

ISI Essential Science Indicators

Forschung im internationalen Vergleich - Wissenschaftsindikatoren auf Zitationsbasis

Bewertung wissenschaftlicher Forschungsergebnisse aus einer elektronischen Datenbank heraus? Rangordnungen der wichtigsten Institutionen, Wissenschaftler, Zeitschriften und sogar Länder in Fachdisziplinen nach Einfluss? Markierung "heißer", hochaktueller Artikel? Auflisten der hochzitierten Forschungsfronten in den einzelnen Wissenschaftsdisziplinen? Und das alles auf Knopfdruck und nicht mittels umständlicher szientometrischer Verfahren? Geht so etwas überhaupt?

Es geht. Mit den "Essential Science Indicators" (ESI) legt das ISI ein web-basiertes Informationssystem zur Wissenschaftsevaluation vor, das einzigartige Ergebnisse präsentiert und in der Tat ausgesprochen einfach zu bedienen ist. Aber es geht, verglichen mit ausgeklügelten Methoden der empirischen Wissenschaftsforschung, nicht alles. Wo liegen die Grenzen des Systems? Wir werden die Arbeitsweise der ESI, seine Datenbasis, die eingesetzten informatrischen Algorithmen - und deren methodischen Probleme, die Suchoberfläche sowie die Ergebnisdarstellung skizzieren. Als Beispiel dienen uns Aspekte deutscher Forschung. Etwa: In welcher Disziplin haben Deutschlands Forscher den größten internationalen Einfluss? Welches deutsche Institut der Neurowissenschaften kann auf globaler Ebene mitmischen? Oder: Welcher in Deutschland tätige Wissenschaftler führt eine disziplinspezifische Rangordnung an? Letztlich: Wer braucht die "Essential Science Indicators"? - Wir testeten die Essential Science Indicators Mitte Februar 2002 anhand der Version vom 1. Januar

2002, die das Zehn-Jahres-Intervall 1991 bis 2000 sowie die ersten zehn Monate aus 2001 berücksichtigt.

Datenbasis: Hochzitierte Artikel in Fachzeitschriften

Die Essential Science Indicators geben für 22 vordefinierte Wissenschaftsdisziplinen (siehe Tabelle 1) tabellarische und graphisch aufbereitete Informationen über die jeweils in den letzten zehn Jahren meistzitierten

- Artikel
 - Autoren
 - Institutionen
 - Länder
 - Zeitschriften.
- Hinzu kommen
- Hot Papers (hochzitierte Artikel der letzten zwei Jahre)
 - Forschungsfronten (Co-Zitations-Analysen)
 - Baselines (Basisdaten pro Disziplin und Jahr zum Vergleich).

Ergänzt werden die ESI durch redaktionell aufbereitetes Material in den Reihen "in-cites" (Hintergrundberichte zu hochzitierten Autoren, Institutionen oder Ländern), "Special Topics" (Zitationsanalysen mit Expertenkommentaren zu ausgewählten Forschungsbereichen) und "Science Watch" (Newspaper des ISI zu aktuellen Entwicklungen in den Wissenschaften). Abbildung 1 zeigt die Optionen der Essential Science Indicators im Überblick.

Grundlage der Essential Science Indicators sind alle (rund 8.500) Quellenzeitschriften der Zitationsdatenbanken des Institute for Scientific Information mit den darin enthaltenen Publikationen und Zitationen (vgl. Stock 1999). Zugrundegelegt werden jeweils die letzten zehn Jahre sowie der laufende Jahrgang (in Zwei-Monats-Intervallen). Bearbeitet wurden für die aktuelle Version rund sieben Millionen Artikel mit darin enthaltenen 53 Millionen Fußnoten. Alle Zeitschriften werden genau einer von 22 Disziplinen zugeordnet; die in den Periodika enthaltenen Artikel gelten ebenfalls dieser Disziplin zugehörig. ISI folgt dabei

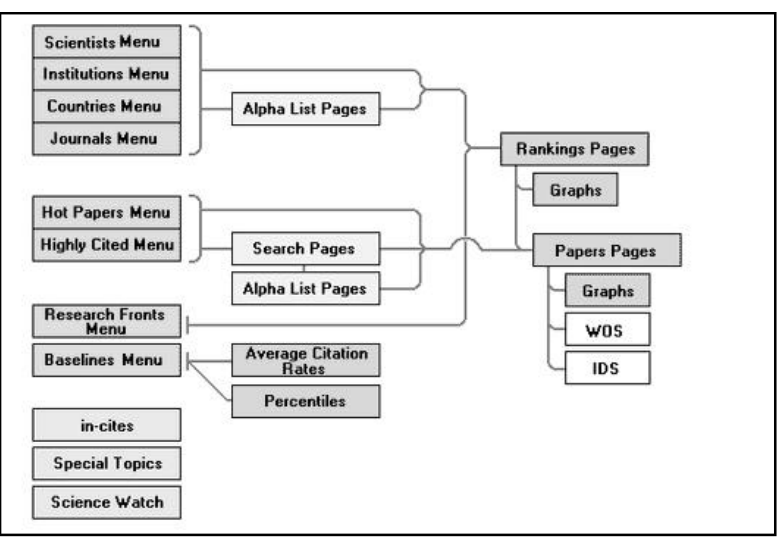


Abbildung 1: Essential Science Indicators. Ablaufdiagramm

seiner Kategorieneinteilung in den Current Contents (CC). Nicht in den CC verzeichnete Zeitschriften werden ad hoc einem der 22 Felder zugeordnet. Für die 22 Disziplinen bildet ISI Perzentile aller Artikel und selektiert für die ESI die Arbeiten im 1%-Perzentil. Es gehen also nur solche Dokumente in die Datenbasis ein, die im Rahmen ihrer Disziplin gebührend wahrgenommen worden sind. ESI verfügt über eine Beobachtungsbasis von rund 70.000 Papers. Abweichungen in der Methode gibt es bei den Hot Papers (Einschränkung auf die letzten zwei Jahre und auf das 0,1%-Perzentil; in der aktuellen Version: 1.400 Hot Papers) und den Forschungsfronten (Einschränkung auf die letzten fünf Jahre sowie mittels zusätzlicher Schwellenwerte).

Von den drei Mio. Namen von Wissenschaftlern, die in den letzten zehn Jahren in den ISI-Daten auftauchen, sind 50.000 in den ESI vertreten; von den im Web of Science ab 1991 registrierten 500.000 Institutionen schaffen 2.800 den Sprung in die ESI. Von den 8.500 Zeitschriften liegen 4.500 in den ESI vor, von 200 Ländern deren 150 (jedoch nicht in allen Feldern, sondern nur dort, wo das Land in das 50%-Perzentil fällt).

Zur Bestimmung der meistzitierten Artikel benutzt das ISI bei den ESI eine Variante des Impact Factor. Die Formel arbeitet mit der relativen Zitationsrate ZR (Zitationen pro Publikation) und ist denkbar einfach:

$$\text{relative Zitationsrate ZR} = \frac{\text{Anzahl der Zitationen (im 10-Jahres-Intervall)}}{\text{Anzahl der Publikationen (im selben 10-Jahres-Intervall)}}$$

Viele Probleme des "klassischen" Impact Factor (vgl. Stock 2001) wie der (mit zwei Jahren) sehr kurze Beobachtungszeitraum fallen weg; dafür kommt ein neues. Methodisch benachteiligt werden nunmehr Artikel, die jüngeren Datums sind, da sie - im Vergleich zu bereits zehnjährigen Artikeln - viel weniger Chancen hatten, zitiert zu werden. Ein nur kurzer Blick in die Rankings kann demnach zu Fehlinterpretationen führen. Detailanalysen, die den Rankings beigegebenen Zeitreihen, den Hot Papers und den Forschungsfronten gebührend Beachtung schenken, werden falsche Schlüsse jedoch weitgehend vermeiden.

Der Wert des Impact Factor ist für unterschiedliche Wissenschaftsdisziplinen verschieden ausgeprägt. Die Differenzen sind so groß, dass Vergleiche über Disziplingrenzen anhand der absoluten Zahlen nicht statthaft sind. Hier helfen die in den ESI enthaltenen Baselines weiter, die die disziplinspezifischen durchschnittlichen Zitationsraten sowie die Perzentile auflisten. Wir erhalten somit Daten zur Konstruktion von "Normalisierungsfaktoren", die wir einführen wollen, um Disziplinen miteinander vergleichbar zu machen. Die durchschnittliche Zitationsrate über alle wissenschaftlichen Gebiete hinweg liegt zwischen 1991 und 2001 bei 7,91, für Biologie/Biochemie z.B. bei 13,31 und für Computerwissenschaft bei 2,23 (siehe Abbildung 2). Der Normalisierungsfaktor NF für Biologie/Biochemie beträgt $7,91:13,31=0,6$, der für Computerwissenschaft 3,6. Vergleichen wir Werte für Institutionen, Personen, Zeitschriften oder Länder in Bezug auf unsere beiden Beispieldisziplinen, so müssen die Werte aus den ESI-Tabellen mit dem entsprechenden Faktor NF multipliziert werden. (Dieses Verfahren könnte noch verfeinert werden, indem man die einzelnen Jahreswerte berücksichtigt und einen jahresspezifischen Normalisierungsfaktor errechnet.)

Wissenschaftsevaluation: Was messen die ESI?

Entscheidend für die Wissenschaftsevaluation sind die Fragen nach der Gültigkeit und der Zuverlässigkeit der Daten. Die Essential Science Indicators entstehen durch mehrere Selektionsschritte, die jeweils die meistzitierten und damit einflussreichsten Artikel wissenschaftlicher Zeitschriften auswählen und diese auf mehreren Aggregationsstufen (Artikel, Autor, Institution, Zeitschrift, Land) in Rangordnungen sortieren.

Zitationen in Büchern sind beim ISI grundsätzlich nicht vorhanden. Damit geht eine wichtige Art wissenschaftlichen Einflusses verloren. Man kann nämlich durchaus behaupten, dass die wichtigsten Ideen, Entdeckungen und Erfindungen eines Spezialgebietes in Lehrbuchwissen transportiert, d.h. in einschlägigen Fachlehrbüchern zitiert werden. Lehrbücher spielen u.a. eine bedeutende Rolle bei der Selbstrepro-

duktion der Wissenschaft, sprich der wissenschaftlichen Ausbildung der Studierenden. Diese Dimension wissenschaftlichen Einflusses auf Lehrbücher erfassen ESI prinzipiell nicht.

ISI ist sich dieser Beschränkung seiner Methode bewusst. "It is important to recognize that the data in ESI are limited to ISI-indexed journal articles only. No books, book chapters, or articles published in journals not indexed by ISI are taken into account here, either in terms of publication or citation counts", lesen wir im ESI-Handbuch.

Beschrieben wird ausschließlich der Einfluss innerhalb der Wissenschaftlergemeinschaft im Rahmen der aktuellen Forschungen und Entwicklungen. Dies ist die Domäne des wissenschaftlichen Zeitschriftenartikels, der bevorzugt Fachkollegen anspricht. Da ISI von rund 100.000 (im engeren Sinne wissenschaftlichen) Periodika nur 8.500 auswertet, sind die Daten niemals komplett, durch die Beschränkung des ISI allgemein auf die meistzitierten Zeitschriften und bei den ESI auf die darin enthaltenen meistzitierten Artikel jedoch mehr oder minder repräsentativ. Diese Repräsentativität steigt mit der Höhe des Aggregationsniveaus von der Betrachtung einzelner Artikel und Wissenschaftler bis hin zu Ländern. Auf dieser hohen Aggregationsebene dürfte der wissenschaftsimmanente, am Forschungsprozess direkt orientierte Einfluss methodisch sauber ablesbar sein.

Baselines: Durchschnittliche Zitationsraten und Perzentile als Vergleichswerte

Die Baselines geben einen hochverdichteten Überblick zur Wissenschaft als "Ganzes" und zu den 22 Disziplinen. In Abbildung 2 sehen wir die durchschnittlichen Zitationsraten pro Artikel in den Wissenschaften für die Gesamtdatenbasis sowie für die einzelnen Berichtsjahre. Ein durchschnittlicher beim ISI erfasster Artikel wird demnach in Intervall zwischen 1991 und 2001 exakt 7,91 mal zitiert. Je älter eine Arbeit ist, desto höher ist die relative Häufigkeit ihrer Zitation. Der laufende Jahrgang (2001) zeigt eine relative Zitationsrate von 0,22. Diese steigt über die Jahre kontinuierlich an bis zu einem Wert von 14,14 für 1991 publizierte Arbeiten.

Average Citation Rates for papers published by field, 1991 - 2001 (How to read this data)

Fields	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	All Years
All Fields	14.14	13.54	13.19	11.76	10.39	8.68	7.24	5.36	3.48	1.48	0.22	7.91
Agricultural Sciences	6.65	6.01	6.21	5.63	5.04	4.38	3.50	2.68	1.63	0.66	0.10	3.83
Biology & Biochemistry	23.04	22.30	20.79	19.24	16.44	13.89	11.85	8.88	5.77	2.56	0.37	13.31
Chemistry	12.17	11.24	11.48	10.14	8.86	7.62	6.35	4.91	3.20	1.42	0.20	6.90
Clinical Medicine	14.93	14.53	14.80	13.13	12.07	9.95	8.12	5.94	3.82	1.57	0.21	8.71
Computer Science	4.05	4.26	3.20	3.38	3.09	2.61	2.10	1.55	0.90	0.32	0.08	2.23
Economics & Business	7.62	7.32	6.00	5.90	4.94	3.54	2.83	1.86	1.00	0.38	0.07	3.57

Abbildung 2: Durchschnittliche Zitationsraten wissenschaftlicher Artikel für ausgewählte Disziplinen

In Abbildung 3 betrachten wir die Perzentile der nach Zitationsraten geordneten Artikel. Um unter die Top 10% der Artikel in den Wissenschaften schlechthin zu kommen, braucht man zwischen 1991 und 2001 genau 20 Zitationen. Um das 1%-Niveau zu erreichen, muss die Arbeit im 10-Jahres-Intervall schon 86mal zitiert werden. Die Spitzengruppe des 0,01%-Perzentils erklimmen Artikel, die immerhin 749 Zitationen und mehr einsammeln konnten. Wie bei den durchschnittlichen Zitationsraten liegen auch bei den Perzentilen große disziplinspezifische Abweichungen vor. Ein mathematischer Artikel erreicht das 1%-Perzentil mit nur 24 Zitationen in zehn Jahren, eine Arbeit der Molekularbiologie braucht dafür mit 227 Zitationen rund das Zehnfache.

In den Tabellen der Baselines werden die unterschiedlichen Zitationsgewohn-

heiten innerhalb der Wissenschaftsdisziplinen eindrucksvoll verdeutlicht. Der Nutzer benötigt die Baselines bei der Interpretation jeder konkreten Zitationsrate, kann man doch so ablesen, wie der jeweilige Messwert im Kontext der Disziplin zu interpretieren ist.

Rangordnungen der Artikel, Autoren, Institutionen, Zeitschriften und Länder

Kernstück der ESI sind Rangordnungen, die nach jeweils einem gleichen Muster gebildet werden. Die Suchbildschirme gestatten einerseits eine Selektion nach Wissenschaftsdisziplinen und bilden innerhalb der Disziplin Rankings (von Artikeln, Autoren, Institutionen, Zeit-

schriftentiteln und Ländern) nach der absoluten Anzahl von Publikationen und Zitationen, nach der durchschnittlichen relativen Zitationsrate sowie nach dem Alphabet. Andererseits kann man nach Institutionen, Ländern usw. suchen und erhält Rankings der Fachgebiete. Artikel werden einer Institution bzw. einem Land zugeordnet, wenn mindestens ein Autor bei Autorenkollektiven aus dem Institut oder dem betreffenden Land stammt.

Eine typische Suchoberfläche, hier zur Auswahl der Länderrangordnungen, wird in Abbildung 4 wiedergegeben. Das Retrievalsystem gestattet im oberen Teil das Ankreuzen genau einer Wissenschaftsdisziplin (bzw. der Wissenschaft als "Ganzes") und im unteren Teil die Eingabe eines Suchterms, wobei Trunkierung und Phraseneingabe, jedoch keine Booleschen Operatoren zugelassen sind. Die Suchterme im

Percentiles for papers published by field, 1991 - 2001 (How to read this data)

All Fields	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	All Years
0.01 %	1142	999	987	923	843	679	529	394	216	100	23	749
0.10 %	436	414	398	340	305	250	198	146	95	43	10	276
1.00 %	140	133	127	113	99	83	68	51	33	16	4	86
10.00 %	34	32	31	28	25	21	18	13	9	4	1	20
Agricultural Sciences	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	All Years
0.01 %	952	225	262	208	276	156	112	74	47	24	9	196
0.10 %	126	115	113	113	88	75	51	50	30	12	4	82
1.00 %	55	48	49	43	39	31	25	21	13	7	2	36
10.00 %	18	16	17	15	14	12	10	7	5	3	1	11

Abbildung 3: Perzentile der Zitationsrate wissenschaftlicher Artikel für "All Fields" und "Agricultural Sciences"

COUNTRIES MENU

BY FIELD	▶	Select a country or region from this field: (All Fields) <input type="text"/>	GO
OR			
BY NAME	▶	Select a country or region from the alphabetical list or enter a name to search.	
<u>A-Z (Countries List)</u>			
<i>Example: FRANCE (more examples)</i>			
			SEARCH

Abbildung 4: Suchbildschirm (Beispiel)

unteren Teil können auch über hinterlegte Listen recherchiert werden.

Das Retrievalergebnis ist stets eine Rangordnung. In Abbildung 5 sehen wir die Top-5-Liste der weltweit nach Zitationen pro Artikel führenden Forschungseinrichtungen der Neurowissenschaften. Mit gut 70 Zitationen pro Artikel liegt das Forschungsunternehmen Athena Neuroscience, Inc. aus San Francisco an der Spitze, dicht gefolgt vom Max-Planck-Institut für medizinische Forschung in Heidelberg.

Bei den Geowissenschaften haben wir für die Wissenschaftler nach der absoluten Anzahl der Zitationen sortiert (Abbildung 6). Hier sehen wir einen über lange Jahre in Deutschland tätigen Wissenschaftler auf dem ersten Rang. Paul J. Crutzen, gebürtiger Niederländer, war zwischen 1980 und 2000 Direktor des Max-Planck-Instituts für Chemie (Otto-Hahn-Institut) in Mainz; er erhielt 1995 den Nobelpreis für Chemie.

In den Abbildungen 5 und 6 finden wir unter "View" zwei weiterführende Links, dargestellt mit dem Icon "Paper" und dem Icon "Graph". Das Graph-Symbol führt auf graphische Darstellungen von drei Zeitreihen zur Anzahl der Papers, der Zitationen und der Zitationen pro Paper über den Beobachtungszeitraum von zehn Jahren in gleitenden Fünf-Jahres-Intervallen (ein Beispiel siehe unten unter Abbildung 10). Das Paper-Symbol leitet zu einer Rangordnung der Artikel (mit bibliographischen Angaben) nach der absoluten Zitationsrate (Abbildung 7). Hier taucht neuerdings ein Graph-Symbol auf. Hinter diesem Link verbirgt sich eine Graphik der Zeitreihe der Zitationen des jeweiligen Artikels. Links sind zum Web of Science gelegt, wo sich (häufig) ein Abstract der Arbeit befindet und wo alle Fußnoten (Referenzen) und Zitationen verzeichnet sind. (Der Link ist ausschließlich für Abonnenten des Web of Science gangbar.) Ein

weiterer Button führt zum Dokument-lieferservice des ISI. Beeindruckend ist die Fülle an Links, die in den Ergebnislisten gelegt sind. Im Nachweis von Abbildung 7 können wir beispielsweise die Werte für die beiden Autoren, die Zeitschrift, die Institutionen und das Land erhalten. Negativ fällt im selben Beispiel die veraltete deutsche Postleitzahl auf. Verzeichnet ein Artikel viele Autoren, so ist es z.T. unmöglich herauszubekommen, welches Institut zu welchem Autor gehört.

Auf ein weiteres Negativum muss an dieser Stelle hingewiesen werden. Die Praxis beim ISI, Personennamen durch den Nachnamen und nur die Initialen der Vornamen anzugeben, lässt unterschiedliche Wissenschaftler mit häufigen Namen zu einer informatrischen Einheit zusammenfallen. "Schmidt M" vereinigt in sich hochzitierte Forscher in Physik, Chemie, Klinischer Medizin, Biologie, Astrophysik und Ingenieurwesen.

INSTITUTION RANKINGS IN NEUROSCIENCE & BEHAVIOR

Sorted by: Citations per Paper **SORT AGAIN**

1 - 20 (of 293) Page 1 of 15

	VIEW		INSTITUTION	PAPERS	CITATIONS	CITATIONS PER PAPER
1			ATHENA NEUROSCI INC.	64	4,516	70.56
2			MAX PLANCK INST MED RES.	70	4,736	67.66
3			ROCHE INST MOLEC BIOL.	73	4,686	64.19
4			BRISTOL MYERS SQUIBB PHARMACEUT RES INST.	70	4,212	60.17
5			REGENERON PHARMACEUT INC.	169	9,061	53.62

Abbildung 5: Rangordnung der Institutionen im Fach Neuroscience & Behavior nach relativer Zitationsrate

SCIENTIST RANKINGS IN GEOSCIENCES						
Sorted by: Citations <input type="button" value="SORT AGAIN"/>						
1 - 20 (of 1629)		<input type="button" value="Home"/> <input type="button" value="Previous"/> <input type="button" value="Next"/> <input type="button" value="Last"/>			Page 1 of 82	
	VIEW		SCIENTIST	PAPERS	CITATIONS	CITATIONS PER PAPER
1			CRUTZEN, PJ	110	3,117	28.34
2			STUIVER, M	29	2,966	102.28
3			SOLOMON, S	78	2,186	28.03
4			BLAKE, DR	102	2,150	21.08
5			PELTIER, WR	92	2,097	22.79
6			JACOB, DJ	83	2,084	25.11
7			SACHSE, GW	96	2,013	20.97
8			ROSSOW, WB	55	1,959	35.62

Abbildung 6: Rangordnung der Wissenschaftler im Fach Geosciences nach Zitationsanzahl

7	Citations: 111	<input type="button" value="GO TO WEB OF SCIENCE"/>	<input type="button" value="GO TO DOCUMENT DELIVERY"/>
Title:	THE ROLE OF CLOUDS IN TROPOSPHERIC PHOTOCHEMISTRY		
Authors:	LELIEVELD J ; CRUTZEN PJ		
Source:	J ATMOS CHEM 12: (3) 229-267 APR 1991		
Addresses:	MAX PLANCK INST CHEM , DIV AIRCHEM, POB 3060, W-6500 MAINZ, GERMANY		

Abbildung 7: Artikelnachweis aus der Liste der hochzitierten Arbeiten von Paul J. Crutzen

Das Wissenschaftsgebiet Deutschlands mit dem höchsten internationalen Einfluss: die Physik

Wo ist Deutschlands Grundlagenforschung international Spitze? Je nachdem, welchen Indikator wir anlegen, erhalten wir andere Werte. Wir nähern uns einer Lösung, indem wir in den ESI zunächst für Deutschland die Disziplinen nach der absoluten Zitationszahl auflisten lassen. Die so entstehende Abbildung 8 zeigt die Klinische Medizin mit über 930.000 Zitationen an der Spitze,

gefolgt von Physik und Chemie. Diese Hitparade ist noch wenig aussagekräftig, da sie die Anzahl der Publikationen außer acht lässt. Für eine Disziplin, in der viel veröffentlicht wird, steigt auch die Wahrscheinlichkeit, dass sie viele Zitationen erhält. Und die deutsche klinische Medizin publiziert mit rund 134.000 Arbeiten sehr viel und dies durchaus erfolgreich (immerhin sind in den ESI nur die meistzitierten 1% pro Disziplin enthalten).

Wir ordnen dieselbe Liste neu. Abbildung 9 wählt als Sortierkriterium die relative Zitationsrate. Dies ist gerechter als in Abbildung 8, da hier die Anzahl der Publikationen voll berücksichtigt wird. Die Hitparade hat sich völlig

gewandelt. Neben der recht ausdruckslosen Klasse "Multidisciplinary" liegen die Molekularbiologie sowie die Immunologie an der Spitze. Aber auch Abbildung 8 enthält eine wesentliche Ungerechtigkeit: die unterschiedlichen Zitationsgewohnheiten der einzelnen Disziplinen.

Aus der Baseline der durchschnittlichen Zitationsraten (Abbildung 2) haben wir für die 22 Disziplinen die Normalisierungsfaktoren errechnet und die Werte aus Abbildung 9 damit multipliziert. Das Produkt aus durchschnittlicher Zitationsrate und Normalisierungsfaktor bezeichnen wir als "normalisierte Zitationsrate". Die in Tabelle 1 entstandene Rangfolge nach der nor-

	VIEW		FIELD	PAPERS	CITATIONS	CITATIONS PER PAPER
1			<u>CLINICAL MEDICINE</u>	134,376	930,593	6.93
2			<u>PHYSICS</u>	87,235	733,386	8.41
3			<u>CHEMISTRY</u>	90,889	712,540	7.84
4			<u>BIOLOGY & BIOCHEMISTRY</u>	40,189	561,594	13.97
5			<u>MOLECULAR BIOLOGY & GENETICS</u>	20,307	463,568	22.83
6			<u>NEUROSCIENCE & BEHAVIOR</u>	19,232	260,625	13.55
7			<u>MULTIDISCIPLINARY</u>	5,831	256,224	43.94
8			<u>PLANT & ANIMAL SCIENCE</u>	31,426	170,294	5.42
9			<u>MICROBIOLOGY</u>	12,885	164,976	12.80
10			<u>IMMUNOLOGY</u>	8,579	135,926	15.84

Abbildung 8: Rangordnung deutscher Artikel nach Zitationsanzahl (Top 10 der Wissenschaftsgebiete)

malisierten Zitationsrate dürfte dem "wahren" internationalen Einfluss der deutschen Forschung am nächsten kommen. Rangplatz 1 nimmt nunmehr die Physik ein, Rang 2 geht an die Ingenieurwissenschaften, Rang 3 teilen sich die Geowissenschaften und die Molekularbiologie (einschließlich der Genetik). Die Spitzenstellung der Physik im Rahmen der deutschen Forschung bestätigt übrigens auch eine Studie des ISI (vgl. ISI 2000).

Zur Verdeutlichung eines Trends in der Wissenschaftsentwicklung geben die ESI graphisch aufbereitete Zeitreihen an (Abbildung 10). Die Blöcke repräsentieren jeweils Fünf-Jahres-Werte. Die Anzahl der Physik-Artikel, an denen mindestens ein Autor mit deutscher Anschrift betei-

ligt ist, steigt von rund 34.000 in den Jahren 1991 bis 1995 auf über 45.000 in den Jahren 1997 bis 2001 an. Bei den Zitationen kann die deutsche Physik zwischen dem ersten und dem letzten Fünf-Jahres-Intervall gut 50.000 Zitationen mehr einfahren. Die durchschnittliche Zitationsrate pro Artikel bleibt demgegenüber im letzten Jahrzehnt bei gut vier in den Fünf-Jahres-Intervallen relativ konstant.

Auch wenn wir mit der Physik die einflussreichste Disziplin Deutschlands ausmachen konnten - zur Weltspitze reicht dieses Ergebnis nicht. Bei der absoluten Anzahl der Zitationen belegen die deutschen Physiker Rangplatz zwei hinter den USA, bei den Zitationen pro Artikel gar nur Rang zehn. Allerdings muss

man hinzufügen, dass die Spitzengruppe recht dicht beieinander liegt (Schweiz 11,63; USA 10,54, Dänemark 10,19 und Deutschland - siehe auch Tabelle 1 - 8,41).

Hot Papers: Wo aktueller wissenschaftlicher Fortschritt sichtbar wird

Der Beobachtungszeitraum von zehn Jahren, wie er sonst - zurecht - in den ESI gewählt wird, ist für die Analyse des aktuellen Randes der Wissenschaftsentwicklung zu lang. Bei den Hot Papers greift das ISI auf ein Intervall von zwei Jahren (plus dem angebrochenen Berichtsjahr) zurück.

	VIEW		FIELD	PAPERS	CITATIONS	CITATIONS PER PAPER
1			<u>MULTIDISCIPLINARY</u>	5,831	256,224	43.94
2			<u>MOLECULAR BIOLOGY & GENETICS</u>	20,307	463,568	22.83
3			<u>IMMUNOLOGY</u>	8,579	135,926	15.84
4			<u>BIOLOGY & BIOCHEMISTRY</u>	40,189	561,594	13.97
5			<u>NEUROSCIENCE & BEHAVIOR</u>	19,232	260,625	13.55
6			<u>MICROBIOLOGY</u>	12,885	164,976	12.80
7			<u>SPACE SCIENCE</u>	12,124	132,107	10.90
8			<u>PHYSICS</u>	87,235	733,386	8.41
9			<u>CHEMISTRY</u>	90,889	712,540	7.84
10			<u>GEOSCIENCES</u>	13,754	103,968	7.56

Abbildung 9: Rangordnung deutscher Artikel nach durchschnittlicher relativer Zitationsrate (Top 10 der Wissenschaftsgebiete)

Disziplin	ZR	NF	normalisierte ZR	Rang
Agricultural Sciences	3,00	2,1	6,3	18
Biology; Biochemistry	13,97	0,6	8,4	11
Chemistry	7,84	1,1	8,6	7
Clinical Medicine	6,93	0,9	6,2	19
Computer Science	2,38	3,6	8,6	7
Ecology; Environment	5,77	1,3	7,5	16
Economics; Business	1,92	2,6	5,0	21
Engineering	3,16	3,0	9,5	2
Geosciences	7,56	1,2	9,1	3
Immunology	15,84	0,5	7,9	14
Material Sciences	3,72	2,3	8,6	7
Mathematics	2,53	3,4	8,6	7
Microbiology	12,80	0,6	7,7	15
Molecular Biology; Genetics	22,83	0,4	9,1	3
Multidisciplinary	43,94	0,2	8,8	5
Neuroscience; Behavior	13,55	0,6	8,1	12
Pharmacology; Toxicology	7,49	0,9	6,7	17
Physics	8,41	1,2	10,1	1
Plant and Animal Science	5,42	1,5	8,1	12
Psychology; Psychiatry	4,75	1,1	5,2	20
Social Sciences, general	1,30	2,7	3,5	22
Space Science	10,90	0,8	8,7	6

Tabelle 1: Die einflussreichsten Wissenschaftsdisziplinen Deutschlands nach

normalisierter Zitationsrate. Quelle: ESI; eigene Berechnungen.

ZR: Zitationsrate gemäß ESI; NF: Normalisierungsfaktor;

Rang: nach normalisierter Zitationsrate

Hot Papers sind solche Artikel, die in diesem Zeitraum innerhalb einer Disziplin am meisten zitiert werden. Der Schwellenwert wird hier von sonst 1% auf das 0,1%-Perzentil hochgesetzt. Bei der graphischen Darstellung wird die Zitationsrate auf ein Zwei-Monats-Intervall festgelegt. Hot Papers sind (wie ebenso die Artikel im Zehn-Jahres-Intervall) nach Disziplin, Titelwort, Name eines Wissenschaftlers, Institution, Land oder Zeitschrift suchbar.

Unsere Recherche betrifft das derzeit "heißeste" Papier in der gesamten Datenbasis (Abbildung 11). Es handelt sich um die Krebsstatistik der amerikanischen Cancer Society, veröffentlicht Anfang 2000 und seitdem bis Ende Oktober 2001 605mal zitiert.

Forschungsfronten

Netze von hochzitierten Artikeln, die häufig gemeinsam zitiert, d.h. co-zitiert werden, gelten als Indikatoren auf eine gemeinsame Forschungsfront. Die Methode zum Erkennen und Analysieren von "Research Fronts" geht auf Arbeiten von Henry Small vom ISI zurück. Die sog. "Clusterkerne" dieser Forschungs-

fronten sind in ihrem Spezialgebiet als bahnbrechende Arbeiten wahrgenommen worden. Wie machen die Essential Science Indicators diese sichtbar? Als erster Schritt werden pro Disziplin für die letzten fünf Jahre (plus das angebrochene Jahr) die jeweils höchstzitierten Artikel selektiert. Ausgezählt wird die Co-Zitationshäufigkeit für alle in dieser Datenbasis vorkommenden Dokumente, d.h. die Anzahl derjenigen zitierenden Artikel, die zwei Arbeiten A und B gemeinsam im Anmerkungsapparat aufzuführen. Die ermittelte Co-Zitationshäufigkeit von A und B wird mit einem zweiten Schwellenwert (derzeit bei größer/gleich 2) verglichen. Der nächste Schritt berechnet die "normalisierte Co-Zitation" als Größe des Abstandes bzw. der Nähe nach der Formel:

Co-Zitationshäufigkeit [A, B]

$$\frac{\text{Co-Zitationshäufigkeit [A, B]}}{(\text{Zitationsanzahl [A]} * \text{Zitationsanzahl [B]})^{1/2}}$$

d.h. es wird ein Quotient aus der Co-Zitationshäufigkeit von A und B und der Quadratwurzel des Produktes der Zitationszahlen der beiden Artikel A und B gebildet. Alle Paare, die einen weiteren Schwellenwert (derzeit bei größer/gleich

0,3) überschreiten, gehen in die Clusteranalyse ein. Von jedem co-zitierten Paar wird analysiert, ob einer der beiden Artikel in einem weiteren Co-Zitationspaar vorkommt usw., bis alle Paare abgearbeitet sind. Dieses Verfahren ist in der Clusteranalyse als Single Linkage-Verfahren bekannt. Die entstehenden Cluster haben als Minimum zwei Dokumente als Elemente, können aber auch weitaus größer sein. Ihnen werden aus den Titeln der co-zitierten Artikel Themen zugeordnet.

Die Essential Science Indicators geben für jedes Cluster die Anzahl der Elemente im Clusterkern, deren kumulierte Zitationsrate, die durchschnittliche Zitationsrate pro Artikel sowie das Durchschnittsjahr der Veröffentlichung der Kern-Artikel an.

Die Suche nach den Forschungsfronten geschieht entweder über die Auswahl einer der 22 Disziplinen (so unser Beispiel in Abbildung 12 nach Computer Science) oder nach einem Wort innerhalb der Themen der Forschungsfronten. Angezeigt werden die entsprechenden Cluster als Rangordnung mit fünf Sortieroptionen (siehe geöffnetes Sortierfeld in Abbildung 12). Unter "View" kommen wir zum einen zur Verteilung der Artikel im Clusterkern nach deren Veröffentlichungsdatum und zum anderen zu den Daten der einzelnen Artikel (analog zu Abbildung 7).

Das Angebot der Forschungsfronten im Rahmen der ESI ist nahezu ideal. Wünschenswert wäre eigentlich nur noch die graphische Darstellung der Cluster (als semantische Netze oder als Knowledge Trees).

Preis

Die Preisgestaltung beim ISI orientiert sich auch an der Größe der abonnierenden Institution bzw. des jeweiligen Konsortiums. Insofern ist der folgende Preis nur ein Anhaltswert. Man muss mit rund US-\$ 15.000 pro Jahr für eine Lizenz rechnen, die sich auf ein gesamtes Netzwerk (etwa einen Campus) bezieht. Da es sich bei den Essential Science Indicators um ein Monopolprodukt handelt, liegt der Preis durchaus im erwarteten (Hochpreis-) Segment.

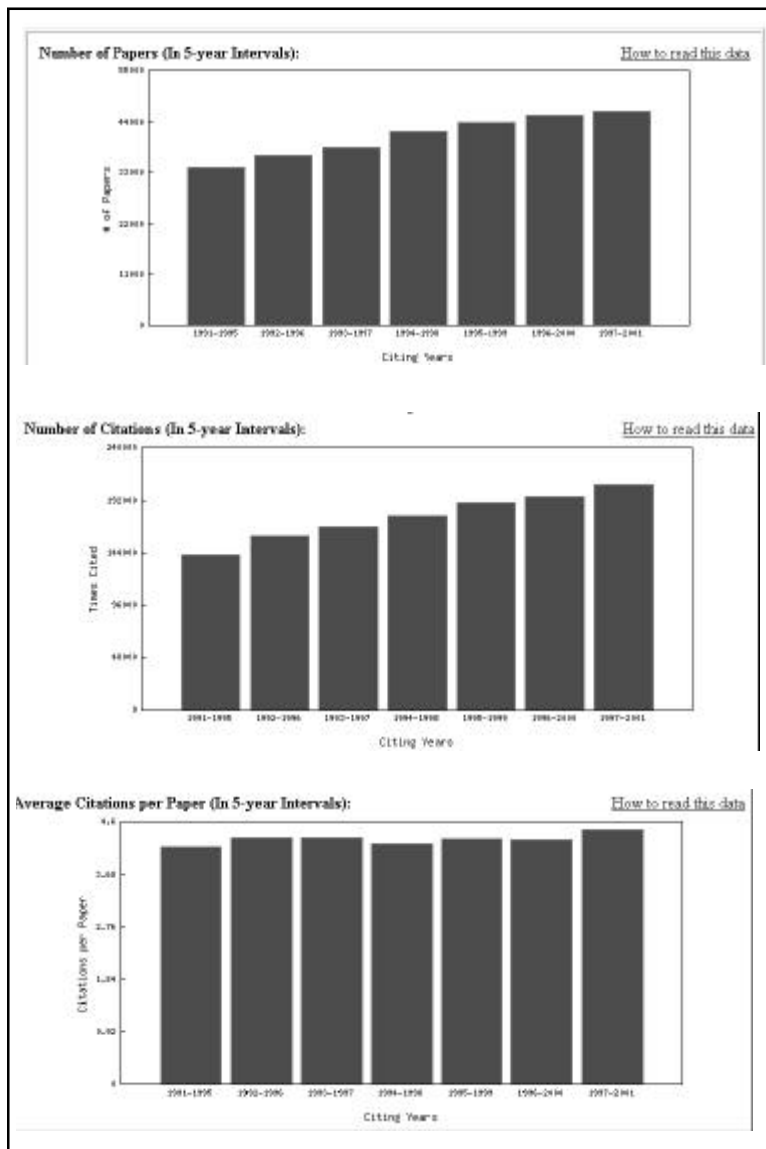


Abbildung 10: Die deutsche physikalische Forschung: Publikationszahlen, Zitationszahlen und durchschnittliche Zitationsraten

Zielgruppen

Wer braucht ESI? Zwingend notwendig scheint mir der Einsatz dieses Analyseinstruments für alle Einrichtungen der Wissenschaftspolitik, also der Forschungsministerien des Bundes und der Länder, weiterhin für Institutionen der Wissenschaftsförderung (etwa die Deutsche Forschungsgemeinschaft, die Volkswagen-Stiftung oder die Alexander von Humboldt-Stiftung) sowie für Einrichtungen der Wissenschaftsstatistik (wie die Wissenschaftsstatistik beim Stifterverband für die Deutsche Wissenschaft bzw. das Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie bei der Erstellung des "Bundesberichts Forschung"). Hier erhält die Wissen-

schaftspolitik und -förderung notwendige Informationen zum wissenschaftlichen Output, den wissenschaftlichen Leistungen und ihren Wirkungen, die sich in wissenschaftlichen Artikelpublikationen und ihren Zitationen niederschlagen.

Sinnvoll einsetzbar sind die ESI bei Großforschungseinrichtungen und Universitäten, können doch so deren Institute und Mitarbeiter ihre Stellung in der internationalen Wissenschaft verfolgen. Bei Förderanträgen lassen sich sicherlich gute Argumente aus (insbesondere positiven) ESI-Daten ableiten.

Weniger sinnvoll dürfte der Einsatz der ESI in der Wissenschaftsforschung sein. Hier werden i.d.R. detaillierte Analysemethoden eingesetzt, die den Szientometriker entweder zu Hosts mit elabo-

rierter Software (u.a. zu den ISI-Datenbanken bei DIALOG) oder - bei noch aufwändigeren Studien - direkt zur Research Services Group beim ISI in Philadelphia leiten.

Außerhalb des engeren Bereiches der Forschung und der Wissenschaftsförderung sehe ich einen Einsatzbereich bei Zeitungen und populärwissenschaftlichen Zeitschriften in den Wissenschaftsredaktionen, da die Redakteure auf diese Art Anregungen über "heiße" Forschungsthemen sowie Material zu Wissenschaftlern und Institutionen bekommen.

Fazit

Die Essential Science Indicators machen Wissenschaftsevaluation denkbar einfach. Ist man mit den Rangordnungen nach Artikeln, Ländern, Institutionen, Wissenschaftlern und Zeitschriften zufrieden, so ergeben sich Indikatoren zur evaluierenden Wissenschaftsstatistik "auf Knopfdruck". Die Wissenschaftsstatistik erhält im Rahmen ihrer Outputindikatoren quantitative Angaben darüber, was aus (nationaler oder institutioneller) Forschung herauskommt und wie dieser Output - im Weltmaßstab - "angekommen" ist.

Einige Probleme müssen angesprochen werden. Zunächst erscheint die Einteilung des Wissenschaftsbereiches in 22 Felder etwas grob, die Zuordnung der einzelnen Zeitschriften zu den Feldern zumindest fragwürdig. Schlimm ist aber, dass das Retrievalsystem keine Optionen bietet, dieses Problem zu umgehen und keine benutzerspezifische Selektion zulässt. Eine weitere Schwierigkeit bereitet das Retrievalsystem bei Rangordnungen, die über den Bezug auf Disziplinen hinausgehen. So sind derzeit Rangordnungen nach Wissenschaftlern oder Institutionen innerhalb eines Landes nicht ableitbar. Bei den Suchbildschirmen hat der Nutzer die Möglichkeit, aus zwei Retrievaloptionen auszuwählen (bei den Forschungsfeldern etwa nach Disziplin oder nach einem Term aus den Clustertiteln), er kann aber nicht beide Optionen verbinden, obgleich dies sinnvoll wäre (es geht also nicht: gewisser Term nur in einem bestimmten Feld). Der naheliegende Wunsch nach Optimierung der nutzerdefinierten Suchoptionen ist derzeit aus technischen Gründen kaum rea-

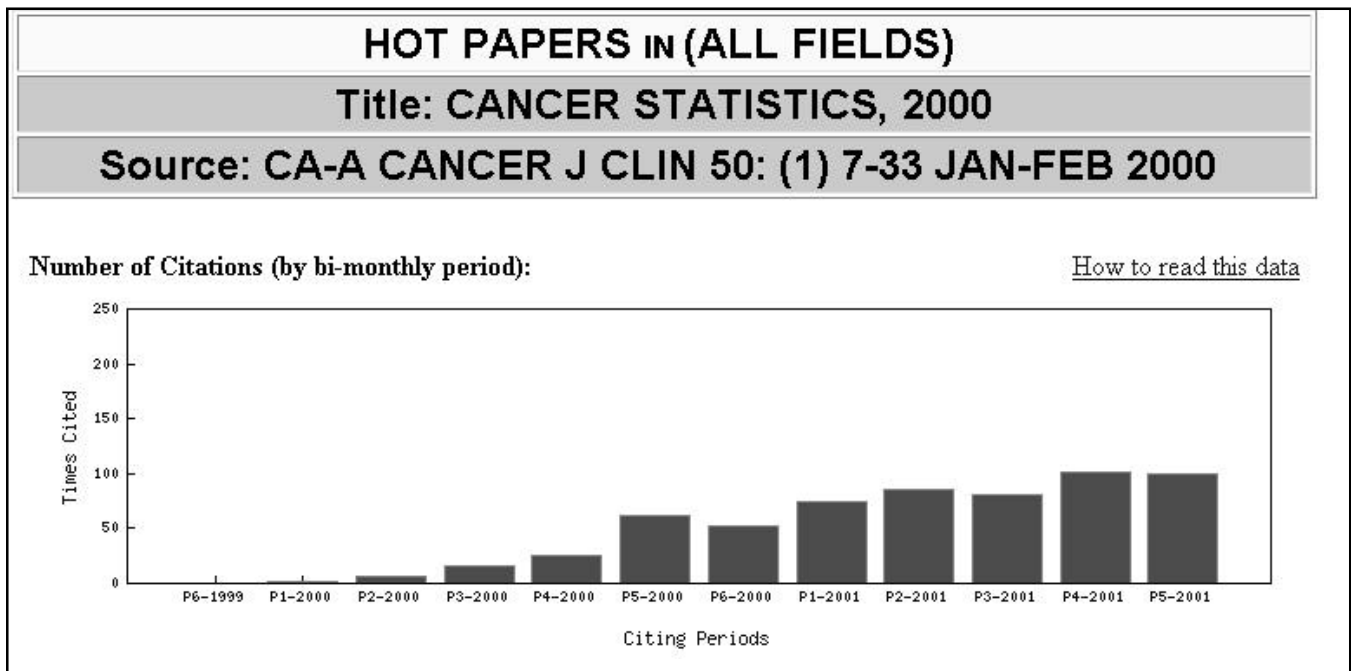


Abbildung 11: Zitationsgeschichte eines Hot Paper. Das derzeit höchstzitierte aktuelle Dokument in den ESI

RESEARCH FRONTS RANKINGS IN COMPUTER SCIENCE

Sorted by: Citations

1 - 20 (of 108) Page 1 of 6

VIEW	FRONTS	PAPERS	CITATIONS	CITATIONS PER PAPER	MEAN YEAR
1	QUANTUM ERROR CORRECTION; ENTANGLEMENT; GOOD QUANTUM ERROR-CORRECTING CODES EXIST; PERFECT QUANTUM ERROR CORRECTING CODE; QUANTUM ERROR-CORRECTING CODES SATURATING	39	3,282	84.15	1997.4
2	RADIATIVE ENERGY LOSS; NON-ABELIAN ENERGY LOSS; SMALL X	35	1,176	33.60	1999.2
3	ELEMENT-FREE GALERKIN METHOD; MESHLESS LOCAL BOUNDARY INTEGRAL EQUATION (LBIE) METHOD; MESHLESS METHODS; HYBRID BOUNDARY ELEMENT METHODS USING MODIFIED VARIATIONAL FORMULATION	24	825	34.38	1997.5

Abbildung 12: Forschungsfronten. Anzeige der Rangordnung für die Computerwissenschaft

lissierbar, da die Updateläufe (mit den Sortierläufen für die Perzentile sowie der aufwändigen Präparation der Tabellen) schon jetzt rund drei Wochen Rechenzeit benötigen. ISI hat signalisiert, die Funktionalität in Richtung komplexerer Recherchen auszubauen.

Das System arbeitet nahezu selbsterklärend. Positiv hervorzuheben sind die diversen Graphiken mit Zeitreihen zu Publikations- und Zitationsraten sowie die vielfältigen Links. Wir vergeben für den Bedienungskomfort vier (von fünf möglichen) Punkten.

In seiner Preisgestaltung geht ISI grundsätzlich ins Hochpreissegment. Das Produkt ist sehr gut und - sehr teuer. Hier müsste doch ein Spielraum nach unten drin sein! Allerdings müssen wir zugeben, dass der Aufwand bei der ESI-Produktion nicht unerheblich ist, und der Preis somit nicht überzogen erscheint. Aus Nutzersicht ist zu bedenken, dass die Erstellung ähnlicher Outputindikatoren ohne ESI weitaus teurer und zeitintensiver wird. Wenn beispielsweise eine Institution der Wissenschaftspolitik eine Studie zur Evaluation deutscher Forschung

in Auftrag gäbe, wäre ein solches Projekt auch nicht billig. Deshalb: vier Punkte.

Bei der Qualität der Inhalte müssen wir die Höchstnote vergeben. Insbesondere auf den hochaggregierten Ebenen wie Länder und Institutionen liegen valide Ergebnisse vor, mit den Hot Papers und den Forschungsfronten wird auch aktuellen Wissenschaftsentwicklungen Rechnung getragen. Die Daten der Essential Science Indicators erweisen sich als nicht umgebares empirisches Material bei der Beschreibung wissenschaftlichen Outputs, Material, das jede Maßnahme

Auf dem ersten Blick

ISI Essential Science Indicators

Preis-Leistungsverhältnis ⚡⚡⚡⚡
 Bedienungskomfort ⚡⚡⚡⚡
 Qualität der Inhalte ⚡⚡⚡⚡

der Wissenschaftsevaluation, der Wissenschaftsberichterstattung und der Wissenschaftspolitik begleiten sollte.

Wolfgang G. Stock

Weiterführende Informationen

ISI Europe
 Ryan Sheppard
 Brunel Sciences Park, Building One
 Uxbridge, UB8 3PQ
 United Kingdom
 E-Mail: Ryan.Sheppard@isinet.co.uk
 Demo-Version:
<http://www.isinet.com/demos/esi/>
 Tutorial: http://sunweb.isinet.com/tutorials/esi/wel_01.htm

Literatur

ISI: Science in Germany, 1995-99. - In: in-cites Sept. 18, 2000. - URL: www.in-cites.com/research/2000/sep-tember_18_2000-1.html.

ISI: ISI Essential Science Indicators Identifies Research Excellence in Germany. - Philadelphia: ISI, May 29, 2001. - URL: www.isinet.com/isi/news/2001/research/8115207/index.html.

ISI: Essential Science Indicators. Quick Reference Card. - Philadelphia: ISI, o.J. [2001].

Wolfgang G. Stock: JCR on the Web. Journal Citation Reports: Ein Impact Factor für Bibliotheken, Verlage und Autoren? - In: Password Nr. 5 (2001), S. 24-39.

Wolfgang G. Stock: Web of Science. Ein Netz wissenschaftlicher Informationen - gesponnen aus Fußnoten. - In: Password Nr. 7+8 (1999), S. 21-25.

Helen Szigeti: Explore the scientific landscape. - In: ISI KnowledgeLink 2001. - URL: www.isiknowledgeline.com/explore.html.

Bielefeld

Karl-Wilhelm Neubauer geht

Bielefeld 2000 Conference zu "Wissensmanagement - Eine neuer Hype der Informatiker?" Mit Dr. Karl-Wilhelm Neubauer, der auf der Bielefeld 2000 Conference verabschiedet wurde, hat nach 16 Jahren als Leiter der Universitätsbibliothek Bielefeld ein weiterer Pionier der Branche die Informationsszene verlassen. Neubauer, 1939 in Theben in der heutigen Slowakei geboren, baute nach Studium und Bibliotheksausbildung die Zeitschriftendatenbank auf, die heute die Grundlage für alle Dokumentenliefererdienste für Zeitschriftenartikel in Deutschland bildet. In Bielefeld begann er 1996 das erste deutsche Bildungsprojekt zur elektronischen Informationsversorgung. Seine alle zwei Jahre stattfindenden von vornherein international angelegten Konferenzen informierten teilweise als erste die deutschen Bibliotheken über neue weltweite Trends, beispielsweise über die Suchmaschinen für das Internet.

Für die letzte von ihm verantwortete Konferenz lässt sich Ähnliches leider nicht sagen. Zwar gab es eine Reihe hochinteressanter Vorträge und wurden auch genügend Strategien vorgestellt, die die Informationsversorgung für alle gewährleisten sollen. Leider wurde zu der Qualität der Informationen und vor allem zu den Kosten zu wenige konkrete Aussagen gemacht. Die selbstgesteckten Ziele der Konferenz - die Teilnehmer über den neuesten Stand der Technik zu informieren, Vorschläge für weitere Entwicklungen vorzubereiten und den künftigen Bedarf für die Nutzung elektronischer Dienstleistungen und Publikationen in und durch Bibliotheken zu besimmen - wurden nicht eingelöst.

Für eine weitergehende Qualitätssicherung der universitären Informationsversorgung.

Aus den Vorträgen: ● **Prof. Dr. Dr. Hermann Maurer** von der Technischen Universität Graz wies auf die neuen Möglichkeiten der Mobilkommunikation hin. Mit dem im Internet verfügbaren Wissen werden wir demnächst überall ein "Zusatz-

hirn" mit uns führen. ● **Dr. Friedrich Bode** vom **Ministerium für Schule, Wissenschaft und Forschung des Landes Nordrhein-Westfalen** forderte eine weitergehende Qualitätssicherung der universitären Informationsversorgung. An konkreten Maßnahmen schlug er vor: einen Ausbau des Angebote elektronischer Inhalte (Teilnahme an Konsortien!) sowie eine Optimierung des Informationssystems durch eine bessere Zusammenarbeit von Fakultäten und auch externen Instituten beim Aufbau spezifischer Informationsangebote wie etwa der Integration der digitalen Informationsversorgung in E-Learning-Systeme. ● Der Repräsentant von **Elsevier** verteidigte die Preispolitik seines Konzerns, indem er zwar zugab, dass elektronische Produkte teurer als die herkömmlichen Printprodukte seien. Gleichwohl seien sie kostensparender, da bei den elektronischen Medien die Kosten für Buchbindearbeiten, Katalogisierung und Archivierung weitgehend entfielen. Diese Argumentation schien die anwesenden Bibliothekare nicht zu überzeugen. ● Der inhaltliche Schwerpunkt "Informationsversorgung im Internet und die Rolle der Suchmaschinen" bestand mehr aus Product Reviews etwa der von großen Volltextanbietern wie Ingenta bereitgestellten Suchmaschinen. Die erhoffte Suchmaschine, die mit größtem Recherche-Komfort möglichst relevante Ergebnisse liefert, wurde auch hier nicht gefunden.

Fortsetzung von Seite 3

Informationspolitik

Derzeit wird unter dem Titel "Wissen schafft Märkte" eine **Verwertungs-offensive** der Bundesregierung, die die Hochschulen in die Lage versetzen soll, zukünftig alle an der Hochschule entstehenden Erfindungen in Anspruch zu nehmen, zu patentieren und zu verwerten. Das derzeit noch dagegenstehende "Hochschullehrerprivileg" entfällt Bis Ende 2003 werden Hochschulen und außeruniversitäre Forschungseinrichtungen in die Lage versetzt, ihr Patent- und Verwertungsgeschäft über professionelle Patent- und Verwertungsagenturen (PVA) abzuwickeln.

Anmerkung: Soll Password diese Ausführungen in einem Satz bewerten? Der Mann sagt etwas aus.