

Wahrheit und Klarheit

Uneinheitliche Namen von Universitäten und ihre Folgen für die Forschungsevaluation

| LUTZ BORNMANN | WERNER MARX |

MATTHIAS WINTERHAGER | Die Bewertung von Forschungsleistungen durch Universitätsrankings ist inzwischen weit verbreitet. Die Ergebnisse von Rankings können jedoch maßgeblich durch uneinheitliche (und inkorrekte) Autorenadressen bzw. Namen von Universitäten beeinflusst werden. In diesem Beitrag werden Vorschläge unterbreitet, wie diesem Problem begegnet werden kann.

Universitäten, wie auch andere Forschungseinrichtungen, sehen sich zunehmend mit öffentlich verbreiteten bibliometrischen Analysen konfrontiert, die sie nicht selbst initiiert haben und bei deren Durchführung sie nicht beteiligt sind. Die Ergebnisse solcher Analysen, insbesondere wenn sie in Form von Ranglisten wie im Leiden-Ranking, dem Shanghai-Ranking oder dem THES-Ranking publiziert werden, können für die betroffenen Institutionen erhebliche forschungspolitische Wirkung entfalten. Den Rankings liegen (wie in vielen bibliometrischen Analysen) Daten aus dem Web of Science (WoS, Thomson Reuters) zugrunde. Die Daten bestehen aus einem Publikationsset, in dem zu jeder Publikation der Titel der Arbeit, die Zeitschrift, die Adressen aller Autoren und weitere Informationen verzeichnet sind. Um für ein Ranking die Ergebnisse zum Publikations-Output und Citation Impact der Universitäten erzeugen zu können, ordnen die Produzenten der Ran-

kings jede Publikation anhand der Angaben im Adressfeld der Datenbank den Universitäten zu.

Nicht erkannte Varianten

Ein Problem bei der Zuordnung der Publikationen zu den Universitäten (oder auch anderen Forschungseinrichtungen) besteht darin, dass die Autorenadressen und insbesondere die darin

»Die Produzenten der Rankings ordnen jede Publikation anhand der Angaben im Adressfeld der Datenbank den Universitäten zu.«

enthaltenden Namen der Universitäten in ganz unterschiedlichen Varianten im Adressfeld der Datenbanken verzeichnet sind. Hauptursache hierfür sind die sehr unterschiedlichen (und teilweise fehlerhaften) Angaben der institutionellen Namen und Adressen auf den Veröffentlichungen, die von den Mitarbeitern einer Universität publiziert werden. Da diese Varianten nur zum geringen Teil von Thomson Reuters (und anderen

Datenbankherstellern) erkannt und für das WoS vereinheitlicht werden, kann die korrekte und einigermaßen vollständige Zusammenführung der Arbeiten einer Universität für bibliometrische Studien sehr aufwändig sein. Die Zusammenführung setzt genaue Kenntnisse über die Universität (Namensänderungen, Namensabkürzungen, Alternativenamen, Außenstellen usw.) voraus. Besonders fehleranfällig sind in Deutschland Standorte, an denen mehrere Universitäten und eine Vielzahl außeruniversitärer Forschungseinrichtungen angesiedelt sind (z.B. Berlin, Hamburg und München). Die Rankingproduzenten weisen manchmal in Methodenberichten auf die Schwierigkeiten bei der Adresszuordnung hin.

In der nachfolgenden Tabelle sind beispielhaft die zwölf häufigsten Varianten des Namens der Eidgenössischen Technischen Hochschule Zürich (abgekürzt als ETH Zürich oder auch ETHZ) im Suchfeld der Autorenadressen des WoS auf der Basis der seit Anfang 2000 erschienenen Publikationen aufgelistet. Es handelt sich um die von Thomson Reuters erzeugten Kurzfassungen der Namen. Der häufigste Eintrag „ETH“ resultiert beispielsweise aus im WoS abgelegten Autorenadressen, wie der folgenden: „ETH, Inst Theoret Comp Sci, CH-8092 Zurich, Switzerland“. Die entsprechende Originaladresse auf der Publikation lautet: „Inst. of Theoretical Computer Science, ETH Zürich, Uni-

AUTOREN

Dr. Lutz Bornmann (li.) arbeitet im Stabsreferat für Wissenschafts- und Innovationsforschung sowie Forschungsanalyse in der Generalverwaltung der Max-Planck-Gesellschaft in München.

Dr. Werner Marx (re.) ist zuständig für den Zentralen Informationsservice für die Max-Planck-Institute der Chemisch-Physikalisch-Technischen Sektion am Max-Planck-Institut für Festkörperforschung in Stuttgart.

Matthias Winterhager (ohne Abb.) ist wissenschaftlicher Mitarbeiter am Institut für Wissenschafts- und Technikforschung (IWT) der Universität Bielefeld. Er leitet dort das in dem Beitrag erwähnte Projekt zur Bereinigung deutscher Adressdaten, das im Rahmen des Kompetenzzentrums Bibliometrie für die deutsche Wissenschaft durchgeführt wird.



versitätstr. 6 – 8092 Zürich, Switzerland“. Wie die Tabelle zeigt, haben einige Autoren ETH und Universität Zürich innerhalb ein und derselben Adresse aufgeführt, z. B. „Institute for Biomedical Engineering, University and ETH Zurich, Switzerland“ (das Institut ist an beide Universitäten angebunden). Im WoS wird in solchen Fällen „UNIV & ETH“ als Name der Institution abgelegt. Die dritte und vierte Spalte in der Tabelle gibt jeweils die Anzahl und den Prozentsatz der Publikationen an. Die Tabelle zeigt, dass die ETH (wie die meisten Institutionen) mit einer Vielzahl von Varianten im WoS vertreten ist.

Aufwändige Datenbereinigung

Inwieweit für bibliometrische Analysen alle Adressvarianten einer Universität berücksichtigt werden, hängt vor allem vom Aufwand ab, mit dem die Datenbereinigung durchgeführt wird. Die nachträgliche Bereinigung der deutschen Adressdatensätze im WoS ist derzeit Gegenstand eines vom Bundesministerium für Bildung und Forschung geförderten Projekts. Ziel ist die Bereitstellung einer möglichst aktuellen und vollständigen Zuordnung des deutschen Publikationsoutputs in den international führenden Fachzeitschriften zu den relevanten Hauptinstitutionen der großen Sektoren (Hochschulen, Max-Planck-Gesellschaft, Helmholtz-Gemeinschaft Deutscher Forschungszentren, Fraunhofer-Gesellschaft, Leibniz-Gemeinschaft). Das Zuordnungssystem kann im Kern auf Relationen von Identifizieren zurückgeführt werden und ist somit anschlussfähig an andere Datenbestände. Es ermöglicht einen Austausch mit interessierten Universitäten (und anderen Forschungseinrichtungen) zu den jeweils sie selbst betreffenden Zuordnungsdaten, etwa zum Abgleich mit hauseigenen Publikationsverzeichnissen. Auf diese Weise können die betroffenen Institutionen (auf Wunsch) in den Prozess der Qualitätsverbesserung einbezogen werden. Eine derart aufwändige Aufbereitung von Adressdaten kann bei Rankings im globalen Maßstab kaum vorausgesetzt werden, da sie schon aus ökonomischen Gründen nicht ohne weiteres realisierbar ist.

Wegen der schiefen Verteilung von Zitierungen auf Publikationen (einige wenige Arbeiten tragen wesentlich und viele Arbeiten kaum zur Gesamtzahl der Zitierungen bei) kann das Fehlen von verhältnismäßig wenigen Publikationen mit weniger üblichen Adressvari-

Die zwölf häufigsten Varianten des Namens der Eidgenössischen Technischen Hochschule (ETH) Zürich im Adressfeld des WoS			
#	Namensvariante	# Papers	% Papers
1	ETH	21306	48.7
2	SWISS FED INST TECHNOL	10839	24.8
3	ETH HONGGERBERG	4071	9.3
4	ETH ZENTRUM	2827	6.5
5	ETH ZURICH	2745	6.3
6	ETHZ	334	0.8
7	SWISS FED INST TECHNOL ZURICH	220	0.5
8	SWISS FED INST TECHNOL ETH	98	0.2
9	UNIV ETH ZURICH	88	0.2
10	SWISS FED INST TECHNOL ZURICH ETHZ	62	0.1
11	SWISS FED INST TECHNOL ETHZ	47	0.1
12	ETH ZURICH HONGGERBERG	44	0.1

anten in einem Publikationsset, das im Rahmen eines Rankings für eine Universität zusammengestellt wurde, zu deutlichen Einbußen beim Citation Impact führen.

Einheitliche Vorgaben und Kontrolle wichtig

Unübliche oder fehlerhafte Adressen bzw. Namen von Universitäten, die früher lediglich ein Schönheitsfehler in einer für den Leser beigefügten Herkunftsinformation der Autoren waren, stellen also eine ernst zu nehmende Gefahr im Rahmen von Evaluierungen auf der Basis bibliometrischer Indikatoren

»Einige Institutionen haben für ihre Mitglieder einheitliche Regeln zur Angabe der Adressen in wissenschaftlichen Publikationen aufgestellt.«

(wie Rankings) dar. Dem versuchen einige Institutionen entgegenzuwirken, indem sie für ihre Mitglieder einheitliche Regeln zur Angabe der Adressen in wissenschaftlichen Publikationen aufstellen. Beispiele dafür sind die Max-Planck-Gesellschaft, die Universität Oslo und das Karolinska-Institut (Stockholm). Eine flächendeckende Durchsetzung solcher Regeln könnte erheblich dazu beitragen, Fehlurteile in bibliometrischen Analysen zu reduzieren. Jede Universität, die den eigenen Publikationsoutput in internationalen Literaturdatenbanken angemessen repräsentiert sehen will, sollte ihren Mitgliedern im Sinne einer *corporate identity* einheitliche Vorgaben für entsprechende deutsch- und englischsprachige Na-

mensvarianten machen. Die Empfehlung sollte sich durchaus auch auf die vollständige Adresse und vor allem nicht nur auf den Namen der Hauptinstitution (einer Universität) beziehen. Neben der Vorgabe von Namensvarianten ist auch deren effektive Umsetzung und Kontrolle innerhalb der Universitäten entscheidend. In einigen Universitäten werden zentrale Dienstleister aus dem eigenen Haus (z.B. die Bibliothek) in den Vorgang des Einreichens von Manuskripten bei Zeitschriften einbezogen. Dabei kann die routinemäßige Überprüfung der korrekten Adressangaben gemäß den vereinbarten Regeln oh-

ne weiteres mit berücksichtigt werden. Dies ist ein guter Ansatzzpunkt bereits im Vorfeld der Veröffentlichung, um die angemessene Sichtbarkeit der eigenen Institution bis in die Literaturdatenbanken hinein nachhaltig sicherzustellen.

Einheitliche Namen bzw. Adressen der Universitäten auf den Manuskripten der Mitarbeiter können maßgeblich dazu beitragen, dass die betreffende Universität bei Rankings bzw. ganz generell bei bibliometrischen Studien zur Forschungsbewertung der tatsächlichen Leistung entsprechend dargestellt wird.