

„U-Multirank“ und „U-Map“ als Ansätze zur Schaffung von Transparenz im europäischen und globalen Hochschulsystem – Konzepte und Erfahrungen

Frank Ziegele, Frans van Vught

Im Zuge der wachsenden Bedeutung weltweiter Hochschulrankings werden zunehmend deren problematische Aspekte thematisiert – etwa ein sich ausweitendes „reputation race“. Ausgehend von der Überzeugung, dass die über Rankings herzustellende Transparenz von grundlegender Bedeutung für verschiedenste Akteure im Hochschulbereich ist, widmen sich die Autoren den Projekten „U-Map“ und „U-Multirank“. Dabei zeichnen sie nach, welche notwendigen Gestaltungsprinzipien sich aus der Betrachtung des theoretischen Hintergrunds von Klassifikationen und Rankings sowie aus den Erfahrungen mit mehreren Rankings ergeben. Hiernach stellen sie dar, wie jene beiden „multi-dimensional“ und „user-driven“ Ansätze sich an den erarbeiteten Anforderungen orientieren, zentrale negative Effekte zu vermeiden versuchen, in komplementären Funktionen zusammenspielen und so einen Beitrag zur adäquaten Schaffung von Transparenz leisten können. Erste praktische Erfahrungen mit den beiden Ansätzen werden aufgegriffen.

1 Einführung

Globale Hochschulrankings haben weitreichende Effekte: 2010 forderten dänische Politiker, einen Abschluss von einer Top-20-Universität zum Kriterium für Einwanderung nach Dänemark zu machen, Russland hat angekündigt, die Ranking-Platzierung als Kriterium zur Anerkennung ausländischer Abschlüsse zu machen, und die indische Regierung erlaubt lokalen Universitäten bestimmte Abkommen nur mit Universitäten aus den Top 500 der Times Higher Education und Shanghai World Rankings (*Holmes 2012*). Wenn politische Sonntagsreden überall in der Welt fordern, dass man im eigenen Land eine gewisse Zahl von „World Class Universities“ benötige, wenn in Dänemark das komplette Wissenschaftssystem durch Bildung großer, fusionierter Einheiten verändert wird und wenn staatliche Exzellenzinitiativen in vielen europäischen Ländern die Spitzenforschung stärken sollen, dann hat das zumindest teilweise auch damit zu tun, dass man sich in weltweiten Rankings an der Spitze platzieren will. Denn große, forschungsstarke Einheiten schneiden in Weltrankings gut ab. Auch unter Hochschulrektoren und Hochschulrektorinnen ist es beliebt, sich zunächst über die mangelnde Aussagefähigkeit von Rankings zu beklagen, im selben Atemzug aber als

wichtiges strategisches Ziel der Hochschule ein Klettern im Ranking auf einen bestimmten Platz zu formulieren. Ein globales, in hohem Maße Ressourcen beanspruchendes „reputation race“ (van Vught 2008, S. 168 f.) ist im Hochschulsystem in vollem Gange.

Viele Akteure sehen darin große Gefahren. Wenn bei Rankings die Reputation der Hochschulen aus Sicht der Kollegen und Kolleginnen im Vordergrund steht, droht dann nicht eine „self-fulfilling prophecy“, indem sich die Leistung von Hochschulen tatsächlich den (Vor-)Urteilen der Peers anpasst? Droht angesichts der Selektivität von Rankingindikatoren die Vernachlässigung wichtiger Aufgaben von Hochschulen, die in den Rankings nicht abgebildet werden? Haben Hochschulen mit geisteswissenschaftlichen Schwerpunkten in den bibliometrisch ausgerichteten Rankings überhaupt eine Chance oder sind die verwendeten Messansätze zu ihren Lasten verzerrt?

Diesen kritischen Punkten in Bezug auf bestehende Rankings steht die generelle Bedeutung von Transparenz im Hochschulsystem gegenüber: Hochschulsysteme werden immer komplexer und für die „Stakeholder“ (Studierende, Partner der Hochschule, Wirtschaft, Steuerzahler etc.) zunehmend undurchschaubarer. Die Anspruchsgruppen können aber nur auf Basis valider Informationen über Hochschulen rationale Entscheidungen treffen: über die Studienortwahl, über die Auswahl von Partnern, über den Vergleich mit Benchmarks usw. Das hat Rückwirkungen auf politische Ziele: Die Mobilität in der „European Higher Education Area“ hängt auch von der Informiertheit der Studierenden über die Bedingungen an Hochschulen in den unterschiedlichen Ländern ab. Wenn sich in einem Hochschulsystem Profile ausdifferenzieren, dann bedarf es auch der Transparenz über die Leistungen der Hochschulen in Bezug auf ihr jeweiliges Profil. Die Befriedigung vielfältiger Bedürfnisse wird auf Basis der Transparenz im Hinblick auf die Differenzen innerhalb der Hochschullandschaft erleichtert. Damit erscheinen Rankings als ein wünschenswertes Instrument – gleichzeitig scheinen die existierenden internationalen Rankings nicht in der Lage zu sein, den Transparenzielen gerecht zu werden, ohne massive Probleme zu erzeugen.

Beide generellen Ansatzpunkte – die Unzufriedenheit mit vorhandenen Rankings und die grundsätzliche Notwendigkeit von Transparenz – haben die Europäische Kommission auf den Plan gerufen. In der Communication zur Modernisierungsagenda aus 2011 (Europäische Kommission 2011) wird formuliert:

„Die Kommission wird die Verbesserung der Faktengrundlage für die politische Entscheidungsfindung in den Schlüsselbereichen in den Mittelpunkt stellen. Die verfügbaren Informationen über die Leistung von Hochschulen konzentrieren sich hauptsächlich auf Universitäten mit dem Schwerpunkt Forschung, so dass nur ein sehr kleiner Teil der europäischen Hochschulen erfasst wird: Es ist wichtig, die Analysen und

Informationen zu erweitern und alle Leistungsaspekte zu erfassen – damit die Studierenden ihre Studien- und Studienortwahl gut informiert treffen können, damit die Hochschulen ihre Stärken ermitteln und ausbauen können, damit die politisch Verantwortlichen bei ihrer strategischen Entscheidungsfindung zur Reform der tertiären Bildungssysteme unterstützt werden können. Es ist erwiesen, dass ein multidimensionales Ranking- und Informationstool machbar ist und bei den Interessenträgern im Bildungsbereich auf breite Zustimmung stößt.“

Beginnend ab 2005 initiierte die Europäische Kommission zunächst mehrere Projekte unter dem Stichwort „U-Map¹“ mit dem Ziel der Entwicklung einer Klassifikation von Hochschulen und in den Jahren 2010 und 2011 eine Machbarkeitsstudie unter dem Titel „U-Multirank²“, in deren Rahmen ein „multi-dimensional, user-driven“ Rankingssystem entwickelt sowie in einem Pilotversuch weltweit getestet wurde. Im Rahmen des U-Multirank-Projekts wurde ein Vorschlag zur Integration und Abstimmung der Klassifikations- und Ranking-Modelle entwickelt.

Im Folgenden werden zunächst die Begriffe und Rollen von Klassifikation und Ranking sowie ihr Zusammenspiel präzisiert und Prinzipien für deren geeignete Gestaltung abgeleitet. Sodann werden die wesentlichen Merkmale und Ergebnisse der genannten europäischen Projekte dargestellt. In einem Fazit werden der Stand der Dinge reflektiert und Zukunftsperspektiven aufgezeigt.

2 Klassifikation und Ranking

Hochschulsysteme sind gekennzeichnet durch Diversität und Differenzierung (*van Vught 2008*): Diversität beschreibt den Grad der Unterschiedlichkeit zu einem bestimmten Zeitpunkt, Differenzierung steht für die Veränderung der Vielfalt, zum Beispiel durch Ausdifferenzierung von Profilen oder der Gründung neuer Hochschulen und Hochschultypen. Diversität ist demnach ein statisches, Differenzierung ein dynamisches Konzept. Zu unterscheiden sind zwei verschiedene Arten von Diversität: horizontale und vertikale Diversität (*Teichler 2007*). Während die horizontale Diversität Unterschiede beschreibt, ohne Aussagen über „besser“ oder „schlechter“ zu machen, bezieht sich die vertikale Diversität auf die Unterschiede in Leistung und Reputation.

Eine normative Grundüberlegung hinter Klassifikations- und Rankingsystemen, die auch im Rahmen der Experten- und Stakeholder-Konsultationen im Umfeld von U-Map und U-Multirank immer wieder aufgetaucht ist, ist folgende: Diversität, horizontal wie

¹ Ein Überblick über das U-Map-Projekt sowie Berichte zu dessen Entwicklung und Implementierung finden sich auf <http://www.u-map.eu/>.

² Ein Überblick über das U-Multirank-Projekt sowie Berichte zu dessen Entwicklung und Implementierung finden sich auf <http://www.u-multirank.eu/>.

vertikal, ist positiv, denn sie schafft Wahlmöglichkeiten, erlaubt Spezialisierung, stärkt Wettbewerb und Innovationsfähigkeit (vgl. *Schmücker 2011, S. 19 f.*), ermöglicht flexible Korrekturen von Fehlentwicklungen und eröffnet die Möglichkeit, Massifizierung und Qualitätsförderung im Hochschul- und Forschungssystem zu kombinieren. Wird Diversität transparent, dann wird sie (und werden die positiven Effekte) dadurch gefördert – „transparency tools“ kommt deshalb eine zentrale Bedeutung für die Entwicklung von Hochschulsystemen zu. Wichtig ist, dass die Werkzeuge sich jeweils klar auf die beiden Diversitätsaspekte beziehen: Die Klassifikation ist das Instrument zur Abbildung von horizontaler, das Ranking zur Abbildung von vertikaler Diversität. Sie sind damit komplementär, bilden aber verschiedene Sachverhalte ab. Soweit die Idealvorstellung von Transparenzansätzen. Eingangs wurden bereits Zweifel dargestellt, ob es vor allem Rankings wirklich gelingt, diese Vielfalt auch tatsächlich valide abzubilden (eine umfassende Darstellung der Ranking-Kritik findet sich bei *Hazelkorn 2011*).

Der traditionelle Ansatz der Klassifikation ist die Legaldefinition von Hochschultypen, beispielsweise unsere deutsche Unterscheidung von Universität und Fachhochschule. Ein Problem dabei ist, dass diese national ausgerichtete Typisierung in einem europäischen oder globalen Zusammenhang nicht mehr funktioniert. Die Möglichkeit, auch im internationalen Kontext zu einer Klassifikation zu kommen, ist ein empirischer Ansatz unter Verwendung von Indikatoren, die Unterschiede beschreiben. Dabei sind mit „typology“ und „mapping“ nochmals zwei Varianten von Klassifikation zu unterscheiden: Bei der Typologisierung führt eine bestimmte Kombination von Indikatorwerten zur Zuordnung zu einem definierten Typ, zum Beispiel einer „teaching university“. Jede Hochschule gehört dort einem bestimmten Typ an. Beim „mapping“ wird anhand von Indikatoren ein empirisch gemessenes Profil beschrieben, das jeweils spezifisch für die einzelne Hochschule ist und nicht zu einem festgefügteten, mit einem bestimmten Namen versehenen Typ führt. Da in diesem System die Nutzer flexibel entscheiden können sollen, welche der angebotenen Indikatoren sie für eine Gruppierung der Hochschulen wählen, können diese je nach ausgewählten Indikatoren in unterschiedlichen Vergleichsgruppen landen.

Mit den unterschiedlichen Rollen von Klassifikation und Ranking – erstere machen Profilunterschiede bewusst, letztere bilden Leistungsunterschiede ab (oder versuchen es zumindest) – offenbart sich aber auch die Komplementarität der Instrumente: Mit der Klassifikation findet man vergleichbare Hochschulen, mit dem Ranking lassen sich innerhalb dieser Gruppe vergleichbarer Einrichtungen Leistungsunterschiede bewerten (mit Messgrößen, die zum jeweiligen Profil passen). U-Map und U-Multirank zielen genau auf diese Rollenteilung und Verknüpfung ab. Schon aus der Grundlogik der Diversität wird deutlich, dass beide Systeme den Ansatz der „Multidimensionalität“ verfolgen müssen – die horizontale Vielfalt ist größer als eine bloße Unterscheidung in lehr- und forschungsintensiv; und wenn ein Ranking für vielfältige Profile geeignet

sein will, muss es auch über bloße Messung von Forschungsleistung hinausgehen und das gesamte Aufgabenspektrum einer Hochschule mit einer größeren Zahl von Indikatoren abbilden. Im folgenden Abschnitt werden die Gestaltungsprinzipien für U-Map und U-Multirank aus den Problemen bestehender Systeme begründet und im Überblick dargestellt.

3 **Gestaltungsprinzipien für Klassifikationen und Rankings**

Die Gestaltungsprinzipien, die der Entwicklung von U-Multirank und U-Map zugrundegelegt wurden, ergeben sich unmittelbar aus den Problemen und schädlichen Effekten bestehender Systeme. Zu erwähnen sind mit Blick auf bestehende Rankings in aller Kürze insbesondere die Vernachlässigung von Vielfalt durch Fokussierung auf die „world class research university“ und auf Forschungs-/Reputationsindikatoren, die mangelnde Berücksichtigung der Stakeholder-Bedürfnisse, die Irreführung durch Aggregation zu „composite indicators“ mit willkürlichen Gewichtungen, der „league table“-Ansatz mit der Übertreibung von vertikalen Unterschieden sowie die zu geringe Beachtung der Fächerebene, die gerade aus der Sicht von Studierenden eine maßgebliche Rolle spielt (vgl. *Federkeil/van Vught/Westerheijden 2012; Federkeil 2013 in diesem Band*).

Bei Klassifikationen ergibt sich das Problem, dass deren rein deskriptive Kategorien häufig ungewollt doch mit vertikaler Diversität aufgeladen werden. Führt eine Klassifikation zum Beispiel zu einer Unterscheidung von forschungs- und lehrorientierten Hochschulen, dann werden in den normativen Vorstellungen der wissenschaftlichen Community (und darüber hinaus bis hinein in Politik und Medien) tendenziell die ersten als hierarchisch übergeordnet und von höherer Leistungsfähigkeit betrachtet, selbst wenn die Klassifikation nur Unterschiede in den Schwerpunkten beschreibt. Hinzu kommt, dass Klassifikationen Hochschulen allzu leicht in Schubladen packen, aus denen diese dann nicht mehr herauskommen und die stattdessen zu „self-fulfilling prophecies“ werden. Dies kann durch eine Koppelung der staatlichen Mittelvergabe an solche Einteilungen noch verstärkt werden, wonach dann Dynamik im Sinne der Differenzierung behindert und nicht gefördert wird.

Aus solchen Problemlagen heraus, kombiniert mit vorliegenden Erfahrungen vor allem mit der Carnegie Classification³ in den USA (*van Vught/Kaiser/File/Gaethgens/Peter/Westerheijden 2010, S. 14 f.*), wurde U-Map auf Basis folgender Gestaltungsprinzipien entwickelt (*ebd., S. 10 ff.*):

³ Für weitere Informationen zu der Carnegie Classification of Institutions of Higher Education siehe <http://classifications.carnegiefoundation.org/>.

- Die Klassifizierung soll auf reliablen und validen empirischen Daten beruhen (und nicht auf gesetzlichen Regelungen).
- Sie soll mehrere Dimensionen umfassen, die das vielfältige Aufgabenspektrum von Hochschulen weitestgehend abdecken. Innerhalb der Dimensionen soll es mehrere Indikatoren geben, die nebeneinander stehen und zusammen ein Bild in Bezug auf die jeweilige Leistungsdimension ergeben. Die Relevanz der Dimensionen und Indikatoren soll aus Sicht der Stakeholder definiert werden und die Vielfalt ihrer Sichtweisen einbeziehen.
- Sie soll deutlich machen, dass keine Hierarchien von Hochschultypen geschaffen werden, sondern dass alle abgebildeten Profile gleichwertig sind.
- Sie ist deskriptiv und nicht präskriptiv, d. h. es gibt keine definierten „Typen“ und es wird der oben beschriebene „mapping“-Ansatz verfolgt.
- Sie soll offen sein für alle europäischen Hochschulen und alle Arten von Profilen.
- Sie soll soweit wie möglich auf verfügbaren Daten beruhen, um die Last der Datensammlung bei den Hochschulen zu minimieren; soweit aber keine vergleichbaren Daten vorhanden sind, ist dies durch Erhebungen bei den Hochschulen zu überwinden.

Analog dazu wurden Gestaltungsprinzipien für U-Multirank entwickelt, mit dem Ziel, die genannten Probleme zu vermeiden und Erfahrungen mit innovativen Ranking-Systemen (Leiden Ranking, CHE Ranking⁴) sowie die Anforderungen der „Berlin Principles“ der International Ranking Experts Group (*IREG 2006*) an gute Rankings zu berücksichtigen (*Federkeil/van Vught/Westerheijden 2012; Federkeil/Kaiser/van Vught/Westerheijden 2012, S. 85 ff.*):

- Rankings sollten von den verschiedenen Nutzern durch die Auswahl von ihnen als relevant erachteten Größen selbst gestaltet werden („user-driven“). Maßgeblich dafür ist die Tatsache, dass es kein objektives Ranking gibt und somit alle Messsysteme zur Abbildung der Realität immer durch ein normatives Design der Ranking-Konstrukteure gesteuert sind. Bevormundung lässt sich demnach nur durch den aktiven Einbezug der Benutzer vermeiden.
- „User-driven“ impliziert zudem, dass bei der Entwicklung der Dimensionen und Indikatoren deren Relevanz in intensiven Stakeholder-Konsultationen getestet wird, sowie dass auch im Echtbetrieb des Ranking permanent Feedbacks der Stakeholder eingeholt und in Weiterentwicklungen verarbeitet werden müssen.
- Der „user-driven“-Ansatz wird nur möglich durch Mehrdimensionalität, analog zu den U-Map-Prinzipien. Die Dimensionen und Indikatoren von U-Multirank sollten die

⁴ Für Informationen zum Leiden Ranking siehe <http://www.leidenranking.com/>; für Informationen zum CHE Ranking <http://www.che-ranking.de>.

Vielfalt von Funktionen und Aufgaben der Hochschulen abdecken. Sie sollten alle ungewichtet nebeneinander gestellt und nicht zu einem „composite indicator“ verarbeitet werden.

- Bei den Indikatoren soll es um tatsächliche Leistung statt um Reputation gehen.
- Ein Ranking sollte nur vergleichbare Hochschulen einbeziehen, da Leistung nur dann vergleichbar ist, wenn auch die Profile der betrachteten Hochschulen vergleichbar sind.
- Ein Ranking sollte die institutionelle Ebene und die Fächerebene abbilden. Es gibt nur sehr wenige Hochschulen, die in allen Fächern gleich gute Leistungen erbringen; eine bloße Abbildung der Ebene der Gesamtinstitution würde die Unterschiede zwischen den Fächern über eine Durchschnittsbildung verwischen. Zudem sind für unterschiedliche Stakeholder-Gruppen verschiedene Ebenen relevant.
- An die Stelle der „league tables“ mit Einzelplatzierungen der Hochschulen soll ein Ranking mit Gruppenbildung treten (vgl. *Usher/Medow 2009, S. 14*). Eine begrenzte Anzahl an Leistungsgruppen, die sich im Leistungsniveau deutlich unterscheiden, sollte für jeden Indikator identifiziert und ausgewiesen werden. Damit wird ein Ranking zwar „gröber“, vermeidet aber die übertriebene Darstellung von Leistungsunterschieden, die faktisch nur sehr klein sind.
- Rankings sollten auf einer ausgewogenen Vielfalt von Datenformen basieren, damit sich im multidimensionalen Kontext durch die Kombination verschiedener Indikatoren ein aussagekräftiges Bild ergibt. Deswegen sollten sowohl Fakten als auch subjektive Urteile durch verschiedene Datenbankauswertungen und Befragungen erhoben werden.
- Rankings sollten die von existierenden Rankings bekannten methodischen Fehler vermeiden, zum Beispiel die unzureichende Berücksichtigung von Fächerdifferenzen in bibliometrischen Messverfahren (*Moed 2005; Federkeil/van Vught/Westerheijden 2012*).
- Die Minimierung des Erhebungsaufwands empirischer Daten gilt analog zum entsprechenden Prinzip bei U-Map, wozu ebenfalls das Prinzip des „pre-filling“ herangezogen wird.

Die Gestaltungsprinzipien machen die komplementären Rollen von U-Map und U-Multirank deutlich: Beide können miteinander verknüpft werden, weil sie mit der Multidimensionalität und dem „user-driven“-Ansatz derselben Methodik folgen. Verbunden werden sie nach der in Abschnitt 2 erläuterten Logik: U-Map zeigt Aktivitätsprofile von Hochschulen und damit ihre horizontalen Unterschiede, U-Multirank hingegen offenbart Leistungsprofile, also vertikale Diversität. Ein sinnvolles Ranking in U-Multirank wird nur möglich, wenn man zuvor mit U-Map oder einer ähnlich ausgerichteten Form der Klassifikation die vergleichbaren Hochschulen gefunden hat.

4 U-Map: Merkmale und Erfahrungen

Die Umsetzung der im vorigen Kapitel skizzierten Gestaltungsprinzipien erfolgte bei dem von der Europäischen Kommission unter dem „Lifelong Learning“-Programm finanzierten U-Map-Projekt im Rahmen eines mehrstufigen, iterativen Design-Prozesses (Kaiser/van Vught 2009; van Vught/Kaiser/File/Peter/Gaehdgens/Westerheijden 2010). Dessen Ausgangspunkt bildete die zu Projektbeginn identifizierte Hauptaufgabe, an den Bedürfnissen verschiedener Stakeholder ausgerichtete Möglichkeiten der Abbildung differierender Hochschulprofile bereitzustellen. Von 2005 bis 2009 wurde in drei größeren Projektphasen ein Klassifikationssystem entwickelt und stetig überarbeitet.

Der Design-Prozess lässt sich grob in fünf Schritte einteilen: die Klärung der theoretischen und konzeptionellen Grundlagen, etwa der Abgrenzung der zu klassifizierenden Institutionen; die Identifizierung der relevanten Dimensionen der Klassifizierung; die Entwicklung geeigneter Indikatoren für die Dimensionen; die Prüfung der zur Verfügung stehenden Daten und möglicher neuer Datenquellen; sowie eine erste konkrete Anwendung der Klassifikation. Quer zu diesem „idealen“ Ablauf stand die bei den mit jedem Schritt neu gewonnenen Erkenntnissen ansetzende Überprüfung der Qualität des Herzstücks der Klassifikation – der Dimensionen und der Indikatoren.

Die Herausforderung bei der Wahl der Dimensionen war die Vermittlung zwischen den erwünschten, mit einer höheren Zahl an Dimensionen verbundenen Wahlmöglichkeiten und der angestrebten, auf eine eher geringe Zahl an Dimensionen angewiesenen Komplexitätsreduktion. Zuerst wurde mittels Literaturrecherchen eine längere Liste an Indikatoren erstellt, die über Stakeholder-Konsultationen auf 14 reduziert wurde. Nach weiteren Untersuchungen und ersten Datenerhebungen wurden schließlich die aktuell in U-Map implementierten sechs Dimensionen ausgewählt. In einem ähnlichen Verfahren wurden etwas zeitversetzt dazu die einzelnen Indikatoren entwickelt, auf ihre Qualität hin untersucht und, falls notwendig, überarbeitet. Der Schwerpunkt hierbei lag auf der Relevanz für die jeweilige Dimension aus Sicht der Stakeholder, der Qualität des Indikators sowie den Möglichkeiten und den Kosten der Erhebung der benötigten Daten.

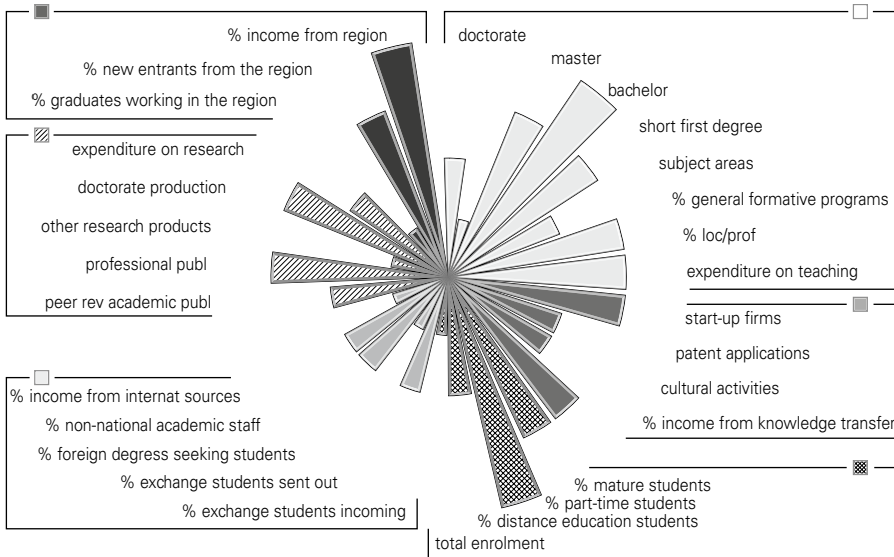
Ein zentrales Element der Entwicklung der Dimensionen und Indikatoren war die Interaktion mit den Stakeholdern. So wurde der gesamte Prozess von einem mit verschiedenen Stakeholdern besetzten Beirat begleitet. Zudem wurden in jeder Phase die (Zwischen-)Ergebnisse im Rahmen von Tagungen, Konferenzen und Treffen vorgestellt und diskutiert. Mit zunehmender Konkretion von U-Map wurden diese Formen der Auseinandersetzung ebenfalls konkreter. Mit der Liste der zunächst 14 Dimensionen und den dazugehörigen Indikatoren wurden acht Fallstudien an Hochschulen

durchgeführt – zwei davon mit Besuchen des Projektteams vor Ort – im Rahmen derer jene eingehend überprüft wurden. Außerdem wurden zwei Fragebögen entwickelt, mittels derer eine europaweite Untersuchung der Relevanz der Dimensionen und der Qualität der Indikatoren an mehr als 60 Hochschulen durchgeführt wurde. Diese Fragebögen zielten explizit auf die Relevanz der ausgewählten Dimensionen, die Validität und Reliabilität der Indikatoren und die Verfügbarkeit der Daten. Zusätzlich zu der Überarbeitung der Dimensionen und Indikatoren anhand der generierten Erkenntnisse wurde auf der Internetseite des Projekts ein in sechs Bereiche aufgeteiltes Diskussionsforum eingerichtet, um Auseinandersetzungen mit den Indikatoren über die Konferenzen hinaus zu ermöglichen.

In der aktuellen Version basiert U-Map auf sechs Dimensionen mit jeweils drei bis fünf Indikatoren (*siehe Anlage 1*). Bei den Dimensionen handelt es sich um das Profil von Lehre und Lernen, das Profil der Studierendenschaft, das Ausmaß der Forschungsaktivitäten, die Ausprägung des regionalen Engagements, die Intensität der Wissenstransfer-Aktivitäten und die Ausprägung der internationalen Ausrichtung. Insbesondere die dazugehörigen Indikatoren machen deutlich, dass hier horizontale Diversität und keine Leistungsbewertung abgebildet werden soll: Ausgabenanteile für Lehre und Forschung, Breite der Disziplinen, Anteile verschiedener Abschlussarten, Zusammensetzung der Studierendenschaft – solche Indikatoren beschreiben Unterschiede, ohne sie qualitativ zu werten. Beispiele für die Operationalisierungen der Indikatoren sind: „Mature students“ – Anteil der Studierenden über 30 Jahre (Kopfzahlen, alle Programme) an der Gesamtzahl eingeschriebener Studierender; „International academic staff“ – Anzahl der Mitarbeiter mit ausländischer Nationalität, festangestellt oder auf Austauschbasis. Dem offenen Prinzip von U-Map nach sind aber weder die Dimensionen noch die Indikatoren und ihre Operationalisierung unveränderbar. Sowohl Entwicklungen im Hochschulbereich als auch eine wachsende Menge an verfügbaren Daten können über einen stetigen Entwicklungsprozess in die Klassifikation einfließen.

Mit den einzelnen Indikatorwerten wird schließlich das „Aktivitätenprofil“ einer Hochschule erstellt. Neben den bloßen quantitativen Daten leistet U-Map eine Visualisierung der Hochschulprofile mit dem sogenannten „sunburst chart“ (*siehe Abbildung 1; van Vught 2010, S. 33 ff.*).

Abbildung 1: Beispiel für das Aktivitätenprofil einer Hochschule anhand der U-Map-Dimensionen und -Indikatoren



Die Graphik illustriert jede Dimension durch eine unterschiedliche Schattierung, jeder Strahl steht für einen Indikator. Die Indikatorwerte werden in Quartile geordnet und es gibt entsprechend der Gruppenzuordnung vier Strahlängen, je kleiner der Indikatorwert desto kürzer der Strahl. Das dargestellte Beispiel zeigt das Profil einer regional- und marktorientierten Hochschule mit intensiver Transferaktivität, einer lokalen aber diversen Studierendenschaft, geringer Internationalität und einem Fokus auf „undergraduates“. Da es um die Abbildung der horizontalen Diversität geht, bedeutet für die Strahlen „länger“ nicht gleich „besser“. Würde eine Hochschule nur in einer der Dimensionen lange Strahlen aufweisen, dann bedeutet das lediglich, dass sie auf diese Funktion einer Hochschule spezialisiert ist, nicht, dass ihre Leistung geringer ist.

Die Graphik ist als Teil eines Webtools konzipiert, dessen Ziel es ist, die Daten von U-Map über ein flexibles, nutzergesteuertes Internetangebot verfügbar zu machen. Dahinter steht die Idee, dass ein Nutzer einerseits die Profile bestimmter Hochschulen oder der Hochschulen in einem Land ermitteln können soll. Darüber hinaus sollen Nutzer andererseits aber auch über den „profile finder“ die sie interessierenden Ausprägungen der Dimensionen abfragen können und sodann einen Überblick über alle Hochschulen erhalten, die dem gewünschten Profil entsprechen (falls Nutzer die Indikatoren entsprechend auswählen, könnten sie sich beispielsweise auch nur auf eine Dimension beziehen).

Hieran wird deutlich, wie wichtig die Frage der Datenaufbereitung und Visualisierung für die Aussagekraft von Transparenzinstrumenten ist. Die Komplexität der Datenfülle durch eine Umsetzung in Graphiken zu reduzieren, ist einerseits notwendig, um die Informationen „verdaubar“ zu machen. Andererseits stecken darin stets auch Gefahren des Informationsverlusts und der zu starken Vereinfachung. Das „sunburst chart“ ist ein Kompromiss: Es zeigt alle Indikatoren ohne Aggregation, wobei diese durch die graphische Umsetzung auf einen Blick erfassbar werden. Restrisiko bleibt, dass die Intuition der Graphik „je länger, desto besser“ nahelegt, was einen Erklärungsbedarf der Graphik auslöst. Die Nutzersteuerung durch Auswählen von zu betrachtenden Dimensionen und Indikatoren erhöht den Nutzen des Instruments deutlich, denn es passt sich flexibel an Informationsbedürfnisse an.

Die intendierte Reichweite von U-Map ist Europa. Im Augenblick erfolgt der „roll out“ länderweise in den Staaten, in denen sich die Regierung für eine Implementierung entschieden hat. Derzeit sind das die Niederlande und Flandern, Portugal, Estland und die „nordic countries“ (Norwegen, Finnland, Schweden, Dänemark). Daneben können einzelne Hochschulen ihre Daten auf freiwilliger Basis liefern. Auch in den genannten Staaten ist die Teilnahme teilweise freiwillig, so dass nicht alle Hochschulen mitmachen, so in den nordischen Ländern. Die Implementierung von U-Map hat neben der Weiterentwicklung des Herzstücks der Klassifikation eine Reihe von Erfahrungen erbracht (*Kaiser/Faber/Jongbloed/File/van Vught o.J.; U-Map o.J.; U-Map/Norden o.J.*):

■ Die Messkonzepte und Erhebungsinstrumente funktionieren im internationalen Kontext. Für die Einführung hat sich ein standardisiertes Vorgehen bewährt, bei dem in einem Kick-off-Treffen zunächst der Nutzen des Ansatzes vermittelt wird, dann nach „pre-filling“-Möglichkeiten gesucht und in einem technischen Workshop mit den Hochschulen Erhebungsfragen diskutiert werden. Nach der Online-Datenerhebung, die sich mittlerweile nicht mehr an den einzelnen Dimensionen, sondern an der Art und Weise, wie die Daten in den Hochschulen vorliegen, orientiert, erfolgt eine intensive Phase der Datenverifizierung (Analyse von Ausreißern, Inkonsistenzen, unerwarteten Ergebnissen) mit Rückfragen an die Hochschulen sowie eine Überarbeitung, Analyse und Aufbereitung der Daten.

■ Die Möglichkeiten zum „pre-filling“ sind länderspezifisch unterschiedlich, aber etwas begrenzt. Nationale Hochschulstatistiken in Europa bieten bislang keine hinreichende Vergleichbarkeit von Daten, um eine Grundlage für Systeme wie U-Map (oder U-Multi-rank) zu schaffen. Das Beispiel von Norwegen zeigt aber zumindest das Potential dieses Ansatzes. Mit den dort vorhandenen Datenquellen wurde eine Fallstudie zu den Möglichkeiten des „pre-filling“ durchgeführt. Zwar war der Aufwand des ersten Durchlaufs durchaus hoch, da Operationalisierungen konkretisiert und weitere methodische Fragen geklärt werden mussten, nach Abschluss des Testlaufs wurde der Aufwand für eine weitere Erhebung von den nationalen Behörden jedoch als gering eingeschätzt.

■ Der Ansatz des „mapping“ anstelle einer Typologie hat sich als tauglich erwiesen. In den nationalen Auswertungs-Workshops zu U-Map wurden Stakeholder gefragt, welche Hochschultypen sie angesichts der ermittelten Profile definieren würden. Parallele Arbeitsgruppen kamen jeweils zu anderen Ergebnissen der Typenbildung. Damit wurde den Beteiligten deutlich, dass eine feste Definition von Typen willkürlich erscheint und es besser ist, die nutzerorientierte Kombination von Merkmalen zuzulassen.

■ Es ist wichtig, den politischen Umgang mit Klassifikationen in die richtige Richtung zu lenken. Ein positiver Effekt ist beispielsweise, wenn mit Hilfe der Klassifikation eine Regierung reflektiert, ob im Land ein hinreichendes Spektrum an unterschiedlichen Profilen vorhanden ist oder bestimmte Profilausprägungen gestärkt werden sollten, zum Beispiel, indem durch neue Hochschulen Lücken geschlossen werden. Im Zuge der Untersuchung der Bedürfnisse und Erwartungen einzelner Hochschulen in der zweiten Phase des Design-Prozesses wurde festgestellt, dass die Institutionen U-Map als ein geeignetes Mittel für ihre strategische Planung ansehen, sowohl für die Schärfung ihres Profils und ihrer Mission als auch für Vergleiche mit anderen Institutionen, die ein ähnliches Profil aufweisen.

■ Horizontale und vertikale Diversität sind in ihrer empirischen Messung zum Teil schwer auseinanderzuhalten. Sind hohe Ausländeranteile oder viele Spin-offs nur eine Beschreibung von Unterschieden oder zeigen sie Leistungsunterschiede an? Besonders virulent wird das Problem bei den in U-Map verwendeten Forschungsindikatoren: Die verwendeten bibliometrischen Ansätze werden üblicherweise auch als Leistungsindikatoren eingesetzt. Es wird dadurch schwieriger, die Nicht-Hierarchie unterschiedlicher Profile in einer Klassifikation zu kommunizieren.

Aus dem letztgenannten Problem resultiert eine wichtige Schlussfolgerung für die Verbindung von Klassifikation und Ranking: Wenn im Rahmen der zukünftigen Implementierung von U-Multirank eine Klassifikation integriert wird, mit der vergleichbare Hochschulen ausgewählt werden, dann sollte (im Gegensatz zum Pilottest) nicht direkt U-Map verwendet werden. Stattdessen sollte auf ein reduziertes Indikatorenset zurückgegriffen werden, bei dem die Gefahr der Überlappung mit der Leistungsmessung nicht besteht und beispielsweise bibliometrische Maße nicht enthalten sind.

5 U-Multirank: Merkmale und Ergebnisse der Machbarkeitsstudie

Der Entwicklungsprozess des ebenfalls von der Europäischen Kommission geförderten U-Multirank-Projekts weist bezüglich der eingangs skizzierten Grundüberlegungen einige Parallelen zu demjenigen von U-Map auf. Ein mehrstufiger, iterativer Designprozess unter Einbezug der Stakeholder ist hier ebenfalls zentral. Bei U-Multirank wurde jedoch im Gegensatz zu U-Map die Implementierungsphase noch nicht abgeschlossen, sondern bisher nur die Machbarkeitsstudie durchgeführt (Indikatorenliste zum Ende

der Machbarkeitsstudie *siehe Anlage 2*). Diese Machbarkeitsstudie baut auf einem ersten Design eines multidimensionalen Rankings durch das U-Multirank-Konsortium auf.

U-Multirank greift auf Hochschul- und Fächerebene fünf von den zentralen Funktionen der Hochschulen und ihren verschiedenen Zielgruppen abgeleitete Dimensionen auf: Lehre, Forschung, Wissenstransfer, internationale Orientierung und regionales Engagement – jeweils mit mehreren Indikatoren operationalisiert. Als die entscheidenden Faktoren für die Qualität eines Rankings standen beim Design-Prozess von U-Multirank die Indikatoren und deren fundierte Ableitung im Zentrum der Aufmerksamkeit. Wie Abbildung 2 am Beispiel der Internationalitäts-Dimension zeigt, wurde dazu ein systematischer Prozess zur Indikatorenfindung realisiert.

Abbildung 2: Bewertung für die fachbezogenen Indikatoren der Dimension „internationale Orientierung“ von U-Multirank

International Orientation	Rating of indicators (pre-pilot)						Feasibility score (post-pilot)				
	Relevance	Concept/construct validity	Face validity	Robustness	Availability	Preliminary rating	Feasibility score	Data availability	Conceptual clarity	Data consistency	Recommendation
Field-based ranking											
Percentage of international students	■	▲	■	■	▲	A	A	▲	▲	▲	
Incoming and outgoing students	▲	▲	■	■	▲	A	A-B	■	■	▲	
Opportunities to study abroad (student satisfaction)	■	■	▼	■	▲	A	B	■	▲	▲	
International orientation of programs	▲	▲	▼	■	■	A	B	■	▲	▲	
International academic staff	▲	■	■	■	■	B	A-B	■	▲	■	
International joint research publications	▲	■	▼	▲	▲	B	A	▲	▲	▲	
International research grants	▼	■	■	■	▲	B	B	■	▲	▲	
International doctorate graduation rate	▲	■	▼	■	■	B	A	▲	▲	▲	

Quelle: van Vught/Ziegele 2011, S. 130⁵

⁵Wie schon bei Abbildung 1 konnten auch hier die Farben aufgrund des Drucks nicht wiedergegeben werden. Im Original besitzen die nach oben weisenden Pfeile die Farbe Grün, die Rechtecke die Farbe Gelb und die nach unten weisenden Pfeile die Farbe Rot.

Zunächst wurde eine relativ umfassende Liste von Indikatoren für die jeweiligen Dimensionen entwickelt, basierend auf Literatur und Praxis der Leistungsmessung an Hochschulen, plus einige Neuentwicklungen bezogen auf die für Leistungsvergleiche eher „neuen“ Dimensionen.⁶ Versucht wurde insbesondere, auch über reine Input- und Prozess-Indikatoren hinaus, die „outcomes“ erfassbar zu machen. Diese Indikatoren wurden einem ersten („pre-pilot“) Experten-Check unterzogen im Hinblick auf Validität, Verbreitung in vorhandenen Systemen („face validity“), Robustheit (bestehend aus Reliabilität und Vergleichbarkeit im internationalen Kontext) und Verfügbarkeit. Die Symbole in der Tabelle signalisieren eine positive, negative oder neutrale/mittlere Einschätzung der Eignung. Zudem wurden sie in Workshops und einem Online-Feedback Stakeholdern zur Beurteilung der Relevanz vorgelegt. Daraus ergab sich ein erstes Gesamturteil über die Verwendbarkeit des Indikators, ausgedrückt in einem ABC-Rating (A für „problemlos verwendbar“, B für „wünschenswert, aber noch Nachbesserungsbedarf“ und C für „nicht verwendbar“). Mit den A- und B-Indikatoren wurde dann der Test mit Pilothochschulen durchgeführt, der zusätzliche („post-pilot“) Erkenntnisse in Bezug auf Verfügbarkeit der Daten, Verständlichkeit des Konzepts für die Hochschulen und die Konsistenz der Antworten in den Fragebögen erbrachte. Die Einstufung in A, B oder C wurde abschließend nochmals vorgenommen.

Wesentlicher Teil des U-Multirank-Ansatzes ist die Umsetzung in ein interaktives Webtool mit entsprechenden Präsentationsformen. Im Rahmen von U-Multirank ist keine Veröffentlichung einer festen Ranking-Liste geplant, sondern die Einführung eines nutzergesteuerten, internetbasierten Instruments der Auswahl von Merkmalen und Leistungsindikatoren von Hochschulen, wobei, wie bereits erläutert, eine vereinfachte Klassifikation als Vorstufe für das Ranking integriert ist. Die endgültige Konzeption des Webtools war nicht Gegenstand der Machbarkeitsstudie, wohl aber eine Grundidee zu seiner Gestaltung: Ein Nutzer soll zunächst wählen können, ob er das institutionelle oder das fachbezogene Ranking durchführen will (inklusive Handlungs- und Interpretationsanleitungen, welche Informationen für welche Art von Zielgruppe oder Entscheidung nützlich sein könnten). Sodann kann er die gewünschten Merkmale der Hochschule auswählen (das „mapping“). Daraufhin erhält er eine standardisierte Übersichtstabelle wie im Beispiel in Abbildung 3, das auf anonymisierten Echtdateien aus dem Pilottest beruht.

⁶ Eine ergänzende, interessante Perspektive auf die Vielfalt möglicher Indikatoren eröffnet eine Studie zu Universitätsprofilen in Deutschland (*Schmücker 2011*). Ungeachtet des anderen Zugangs zu dem Thema Diversität wird dort eine Vielzahl verschiedener Indikatoren diskutiert und auf verfügbare Daten hin geprüft, womit ein Einblick in die mögliche Vielfalt, aber auch die schon auf einem nationalen Level teilweise problematische Verfügbarkeit und Vergleichbarkeit der Daten gegeben wird.

Abbildung 3: Übersichtstabelle zum fächerbezogenen Ranking (Beispiel Maschinenbau)

	Teaching & Learning			Research			Knowledge Transfer			International orientation			Regional engagement		
	student staff ratio	graduation rate	qualification of academic staff	research publication output	external research income	citation index	% income third party funding	CPD courses offered	startup firms	international academic staff	% international students	joint international publ.	graduates working in the region	student internships in	regional co-publication
Institution 4	○	●	●	○	●	●	●	●	●	○	●	●	○	●	●
Institution 8	○	●	●	●	-	●	●	●	●	●	●	-	●	○	○
Institution 3	●	○	○	●	○	●	●	○	●	●	○	●	●	-	●
Institution 5	●	○	○	●	○	○	●	○	●	●	●	○	●	○	-
Institution 1	●	-	-	●	○	○	○	●	○	●	●	○	○	●	○
Institution 9	●	●	●	●	●	●	●	●	●	○	●	●	●	●	●
Institution 7	●	●	●	●	-	●	○	-	○	-	●	●	●	●	●
Institution 2	●	●	●	○	○	○	-	●	●	●	●	●	●	●	●
Institution 6	●	●	●	●	-	●	○	●	○	●	○	○	○	●	○

Quelle: van Vught/Ziegele 2011: S. 20

Für alle Dimensionen werden zunächst drei Indikatoren angeboten. Die Punkte symbolisieren angelehnt an das Ampelsystem die Zugehörigkeit zur Top-, Mittel- oder Schlussgruppe in Bezug auf den jeweiligen Indikator.⁷ Entsprechend dem Prinzip der Multidimensionalität werden die Indikatoren nicht aggregiert, auch nicht innerhalb der Dimensionen, sondern im Überblick dargestellt. Das Beispiel zeigt den Nutzen der Mehrdimensionalität: Zum einen findet sich keine Hochschule, die in allen Dimensionen gleich stark ist, hinter einem „composite indicator“ würden diese Differenzen verschwinden. Zum anderen zeigt sich, dass Hochschulen, die zum Beispiel in der Forschung schwach sind und daher in traditionellen Rankings keine Chance hätten, durchaus erstklassige Leistung in anderen Dimensionen erbringen können. Gerade die Visualisierung fördert damit die Wahrnehmung, dass Hochschulen gute Leistungen auf vielfältige Weise erbringen und in verschiedenen Bereichen Stärken und Schwächen haben können.

Aus der standardisierten Tabelle soll ein Direkteinstieg in die Interaktivität möglich sein: Zum Beispiel sollte man nach der Gruppierung bei einzelnen Indikatoren sortieren oder Indikatoren in der Liste durch andere aus dem Gesamtkatalog austauschen

⁷ Die Vorgehensweise der Zuordnung der einzelnen Institutionen zu einer von drei Gruppen und die an das Ampelsystem angelehnte Form der Visualisierung entstammen dem CHE Hochschulranking. Dort wird jedoch mittlerweile aus methodischen Gründen die Schlussgruppe blau symbolisiert (siehe <http://www.che-ranking.de/methodenwiki/index.php/Hauptseite>).

können; in einem personalisierten Ranking sollte eine individuelle Auswahl von Dimensionen/Indikatoren möglich sein, eventuell auch mit der Möglichkeit, selbst zu gewichten. Neben der tabellarischen sind auch graphische Darstellungen angedacht, möglicherweise ähnlich zum bei U-Map verwendeten „sunburst chart“, um ein Leistungsprofil einer Hochschule auf einen Blick zeigen zu können.

U-Multirank wurde in seiner Gesamtheit in einer Pilotstudie getestet (*Callaert/Epping/Federkeil/File/Jongbloed/Kaiser/Roessler/Tijssen/van Vught/Ziegele 2012*). 316 Hochschulen wurden weltweit angefragt, 159 haben an dem Pilotranking teilgenommen und zwischen November 2010 und März 2011 die empirischen Instrumente bearbeitet (institutionelle Fragebögen auf Leitungs- und Fächerebene, Studierendenbefragung). Für diese Hochschulen wurden auch bibliometrische und Patentdaten erhoben. Die Systeme und Prozesse rund um eine derart komplexe Erhebung wurden entwickelt, darunter: Datenbanken, Online-Fragebögen, ein Helpdesk-Service und Kommunikationsstrukturen, FAQs und Anleitungen, Mechanismen zur Prüfung der Datenqualität und anknüpfende Feedbackschleifen mit den Hochschulen. 109 Hochschulen beantworteten die Fragen zum institutionellen Ranking, 57 im Pilotfach Betriebswirtschaftslehre und 108 in den beiden ingenieurwissenschaftlichen Pilotfächern. Fast ein Drittel der beteiligten Hochschulen kamen aus Nicht-EU-Staaten. Knapp 6000 Studierendenfragebögen wurden in die Analyse einbezogen.

Der Pilotversuch brachte eine Reihe von Erkenntnissen, die für die weitere Entwicklung von U-Multirank von erheblicher Bedeutung sind:

■ In drei der fünf Dimensionen ist es kein Problem, genügend relevante, aussagekräftige und praktikable Indikatoren zu finden. Abbildung 4 zeigt, wie viele Indikatoren in den Dimensionen getestet wurden und welcher Anteil davon problemlos umgesetzt werden kann (Kategorie A), weitere Bearbeitung erfordert (B) oder nicht machbar ist (C). Ein noch zu geringer Anteil an A-Indikatoren findet sich beim Wissenstransfer und beim regionalen Engagement – nicht zufällig liegen die größten Probleme bei den bisher in internationalen Rankings kaum abgebildeten Dimensionen, was sich selbst bei einer nationalen Betrachtung wie derjenigen in Deutschland zeigt (*Schmücker 2011, S. 95*).

Abbildung 4: Ergebnisse des Pilotversuchs im Hinblick auf die Machbarkeit der U-Multirank-Indikatoren

Dimension	Gesamtzahl der getesteten Indikatoren	Indikatorbewertung nach dem Pilotversuch		
		A: Indikatoren benötigen keine oder kleine Veränderungen	B: Indikatoren müssen weiter bearbeitet werden	C: Indikatoren, die nicht verwendet werden
Lehre	23	61 %	39 %	0 %
Forschung	16	56 %	38 %	6 %
Wissenstransfer	15	27 %	47 %	27 %
Internationale Orientierung	16	68 %	31 %	0 %
Regionales Engagement	11	18 %	64 %	18 %
Total	81	49%	42%	9%

Quelle: eigene Darstellung

■ Die Probleme mit der Machbarkeit von Indikatoren sind sehr spezifisch für einzelne Messgrößen, es ist kein übergeordnetes generelles Problem zu verzeichnen. Dazu einige Beispiele: Patente sind auf Fächerebene nicht zu ermitteln, weil die Fächerstrukturen in den verwendbaren Patentdatenbanken nicht abbildbar sind; bei den regionalen Indikatoren ist es schwer zu fassen, was eine Region ist (in Europa ist die NUTS-Regioneneinteilung hilfreich, aber darüber hinaus trägt dieses Instrument nicht); Daten über die Beschäftigung von Absolventen werden in den Staaten nicht oder nicht in vergleichbarer Weise bereitgestellt.

■ Ein interessantes Ergebnis ist die Entwicklung neuer bibliometrischer Indikatoren, die die Herkunft von Co-Autoren einer Publikation analysieren. Sind die Autoren aus unterschiedlichen Staaten, Ländern oder aus Hochschule und Wirtschaft, ergeben sich daraus Messwerte für Internationalität, Regionalität und Wissenstransfer. Bibliometrie wird dadurch über die Anwendung auf die Dimension Forschung hinaus anwendbar – ein wichtiger Fortschritt gegenüber den bestehenden globalen Rankings.

■ Eine besondere methodische Herausforderung ist die Studierendenbefragung im Hinblick auf den internationalen Kontext. Sind Zufriedenheitswerte von Studierenden über Länder- und Kulturgrenzen hinweg tatsächlich vergleichbar, so dass sie in ein

Ranking einfließen können? Unter Verwendung sogenannter „anchoring vignettes“⁸ wurden in der Pilotstudie die Erwartungsniveaus der Studierenden ermittelt. Wenn sich in verschiedenen Ländern die Erwartungen systematisch unterscheiden würden, dann wäre die Vergleichbarkeit der Studierendenzufriedenheit nicht gegeben (beziehungsweise müsste durch einen Korrekturfaktor entsprechend unterschiedlicher Erwartungsniveaus normiert werden). In der Datenbasis des Pilottests waren systematisch länderspezifische Erwartungsdifferenzen jedoch nicht festzustellen, so dass die Studierendurteile verwendbar sind. Im Rahmen einer Einführung von U-Multirank sind solche Tests der Datenvergleichbarkeit fortzuführen.

■ Die Bereitschaft zur Teilnahme am Pilotversuch war zufriedenstellend, aber ungleich verteilt. Besonders hohe Bereitschaft bestand bei Hochschulen, die in traditionellen Weltrankings aufgrund ihres Profils keine Chance haben, sichtbar zu werden, darunter auch Institutionen aus Afrika, Lateinamerika und dem Nahen Osten, die ursprünglich nicht in das Ranking implementiert werden sollten. Am geringsten war die Mitwirkungsbereitschaft in den USA und in China.

■ Die Mitwirkungsbereitschaft der Hochschulen ist erforderlich, da eine komplette Erhebung so umfassender Daten aus bestehenden Datensammlungen nicht realistisch ist. U-Multirank ist damit auf die institutionellen Datenlieferungen angewiesen. Ohne permanente Akzeptanz- und Überzeugungsarbeit würde ein Ranking dieser Art nicht funktionieren. Deutlich wurde, dass für den Teilnahmewillen der Hochschulen zwei Faktoren eine entscheidende Rolle spielen: Erstens, wie bereits in den Gestaltungsprinzipien gefordert, die intensiven „stakeholder consultations“ und die Responsivität des Rankings auf Anforderungen und Rückmeldungen aus dem Hochschulsystem. Zweitens die positive Kosten-Nutzen-Relation aus Sicht der jeweiligen Hochschulen. Die Gestaltung des Rankings kann sowohl die Kosten- wie die Nutzenseite beeinflussen: Die Kosten sinken durch „pre-filling“ oder durch Koordination mit nationalen Rankings, der Nutzen steigt, wenn die Hochschulen aus den Ranking-Daten Sonderauswertungen für eigene, interne Stärken-Schwächen-Analysen erhalten oder wenn sie aus Rankings Vorteile für das Marketing ziehen können (was bei Multidimensionalität eher der Fall ist, weil man gezielt auf vielfältige Stärken verweisen kann).

Alles in allem sprechen weder die Bereitschaft einzelner Institutionen, noch umfassendere Probleme mit den Indikatoren, noch logistische, organisatorische oder finanzielle Gründe gegen eine Ausweitung von U-Multirank auf andere (Welt-)Regionen

⁸ Die Verwendung von „anchoring vignettes“ zielt in Befragungen zu Einschätzungen darauf ab, potentielle systematische Unterschiede im Antwortverhalten verschiedener Personengruppen zu messen, um die Vergleichbarkeit der Antworten sicherzustellen. Dazu werden allen Befragten analog zu den eigentlichen Untersuchungsfragen dieselben hypothetischen Fälle vorgelegt – die „anchoring vignettes“. Über den Vergleich der auf die „anchoring vignettes“ bezogenen Einschätzungen kann dann entweder das Vorhandensein von Unterschieden – zum Beispiel im Hinblick auf Erwartungsniveaus – zwischen den Gruppen ausgeschlossen oder eine Bereinigung der Antworten vorgenommen werden. (King/Murray/Salomon/Tandon 2004; Rice/Robone/Smith 2010)

oder auf andere Fachdisziplinen (vgl. *Callaert/Epping/Federkeil/File/Jongbloed/Kaiser/Roessler/Tijssen/van Vught/Ziegele 2012, S. 164 f.*). Ferner gilt auch für U-Multirank, und dies nicht nur für eine Ausweitung auf andere Disziplinen, dass es sich hierbei um ein stetig weiterzuentwickelndes Projekt handelt, das sich vor allem an die neuen Anforderungen im Hochschulbereich und die zunehmenden Möglichkeiten der Datengenerierung anpasst.

6 Fazit und Ausblick

Transparenz im europäischen und weltweiten Hochschulsystem ist eine wichtige Voraussetzung für Mobilität, präferenzgerechte Angebote, Wettbewerb und Teilhabe an Hochschulbildung. Dafür bedarf es innovativer Indikatorensysteme, die weit über traditionelle internationale Hochschulrankings hinausgehen. „Multi-dimensional, user-driven“ Klassifikations- und Rankingsysteme können hier einen wichtigen Beitrag leisten. Dies hat die Europäische Kommission erkannt und die Projekte „U-Map“ und „U-Multirank“ initiiert. Bis diese Systeme ihre volle Funktionsfähigkeit erreichen, sind allerdings noch etliche Detailprobleme zu lösen. Die Systeme haben einen besonderen Spagat zu leisten: Sie müssen für ihre Zielgruppe, insbesondere Studieninteressenten, die in Bezug auf Hochschulen zunächst „Laiennutzer“ sind, verständliche Information bieten. Gleichzeitig müssen sie die Hochschulen überzeugen, dass die bereitgestellte Information hinreichenden Nutzen für Hochschulen erzeugt, um die Kosten der institutionellen Datenerhebung aufzuwiegen.

Im Dezember 2012 hat nach einer Ausschreibung des Projekts durch die EU-Kommission die Phase der Implementierung von U-Multirank begonnen, die sich auf zwei plus weitere zwei Jahre bezieht (*Europäische Kommission 2012*). Ein internationales Konsortium unter Führung des Centrum für Hochschulentwicklung (CHE) und des Center for Higher Education Policy Studies (CHEPS) soll die Fragen der operativen Einführung klären: die Implementierung der Erhebungsroutinen, die technische Entwicklung des Webtools, die Erarbeitung eines Business-Plans zur langfristigen Governance und Finanzierung von U-Multirank sowie die Entwicklung einer Kommunikationsstrategie, um die Nutzung und Beteiligung zu gewährleisten. Das alles erfolgt in enger Zusammenarbeit mit Stakeholder-Gruppen, zum Beispiel arbeiten die European Student Union und Hochschulnetzwerke (das Universities of Applied Sciences Network, die Conference of European Schools for Advanced Engineering Education and Research und das International Research Universities Network) als Konsortialpartner direkt an der Entwicklung mit. Ein erstes Ranking soll Anfang 2014 erscheinen. Datenqualität, Nutzen für die Hochschulen und Relevanz sowie Aufbereitung der Daten für die Nutzer werden die zentralen Erfolgsfaktoren der Implementierung sein.

Literatur

Callaert, Julie; Epping, Elisabeth; Federkeil, Gero; File, Jon; Jongbloed, Ben; Kaiser, Frans; Roessler, Isabel; Tijssen, Robert; van Vught, Frans; Ziegele, Frank (2012): The Pilot Test and Its Outcomes. In: van Vught u. a. (Hrsg.): *Multidimensional Ranking. The Design and Development of U-Multirank*. Heidelberg/London/New York, S. 135–166

Europäische Kommission (2011): Wachstum und Beschäftigung unterstützen – eine Agenda für die Modernisierung von Europas Hochschulsystemen. Mitteilung der Kommission an das Europäische Parlament, den Rat, den Europäischen Wirtschafts- und Sozialausschuss und den Ausschuss der Regionen. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2011:0567:FIN:DE:PDF> (Zugriff: 14. Januar 2013)

Europäische Kommission (2012): New ranking targets 500 universities. Press release. http://europa.eu/rapid/press-release_IP-12-1373_en.htm#PR_metaPressRelease_bottom (Zugriff: 14. Januar 2013)

Federkeil, Gero (2013): Internationale Hochschulrankings – Eine kritische Bestandsaufnahme. In: *Beiträge zur Hochschulforschung* 2/2013, S. 34–48

Federkeil, Gero; Kaiser, Frans; van Vught, Frans; Westerheijden, Don F. (2012): Background and Design. In: van Vught, Frans u. a. (Hrsg.): *Multidimensional Ranking. The Design and Development of U-Multirank*. Heidelberg/London/New York, S. 85–96

Federkeil, Gero; van Vught, Frans; Westerheijden, Don (2012): An Evaluation and Critique of Current Rankings. In: van Vught, Frans u. a. (Hrsg.): *Multidimensional Ranking. The Design and Development of U-Multirank*. Heidelberg/London/New York, S. 39–70

Hazelkorn, Ellen (2011): Rankings and the reshaping of higher education: The battle for world-class excellence. London

Holmes, Richard (2012): Power and responsibility – The growing influence of global rankings. <http://www.universityworldnews.com/article.php?story=20120630083447805> (Zugriff: 08. Januar 2013)

IREG Observatory on Academic Ranking and Excellence (IREG) (2006): Berlin Principles on Ranking of Higher Education Institutions. http://www.ireg-observatory.org/index.php?option=com_content&task=view&id=107&Itemid=99 (Zugriff: 09. Januar 2013)

Kaiser, Frans; van Vught, Frans (2009): The European Higher Education Classification: The Design Process. In: van Vught, Frans A. (Hrsg.): *Mapping the Higher Education Landscape. Towards a European Classification of Higher Education*. o.O., S. 87–103

Kaiser, Frans; Faber, Marike; Jongbloed, Ben; File, Jon; van Vught, Frans (o.J.): Implementing U-Map in Estonia. 2011 Case Study Report. <http://www.u-map.eu/estonia/UMapReportEstonia.pdf> (Zugriff: 08. Januar 2013)

King, Gary; Murray, Christopher J.L.; Salomon, Joshua A.; Tandon, Ajay (2004): Enhancing the Validity and Cross-Cultural Comparability of Measurement in Survey Research. In: *American Political Science Review* 98, 2004, 1, S. 191–207

Moed, Henk F. (2005): Citation analysis in research evaluation. Dordrecht

Rice, Nigel; Robone, Silvana; Smith, Peter C. (2010): International Comparison of Public Sector Performance: The Use of Anchoring Vignettes to adjust Self-Reported Data. In: *Evaluation* 16, 2010, 1, S. 83–101

Schmücker, Stefanie (2011): Universitätsprofile – Konzeption, Komponenten sowie empirische Umsetzung an deutschen Universitäten. München

Teichler, Ulrich (2007): Higher education systems. Conceptual frameworks, comparative perspectives, empirical findings. Rotterdam

U-Map (o.J.): U-Map Portugal Dissemination Seminar. 10 October 2011, Aveiro (Portugal). <http://www.u-map.eu/portugal/PortugalReportDisseminationSeminar.pdf> (Zugriff: 08. Januar 2013)

U-Map; Nordic Council of Ministers (Norden) (o.J.): U-Map Nordic Countries Dissemination Seminar. 22 October 2012, Copenhagen (Denmark). http://www.u-map.eu/nordic/Report%20Umap%20seminar_191112.pdf (Zugriff: 08. Januar 2013)

Usher, Alex; Medow, John (2009): A Global Survey of University Rankings and League Tables. In: Kehm, Barbara M. u. a. (Hrsg.): *University Rankings, Diversity, and the New Landscape of Higher Education*. Rotterdam/Boston/Taipei, S. 3–18

van Vught, Frans (2008): Mission diversity and reputation in higher education. In: *Higher Education Policy* 21, 2008, 2, S. 151–174

van Vught, Frans; Kaiser, Frans; File, Jon; Gaehtgens, Christiane; Peter, Rolf ; Westerbeijden, Don (2010): The European Classification of Higher Education Institutions. http://www.u-map.org/U-MAP_report.pdf (Zugriff: 08. Januar 2013)

van Vught, Frans; Ziegele, Frank (Hrsg.) (2011): Design and Testing the Feasibility of a Multidimensional Global University Ranking. Final Report. http://ec.europa.eu/education/higher-education/doc/multirank_en.pdf (Zugriff: 09. Januar 2013)

van Vught, Frans; Ziegele, Frank (Hrsg.) (2012): Multidimensional Ranking. The Design and Development of U-Multirank. Heidelberg/London/New York

Anlage 1: U-Map Dimensionen und Indikatoren

Teaching and learning profile	
1.	Orientation of degrees
2.	Subject areas covered
3.	Degree level focus
4.	Expenditure on teaching
Student profile	
1.	Mature or adult learners
2.	Students enrolled (headcount)
3.	Part-time students
4.	Students enrolled in distance learning programs
Research involvement	
1.	Peer reviewed academic publications
2.	Professional publications
3.	Doctorate production
4.	Expenditure on research
5.	Other research products
Regional engagement	
1.	First year bachelor students from the region
2.	Importance of local/regional income sources
3.	Graduates working in the region
Involvement in knowledge exchange	
1.	Cultural activities
2.	Income from knowledge exchange activities
3.	Patent applications filed
4.	Start up firms
International orientation	
1.	Foreign degree seeking students
2.	Importance of international sources of income
3.	Students sent out in European and other international exchange programs
4.	Incoming students in European and other international exchange programs
5.	Non-national teaching and research staff

Anlage 2: U-Multirank Dimensionen und Indikatoren

Teaching & Learning		
Institutional Ranking	1.	Expenditure on teaching
	2.	Graduation rate
	3.	Interdisciplinarity of programs
	4.	Relative rate of graduate (un)employment
	5.	Time to degree
Field-based Ranking	1.	Student-staff ratio
	2.	Graduation rate
	3.	Percentage graduating within norm period
	4.	Qualification of academic staff
	5.	Relative rate of graduate (un)employment
	6.	Interdisciplinarity of programs
	7.	Gender balance
	8.	Inclusion of work experience
Field-based Ranking: Student satisfaction indicators	1.	Student satisfaction: Overall judgment
	2.	Student satisfaction: Evaluation of teaching
	3.	Student satisfaction: Inclusion of work experience
	4.	Student satisfaction: Organisation of program
	5.	Student satisfaction: Libraries
	6.	Student satisfaction: Laboratories
	7.	Student satisfaction: Quality of courses
	8.	Student satisfaction: Social climate
	9.	Student satisfaction: Support by teachers
	10.	Student satisfaction: Computer facilities
Research		
Institutional Ranking	1.	Percentage of expenditure on research
	2.	Percentage of research income from competitive sources
	3.	Total publication output
	4.	Post-docs per FTE academic staff
	5.	Interdisciplinary research activities
	6.	Field-normalised citation rate
	7.	Highly cited research publications
	8.	Art-related outputs per FTE academic staff
Field-based Ranking	1.	Highly cited research publications
	2.	Field-normalized citation rate
	3.	External research income
	4.	Total publication output
	5.	Doctorate productivity
	6.	Student satisfaction: Research orientation of program
	7.	Post-docs per PhD completed

Knowledge Transfer		
Institutional Ranking	1.	Incentives for knowledge transfer
	2.	Percentage of income from third party funding
	3.	University-industry joint research publications
	4.	Patents awarded
	5.	Technology transfer office staff per FTE academic staff
	6.	CPD courses offered per FTE academic staff
	7.	Co-patenting
	8.	Start-ups per FTE academic staff
Field-based Ranking	1.	Academic staff with work experience outside higher education
	2.	Joint research contracts with private sector
	3.	University-industry joint research publications
International orientation		
Institutional Ranking	1.	Percentage of programs in foreign language
	2.	Percentage of international academic staff
	3.	International doctorate graduation rate
	4.	International joint research publications
	5.	Percentage of students in international joint degrees
	6.	Percentage foreign degree-seeking students
	7.	Percentage students coming in on exchanges
	8.	Percentage students sent out on exchanges
Field-based Ranking	1.	Incoming and outgoing students
	2.	International orientation of programs
	3.	International academic staff
	4.	International research grants
	5.	International joint research publications
	6.	Percentage of international students
	7.	International doctorate graduation rate
	8.	Student satisfaction: Opportunities to study abroad
Regional engagement		
Institutional Ranking	1.	Percentage of graduates working in the region
	2.	Percentage of income from regional sources
	3.	Regional joint research publications
	4.	Research contracts with regional partners
	5.	Percentage of students in internships in local enterprises
Field-based Ranking	1.	Degree theses in co-operation with regional enterprises
	2.	Graduates working in the region
	3.	Student internships in local enterprises
	4.	Regional joint research publications

Anschriften der Verfasser:

Professor Dr. Frank Ziegele
Verler Straße 6
33332 Gütersloh
E-Mail: frank.ziegele@che.de

Frans van Vught
Rue Montoyer 31
B-1000 Brüssel
E-Mail: f.a.vanvught@utwente.nl

Frank Ziegele ist Geschäftsführer des CHE Centrum für Hochschulentwicklung und Professor für Hochschul- und Wissenschaftsmanagement an der Hochschule Osnabrück.

Frans van Vught ist High Level Policy Advisor für die Europäische Kommission und Präsident des European Centre for Strategic Management of Universities (ESMU), zuvor war er Rektor und Präsident der Universität Twente in den Niederlanden.