

Zur Evaluation der Informatik mittels bibliometrischer Analyse

Friedemann Mattern

Nicht alles was zählt, kann gezählt werden, und nicht alles was gezählt werden kann, zählt!

Albert Einstein

Zur Bewertung von Forschungsinstitutionen oder einzelnen Wissenschaftlern werden zunehmend bibliometrische Analysen eingesetzt. Gerade die Informatik eignet sich jedoch kaum dazu, mit klassischen bibliometrischen Methoden, etwa basierend auf dem ISI Science Citation Index, analysiert zu werden; auch eine weitgehend automatische Auswertung von bibliographischen Quellen im WWW liefert im Einzelfall eher zweifelhafte Evaluationsergebnisse.

Unter einer *bibliometrischen Analyse* wird die statistische Auswertung wissenschaftlicher Publikationen verstanden, typischerweise zum Zweck, die wissenschaftliche Leistung von Institutionen, Forschungsgruppen oder gar einzelnen Forschern zu messen. Die Bibliometrie ist damit ein Mittel zur Evaluation von Forschungsleistung und somit auch zur Bewertung der Qualität von Wissenschaftlern. Solche Instrumente gewinnen an Bedeutung, da Politik und Öffentlichkeit verstärkt auf eine einfach handhabbare Leistungskontrolle der Forschung drängen und daher immer öfter gefordert wird, wissenschaftliche Qualität und Effizienz quantifizierbar zu machen. Ein Indiz in dieser Hinsicht stellt auch die zunehmende Popularität von bibliometrisch fundierten Ranglisten dar, die mittlerweile auch für die Informatik existieren [15].

1 Wozu Bibliometrie?

Die Gründe für den Trend hin zu mehr Qualitätskontrolle im Wissenschaftsbereich sind vielfältig. Sie reichen vom steigenden Rechtfertigungsdruck von Entscheidungsträgern (verbunden mit a priori positiv unterlegten Eigenschaften wie Transparenz, Nützlichkeit oder Leistungsorientierung) über die um sich greifende Managementmethodik der Zielvereinbarung (welche nachprüfbare Leistungsindikatoren zwingend nach sich zieht – „you can’t manage what you can’t measure“) bis hin zur Erkenntnis, dass im Zeitalter der Globalisierung eine Steuerungsfunktion nur aufgrund objektiver und möglichst breit vergleichbarer Daten wahrgenommen werden kann.

Auf seine sarkastisch-nette Art drückte Gunter Dueck dies kürzlich im Informatik-Spektrum so aus [3]: „Damit der Wert der Science einer Universität berechnet werden kann, braucht man Messzahlen, die möglichst objektiv erscheinen sollen, damit nicht über sie diskutiert werden muss, die leicht zu messen sein müssen ... und unter deren Joch die als faul angenommenen Wissenschaftler das Arbeiten lernen.“ Vielleicht steckt hinter der Evaluationseuphorie manchmal auch einfach nur Neugierde oder der Wunsch der Öffentlichkeit, in summarischer Form orientiert zu werden und Auskunft über die Produktivität und Nützlichkeit (oder gar den Sinn) einer schwer durchschaubaren und teuren Angelegenheit zu erhalten, und mitunter mag es sich beim Ruf nach bibliometrischen Verfahren auch lediglich um eine Verlegenheitslösung mangels anderer einfach anwendbarer Verfahren zur Qualitätsfeststellung handeln.

Friedemann Mattern

Institut für Informationssysteme, ETH Zürich, CH-8092 Zürich

Es ist natürlich eine Illusion, wissenschaftliche Leistung möglichst objektiv auf einer eindimensionalen Skala anordnen zu wollen, so dass Forscher X etwa die Qualität 4.3, Wissenschaftlerin Y dagegen 7.2 auf dem (nach oben offenen?) Qualitätsmaßstab bekäme. Dennoch kann man sich gelegentlich des Eindrucks nicht erwehren, dass dieser reduktionistische Wunsch insgeheim bei Rankingverfahren mitschwingt, so ausgeklügelt sie im Einzelfall auch sein mögen. Jedenfalls scheint es für Wissenschaftsbürokraten sehr verführerisch zu sein, die Leistung von Wissenschaftlern auch ohne Fachkenntnis beurteilen können zu wollen. Als Analysegegenstand bieten sich hierfür Publikationen als der am leichtesten zugängliche Output von Forschern unmittelbar an; außerdem stellt das Veröffentlichende ja eine wesentliche – und daher in einer Beurteilung in jedem Fall zu berücksichtigende – Tätigkeit eines Wissenschaftlers dar, da es sich dabei um die wichtigste Möglichkeit handelt, Forschungserkenntnisse weiterzugeben.

Die Zahl der Publikationen alleine taugt aber offensichtlich nicht für eine Evaluation, denn sie ist lediglich ein Maß für die Produktivität eines Autors, „für seinen Fleiß, seine Zielstrebigkeit, sowie seine Fähigkeit und Bereitschaft, seine Forschungsergebnisse zu präsentieren“ [9] – dies sind aber keine primären Qualitätseigenschaften von Wissenschaftlern. Entscheidender als die Zahl von Veröffentlichungen ist sicherlich deren Qualität. Ob ein wissenschaftlicher Aufsatz gut und originell ist, kann allerdings eigentlich nur ein Experte beurteilen, und auch dies stellt dann ein subjektives und oft keineswegs unstrittiges Werturteil dar. Strebt man eine möglichst „sachkenntnisfreie“ Bewertung an, müssen infolgedessen andere Kriterien als Approximation für inhaltliche Qualität herhalten, wie z.B. die Beachtung, die eine Arbeit in Fachkreisen erhält. Diese lässt sich etwa über den *impact factor* der veröffentlichenden Fachzeitschrift oder die Häufigkeit, mit der ein Autor zitiert wird, abschätzen. Die Zitationshäufigkeit wiederum lässt sich mehr oder weniger gut automatisch aus Literaturdatenbanken oder neuerdings auch mittels spezialisierter Suchmaschinen im WWW ermitteln.

Ausgehend von solchen Überlegungen wurden in den vergangenen Jahren verfeinerte bibliometrische Methoden entwickelt, die aus der lexikalischen und statistischen Auswertung von wissenschaftlichen Publikationen scheinbar „objektivierte“ Indikatoren zur Beurteilung der Forschungsleistung ermitteln. Diese Verfahren werden entsprechend vermarktet, da sich die Evaluation von Wissenschaft zu einem Geschäft entwickelt hat und Evaluationsagenturen dankbare Abnehmer von solchen Methoden sind, die gerade keine oder nur wenig fachspezifische Expertise voraussetzen. So preist etwa das *Centre for Science and Technology Studies* in Leiden [2] seine Methode folgendermaßen an: „After years of research, we succeeded in creating a world-wide unique bibliometric data-system. This data-system is not only a very large collection of the data of hundreds of millions published, citing, and cited publications. A long period of research on organisational and cognitive structures in science, characteristics of scientific communication, and on concepts of performance, perceptions of quality and evaluation, resulted in knowledge which is ‘encoded’ in additional information and computer algorithms.“

Leider ist dies aber nicht so einfach. Immerhin wird von den Methodenentwicklern zugestanden, dass es sich nur um ein durchaus mit Schwächen behaftetes Hilfsmittel zur Evaluation handelt, das Peer-Reviewing nicht ersetzt sondern höchstens ergänzen kann. So heißt es etwa in [2]: „In the application of our bibliometric method, the ‘primacy’ of peer review is fully recognised.“ Und in [12]: „Impact and scientific quality are by no means identical concepts ... A final quality judgement on a research group or programme can be given by peers only.“

Oft wird das Heil dann allerdings in der Kombination von Peer-Reviewing und bibliometrischen Analysen gesehen, oder die bibliometrischen Methoden werden zumindest als Hilfsmittel für menschliche Experten empfohlen. Die Tatsache, dass die Verfahren sehr differenziert betrachtet werden müssen, nur eine eingeschränkte Aussagekraft haben und mitunter (abhängig vom Fach oder der Datenlage) aus prinzipiellen Gründen völlig versagen können, wird von den Auftraggebern bibliometrischer Evaluationen allerdings kaum wahrgenommen – zu stark ist offenbar der Wunsch nach einer automatischen Leistungsbeurteilung ohne Fachkenntnis, so dass Kritik an den Verfahren schnell als Nörgelei am Detail abgetan wird und die

Meinung vertreten wird, dass bibliometrische Methoden jedenfalls besser als nichts seien und schon „irgendetwas“ aussagen würden.

2 Die Aussagekraft von Zitierhäufigkeiten

Über die Problematik, Publikationen als objektivierbare Kriterien für Leistung in der Forschung heranzuziehen, ist – etwa in Hinblick auf Selbstzitate und Zitierkartelle – schon viel geschrieben worden. Eine umfassende Studie „Bibliometric Profiles of Academic Physics Research in the Netherlands“ [12] stellt z.B. fest, dass (im Bereich der Physik) alleine die direkten Selbstzitate schon ca. 30 Prozent aller Referenzen ausmachen. Einige Kernaussagen zur Bedeutung von Publikationslisten und Zitierhäufigkeiten mögen daher an dieser Stelle genügen.

Die absolute Zahl von Publikationen innerhalb eines Zeitraumes scheidet, wie oben bereits dargelegt, als Kriterium jedenfalls aus: Rankings unterscheiden oft nicht zwischen Einzel- und Koautorenschaften und generell befinden sich unter den meistzitierten Wissenschaftlern jeder Disziplin regelmäßig solche mit wenigen, aber viel zitierten, und solche mit vielen, aber wenig zitierten Publikationen [4]. Klaus Fischer, Professor an der Universität Trier und ein Spezialist für Scientometrie, stellt allerdings auch das feinere Maß der *Zitierhäufigkeit* als Kriterium in Frage, da dies kein Gradmesser wissenschaftlicher Leistung, sondern allenfalls sozialer Anerkennung sei. Seiner Meinung nach belohnt dieser Indikator u.a. die Beteiligung an großen und „boomenden“ Forschungsfeldern sowie die Benutzung des Englischen als dominierende Wissenschaftssprache [4]. In einem von der Max-Planck-Gesellschaft veröffentlichten Artikel *Kann man Forschungsqualität messen?* [14] heißt es dazu: „In diesem Zusammenhang sollte streng unterschieden werden zwischen Resonanz oder Wirkung einerseits und Bedeutung oder Qualität andererseits. Letztere Begriffe sind wesentlich durch Interessen und Wertvorstellungen geprägt und als relative Kategorien nicht objektiv und quantitativ messbar.“

Dass die Zitierhäufigkeit dennoch im Sinne eines qualitativen Kriteriums große Beachtung findet, wird an der Popularität des *Science Citation Index* [6, 7, 23] vom *Institute for Scientific Information* (ISI) erkennbar – obwohl bereits der Erfinder dieses Indexes, Eugene Garfield, betont hat, dass damit nicht die *Qualität*, sondern nur die *Nützlichkeit* einer Arbeit für die laufende Forschung anderer gemessen wird [9]. Aus dem Index geht jedenfalls hervor, dass nur äußerst wenige Veröffentlichungen oft zitiert werden und sehr viele nur selten (1-10 Mal) oder überhaupt nicht. Eine genauere Analyse der Jahre 1981-1985 zeigt, dass fast 50% aller Arbeiten 5 Jahre nach dem Erscheinen kein einziges Mal von anderen Autoren zitiert wurden und nur ca. 20% mehr als ein Mal zitiert wurden [14]. Wenn jedoch die Arbeiten eines Autors überhaupt nicht zitiert werden, sind diese möglicherweise nicht besonders wichtig – oder aber ihrer Zeit voraus, wie Kutzelnigg in seinem Aufsatz *Kann man wissenschaftliche Leistung messen?* [9] süffisant bemerkt. Wie in [14] dargelegt wird, erreichen hochzitierte Publikationen auch häufig erst nach einigen Jahren oder gar Jahrzehnten das Maximum der Zitierungen. Die Autoren folgern daher: „Herausragende Publikationen können dementsprechend kaum bereits einige Jahre nach ihrer Entstehung mittels Zitierungsanalyse entdeckt werden. Dazu bedarf es schon Experten mit außergewöhnlichem Sachverstand.“

Kutzelnigg [9] führt noch einige weitere bedenkenswerte Aspekte an. So unterscheiden die meisten Auswertungen von Zitierhäufigkeiten nicht zwischen Selbstziten und Fremdziten; nur Letztere sind aber ein Maß für die Beachtung, die ein Autor tatsächlich findet. Weiterhin werden bekannte Autoren ohne Rücksicht auf die Bedeutung der betreffenden Arbeit häufiger zitiert als unbekannte: wer viele Freunde hat, schneidet besser ab als ein Einzelgänger. Schliesslich werden zum Vorteil bereits etablierter Autoren Zitate gelegentlich einfach unbelesen aus anderen Quellen übernommen – offenbar eine alte Technik, denn Cervantes spottete darüber bereits 1605 in seinem *Don Quijote*: „Kommen wir nun zu den Zitaten. Ihr habt nichts weiter zu tun als ein Buch herbeizusuchen, das sie alle bereits angeführt hat. Nun wohl, dies Verzeichnis setzt Ihr in Euer Buch; vielleicht gibt's immerhin einen Einfältigen, der

glaubt, Ihr hättet die ganzen Autoren in Eurem Werk benutzt! Und wenn auch zu weiter nichts, so wird jener große Autorenkatalog wenigstens dazu dienen, Eurem Buch Ansehen zu verschaffen!“

Arbeiten viele Wissenschaftler auf einem Gebiet, werden wichtige Arbeiten naturgemäß öfter zitiert, als wenn die Thematik nur einen kleinen Kreis beschäftigt – wodurch ausgelagte Gebiete bevorzugt, innovative Forschungsthemen aber benachteiligt werden. Oft werden, ganz im Interesse der Leser, Übersichtsartikel statt Originalarbeiten zitiert – auch dadurch werden innovative Forscher eher benachteiligt. Schließlich muss, wenn eine Arbeit zitiert wird, das nicht unbedingt in einem anerkennenden Sinn geschehen: auch wenn man eine Arbeit kritisiert oder zerpfückt, zitiert man sie. Kutzelnigg schreibt weiter: „Eine Rangordnung, basierend auf Zahl der Publikationen oder Zahl der Zitate, sollte man als ... ein Beispiel dafür [ansehen], dass das, was technisch möglich ist (zumal wenn es im wirtschaftlichen Interesse des ‚Institute for Scientific Information‘ liegt), auch gemacht wird.“

Um überhaupt zitiert zu werden, muss man Autor einer Veröffentlichung sein. Wer aber zählt als Autor? Über Recht und Sitte, als Autor einer Arbeit genannt zu werden, ist schon viel geschrieben und gestritten worden, bekanntlich gelten je nach Fach unterschiedliche Traditionen hinsichtlich Erstautorschaft und Ehrenautorschaft. Gravierender allerdings scheint, dass dies auch innerhalb eines Faches stark streuen kann. So hat etwa die Deutsche Forschungsgemeinschaft (DFG) festgelegt: „Als Autoren einer wissenschaftlichen Originalveröffentlichung sollen alle diejenigen, aber auch nur diejenigen, firmieren, die ... zur Formulierung des Manuskripts selbst wesentlich beigetragen ... haben ... Mit dieser Definition von Autorschaft werden andere – auch wesentliche – Beiträge wie Verantwortung für die Einwerbung der Förderungsmittel, Unterweisung von Mitautoren, ... Leitung einer Institution oder Organisationseinheit, in der die Publikation entstanden ist, für sich allein nicht als hinreichend erachtet, Autorschaft zu rechtfertigen. Eine Ehrenautorschaft ist ... keinesfalls akzeptabel.“ Dagegen wird in der Stellungnahme *Das Namensnennungsrecht bei wissenschaftlichen Veröffentlichungen* der Rechtsabteilung der ETH Zürich die Ehrenautorschaft sogar explizit geregelt: „An erster Stelle der Hauptautor, an letzter Stelle der Professor oder der so genannte group leader“. Nun bedeutet dies in der Praxis natürlich keineswegs, dass an der ETH immer nach der einen Regel und bei DFG-Projekten immer nach der anderen Regel verfahren wird. Es zeigt aber, dass je nach Verfahrensweise (oder besser: Redlichkeit) ein Wissenschaftler weitaus mehr Veröffentlichungen und Zitate als ein gleich produktiver anderer Wissenschaftler haben kann; eine Vergleichbarkeit auch innerhalb eines Faches ist also keineswegs gegeben.

Ist schon die Frage, wann jemand als Autor gilt, im Grunde genommen offen, so ergibt sich ein mindestens ebenso großer Unsicherheitsfaktor bei der Frage, was eine „richtige“ Publikation ist und ob man bei einer bibliometrischen Analyse aufgrund der Datenlage garantieren kann, die relevanten Zitate in gerechter Weise zu erfassen. Gerade beim Fach Informatik sind hier Zweifel angebracht, wie nachfolgend skizziert wird.

3 Die Besonderheiten der Informatik

Die Informatik ist eine Querschnittswissenschaft, die außergewöhnlich heterogen in ihrer Fachkultur ist. Nahe der Mathematik stehende und ihrer Tradition verbundene Teilgebiete sind ebenso vertreten wie ingenieurwissenschaftlich geprägte oder der Wirtschaftswissenschaft nahe stehende Teildisziplinen. Dies impliziert teilweise ganz unterschiedliche Wertvorstellungen hinsichtlich Zitierweisen, der Bedeutung von Fachjournalen relativ zu Konferenzbeiträgen oder der Halbwertszeit des Wissens und damit sowohl des *impact factor* als auch der Aktualität (und damit der Zitierwahrscheinlichkeit) von Aufsätzen. Die Ergebnisse anderer Fächer lassen sich nicht ohne weiteres auf die Informatik als Ganzes übertragen.

Dass hinsichtlich der Zitationshäufigkeit eine Vergleichbarkeit selbst innerhalb eines Faches kaum gegeben ist, stellt schon die empirische Untersuchung [12] fest: „Large differences exist in citation practices among the various subfields of physics“ – dies dürfte umso mehr für die

Informatik gelten. So mag es im Bereich der Physik aufgrund ihres universellen Gültigkeitsanspruches i.A. wenig sinnvoll sein, in einer anderen Sprache als dem Englischen zu publizieren. Dagegen hat z.B. die Wirtschaftsinformatik als anwendungsorientierte Disziplin eine starke regionale Komponente, so dass hier landessprachliche Publikationen (die dann allerdings in der Regel eine geringere internationale Visibilität besitzen und folglich wesentlich seltener zitiert werden) zumindest im Sinne einer unmittelbaren Wirkungsentfaltung von Vorteil sein können.

Typisch für die Informatik ist das Lösen von Problemen – sei es durch die Veröffentlichung neuer Algorithmen und Lösungsverfahren oder durch die Angabe geeigneter Systemarchitekturen oder gar Prototypimplementierungen. Schwierige Informatik-Probleme kann man aber fast immer auf unterschiedliche Art, und damit gut oder weniger gut lösen. Eine naive oder eher schlechte Problemlösung, die bei vielen den Eindruck hervorruft, es besser zu können, führt dann fast regelmäßig zu vielen Zitaten der Arbeit. Hingegen wird eine gute Arbeit, die ein Problem abschließend löst und keine Verbesserungen mehr ermöglicht, eher selten zitiert, wenn sie nicht gerade einen wissenschaftlichen Durchbruch darstellt. Dieses Phänomen lässt sich in der Informatik jedenfalls öfter beobachten und invertiert geradezu die Bedeutung der Zitationshäufigkeit für die Qualität einer Arbeit!

Auch die große Bandbreite zwischen Theorie und Praxis ist eine hier relevante Eigenschaft der Informatik. Die primären Ergebnisse eines Theoretikers mögen z.B. Veröffentlichungen sein, die den fortwährenden Erkenntnisgewinn dokumentieren, er produziert dementsprechend relativ viel. Einem Praktiker mag es dagegen wichtig sein, größere Projekte durchzuführen und Prototypen zu entwickeln, um damit z.B. zum Wissenstransfer beizutragen und faktische Erfahrungen zu sammeln – da steht eine Veröffentlichung vielleicht erst ganz am Ende und ist nicht das Ziel an sich. Der bekannte Informatiker Roger Needham, Direktor des Microsoft Research Laboratory in Cambridge, schrieb zum Problem der Wahrnehmung praktischer Informatik einmal Folgendes: „Although for most of my career I was a practical builder of systems, the things I’m best known for are two papers, both of a theoretical nature and both done when I was on sabbatical leave. So you can work on a complicated system for seven years, and nobody remembers that“ [17]. An sich würde man sich wünschen, dass praktische Informatiker und Software-Architekten vor allem nach den Systemen, die sie „schreiben“ und verbreiten, beurteilt werden (schließlich werden „normale“ Architekten ja auch an ihren Bauwerken gemessen!), das scheint jedoch kaum durchführbar zu sein.

Naturgemäß werden in der Informatik schon seit längerem viele (aber längst nicht alle) Artikel zusätzlich in elektronischer Weise publiziert. Dies hat gleich in mehrerlei Hinsicht Konsequenzen. Zum einen verlieren dadurch klassische Publikationsmedien, insbesondere Fachjournale, stark an Bedeutung, da sie, abhängig von der Teildisziplin, teilweise nur noch als „Write only-Medium“ für Dokumentationszwecke genutzt werden. Zum anderen werden elektronisch verfügbare Artikel wesentlich häufiger zitiert, wie Steve Lawrence bei einer umfassenden Analyse herausfand: „The results are dramatic. There is a clear correlation between the number of times an article is cited, and the probability that the article is online. The mean number of citations to offline articles is 2.74, and the mean number of citations to online articles is 7.03“ [11].

Diese vom *NEC Research Institute* durchgeführte Analyse ist insofern bemerkenswert, als dass sie auf einer einzigartigen Datensammlung von informatikbezogenen Veröffentlichungen beruht, die Autoren über ihre Homepage im Web öffentlich zugänglich gemacht haben und die mit einer speziellen Suchmaschine aufgespürt wurden [8, 18]. In [11] heißt es dazu: „Research Index gathers fragmented scientific resources from around the web, and automatically organizes them within a citation index ... The prototype form of Research Index is being applied to the computer sciences. Its archive of papers in this subject alone, at 270000 articles, is bigger than leading online scientific archives ... The engine already has an enthusiastic following among computer scientists.“ Im April 2001 enthielt der Index außer den genannten ca. 270000 Artikeln insgesamt über eine halbe Million Autoren und über 3 Millionen automa-

tisch ermittelte Zitate und stellt damit einen außergewöhnlichen Fundus sowohl für die Informatik selbst als auch für die bibliometrische Forschung dar.

Selbst wenn der *Research Index* (der auch unter dem Namen *CiteSeer* bekannt wurde) für das Gebiet der Informatik umfassender als beispielsweise der bekannte *ISI Science Citation Index* ist, muss man genau hinsehen, wenn man die Ergebnisse interpretieren will. Leider sind nur wenige Aspekte über die Suchmaschinen und das Analyseverfahren des automatischen Zitierungsindex öffentlich bekannt [10]. Hinsichtlich derjenigen Teildisziplinen, die ihre Publikationen traditionell im Postscript- oder pdf-Format im WWW bereitstellen, erscheint der Index allerdings relativ vollständig, auch wenn man beachten muss, dass die Bedeutung einer Veröffentlichung nicht (etwa entsprechend dem *impact factor* einer Zeitschrift) gewichtet wird und sich auch technische Berichte und andere nicht durch ein Herausgebergremium gefilterte Aufsätze auf solchen WWW-Seiten finden. Ferner ist aufgrund unterschiedlicher Fachkulturen und der regional unterschiedlich schnellen Adaption und Nutzung des WWW die theoretische Informatik einerseits und die US-amerikanische Informatik andererseits stark überbewertet. Neben anderen kultur- und länderspezifischen Faktoren (wie Sprache, Dominanz von Herausgebergremien, Relevanz der Sichtbarkeit für die eigene Karriere etc.) ist dies vermutlich der Grund dafür, dass sich im Jahr 2000 unter den 1000 meistzitierten Autoren weniger als 10 Deutsche identifizieren ließen, was innerhalb der Gesellschaft für Informatik für eine gewisse Irritation gesorgt hat.

Dass die automatische Analyse des *Research Index* ihre Grenzen hat, äußert sich u.a. darin, dass es Probleme mit Namen gibt, die Umlaute enthalten [15], und dass in der Hitliste der *most cited authors in Computer Science* [19] der „Autor“ *P. Thesis* immerhin den Rang 383 einnimmt! Offenbar lassen sich mit rein lexikalischen Methoden die Autoren nicht immer aus den in unterschiedlichem Format gespeicherten elektronischen Dokumenten eindeutig ermitteln, und da bei der Identifikation keine Vornamen sondern lediglich Namenskürzel verwendet werden, sind nicht nur Aussagen über *J. Smith* oder *P. Brown* unmöglich, sondern tatsächlich zu den meisten Einzelpersonen mit „gängigen“ Nachnamen sinnlos – so haben beispielsweise die geschätzten Kolleg(innen) *M. Gross* und *D. Wagner* alleine in der bekannten DBLP-Datenbank zur Informatik-Fachliteratur [13, 21] jeweils 6 Namensvettern.

Es ist allgemein anerkannt, dass Konferenzbeiträge in der Informatik einen hohen Stellenwert genießen. Oft wird ihnen, im Unterschied zu manch anderen Fächern wie z.B. der Physik, sogar eine größere Bedeutung zugemessen als Artikeln in Fachzeitschriften. Die Analyse des *Research Index* [16] bestätigt dies in eindrucksvoller Weise: Danach befindet sich unter den nach Zitationshäufigkeit sortierten ersten zehn Informatik-Publikationsorganen z.B. nur eine einzige Zeitschrift (*ACM Transactions on Computer Systems*), alles andere sind Konferenzen. In einer begleitenden Untersuchung [8] heißt es dazu wörtlich: „We find that conference papers and technical reports play a very important role in computer science research, especially regarding access to the very latest research.“ Der *Research Index* selbst besteht immerhin zu 49% aus Konferenzbeiträgen, zu 23% aus Zeitschriftenartikeln und zu 8% aus Büchern und Buchbeiträgen (und zu 20% aus sonstigem, z.B. technischen Berichten). Allerdings besteht bei den Zitaten naturgemäß die Neigung zu den „stabileren“ Referenzen: bei 41% aller Zitate handelt es sich um Zeitschriftenartikel, bei 19% um Bücher und Buchbeiträge und nur bei 28% um Konferenzbeiträge.

Die Rangliste [16] enthüllt übrigens auch, dass einige Zeitschriften mit klangvollem Namen (wie z.B. *Wirtschaftsinformatik*, *International Journal of Network Management*, *Journal of Visual Languages and Computing*) praktisch kaum zitierte Artikel enthalten, wenn man die Selbstzitate von Autoren eliminiert. Tatsächlich sind in vielen (aber nicht allen!) Teildisziplinen der Informatik Veröffentlichungen in Zeitschriften in den letzten zehn bis zwanzig Jahren tendenziell eher unwichtiger geworden. Die Beobachtung zeigt jedenfalls, dass die Anzahl der Einreichungen bei vielen klassischen Fachjournals zurückgeht, die Auswahlrate daher oft nur noch 1 aus 2 beträgt. Folglich sind Zeitschriftenaufsätze in der Informatik, bis auf wenige aber wichtige Ausnahmen, als generelles Qualitätskriterium nicht mehr gut geeignet.

Aus nahe liegenden Gründen beharren indes kommerzielle Bibliometriker dennoch auf der weitgehend alleinigen Berücksichtigung von Zeitschriften, da nur diese in traditioneller Weise bibliographisch in Datenbanken wie dem *ISI Science Citation Index* erfasst sind und damit eine weitestgehend automatisierte (und damit fachkenntnisarme und billige) Analyse ermöglichen, obwohl sie sich der Problematik zumindest teilweise bewusst sind. So heißt es zu den nicht vom Science Citation Index erfassten Publikationen, die auch die Konferenzberichte umfassen, etwa bei [2]: „It is our experience that these publications may play an important role in the international impact of a research group. We expect that this will be to a lesser extent the case for most natural and life sciences, and more important for mathematics and biology, as well as for engineering, social and behavioural sciences, and for the humanities.“

Die Berücksichtigung von Tagungen und Workshops ist in der Informatik jedenfalls unerlässlich. Die Tatsache, dass viele wichtige Konferenzen streng begutachtet werden und mit einem Annahme / Ablehnungsverhältnis von bis zu 1 zu 10 sogar kompetitiver als die meisten Fachzeitschriften sind (selbst wenn man einkalkuliert, dass die Einreichungsschwelle bei Konferenzen im Allgemeinen niedriger liegt), bedeutet, dass eine Beschränkung einer bibliometrischen Analyse auf Fachzeitschriften – was in anderen Fächern, wie etwa der Physik, adäquat sein kann – für die Informatik ein völlig verzerrtes Bild liefert. Eine solche Analyse hat kaum einen Aussagewert und kann unter Umständen sogar „schlechte“ Informatik positiv darstellen und umgekehrt!

Natürlich existieren auch in der Informatik viele Konferenzen, die nahezu alles, was eingereicht wird, annehmen. Zudem werden Konferenzbeiträge schon aus Zeitgründen in der Regel nicht so gründlich begutachtet wie Zeitschriftenartikel und oft wird bei Konferenzen wissenschaftliche Inzucht betrieben, indem dem Programmkomitee nahe stehende Autoren bevorzugt werden. Nur „Insider“ können daher entscheiden, ob es sich bei einer Konferenz oder einem Workshop um eine qualitativ hochwertige Veranstaltung handelt oder nicht. Eine Qualitätsbeurteilung ohne Fachkenntnis ist bei Informatik-Veröffentlichungen jedenfalls zum Scheitern verurteilt.

4 Die Problematik bibliometrischer Datenquellen und des „impact factor“

Den meisten bibliometrischen Analysen liegt die Datenbank des *Science Citation Index* (SCI) zugrunde. Dieser Index wird vom *Institute for Scientific Information* in Philadelphia herausgegeben und berücksichtigt etwa 3500 Fachzeitschriften mit Schwerpunkt Natur- und Lebenswissenschaften, jedoch keine Konferenzberichte oder Buchbeiträge.

Es ist allgemein bekannt, dass der SCI schlecht für die Mathematik und nur sehr eingeschränkt brauchbar für die Informatik ist. So relativiert z.B. [12] die Bedeutung des SCI: „In general, SCI coverage of scientific journals is less adequate in the more technical or applied fields of science than in basic natural and life science disciplines. Furthermore, proceedings of international conferences and research reports are more important in applied fields than they are in basic disciplines ... In some programmes, the most important way of publishing may not be in refereed journals, but in conference proceedings or other media. For instance, in some application- and problem-oriented fields proceedings, books and reports play a more important role in the communication of research results than elsewhere. Moreover, in specific subfields, especially in the technical and applied sciences, the coverage of the Science Citation Index is less good. Several technical or applied journals are not included as source journals in the SCI.“

Letzteres lässt sich für die Informatik bestätigen. Beispielsweise ist vom renommierten Springer-Verlag seit Mitte der 1980er-Jahre herausgegebene Zeitschrift *Distributed Computing* (ISSN 0178-2770) nicht im SCI, obwohl sie entsprechend den Daten des *Research Index* hinsichtlich der Zitate Position 57 von nahezu 1000 betrachteten Konferenzen und Zeitschriften einnimmt und in dieser Hinsicht zu den wichtigsten Fachjournalen zählt [16]. Viele weitere solche Beispiele lassen sich angeben.

Dies bedeutet, dass nicht nur die für die Informatik unerlässlichen Konferenzberichte fehlen, sondern es sind auch entscheidende und möglicherweise ganze Teilgebiete prägende Fachzeitschriften nicht berücksichtigt. Einige Informatik-Teildisziplinen schneiden damit bei auf dem SCI beruhenden bibliometrischen Analysen notorisch schlecht ab und ob ein Autor gut oder schlecht wekommt, hängt nicht davon ab, ob er eine fachlich adäquate oder renommierte Zeitschrift gewählt hat oder nicht, sondern ob es sich dabei um eine SCI-Zeitschrift handelt. Wie unter diesen Bedingungen bibliometrische Analysemethoden einen Vergleich der Forschungsaktivitäten einer Einheit mit dem internationalen Durchschnitt im Forschungsgebiet erlauben sollen, wie dies beispielsweise vom *Centre for Science and Technology Studies* behauptet wird („whether or not the performance of the group or institute is above or below the western world-wide average of its own research field“ [2]), bleibt jedenfalls schleierhaft.

Auch wenn Analysen (wie z.B. die oben erwähnte *Bibliometric Profiles of Academic Physics Research in the Netherlands*) einen Disclaimer der Art „we stress that the indicators presented in this study do not give information on the part of the results which are published in other media than the journals included in the Science Citation Index“ [12] enthalten, besteht die Gefahr, dass von den Auftraggebern und Nutzern einer Studie die weitreichenden Konsequenzen daraus, die eine solche Analyse unter Umständen wertlos oder bei leichtfertiger Interpretation sogar schädlich werden lassen, praktisch nicht gesehen werden.

Eine weitgehend manuell gepflegte Datenbasis der Informatik-Fachliteratur stellt die *DBLP-Bibliographie* dar [13, 21]. Dieser ursprünglich auf die Themen Datenbanksysteme und Logikprogrammierung ausgerichtete Literaturindex wurde im Laufe der Zeit auf weitere Teilgebiete der Informatik ausgedehnt und wies im März 2001 ca. 200000 Artikel nach. Allerdings sind einige Teilgebiete, wie etwa die technische Informatik, unterrepräsentiert, auch sonst erscheint er stellenweise lückenhaft.

Die Vollständigkeit von Bibliographien und Zitationsindizes wird spätestens dann relevant, wenn diese zur Erstellung von Ranglisten herangezogen werden, wie etwa bei der aufschlussreichen Analyse von Nebel [15]. Dort wird versucht, Indikatoren zur wissenschaftlichen Produktivität von universitären Informatik-Instituten aus frei verfügbaren WWW-Quellen wie dem *Research Index* und der *DBLP-Bibliographie* zu bestimmen. Als Reaktion darauf schrieb ein Leser allerdings: „Eine Suche in der im Artikel angegebenen *DBLP-Datenbasis* ergab in meinem Fall 9 ziemlich willkürliche Einträge (statt mittlerweile 80 referierte Papers und 5 Bücher).“ Auch eine Stichprobe des Autors ergab, dass seine mit Abstand am häufigsten zitierte Arbeit (im *Research Index* Rang 59 von ca. 270000 [20]) in der *DBLP-Datenbasis* überhaupt nicht auftaucht. Es ist zu vermuten, dass dies keine Einzelfälle sind und der *DBLP-Bibliographie* daher zumindest bei der Bewertung oder dem Vergleich einzelner Personen nicht blind vertraut werden darf, da unter Umständen ein völlig schiefes Bild entstehen könnte.

So sinnvoll fachbezogene Bibliographien [21] und Zitationsindizes [18] als Hilfsmittel zur Fachinformation auch sein mögen, so bergen sie bei der Beurteilung von einzelnen Personen gerade deswegen große Gefahren: Jemand, der (vielleicht nur aus organisatorischen Gründen) einem zu evaluierenden Informatik-Institut angehört, aber nicht zentral innerhalb der Informatik publiziert, erhält damit zwangsläufig schlechte Publikations- und Zitierungswerte [15].

Oft wird zur Beurteilung von Veröffentlichungen und Autoren auch der so genannte *impact factor* herangezogen: „The impact factor has moved in recent years from an obscure bibliometric indicator to become the chief quantitative measure of the quality of a journal, its research papers, the researchers who wrote those papers“ [1]. Bei diesem *ISI journal citation reports impact factor* handelt es sich um eine Maßzahl für eine Fachzeitschrift, die ausdrückt, wie viele Zitate jeder Artikel der Zeitschrift im Durchschnitt jährlich im zweiten und dritten Jahr nach seinem Erscheinen erhält. Diese Maßzahl ist allerdings wieder fachspezifisch und eignet sich daher nicht als fachübergreifende Maßeinheit. Amin und Mabe [1] schreiben dazu: „Fundamental and pure subject areas have higher average impact factors than specialized or

applied ones. The variation is so significant that the top journal in one field may have an impact factor lower than the bottom journal in another area.“

Nach ihren Untersuchungen liegt der mittlere *impact factor* von Zeitschriften im Bereich der Mathematik und Informatik bei ca. 0.5, derjenige von Physik-Zeitschriften ist dagegen mehr als 3 Mal so hoch, im Bereich Biowissenschaften ist der mittlere *impact factor* sogar 6 Mal so hoch. Hierfür mag es eine Reihe von Gründen geben, z.B. enthalten in der Biochemie Artikel im Durchschnitt über 30, in der Mathematik aber weniger als 10 Referenzen. Vielleicht hat dies aber auch mit der längeren „Halbwertszeit“ fundamentaler Resultate der Mathematik und Informatik im Vergleich zu aktuell wichtigen Entdeckungen der Physik und Biowissenschaften zu tun, und möglicherweise ist es auch ein Indiz dafür, dass zur Publikation aktueller Resultate Fachzeitschriften in der Informatik nur eine untergeordnete Rolle im Vergleich zu Konferenzbeiträgen spielen. Jedenfalls unterstreicht dies einmal mehr die starken fachbezogenen Unterschiede hinsichtlich des Zitierwesens.

Amin und Mabe kommen in ihrer Analyse zu folgendem Ergebnis: „The use of the absolute value of an impact factor to measure quality should be strongly avoided...The use of journal impact factors for evaluating individual scientists is even more dubious. Extending the use of the journal impact factor from the journal to the authors of papers in the journal is highly suspect; the error margins can become so high as to make any value meaningless“ [1]. Sie schließen mit der generellen Warnung „impact factors are not a direct measure of quality and must be used with considerable care.“

5 Konsequenzen?

Bibliometrische Analysen dienen verschiedenen Zwecken, vor allem aber werden sie als Evaluationshilfsmittel verwendet. Bei einer Evaluation möchte man Aussagen über die Qualität einer wissenschaftlichen Institution und ihrer beteiligten Personen erhalten – was aber macht die Qualität eines Wissenschaftlers aus? Sind Publikationen dafür überhaupt entscheidend? Bernhard Nebel stellt für die Informatik dazu jedenfalls kritisch fest: „Die nichtpublizierten Ergebnisse in Form von Software-Systemen haben mit großer Sicherheit einen viel größeren gesellschaftlichen Effekt als Publikationen, die unter Umständen nur von einer kleinen Zahl von Spezialisten gelesen wird“ [15].

Aber selbst wenn man sich auf die Publikationen als Indikator konzentriert, sollte es einleuchtend sein, dass sich ein Fach nicht ohne Fachkenntnis evaluieren lässt, da auch statistisch gewonnene Angaben zu Veröffentlichungen und Zitaten fachbezogen interpretiert werden müssen und die Vollständigkeit und Adäquatheit der zugrunde liegenden Datenbasis sowie die jeweilige Fachkultur nur von einem „Insider“ richtig gewürdigt werden kann.

Neben Publikationszahlen lassen sich natürlich auch andere mögliche Indikatoren für die wissenschaftliche Produktivität (wenn auch nicht direkt für die Qualität!) von Institutionen oder Einzelpersonen ausmachen; Nebel untersucht dazu u.a. die Promotionsrate und die Drittmittelhöhe pro Wissenschaftler(in) [15]. Generell stellt er fest, dass die Variationsbreite aller Indikatoren, insbesondere der publikationsbezogenen, sehr groß ist. Ein besonders bemerkenswertes Ergebnis der Studie ist, dass (bezüglich der Informatik an deutschen Universitäten) das Drittmittelaufkommen und die Publikationszahlen weitgehend unabhängig voneinander sind; tatsächlich weisen die beiden Kriterien nur einen Korrelationskoeffizienten von 0.07 auf! Dies ist insofern interessant, als dass auf den ersten Blick beide Indikatoren für sich betrachtet sinnvoll erscheinen. Aber wie aussagekräftig sind verschiedene Leistungsindikatoren, die nicht miteinander korrelieren, tatsächlich? Wäre eine wie auch immer geartete Verquickung der beiden Indikatoren, um sowohl die mathematisch orientierten als auch die ingenieurwissenschaftlich orientierten Forschungsansätze gerecht zu berücksichtigen, überhaupt noch sinnvoll interpretierbar?

Ungeachtet aller Probleme und Kritikpunkte scheint die automatische Verarbeitungsmöglichkeit maschinenlesbarer Datenquellen neuerdings zu noch weitergehenden zweifelhaften „Ex-

perimenten“ zu verführen. Ein Beispiel dafür stellen Mapping-Studien dar, die auf der Ko-Zitationsanalyse oder neuerdings auch auf der Ko-Wortanalyse von Titeln und Deskriptoren von Artikeln beruhen. Bei letzterem Verfahren werden, basierend auf gleichen oder assoziierten Stichwörtern, automatisch Cluster von miteinander in Beziehung stehenden Veröffentlichungen gebildet und diese thematischen Cluster hinsichtlich ihrer zeitlichen Dynamik (z.B. relative Zu- oder Abnahme ihres Umfangs) oder bezüglich ihrer Bedeutung und Entwicklung im Vergleich zu entsprechenden Clustern anderer Länder oder Institutionen analysiert und typischerweise in graphisch aufbereiteter Weise plakativ dargestellt.

Beim Versuch, im Rahmen einer derartigen vom Schweizer Wissenschaftsrat in Auftrag gegebenen Studie die Schweizer Informatik- und Mathematikforschung automatisch auf ca. 60 Cluster abzubilden, wurde die gesamte Informatik durch nur 6 thematische Cluster abgedeckt. So entstanden inhaltlich völlig unsinnige Cluster, bei denen z.B. die drei grundverschiedenen Forschungsgebiete „expert systems“, „computer simulations“ und „computational complexity“ in einen Topf geworfen wurden. Man mag darüber nun staunen, sich amüsieren oder ärgern – was aber sind die Konsequenzen? Interessiert sich jemand dafür? Was nützt es, wenn man weiß, dass man in diesem „Gebiet“ besser oder schlechter als der internationale Durchschnitt dasteht?

Die Gläubigkeit an die Aussagekraft und Wirksamkeit bibliometrischer Analysen treibt manchmal tatsächlich seltsame Blüten. So hat kürzlich erst die Verwaltung einer süddeutschen Universität für einige tausend Euro die CD-Version des Citation Index angeschafft und zur gefälligen Beachtung an die Fachbereiche verteilt, nachdem die Universitätsbibliothek gerade kurz zuvor diesen Index aus Kostengründen abbestellen musste!

Wird die Bibliometrie nun zum entscheidenden Qualitätsaspekt eines Wissenschaftlers und Hochschullehrers? Interessant ist in diesem Zusammenhang jedenfalls, dass das *Institute for Scientific Information* seit kurzem Listen der „100 einflussreichsten Forscher“ in verschiedenen Fachgebieten veröffentlicht [22]. Einfluss wird hier bemessen nach der Häufigkeit der Zitate, die einem Wissenschaftler entsprechend den SCI-Datenbanken der Jahre 1981-1999 zu Teil werden. Der Anfang ist mit den Gebieten Physik, Chemie, Neurowissenschaften, Ingenieurwissenschaften und Molekularbiologie bereits gemacht, die Informatik sowie einige weitere Fachgebiete sollen demnächst dazukommen.

Aber ist es tatsächlich spannend, Zitate zu zählen? Sollte dies überhand nehmen, sind Anpassungen und Gegenreaktionen jedenfalls vorprogrammiert. Gerhard Fröhlich drückt dies auf nette Weise so aus [5]: „So hat die Logik des rein Quantitativen, des Messens von Publikations- und Zitationshäufigkeiten, in ihren Effekten fatale (inverse) Ähnlichkeit mit der sowjetischen Planwirtschaftslogik: Führte Letztere dazu, dass die Messung des Plansolls etwa von Weihnachtsbaumständern nach Tonnen zur Produktion möglichst klobiger Exemplare führte (um so leicht und schnell das Soll zu erfüllen und die Planüberschreitungsprämie zu kassieren), verleitet Ersteres dazu, möglichst viele, möglichst kurze Beiträge in Journalen mit möglichst hohem Impact abzusondern, in so genannter ‚Salamipublikationstaktik‘: die Zerteilung des Forschungsertrags in möglichst viele, möglichst hauchdünne Scheibchen.“

Wenn es weniger auf den Inhalt als auf die Zahl der Publikationen und Zitate ankommt, wird tendenziell vermutlich weniger Sorgfalt bei einzelnen Artikeln angewendet werden und es wird der wiederholten Veröffentlichung ähnlicher Ideen Vorschub geleistet. Zu den möglichen Folgen der Verwendung von Zitierlisten als Qualitätskriterium bemerkt Kutzelnigg: „Ähnlich wie die Überbewertung der Zahl der Publikationen zu einem geänderten Publikationsverhalten führen wird und teilweise schon geführt hat, z.B. mit einem Trend zur ‚least publishable unit‘ oder zu Publikationen, bei denen die Autorenliste länger als der eigentliche Text ist, wird auch das ‚ranking‘ nach der Zahl der Zitate einen erheblichen Einfluss auf das Zitieren haben. Die Zahl der Selbstzitate (vor allem der verschleierte) wird weiterhin steigen, befreundete Gruppen werden sich zu Zitierkartellen zusammenschließen“ [9].

Mehr Sorge sollten indessen die langfristigen und weniger offensichtlichen sowie die eher indirekten Konsequenzen derartiger Erfolgskriterien bereiten: So erschwert die Überbetonung

von Veröffentlichungen beispielsweise den Gebietswechsel von Wissenschaftlern, da es einfacher ist, in seinem alten, etablierten Gebiet, in dem man bekannt und anerkannt ist, weitere messbare Erfolge zu erzielen als sich auf etwas neues, ungewisses einzulassen. Desgleichen würde man dann kaum in neuen, noch unbekanntenen Journalen etwas veröffentlichen oder bei einer neuen Konferenz etwas einreichen. Gehemmt wird dadurch die für die Wissenschaft so wichtige Innovationsfreude und das unkonventionelle Denken, das oft erst nach jahrelanger Verzögerung von anderen Forschern aufgegriffen wird und nicht schnell mit hohen Ziterraten belohnt wird.

Ferner leiden zugunsten von messbaren Aktivitäten (z.B. „Produktion“ von Publikationen, Vorträgen, Patenten und Spin-Off-Firmen) wichtige andere, in ihrer Wirkung nicht unmittelbar quantifizierbare Aspekte: Gerade Hochschullehrern sollte aber auch an Dingen gelegen sein wie zum Beispiel gut aufbereiteter und vorbereiteter Lehre, fundierten Gutachten aller Art, intensiver Betreuung von Doktoranden, längeren Gesprächen mit Studierenden, der Vertretung ihres Faches in der Öffentlichkeit und an vielem mehr. Bleibt dafür dann aber noch genügend Zeit und Motivation, wenn sich solche Aktivitäten nicht gut nachweisen und rechtfertigen lassen?

Dazu abschließend nochmals Kutzelnigg: „Vielleicht kann man sich darauf verständigen, dass es gewisse Aspekte wissenschaftlicher Leistung gibt, die sich quantifizieren lassen, das sind gewissermaßen die Sekundärtugenden eines Wissenschaftlers wie Fleiß ..., während die Aspekte, auf die es wirklich ankommt, wie Redlichkeit, Kreativität, Originalität, Witz, Eleganz, Tiefgang, Weitblick und – warum eigentlich nicht? – Genialität, sich einer quantitativen Erfassung naturgemäß entziehen“ [9]. Man mag dem entgegenhalten, dass letzteres Eigenschaften sind, die wir ganz generell an Menschen schätzen – aber unterscheidet nicht gerade das (paradoxaerweise eben kaum messbare!) Maß an Kreativität, Originalität und Genialität einen wirklich guten Wissenschaftler von einem nur durchschnittlichen?

6 Fazit

Wenn die Fernsehsendung mit den meisten Zuschauerzahlen die beste wäre, dann stünde „Big Brother“ auf Rang eins. Qualität ist aber keine eindimensionale Größe und nicht direkt messbar. Die Beschränkung auf das Quantifizierbare und die Reduktion auf outputorientierte Indikatoren stellt eine Verarmung dar und wird generell der Wissenschaft nicht gerecht. Zudem liefern in der Informatik die gängigen bibliometrischen Analysen aufgrund der lückenhaften Datenquellen und der Heterogenität des Faches ein völlig schiefes Bild und dürften damit in der Regel unbrauchbar sein. Würden einige wenige direkt messbare Leistungsindikatoren (wie etwa die Zitationshäufigkeit) an Bedeutung gewinnen, hätte dies längerfristig eher negative Konsequenzen, da dann für andere wichtige Aspekte wissenschaftlicher Betätigung der Anreiz fehlt.

Eine Evaluation der Informatik, die auf einer bibliometrischen Analyse aufbaut, erscheint so gesehen als ein zunächst zwar attraktiver, letztendlich aber untauglicher Versuch, die Leistung des Faches ohne Fachkenntnis beurteilen zu wollen. Zum Peer-Reviewing, das eine fachlich fundierte und ganzheitliche Begutachtung ermöglicht, dürfte es keine Alternative geben – eine sachkenntnislose und scheinbar objektivere maschinelle Bewertung bleibt vorerst jedenfalls eine Illusion.

Literatur

1. Mayur Amin, Michael Mabe: Impact Factors: Use and Abuse. Perspectives in Publishing 1, pp. 1-6, Oct. 2000
2. Centre for Science and Technology Studies (CWTS), Universität Leiden: Evaluation of Research Activities by Advanced Bibliometric Methods. 1999

3. Gunter Dueck: Intuition, E-Man und Drittmittel. Informatik-Spektrum 5:24, pp. 315-321, Okt. 2001
4. Klaus Fischer: Die Universität zwischen Kreativität und Steuerung. Forschung & Lehre 5/2001, pp. 240-245 (Zitate aus der Langfassung des Vortrags, gehalten anlässlich des 51. Hochschulverbandstages im März 2001 in Saarbrücken)
5. Gerhard Fröhlich: Das Messen des leicht Messbaren. Online-Mitteilungen 65 der Vereinigung Österreichischer Bibliothekarinnen und Bibliothekare, Dez. 1999, <http://www.uibk.ac.at/sci-org/voeb/om65.html#gf>
6. Eugene Garfield: Citation Indexing for Studying Science. Nature, 227:669, 1970
7. Eugene Garfield: Citation Indexing: Its Theory and Application in Science, Technology, and Humanities. Wiley, New York, 1979
8. Abby A. Goodrum, Katherine W. McCain, Steve Lawrence, C. Lee Giles: Scholarly publishing in the Internet age: a citation analysis of computer science literature. Information Processing and Management 37:6, pp. 661-675, 2001
9. Werner Kutzelnigg: Kann man wissenschaftliche Leistung messen? Forschung & Lehre 6/2001, pp. 302-305
10. Steve Lawrence, C. Lee Giles, Kurt Bollacker: Digital Libraries and Autonomous Citation Indexing. IEEE Computer, 6:32, pp. 67-71, 1999
11. Steve Lawrence: Online or Invisible? Nature, 411:6837, p. 521, 2001
12. Th.N. van Leeuwen, E.J. Rinia, A.F.J. van Raan: Bibliometric Profiles of Academic Physics Research in the Netherlands. Research Report to the Foundation for Fundamental Research on Matter, Centre for Science and Technology Studies (CWTS), Dec. 1996
13. Michael Ley: Die Trierer Informatik-Bibliographie DBLP. In: M. Janke, K. Pasedach, K. Pohl (Hrsg.): Informatik 97, 27. Jahrestagung der Gesellschaft für Informatik. Berlin Heidelberg New York: Springer, pp. 257-266, 1997
14. Werner Marx, Hermann Schier, Michael Wanitschek: Citation analysis using online databases: Feasibilities and shortcomings. Scientometrics 52:1, pp. 59-82, 2001 (deutsche Version: Kann man Forschungsqualität messen? MPG-Spiegel 3, 1998)
15. Bernhard Nebel: Ranking? Informatik-Spektrum 4:24, pp. 234-249, Aug. 2001
16. NEC Research Institute: Estimated Impact of Publication Venues in Computer Science. <http://citeseer.nj.nec.com/impact.html>, April 2001
17. Tope Omitola: ACM Fellow Roger Needham. ACM Software Engineering Notes, 1:26, pp. 7-10, Jan. 2001
18. <http://citeseer.nj.nec.com/> bzw. <http://researchindex.org>
19. <http://citeseer.nj.nec.com/allcited.html>
20. <http://citeseer.nj.nec.com/source.html>
21. <http://dblp.uni-trier.de/>
22. <http://isihighlycited.com/>
23. <http://www.isinet.com/>