass es auf Daten basiert, die von Forschungsförderern direkt zur Verfügung gestellt werden. Daher würden Hochschulprofessor/innen oder Verwaltungsmitarbeiter/innen keine Zeit damit verlieren, Daten zu erheben oder umfangreiche Fragebögen auszufüllen, sondern sie können sich auf ihre Arbeit konzentrieren, während die DFG das Ranking als Serviceleistung erstellt. Die DFG verfügt über eine sehr gute eigene Datenbasis in diesem Bereich, ebenso die Kooperationspartner BMBF, DAAD etc. Die Zielgruppe habe also keinerlei Aufwand bei der Erstellung der Rankings. Ein weiterer Vorteil sei, dass auch Daten über die außeruniversitäre Forschung bereitgestellt werden. Die Komplexität und Informationsfülle des Materials mache es allerdings erforderlich, viel Energie in die Erklärung und Interpretation der vorgenommenen Analysen zu stecken. Die veröffentlichten Rankingergebnisse umfassen 200 Seiten mit einem umfangreichen Erklärungsapparat, da bei einer reinen Veröffentlichung von Tabellen Missverständnisse vorprogrammiert seien.

Einen Nachteil sieht Güdler darin, dass man eine Passung herstellen muss, wenn Daten aus unterschiedlichen Quellen zusammengeführt werden, die nach unterschiedlichen Systematiken (z. B. Fächersystematiken) erhoben wurden. Dafür wurden Darstellungsformen gefunden, die zwar komplex seien, aber deutlich übersichtlicher als die seitenlange Aufführung von Tabellen. Für die Hochschulen sei letztlich weniger die Frage entscheidend, wie man weiter „nach oben“ kommt oder den „Abstieg“ verhindert, sondern vielmehr, ob und in welche Richtung sich das Profil einer Hochschule entwickelt – das könnten die DFG-Rankings recht gut zeigen.

Perspektive. Künftig soll in den DFG-Rankings nach Abstimmung mit der HRK das Thema der personellen Relativierung (z. B. Drittmittel je Professur) gestärkt werden. Um hier auch fachlich differenzierte Aussagen treffen zu können, die mit den bisher genutzten Daten der Statistischen Ämter nur unzureichend zu bedienen seien, werde diesbezüglich mittelfristig eine engere

6 Was insbesondere der Datenlage geschuldet ist: Vor allem internationale Rankings fokussieren maßgeblich auf Publikationen (und Zitationen), die in Datenbanken erfasst sind, die ausschließlich Aufsätze in internationalen Fachzeitschriften berücksichtigen. Eine solche Basis wird weder den Geistes- und Sozialwissenschaften noch den Ingenieurwissenschaften mit ihren je eigenen Publikationskulturen gerecht.

Abstimmung mit den DFG-Mitgliedshochschulen erforderlich sein. Außerdem ist eine Verbesserung des Online-Angebots geplant, insbesondere in Bezug auf die Detailtiefe und kartografische Darstellungen. Um das Förder-Ranking auch im internationalen Marketing für den „Forschungsstandort Deutschland“ besser einsetzen zu können, wird schließlich für die seit 2006 angebotene englische Fassung „Funding Ranking“ ein kürzeres, fokussiertes, besser auf internationale Zielgruppen zugeschnittenes Format entwickelt.

Kontrovers diskutiert: Bedeutung des Indikators Drittmittel?

Bedeutung variiert in den Fächern: Die Aussage, dass der Indikator Drittmittel der aussagekräftigste bzw. belastbarste sei, stieß auf Widerspruch: Dabei werde übersehen, dass die Relevanz der Drittmittelinwerbung in den verschiedenen Wissenschaftsbereichen sehr unterschiedlich sei (in den Naturwissenschaften sehr wichtig, in den Geisteswissenschaften eher untergeordnet). Deshalb könnten auf diesen Indikator gestützte Rankings keine aussagekräftigen Ergebnisse über die gesamte Hochschullandschaft erbringen. Dem wurde wiederum entgegengehalten, dass DFG-Rankings zwar nicht für alle Fachbereiche wichtig seien, faktisch aber erstaunliche Korrelationen zwischen eingeworbenen Drittmitteln, Landeszuschüssen und dem Wissen, wo „die Besten“ forschen, festzustellen seien. Diese Übereinstimmung sei sicher auch darauf zurückzuführen, dass bei der Drittmittelvergabe qualitative Faktoren eine große Rolle spielen, da die Forschungsleistung im Einzelfall bewertet wird, z. B. auf der Basis eines genauen Begutachtungsprozesses durch Fachexperten bei DFG-Drittmitteln. Eine solche externe Bewertung von Forschungsqualität könne den Hochschulleitungen Entscheidungen erleichtern. Viel schwieriger sei die Einschätzung von Forschungsleistungen, die nicht durch die DFG begutachtet wurden. Auch wenn der Indikator Drittmittel in einigen Wissenschaftsbereichen nicht so stark gewichtet werden könne wie in anderen, sei doch für viele Fächer eine gewisse Aussagekraft gegeben.

Ein Indikator unter mehreren: In verschiedenen Beiträgen kam zum Ausdruck, dass kein Indikator „der beste“ sein könne, da man grundsätzlich immer viele Indikatoren brauche, in denen sich Unterschiedliches abbildet. Nur dann könnten aussagekräftige Ergebnisse über Forschungsleistungen erzielt werden. Güdler stimmte zu, dass mehrere Kennzahlen herangezogen werden müssten und der Indikator Drittmittel nicht alleinseligmachend sei, doch sei es derzeit „der beste, den wir haben“. Wolle man die Forschungs-

leistung z. B. anhand von Publikationen bewerten, müsse man sehr großen Aufwand betreiben, was beim Rating des Wissenschaftsrats ja deutlich geworden sei. Schon in kleinsten Fächern sei allein die Erstellung einer Publikationsübersicht enorm aufwendig. Hinzu komme das Problem, dass Publikationen und Zitationen nur das abbilden, was in der Vergangenheit geleistet wurde, während Drittmittel das Gegenwärtige widerspiegeln.

In der Diskussion wurde die grundsätzliche Problematik des Drittmittelindikators deutlich: Die Hochschulen sind aufgrund der sinkenden Grundfinanzierung dazu gezwungen, die Forschung zunehmend über Drittmittel zu finanzieren, was die Freiheit von Forschung und Lehre gefährden kann. DFG-Drittmittel werden zwar als Ausdruck einer hohen Qualität von Forschungsprojekten gesehen, doch wäre es gefährlich, sie als Kriterium zu stark zu machen. Man müsse dringend mehrere Indikatoren heranziehen, um gute Entscheidungen zu treffen und Fehlsteuerungen in der Wissenschaft verhindern zu können.

Zum Verhältnis zwischen Drittmitteln und dem Grundhaushalt einer Hochschule merkte Güdler an, dass an deutschen Hochschulen Drittmittel nur etwa 20 Prozent des Gesamtetats ausmachen, 80 Prozent seien immer noch Grundausrüstung. Bei der Steuerung einer Hochschule könne man auf Drittmittel-Kennzahlen zurückgreifen, doch halte er es für einen Fehler, die Grundausrüstung an Drittmitteleinwerbungen zu koppeln. Wenn dies dennoch passiere, sei das nicht die Schuld des Indikators Drittmittel, sondern Folge einer Naivität im Umgang mit dieser Kennzahl. Jede Hochschulleitung stehe in der Verantwortung, Drittmittel-Zahlen im DFG-Ranking nicht absolut zu setzen, sondern nur als einen Beitrag zur Entscheidungsfindung zu betrachten.

Problematische politische Effekte: Als wenig überzeugend wurde die Aussage gewertet, die Bedeutung des Indikators Drittmittel werde dadurch untermauert, dass ein hoher Prozentsatz der Wissenschaftler/innen Drittmittel beantragt. Die große Zahl der Anträge sei lediglich Ausdruck einer problematischen Entwicklung: Viele Hochschulprofessor/innen meinten, ohne Drittmittel keine vernünftige Forschung mehr betreiben zu können und jeder, der sich nicht an Drittmitteleinwerbungen beteilige, stehe unter großem Rechtfertigungsdruck. Inzwischen hätten sich negative politische Effekte eingestellt, indem Drittmittel zu einem entscheidenden Indikator für jegliche Mittelverteilung geworden seien – auf Instituts-, Fachbereichs-, Hochschul- und Landesebene. So sei z. B. die Mittelverteilung in Berlin entscheidend auf die Drittmittelverteilung im Rahmen der Exzellenzinitiative zurückzuführen. Mit dieser Tendenz sei eine gravierende Gefahr verbunden:

Wenn man Drittmittel so wichtig nehme – unabhängig davon, welche Forschungsleistungen dabei herauskommen –, und die Verteilung der Normalbudgets immer leistungsabhängiger gestaltet wird, werde eine ausreichende Grundfinanzierung zunehmend in Frage gestellt – und damit genau das, was lange Zeit die hohe Qualität der deutschen Hochschulen ermöglichte. Problematische (wissenschafts-)politische Konsequenzen seien heute schon in den USA sichtbar: Aktuell als „angesagt“ geltende Themen werden wahrscheinlicher gefördert als andere Themen, wobei sich die Konjunkturen der Themen immer wieder verändern. Es sei doch sehr fragwürdig, ob diese Entwicklung langfristig gut für Wissenschaft und Forschung ist.

Nutzen von Rankings

Folgende Aspekte wurden in der Diskussion als nützlich hervorgehoben:

Verschiedene Rankings für unterschiedliche Ziele und Zielgruppen: In der Diskussion dominierte die Auffassung, dass Rankings offensichtlich bestehende Informationsbedürfnisse befriedigen. Deshalb ist es auch nicht sinnvoll, sie grundsätzlich abzulehnen. Vielmehr sollte nach den besten Kriterien, Indikatoren und Methoden für Rankings gesucht werden, um den größtmöglichen Nutzen für bestimmte Zwecke zu erreichen. Es kann nicht das *eine* richtige Ranking geben, aber verschiedene Rankings für spezifische Bedarfe, unterschiedliche Ziele und Adressaten.

Einfluss auf Studienwahl: Auch wenn Studieninteressierte ihren Studienort auf der Basis verschiedener Faktoren auswählen, spielen Rankings als Entscheidungshilfe für viele eine wichtige Rolle. Dies zeigt sich z. B. an der intensiven Nutzung des CHE-Hochschulrankings durch deutsche Studierende, aber auch bei internationalen Studierenden, die sich nach Untersuchungen des DAAD bei ihrer Hochschulwahl stark von Rankings beeinflussen lassen.

Entscheidungsgrundlage für Hochschulleitungen: Viele Führungspersonen in Hochschulen wünschen sich extern ermittelte Ergebnisse und Einschätzungen, um strategische Entscheidungen und Zielvereinbarungen auf breiter fundierter Basis treffen zu können. Dieses Ziel erfordert allerdings, dass die Daten sehr differenziert und fachspezifisch erfasst werden, wie es z. B. beim Forschungsrating des Wissenschaftsrats der Fall ist.

Instrument für Transparenz und Leistungsmessung: Rankings ermöglichen mehr Transparenz, indem durch Leistungsmessungen deutlich wird,

wo die einzelne Hochschule und das deutsche Hochschulsystem im nationalen und internationalen Vergleich stehen. Bei aller Kritik an Rankings *sollte nicht vergessen werden, dass in der Zeit vor der Leistungsmessung im Hochschulsystem häufig Entscheidungen und Mittelverteilungen auf irrationaler Basis und auf intransparenten Wegen getroffen wurden.*

Enorme Steuerungswirkungen: Rankings können im wissenschaftspolitischen Bereich große Steuerungswirkungen entfalten, selbst wenn sie gravierende methodische Fehler aufweisen. Das zeigt das Beispiel Shanghai-Ranking: Die fragwürdig erhobenen Daten wurden zwar nicht zur maßgeblichen Grundlage für Entscheidungen in Hochschulen, doch beeinflusste das schlechte Abschneiden deutscher Hochschulen die politische Diskussion und verstärkte den Wettbewerb im Hochschulsystem.

Ranking für Lehre: Die bisherigen Rankings bewerten überwiegend Forschungsleistungen, die Lehre wird nicht erfasst. Rankings für die Lehre könnten positive Steuerungswirkungen entfalten, wenn dadurch mittelfristig Qualitätsverbesserungen in der Lehre initiiert werden. Allerdings müssten geeignete Kriterien und Methoden entwickelt werden, die durch Studierendenbefragungen zu ergänzen wären.

Kritik an Rankings

Folgende Kritikpunkte standen im Vordergrund:

Vorhersehbare Ergebnisse von Rankings: Der Sinn von Rankings wurde wiederholt in Frage gestellt: Wenn sich so häufig bestätigt, dass die „Insider“ bzw. Fachwissenschaftler/innen sowieso schon vorher wissen, wo die „guten Leute“ der jeweiligen Fächer zu finden sind und wohin die meisten Fördermittel fließen (wie in den Vorträgen zum Ausdruck gekommen sei), sind durch Rankings wenig neue Erkenntnisse zu erwarten. Daraus ergibt sich die grundsätzliche Frage, wozu man angesichts des großen Aufwands und der hohen Kosten überhaupt Rankings braucht bzw. ob der Ertrag in angemessenem Verhältnis zum Aufwand steht. Sollte man knappe öffentliche Gelder nicht produktiver für die Förderung der Wissenschaft ausgeben als für ihre Bewertung in Rankinglisten?

Beeinflussung der Ergebnisse durch gewählte Indikatoren und Methoden: Die Ergebnisse von Rankings werden durch die Methodik sowie durch die Auswahl und Gewichtung der Indikatoren bestimmt: Wer das Verfahren definiert, beeinflusst auch das Ergebnis. Das sei auch beim CHE-

HochschulRanking nicht anders, wo z. B. für das Fach Soziologie die Methoden explizit ausgewiesen werden, die Theorien jedoch nicht. Dies hätte man genauso gut anders machen können, was dann aber auch zu anderen Ergebnissen geführt hätte. Diese Problematik wird bei allen Rankings gesehen, auch beim Rating des Wissenschaftsrates.

Gefahr normierender Wirkungen: Die „Normierung durch Zahlen“ in Rankings birgt die Gefahr, dass die (nicht quantifizierbare) Qualität einer Leistung nicht adäquat erfasst wird. Das wird zum Beispiel in Österreich deutlich, wo die Hochschulen Wissensbilanzen erstellen müssen, in denen alle Leistungen in Zahlen/Normen aufzuführen sind. Es ist jedoch problematisch, qualitative Ereignisse in quantitativen Zahlen auszudrücken. Skepsis wurde auch gegenüber dem OECD-Vorhaben AHELO geäußert, das die Ziele der PISA-Studie auf den Hochschulbereich übertragen möchte, indem die Kompetenzen und Kenntnisse von Absolvent/innen zur Grundlage eines europäischen Vergleichs gemacht werden. Ein weltweites Kompetenzranking dieser Art könnte eine normierende Wirkung im Hinblick auf die vorhandenen Indikatoren haben und zu einer Angleichung der Fächer bzw. der Lehrinhalte führen, was wiederum das Ende der akademischen Freiheit bedeuten und zu einer Verflachung der Wissenschaftslandschaft führen könnte.

Assessment of Higher Education Learning Outcomes (AHELO)

- Projekt der OECD
- Ziel: Leistungsfähigkeit der Hochschulen weltweit an der Qualität der Lehre messen (Forschungsergebnisse und wissenschaftliche Reputation einer Hochschule sind zweitrangig)
- getestet wird nicht das Angebot der Hochschule (Inhalt der Lehrveranstaltung), sondern die Leistungsfähigkeit der Hochschulabsolvent/innen als Ergebnis der Lehre: allgemeine Kompetenzen (Problemlösungsfähigkeit, analytisches und kritisches Denken, kommunikative Fähigkeiten etc.) sowie auf die Disziplin bezogene Kompetenzen (Fachwissen)
- derzeit Machbarkeitsstudie (Beteiligung von 15 Ländern, bisher noch ohne Deutschland); Beginn mit den Fächern Wirtschaftswissenschaften und Ingenieurwissenschaften; Ergebnisse und Schlusskonferenz geplant für Ende 2012

Angesichts des globalen Wettbewerbs der Hochschulen sollen durch

AHELO die Wettbewerbsvorteile der jeweiligen Hochschulen deutlich werden: „... die OECD will nicht am Ende sagen: Das ist die beste Hochschule der Welt. Wir wollen vielmehr jeder Hochschule ein Profil aufzeigen, ihre Stärken und Schwächen, damit sie sich ihre Position am Markt suchen kann.“ (OECD-Koordinator Andreas Schleicher 2009)

Quellen: Interview mit Andreas Schleicher, Spiegel Online, 07.04.2009: „Entscheidend ist, was bei Studenten ankommt“, <http://www.spiegel.de/unispiegel/studium/0,1518,617897,00.html>; OECD: AHELO 2010–2011, <http://www.oecd.org/dataoecd/37/49/45755875.pdf> (31.05.2011).

Rückwärtsgerichtet statt Zukunftsorientierung: Angemerkt wurde, dass Rankings etwas Museales, Rückwärtsgerichtetes haben, da sie zurückblicken und sich nicht durch Zukunftsorientierung und Innovation auszeichnen. Theoretisch kann eine Institution einen vorderen Platz in einem Ranking belegen und es dennoch nicht schaffen, aus ihren Studierenden *change agents* der Zukunft zu machen. Die Aufgabe der Hochschulen ist jedoch, Antworten auf die Herausforderungen der Zukunft zu finden.

Problematische Vermarktung – Trendabhängigkeit statt Trendsetting: Gegenwärtig entsteht bei der Vermarktung von Rankings häufig der Eindruck, es geht nicht primär um wissenschaftliche Leistungen und eine gute akademische Ausbildung, sondern vor allem darum, die anderen Hochschulen im Wettbewerb auszusteichen. Wenn Hochschulen sich an diesem Denken orientieren, werden sie zu Getriebenen, die aktuell erfolgreichen Trends hinterherlaufen statt selbst innovative Wege zu erproben und zukünftige Trends zu setzen. Ziel sollte es aber sein, dass die Hochschulen im internationalen Vergleich ihre eigenen Potenziale erkennen und ihr individuelles Profil entwickeln.

Notwendigkeit politischer Entscheidungen: Rankings sollen unter anderem dazu dienen, Hochschulleitungen und Politik die Entscheidungsfindung durch die Bereitstellung empirischer Daten zu erleichtern. Doch unabhängig davon, wie differenziert und methodisch genau man vorgeht: Letztlich erzeugt jedes Ranking eine klare Hierarchie, die in der Eigenlogik der Bewertung bestimmte Entscheidungen zwingend macht. Wenn – wie in den letzten Jahren – den Hochschulen immer weniger Geld zur Verfügung steht, ergeben sich zwangsläufig Verteilungsprobleme, bei denen Rankings einen unheilvollen, weil ungesteuerten Automatismus auslösen können. Eine verantwortungsvolle Hochschulplanung sollte aber bewusste (wissenschafts-)politische Entscheidungen treffen, statt sich einem Diktat der Zahlen anzuliefern.

Schwere Erfassbarkeit von neuen Wissenschaftsbereichen: Es wurde angezweifelt, dass Rankings methodisch überhaupt geeignet sind, das Potenzial von innovativen und interdisziplinären Ansätzen und Wissenschaftsfeldern angemessen zu erfassen, durch die sich hervorragende Forschung vielfach auszeichnet. Hierzu wurde angemerkt, dass ein Vergleich naturgemäß auf der Bildung von Vergleichsgruppen beruht und – wie in den Wissenschaften üblich – eine Orientierung an Fächern voraussetzt. Neue Wissenschaftsfelder und interdisziplinäre Ansätze sind hingegen meist Einzelfälle, die nicht vergleichbar und per Ranking erfassbar sind. Beim Rating des Wissenschaftsrats kann dieses Defizit teilweise dadurch aufgefangen werden, dass auch die Forschungsstrategie selbst erfasst und bewertet wird. Grundsätzlich sind Rankings aber per definitionem Bewertungsverfahren für vergleichbare Untersuchungsgegenstände, die keinen differenzierten Qualitätsblick auf jedes Detail der gesamten Hochschullandschaft erlauben.

U-MULTIRANK – EIN ALTERNATIVES KONZEPT FÜR EIN MULTI-DIMENSIONALES INTERNATIONALES RANKING

CHE Centrum für Hochschulentwicklung

Hintergrund

Die im letzten Jahrzehnt entstandenen globalen Hochschulrankings, insbesondere das sogenannte Shanghai-Ranking, das *Academic Ranking of World Universities* (ARWU) der Jiaotong Universität Shanghai⁷ und das *QS World University Ranking*⁸, haben die Hochschulwelt verändert. In vielen Ländern haben sie eine Diskussion über die Leistungsfähigkeit der nationalen Universitäten und des Wissenschaftssystems angestoßen, die zu speziellen Programmen zur Förderung von Spitzenhochschulen führten (z. B. in Korea, China). Auch die Diskussion um die Exzellenzinitiative war davon nicht unberührt, selbst wenn ihre Entstehung nicht unmittelbar auf die Rankings zurückzuführen sein mag.

Das Entstehen dieser Rankings und insbesondere die hohe Aufmerksamkeit, die sie gefunden haben, ist auf der einen Seite Ausdruck eines verstärkten globalen Wettbewerbs der Universitäten; auf der anderen Seite verstärken die Rankings diesen Wettbewerb selbst durch ihre Ergebnisse, sodass man von einer Dualität von Rankings und Wettbewerb sprechen kann. In der Folge ist ein „lukrativer Markt“ (Hazelkorn 2011) entstanden, mit einer Serie internationaler Konferenzen über die sogenannten *World Class Universities* und einer Welttournee einzelner Ranking-Anbieter.

Nun kommt mit dem von der Europäischen Kommission geförderten U-Multirank-Projekt noch ein weiteres Ranking hinzu – auch wenn es zunächst nur um die Entwicklung eines Konzeptes und eine Machbar-

⁷ Vgl. www.arwu.org.

⁸ Vgl. www.topuniversities.com. Bis 2009 gab es ein gemeinsames Ranking von QS Quacquarelli Symonds und Times Higher Education (THE). Nach internen Auseinandersetzungen über die Methodik haben sich beide Partner getrennt; QS führt das Ranking in unveränderter Form fort, während THE zusammen mit Thomson Reuters, dem Eigentümer einer der beiden großen bibliometrischen Datenbanken, ein neues globales Ranking entwickelt hat (<http://www.timeshighereducation.co.uk/world-university-rankings/>).

keitsstudie geht. Die politischen Hintergründe des Projektes sind vielfältig: Zum einen gibt es im Kontext des Bologna-Prozesses einen Bedarf an Transparenz über den entstehenden Europäischen Hochschulraum, vom dem wir bisher nicht einmal die exakte Zahl der Hochschulen kennen. Zum anderen spielt auch die „Lissabon Strategy“ der Europäischen Union eine Rolle, die die EU bis 2020 zum wettbewerbsstärksten Wirtschaftsraum der Welt machen will. Hierbei spielen Forschung, Entwicklung und ein hochqualifiziertes Humanpotenzial eine entscheidende Rolle. Gleichzeitig erklärt dies auch, weshalb das U-Multirank-Projekt nicht auf Europa begrenzt ist, sondern einen globalen Vergleich ermöglichen soll. Ein immanenter Anstoß für das Projekt ist die berechtigte Kritik an den vorhandenen globalen Rankings. Zusätzliche Dynamik hat die Entwicklung eines alternativen Konzeptes eines europäischen und internationalen Rankings durch die französische EU-Ratspräsidentschaft im Jahr 2009 erhalten, die dann zur Ausschreibung des Projektes durch die Kommission geführt hat.

Bei dem U-Multirank-Projekt, das im Juni 2011 abgeschlossen wurde, handelt es sich zunächst um eine Machbarkeitsstudie; am Ende des Projektes wird kein veröffentlichtes Ranking stehen, sondern ein in einem Pilotprojekt getestetes Konzept mit einem bewährten Set von Indikatoren sowie ein Konzept für die künftige Implementierung. Eine dauerhafte Implementierung eines Ranking-Systems ist beabsichtigt.

Die Kritik an den bestehenden globalen Rankings

Wenngleich globale Rankings weltweit eine hohe Beachtung finden und kaum eine Universität darauf verzichtet, ihre Position in den Rankings auf ihrer Homepage zu vermelden, sind sie dennoch Gegenstand methodischer Kontroversen und Kritik.⁹ Die Kritik betrifft sowohl die hochschulpolitischen Auswirkungen und Implikationen der Rankings als auch ihre Methodik und Indikatoren.

Die drei einflussreichen globalen Rankings umfassen weitgehend Indikatoren, die die Forschungsleistung messen sollen; Indikatoren zu Studium und Lehre oder anderen wichtigen Teilaspekten der Leistungen von Hochschulen (Technologie- und Wissenstransfer, regionale Einbettung) werden kaum berücksichtigt. Darüber hinaus hat in den Rankings von

⁹ Zusammenfassend vgl. Hazelkorn (2011).

QS und THE die Reputation der Hochschulen ein großes Gewicht (mit 50 bzw. 34,5 % des Gesamtwertes). Reputation ist jedoch kein Leistungsindikator und es bestehen erhebliche Zweifel an der Reliabilität internationaler Reputationsmessungen (Federkeil 2009).

Die bislang existierenden globalen Rankings weisen eine Reihe gravierender methodischer Mängel auf:

- Die globalen Rankings sind alle als Rankings ganzer Hochschulen über die Fächer hinweg begonnen worden. Nun wissen wir, dass sich die einzelnen Fächer einer Hochschule in ihrer Leistungsfähigkeit deutlich unterscheiden können. Eine Hochschule mit einigen sehr guten und einigen schwächeren Fächern wird in einem Ranking der ganzen Hochschule zu einer mittleren bzw. durchschnittlichen Hochschule. Diese Information hat für die meisten Nutzer von Rankings jedoch keinen Aussagewert. Sie hilft weder einem (prospektiven) Studierenden, der an Informationen über ein bestimmtes Studienfach interessiert ist, noch einem Forscher, der sich mit Kollegen und Kolleginnen seines Faches vergleichen will. Hinzu kommt, dass sich die Fachkulturen, z. B. in Bezug auf Publikationen und Zitationen, deutlich unterscheiden. Dadurch laufen institutionelle Rankings letztlich Gefahr, nicht die Leistungsfähigkeit von Hochschulen, sondern primär ihre Fächerstruktur zu messen. Als Reaktion auf diese Kritik haben in den letzten Jahren sowohl das QS als auch das Shanghai-Ranking begonnen, zusätzlich zum institutionellen Ranking auch Rankings nach Fächergruppen und einzelnen Fächern zu veröffentlichen.
- Hinzu kommt, dass die vorhandenen internationalen bibliometrischen Datenbanken, die die Datenbasis aller globalen Rankings bilden, das Publikationsaufkommen in vielen Fächern – insbesondere den Geistes- und Sozialwissenschaften, aber auch Teilen der Ingenieurwissenschaften – nicht hinreichend erfassen. Sie nehmen weitgehend nur Aufsätze in internationalen referierten Zeitschriften auf; d. h. Fächer, in denen Bücher oder Tagungsbeiträge eine wichtige Rolle spielen, sind eindeutig untererfasst (van Raan 2010). Aus dieser methodischen Schwäche resultiert ein Bias zu Lasten der Geistes- und Sozialwissenschaften bzw. umgekehrt werden Universitäten „bevorzugt“, die eine Ausrichtung auf den Natur- und dort insbesondere den Biowissenschaften haben.
- Ein weiterer Punkt, der zu Kritik Anlass gibt, ist die Berechnung von

Ranglisten wie in einer Bundesligatabelle (*league tables*).¹⁰ Dieses Vorgehen suggeriert, dass eine Universität auf Platz 17 besser ist als eine auf Platz 20 rangierende, und dass Rang 133 besser ist als 137. Eine Analyse der Rankings zeigt jedoch, dass häufig geringe Unterschiede im Zahlenwert eines Indikators zu großen Unterschieden in den Rangplätzen führen. So beträgt die mittlere Differenz des Gesamtwertes, der eine 100er-Skala umfasst, zwischen zwei Rangplätzen innerhalb der ersten 100 Universitäten im QS-Ranking lediglich 0,4 Punkte. In Ranglisten, die einzelne Rangplätze vergeben, werden Unterschiede zwischen den Hochschulen tendenziell übertrieben.

Eine wesentliche methodische Schwäche der existierenden globalen Rankings ist die Verwendung eines zusammenfassenden Gesamtwertes (*composite indicator*), der die Leistungsfähigkeit einer Hochschule insgesamt messen soll. Dieses Vorgehen erfordert die Zuweisung von individuellen Gewichten zu den einzelnen Indikatoren. Es gibt aber weder theoretische noch empirische Gründe, einzelnen Indikatoren ein bestimmtes Gewicht zuzuordnen. Empirische Studien haben gezeigt, dass die *composite indicators* der globalen Weltrangings alles andere als robust sind und die Ergebnisse durch kleinste Veränderungen der Gewichtungsschemata stark variieren (vgl. Saisana et al. 2011). Mit den Gewichten wird von den Rankings ein Konzept von Qualität definiert. Unterschiedliche Nutzer von Rankings können jedoch sehr unterschiedliche Vorstellungen von Qualität haben: Für Studierende sind beispielsweise andere Faktoren ausschlaggebend als für Forscher oder für Arbeitgeber. Die Vergabe vordefinierter Gewichte und die Berechnung von Gesamtindikatoren bevormunden die Nutzer von Rankings.

Die existierenden globalen Rankings weisen eine Reihe methodischer Schwächen auf, die ihre Aussagekraft reduzieren. Unter hochschulpolitischen Aspekten ist die einseitige Orientierung auf Forschungsexzellenz international orientierter Volluniversitäten problematisch. Dies hat international zu einer Obsession für die sogenannten *World class universities* geführt, die zu einer Gefahr für die Diversität von Hochschulen und Hochschulsystemen geworden ist. Exzellente Hochschulsysteme zeichnen sich durch eine breite Vielfalt an Einrichtungen aus, die unterschiedlichen Zielsetzungen und -gruppen gerecht werden.

¹⁰ Im Shanghai-Ranking werden seit einiger Zeit Rangplätze nur noch für die ersten 100 Universitäten ausgewiesen, danach nur noch breitere Gruppen (z. B. Rang 101 bis 150); im QS-Ranking finden sich aber immer noch einzelne Rangplätze bis 399.

Das U-Multirank-Projekt

Die Ausschreibung der EU-Kommission zur „Entwicklung des Konzeptes und zum Test der Machbarkeit eines multi-dimensionalen globalen Hochschulrankings“ wurde von einem Konsortium unter Federführung des CHE und des *Center for Higher Education Policy Studies* (CHEPS) der Universität Twente gewonnen. Weitere Partner sind das *Center for Science and Technology Studies* (CWTS) der Universität Leiden, eine Forschungsgruppe der Katholischen Universität Leuven (INCENTIM) sowie das *Observatoire des Sciences et des Techniques* (OST) in Paris.¹¹

Vielfältige Exzellenz: Das U-Multirank-Konzept knüpft in wesentlichen Elementen an die Methodik des CHE-Rankings an und berücksichtigt die berechnete Kritik an den existierenden globalen Rankings. Die Grundidee von U-Multirank ist die Darstellung der Vielfalt von Hochschulen in einem internationalen, multi-dimensionalen Ranking, das unterschiedliche Bereiche von Exzellenz jenseits einer einseitigen Fokussierung auf Forschungsexzellenz sichtbar machen will. Um nur drei Beispiele zu nennen: Hochschulen können in der Lehre, im Bereich lebenslanges Lernen und Weiterbildung, aber auch in ihrer Rolle für die regionale Entwicklung exzellent sein.

Diese Vielfalt zu zeigen, kann einem Ranking jedoch nur gelingen, wenn es *multi-dimensional* ist, d. h. Indikatoren für mehrere Leistungsbereiche umfasst und sich nicht, wie die beschriebenen Rankings, mehr oder weniger ausschließlich auf die Forschung konzentriert. In U-Multirank wurden den Indikatoren zu fünf Teilbereichen entwickelt:

- Studium und Lehre
- Forschung
- Wissens- und Technologietransfer
- Internationale Orientierung
- Regionale Einbettung

Für jeden dieser fünf Teilbereiche wurde in einem umfassenden Diskussionsprozess mit verschiedenen Gruppen von Stakeholdern sowohl für fachbezogene als auch für ein institutionelles Ranking ein Set von Indikatoren entwickelt. Das primäre Kriterium war entsprechend einer Forderung der *Berlin Principles on Rankings of Higher Education*

¹¹ Vgl. www.u-multirank.eu.

Institutions die Relevanz der Indikatoren (IREG 2006). Zusätzlich spielten bei der Bewertung der Indikatoren methodische Aspekte wie Validität, Reliabilität und internationale Vergleichbarkeit eine Rolle. In einem Pilotprojekt wurde die Eignung und Verfügbarkeit dieser Indikatoren mit einem Sample von rund 140 Hochschulen unterschiedlichen Profils aus Europa und außerhalb Europas getestet. Auf der Grundlage dieses Pilotprojektes steht am Ende des ersten U-Multirank-Projektes ein Set von international getesteten Indikatoren für ein multi-dimensionales Ranking zur Verfügung.

Diversität: Die existierenden globalen Rankings, die vorgeben, die weltweit besten Hochschulen zu küren, transportieren unweigerlich die Botschaft, dass die Hochschulen, die nicht zu den besten 500 gehören, eben schlechter sind als diejenigen, die es in den Kreis der Top 500 geschafft haben. In Wirklichkeit sind die meisten Hochschulen in der Welt jedoch nicht schlechter, sondern schlichtweg anders. Die meisten haben eine andere Zielsetzung und ein anderes institutionelles Profil als die in die Rankings aufgenommenen international orientierten Forschungsuniversitäten. Das gilt z. B. für die deutschen Fachhochschulen, die aufgrund der methodischen Ausrichtung der globalen Rankings keine Chance haben, dort sichtbar zu werden, auch wenn sie eine „Weltklasse“-Ausbildung bieten, die ihren Absolventen/innen beste Berufsmöglichkeiten eröffnet. Das gilt aber auch für profilierte bzw. spezialisierte Universitäten. Als Beispiel aus Deutschland kann die Universität Mannheim mit ihrem Profil in den Wirtschafts- und Sozialwissenschaften genannt werden, die ebenfalls nicht unter den besten 500 Universitäten des Shanghai-Rankings auftaucht.

Nun ist es aber einleuchtend, dass auch ein Ranking, das verschiedene institutionelle Profile abdecken will, im weltweiten Maßstab nicht alle Hochschulen miteinander vergleichen kann. Es macht keinen Sinn, eine regional und auf die Lehre orientierte Hochschule ohne Promotionsrecht mit der Universität Oxford zu vergleichen, ebenso wenig macht es Sinn, eine Kunsthochschule mit einer auf die biomedizinische Forschung spezialisierten Universität zu vergleichen.¹² Ein sinnvolles internationales Ranking muss es ermöglichen, in einem ersten Schritt Gruppen von Hochschulen mit einem vergleichbaren institutionellen Profil zu iden-

¹² Das CHE-Ranking hat in Deutschland bislang die – auch rechtlich maßgebliche – Differenzierung zwischen Universitäten und Fachhochschulen als Unterscheidungsmerkmal für getrennte Rankings übernommen. Mit Blick auf die mögliche Auflösung dieses Differenzierungsmerkmals muss dieses Vorgehen in Zukunft möglicherweise überdacht werden.

tifizieren, die dann in einem zweiten Schritt in einem Ranking sinnvoll miteinander verglichen werden können. In diesem Kontext ist das U-Multirank-Projekt eng mit einem ebenfalls von der EU-Kommission geförderten Projekt zur Entwicklung einer Klassifikation der europäischen Hochschulen verbunden (U-Map; vgl. van Vught et al. 2010).¹³ Auf der Basis der fünf Dimensionen von U-Map, die ebenfalls jeweils mit mehreren Indikatoren unterlegt sind, kann ein multi-dimensionales Profil erstellt werden, das die Grundlage für die Auswahl der Hochschulen bildet, die dann in einem Ranking verglichen werden.

Nutzerorientierung: Dieser Ansatz verdeutlicht bereits, dass U-Multirank sich an den Bedürfnissen der Nutzer/innen orientiert. Diese bestimmen, welche Hochschulen mit welchem institutionellen Profil sie miteinander vergleichen wollen. In einem zweiten Schritt können die Nutzer/innen dann – ähnlich wie im CHE-Ranking – aus dem Gesamtset an Indikatoren diejenigen auswählen, die für sie selbst relevant sind. Dadurch entstehen anstelle vordefinierter Listen, die vorgeben, „Qualität“ von Hochschulen objektiv definieren zu können, personalisierte Rankings, die den Präferenzen und Prioritäten der Nutzer/innen entsprechen. Dies schließt nicht aus, dass (auch von Dritten) für bestimmte institutionelle Profile bestimmte Rankings definiert und publiziert werden.

Implementierung: Im Laufe des Projektes wurden auch Vorstellungen für eine weitere, dauerhafte Implementierung von U-Multirank erarbeitet. U-Multirank sollte künftig mit Blick auf den Bedarf an transparenter, grenzüberschreitender Information über den europäischen Hochschulraum alle europäischen Hochschulen umfassen. Darüber hinaus ist das System auch für nichteuropäische Hochschulen offen, um diesen wie den europäischen Hochschulen einen methodisch fundierteren Vergleich mit Hochschulen aus anderen Regionen zu ermöglichen.

13 Weitere Informationen zu diesem Projekt und ein interaktives Web-Tool zur Identifizierung institutioneller Profile finden sich unter: www.u-map.eu.

Literatur

Federkeil, Gero (2009): Reputation Indicators in Rankings of Higher Education Institutions. In: Barbara M. Kehm/Bjørn Stensaker (Hrsg.): *University Rankings, Diversity, and the New Landscape of Higher Education*. Rotterdam/Boston/Taipei: Sense Publishers, S. 19–34.

Hazelkorn, Ellen (2011): *Rankings and the Reshaping of Higher Education. The Battle for World-Class Excellence*. Basingstoke: Palgrave Macmillan.

IREG (2006): *Berlin Principles on Rankings of Higher Education Institutions*. Berlin (www.ireg-observatory.org; 30.05.2011).

Saisana, Michaela/d'Hombres, Béatrice/Saltelli, Andrea (2011): Rickety Numbers: Volatility of University Rankings and Policy Implications. In: *Research Policy* (40) 1, S. 165–177.

van Raan, Anthony (2010): The Social Sciences and the Ranking of Universities. In: *UNESCO: World Social Science Report 2010. Knowledge Divides*. Paris, S. 237–239 (<http://unesdoc.unesco.org/images/0018/001883/188333e.pdf>; 30.05.2011).

van Vught, Frans/Kaiser, Frans/File, Jon M./Gaethgens, Christiane/Peter, Rolf/Westerheijden, Don F. (2010): *U-Map. The European Classification of Higher Education Institutions*. Enschede: CHEPS.

INTERNATIONALE RANKINGS – WIE KANN MAN SIE BESSER MACHEN?

wissenschaftliche Publizistin, Berlin

Diversität und Profilbildung der Hochschulen berücksichtigen

Internationale Rankings werden von den Medien oft ernst genommen, doch sie basieren auf höchst fraglichen methodischen Ansätzen und berücksichtigen die Besonderheiten der nationalen Hochschulsysteme in keiner Weise – dies betonte Prof. Dr. Margret Wintermantel, Präsidentin der Hochschulrektorenkonferenz (HRK). Außerdem seien sie sehr auf die anglophone Welt zugeschnitten. Um den stark differenzierten europäischen Hochschulraum angemessen darzustellen, müssten die unterschiedlichen Profile der Hochschulen in die Betrachtung einbezogen werden: So sind z. B. die Unterschiede zwischen den einzelnen Hochschulen in Polen, Italien, Großbritannien und Frankreich enorm groß. Für die deutschen Hochschulen könne es im Zuge der Profilbildung nicht darum gehen, ein gestuftes System von „Champions League – Bundesliga – Regionalliga“ mit Siegern und Verlierern zu errichten. Vielmehr sollte jede Hochschule ihre eigene Mission definieren und ihre entsprechenden Ziele auf hohem Niveau umsetzen können.

Für ein adäquates Rankingverfahren wird nach Wintermantel eine Vielzahl an Indikatoren benötigt. Nur dann sei es möglich, die Leistungen der verschiedenen Hochschulen, der einzelnen Disziplinen, Fakultäten und auch Forscher/innen differenziert zu erfassen und daraus notwendige Strategien abzuleiten. Auch sollten die Hochschulen selbst und die Studierenden aktiv daran beteiligt werden. Für Deutschland sei es wichtig, dass die Forschungsleistungen, die Hochschulen in Kooperation mit außeruniversitären Forschungseinrichtungen erbrächten, in die Bewertung mit einbezogen werden. Ein europäisches Ranking, das auch dem deutschen System gerecht werde, müsse fundierte Informationen über die Leistungen in den Hochschulen liefern, sodass diese ihre Leistungsfähigkeit steigern und ihre Potenziale so gut wie möglich entwickeln könnten.

Für die Hochschulleitungen sei es wichtig, auf angemessene Rankings als Führungsinformation zurückgreifen zu können, da man viele Daten benötige, um eine Hochschule erfolgreich zu leiten und Forschung und Lehre kontinuierlich zu verbessern. Aber nicht nur die Hochschulleitungen, auch die Politik brauche Informationen, aus denen Entscheidungen abgeleitet werden können. Gute Rankings könnten wichtige Entscheidungshilfen sein und zugleich Entscheidungen besser legitimierbar machen.

Die Idee bei U-Multirank, ein transparentes System zu entwickeln, das den verschiedenen Kategorien von Hochschulen gerechter wird als die bisherigen allgemeinen Rankings, sei auf jeden Fall zu begrüßen. Doch bestehe in der HRK die Sorge, dass dieses Transparenzinstrument unintendierte Effekte haben könnte, wenn zum Beispiel die Klassifizierung der Hochschulen auf der Basis ihrer Profile zu einer Rangreihung der Hochschultypen verwendet werde. Ein positiver Aspekt sei in jedem Fall, dass die Studierenden bei diesem Ranking einbezogen werden, da die Lehre aus Sicht der HRK einen wichtigen Stellenwert hat.

Fachbezogene Rankings entwickeln

Auch Prof. Dr. Peter Lichter vom Deutschen Krebsforschungszentrum in Heidelberg vertrat die Ansicht, dass die deutschen Hochschulen und Forschungseinrichtungen in den internationalen Rankings nicht adäquat erfasst werden. So werde ein wichtiger Unterschied nicht berücksichtigt: Die Spitzenforschung ist in den US-amerikanischen geprägten Hochschulsystemen in die Universitäten integriert, während sie in Deutschland vorwiegend in oder in Kooperation mit außeruniversitären Einrichtungen stattfindet, die in internationalen Rankings meist nicht berücksichtigt werden. Wichtige Forschungsleistungen würden somit nicht angerechnet, was den Wissenschafts- und Forschungsstandort Deutschland benachteilige.

Grundsätzliche Skepsis ist nach Lichter immer dann angebracht, wenn Rankings die gesamte Institution bewerten. Sinnvolle Rankings könne es nur für die Fachebene geben (wie z. B. das Forschungsrating des Wissenschaftsrats, das aber bisher noch keinen internationalen Vergleich ermöglicht). Auch auf europäischer Ebene wäre letztlich nur ein fachspezifisches Ranking interessant. Angesichts der Ausdifferenzierung der Hochschullandschaft und der Profilbildung der Hochschulen entstehe ein gravierendes Problem, wenn Bewertungen von stärkeren und schwächeren Fächern addiert und dann Durchschnitte für Hochschulen gebil-

det werden. Deshalb ist für Lichter auch ein europäisches Ranking wie U-Multirank fragwürdig, das auf die Bewertung ganzer Hochschulen ziele.

Wenig sinnvoll sei es, wenn die vorhandene Mehrdimensionalität des Forschungsratings in internationalen Rankings in eine Gesamtbewertung einfließt, weil die vom Wissenschaftsrat erhobene, differenzierte Forschungsleistung darin völlig untergeht. Auch die Vorstellung, man könne durch die Bildung von Ranggruppen (statt einzelner Rangplätze) Fehler minimieren, hält Lichter für wenig überzeugend. Vielmehr würde man dadurch unintendierte Effekte erzeugen: Hochschulen könnten sich gezwungen sehen, sich jenen Kategorien zuzuordnen, in die das Geld fließt. Eine solche Wirkung würde der gewünschten Differenzierung der Hochschulen entgegenstehen. Deshalb sollten keine neuen Zwänge durch feste Kategorien geschaffen werden, sondern ein Pluralismus verwirklicht werden, der die Entwicklung verschiedener Hochschultypen erlaubt. Eine Ranggruppenbildung würde die Entstehung neuer Hochschultypen jedoch eher verhindern.

Lichter zeigte sich skeptisch, dass die Datenerhebung auf europäischer Ebene durch Standardisierung vereinfacht werden kann – wie sollte das gelingen, nachdem es schon auf nationaler Ebene sehr schwierig ist? Ein neues europaweites, aber fachspezifisches Rankingsystem würde er sehr begrüßen, doch sieht er in dem neuen europäischen U-Multirank einige Gefahren. Rankings mit solch einem Ansatz seien nicht wirklich weiterführend.

Politische Folgewirkungen beachten

Eine grundsätzliche Kritik an Rankings übte Jan Krüger, Bundesvorstand der Juso-Hochschulgruppen. Prinzipiell müsse man bei der Interpretation von Rankings die zugrunde liegende Auffassung von Hochschulpolitik berücksichtigen. Das Ziel von mehr Transparenz durch Rankings sei durchaus positiv. Der wesentliche Punkt sei jedoch, wie diese Rankings politisch genutzt werden. Krüger skizzierte mögliche negative Folgen: So könnte man z. B. festlegen, dass Hochschulen mit exzellenter Bewertung Studiengebühren erheben dürfen oder Hochschulen mit sehr guter Lehre sich Studierende (und deren Anzahl) aussuchen können. Auf solche Folgewirkungen müsse man sehr genau achten. Eine sozialdemokratische Hochschulpolitik dürfe solche Entwicklungen nicht zulassen.

Auch Rankings für Lehre würden die gravierenden Probleme bzw. un-

absehbaren Konsequenzen von Rankings nicht lösen. Es sei gut, wenn Rankings bei der Datenerhebung auch Studierende beteiligen, aber sicher nicht ausreichend, wenn lediglich Bewertungsbögen ausgefüllt werden. Entscheidend sei die direkte Kommunikation der Lehrenden und Lernenden, die sich über Kriterien guter Lehre verständigen müssten. Diesen Austausch könnten Rankings nicht ersetzen. Auch stelle sich die Frage, ob und wie die Ergebnisse dann tatsächlich zur Verbesserung von Lehre und Studium beitragen können. Wenn ein Ranking verdeutlicht, wo gute Lehr- oder Lernleistungen erzielt werden, könne das zwar mehr Transparenz schaffen, doch komme es letztlich darauf an, welche konkreten Folgerungen zur Verbesserung daraus gezogen werden.

Von der Vorstellung, dass Lehr-Rankings für Studieninteressierte von großer Bedeutung für die Studienwahl sein könnten, zeigte sich Krüger nicht überzeugt. Wichtiger seien meist andere Kriterien, wie z. B. die kulturelle Attraktivität und Atmosphäre einer Stadt, die geografische Nähe zu Elternhaus und Freundeskreis, die Qualität der Hochschulberatung und die Atmosphäre vor Ort, aber auch persönliche Empfehlungen. Wesentliche Aspekte seien, wo die Studieninteressierten mit ihren Fragen auf Verständnis treffen, nützliche Informationen bekommen und „sich aufgehoben“ fühlen. Dies würde die Studienwahlentscheidung viel mehr beeinflussen als der Rangplatz einer Hochschule oder eines Faches in einem Lehr-Ranking. Wenn internationale Studierende Rankings stärker zur Hochschulwahl nutzen, sei dies eher Ausdruck davon, dass ihr Studium häufig mit hohen Kosten verbunden ist. Nach Ansicht von Krüger würden Lehr-Rankings sicherlich Steuerungswirkungen entfalten, doch sei es sehr fraglich, ob diese dann im Interesse der Studierenden wären.

Brauchen wir ein europäisches Ranking?

Die Bedeutung des neuen U-Multirank wurde kontrovers diskutiert. Auf der einen Seite wurde die Ansicht vertreten, dass das neue europäische Projekt vor allem aus zwei Gründen eine hohe exemplarische Bedeutung hat: Zum einen weil es übernational in einem sehr heterogenen Raum gemeinsam diskutiert wird und daher von vornherein ein hohes Maß an Differenzierung notwendig macht, was bei anderen Rankings nicht der Fall ist. Zum anderen könnte eine wissenschaftlich verantwortliche und differenzierte Diskussion über ein europäisches Rankingsystem die untauglichen internationalen Rankings (wie das Shanghai-Ranking) überflüssig machen.

Auf der anderen Seite wurden verschiedene Kritikpunkte gegen U-Multirank vorgebracht:

Ein solches Ranking sei nicht notwendig, da es bereits genug Rankings gebe.

Die außeruniversitären Einrichtungen würden nicht ausreichend berücksichtigt.

Das Ranking führe zu einer Kategorisierung der Hochschulen, die die Herausbildung differenzierter Hochschultypen erschwert.

Das Ranking würde die Hochschulen mit erheblichem Aufwand belasten, ohne dass sich für sie erkennbare Vorteile ergeben.

Es sei zweifelhaft, dass für die europäische Ebene valide Indikatoren entwickelt werden können, wenn dies schon für Deutschland enorm schwierig ist.

Unklar sei, welche Ziele U-Multirank verfolgt und welche Zielgruppen erreicht werden sollen. Als Entscheidungshilfe für Studierende bzw. Nachwuchswissenschaftler/innen unter dem Gesichtspunkt europäischer Mobilität werde es nicht benötigt, weil die Entscheidung für ein bestimmtes Land in den meisten Fällen schon vorher gefällt werde. Deshalb wäre es wenig hilfreich, wenn in einem europäischen Ranking z. B. die Universität Helsinki vor der Universität Florenz steht – nationale Rankings wären dann völlig ausreichend. Auch das Ziel, *regional engagement* oder *peer learning* unter den europäischen Hochschulen zu unterstützen, erscheine wenig realistisch, da solche Aktivitäten massiv vom politischen und institutionellen Umfeld abhängig sind.

Dagegen lässt sich einwenden, dass die Integration außeruniversitärer Einrichtungen im Projekt durchaus eine Rolle spielt, wenn auch die konkrete Umsetzung noch offen ist. Im ersten Schritt der Pilotstudie wurden einzelne außeruniversitäre Einrichtungen einbezogen, um herauszufinden, ob diese eingebunden werden können und wie vergleichbar die Daten sind etc. Mittelfristig habe man, so Gero Federkeil, die Idee – zumindest mit Blick auf die Forschung – den Standort als Betrachtungseinheit zu nehmen und nicht die einzelne Hochschule neben den außeruniversitären Forschungseinrichtungen am gleichen Ort. Allerdings seien damit wieder praktische Probleme der Abgrenzung und Zuordnung verbunden.

Die Vorbehalte gegen Rankings auf Hochschulebene kann Federkeil gut nachvollziehen, doch sei dies eine umzusetzende Vorgabe für das Pilotprojekt gewesen. Die Vorstellung des CHE ist, den Vergleich auf Hochschulebene auf Indikatoren zu beschränken, die sich auf Aspekte des strategischen Managements einer Hochschule beziehen und nicht auf die fachliche Dimension, also Lehre und Forschung. In diesen Bereichen stehe man wirklich vor dem Problem, wie diese Ergebnisse auf Hochschulebene sinnvoll addiert werden können, da eine Durchschnittsbildung keine aussagekräftigen Werte ergebe.

Zur Befürchtung fester „Schubladenbildung“ von Hochschulen, die eine unterschiedliche Wertigkeit (und eventuell auch Geldmittel) zugewiesen bekommen, merkte er an: In U-Multirank würde man sicher keine festen Typen definieren und dann typenspezifische Rankings bilden. Stattdessen sei geplant, ein Webtool zur Verfügung zu stellen, in dem sich die Nutzer/innen – mit Blick auf ihren eigenen Vergleichsbedarf – aus den Profilen zur Klassifizierung bedienen und sich ihr eigenes Ranking aus verschiedenen Profilelementen zusammenstellen können. Solche individuellen Rankings mit spezifischem Profil, die nach eigenem Informationsbedarf für den Einzelfall erstellt werden, hätten bei allen Diskussionen mit den relevanten „stakeholdern“ im Unterschied zu festen Rankings viel Zustimmung gefunden.

Was die Nutzung von Rankings angehe, sei aus verschiedenen Studien bekannt, dass in Deutschland sehr wohl ein erheblicher Teil von Studierwilligen Rankings für die Entscheidungsfindung nutzt. Rankings hätten eine gewisse Aufklärungsfunktion über die Unterschiede zwischen den Hochschulen. Zudem bestehe in anderen Ländern ein sehr großes Interesse an Rankings. Durch Studien des DAAD sei bekannt, dass ausländische Studierende Rankings noch mehr nutzen als deutsche Studierende. Die Nutzung von Rankings in den verschiedenen Fächern sei jedoch sehr unterschiedlich, und im Hinblick auf internationale Mobilität seien Rankings im Master-Bereich sicher wichtiger als im Bachelor-Bereich.

Zu den Zielen und den Zielgruppen von U-Multirank sagte Federkeil, dass laut Ausschreibung der EU zahlreiche, sehr verschiedene Ziele verfolgt und Zielgruppen erreicht werden sollten. In der bisherigen Konzeption des Projekts verfolge man jedoch im Wesentlichen zwei Zielsetzungen: Erstens sollen Personen, die sich für Studienbedingungen und Studiengänge in Europa interessieren, Informationen zur Verfügung gestellt werden, zweitens sollen die Hochschulen dabei unterstützt wer-

den, ihre eigene Hochschule im internationalen Kontext zu verorten. Damit seien auch die beiden wichtigsten Zielgruppen benannt.¹⁴

Herausforderungen im Umgang mit Rankings

In der Diskussion wurden folgende Herausforderungen im Umgang mit Rankings angeführt:

Effektive Nutzung von Daten: Wichtig ist die Frage, wie die Ergebnisse von Rankings zum Nutzen der Institutionen konkret umgesetzt werden können. Man sollte sich schon frühzeitig überlegen, wie die Daten effektiv genutzt werden können, damit die Hochschulen nicht gezwungen sind, ständig neue Daten zu produzieren, die sich am Ende womöglich kaum unterscheiden. Auch sollte überprüft werden, ob bereits erhobene Daten (z. B. von institutionellen Audits, Akkreditierungs- und Evaluationsverfahren) in Rankings verwendet werden können, um zu vermeiden, dass die Hochschulen permanent Zuarbeit leisten müssen, was mit erheblichem Aufwand und hohen Kosten verbunden ist.

Lernprozess im Umgang mit Kennzahlen: Viele Rankings basieren auf Kennzahlen. Diese wurden im Hochschulsystem aber erst vor einiger Zeit eingeführt, sodass die Interpretation und Verwendung von Kennzahlen erst gelernt werden muss. Gegenwärtig ist im Umgang mit Kennzahlen noch häufig Unsicherheit und eine gewisse Naivität festzustellen. Das Hochschulsystem braucht sicher eine gewisse Zeit, um mit solch einem neuen Instrument sinnvoll arbeiten zu können. Alle Hochschulakteure mit Leitungsfunktionen müssen erst einmal lernen, wie man steuert und dabei Kennzahlen sinnvoll heranziehen kann.

Entwicklung geeigneter Instrumente: Problematisch ist, dass jede Hochschule ihre internen Datenbanken nach anderen Parametern entwickelt, sodass Ranking-Organisationen auf kein einheitliches Datenset zurückgreifen können. Deshalb wäre es sinnvoll, wenn HRK, DAAD und andere Institutionen gemeinsam ein Datenkonzept und Qualitätssicherungssystem für zukünftige Rankings entwickeln würden. Dabei müsste geklärt werden, welche Daten einbezogen, wie diese erhoben und vergleichbar gemacht werden können. Dann bestünde die Möglichkeit, die

14 Die Ergebnisse der Machbarkeitsstudie „U-Multirank-Projekt“ wurden am 09. Juni 2011 im Rahmen der Abschlusskonferenz in Brüssel präsentiert und offiziell der EU-Kommission übergeben.

Ergebnisse zum gemeinsamen Bestand von Rankings in der deutschen Wissenschaft und im zweiten Schritt zum Wegweiser europäischer Wissenschaftsrankings zu machen. Auch könnte man nach dem Vorbild der DFG, die Regeln guter wissenschaftlicher Praxis aufgestellt hat, Regeln guter Rankings entwickeln.

Kritischer Blick auf Instrumentalisierung: Rankings und Ratings haben langfristige, zum Teil auch ungeplante und unvorhergesehene Auswirkungen auf Forschungstätigkeit, Infrastruktur, Ausbildung etc. im Hochschulsystem. Das bringt auch große Einflussmöglichkeiten der Ranking-Agenturen mit sich. Unbedingt verhindert werden muss, dass diese Agenturen ihre Kategorien gezielt an eigenen Interessen oder den Interessen ihrer Auftraggeber ausrichten und damit unter Umständen den „Wissenschaftsmarkt“ manipulieren.

Verantwortungsvoller Umgang mit den Ergebnissen: Hervorgehoben wurde, dass der sorgsame Umgang mit Informationen bei Rankings von nicht zu überschätzender Bedeutung ist. Das betrifft nicht nur die Berücksichtigung methodischer Mängel von Rankings, sondern auch die Problematik der interessenbestimmten Interpretation von Daten und weitere Gefahren, wie die Steuerung von Forschungsinhalten und mögliche normierende Wirkungen durch die Ergebnisse. Wenn durch eine Fehlorientierung an Rankings die Freiheit von Forschung und Lehre eingeschränkt wird, könnte das zu einer Demoralisierung von Wissenschaftler/innen und schlechteren Forschungsleistungen führen. Eine große Verantwortung im Umgang mit Ratings und Rankings ergibt sich insbesondere für Hochschulleitungen und politische Entscheidungsträger, wenn sie die Ergebnisse als Entscheidungsbasis heranziehen, da ihre Schlussfolgerungen weitreichende Konsequenzen für das deutsche Hochschul- und Wissenschaftssystem haben können.

ISBN: 978-3-86872-805-7



Die Friedrich-Ebert-Stiftung ist im Qualitätsmanagement zertifiziert nach EFQM
(European Foundation for Quality Management): Committed to Excellence