esss european summer school for scientometrics

Scientometric Indicators in Use: An Overview

Sybille Hinze[1], Wolfgang Glänzel[2]

[1]Institute for Research Information and Quality Assurance (iFQ), Berlin, Germany
[2]Centre for R&D Monitoring and Dept. MSI, KU Leuven, Belgium

Berlin, September 2013

www.scientometrics-school.eu

esss european summer school for scientometrics

Structure of the presentation

1. Introduction
2. Types of indicators
3. Bibliometric indicators
    i. Document types
    ii. Counting schemes
    iii. Subject classification
    iv. Citation based indicators

esss european summer school for scientometrics

Introduction

- Indicators = Proxies
  - represent a highly complex reality
  - empirically ascertainable variables and factors, that are used to reflect aspects that cannot be directly measured
  - terms are rather vaguely defined – quality, performance, progress, usefulness, importance…
- Accommodate the need for “objective“ data but, also the interest to better understand developmental processes and contexts of science itself
  - Indicators used as analytical tools but, also as information to inform science policy decisions

esss european summer school for scientometrics

Introduction

„on the applied side, the demand from science policy for ‚objective' data and specific manipulations of data (‚science and technology indicators') is a continuous driving force for the development of quantitative studies of science and technology. (…) on the basic side, science is a complicated system of knowledge production and knowledge exchange, and the use of empirical methods in which sophisticated data-collection and data-handling techniques play a substantial role, is undoubtedly a prerequisite for the advancement of our understanding"

(van Raan 1988, p. 1)

esss european summer school for scientometrics

Introduction

Systematic development of *quantitative and evaluative science studies* since the mid-20th centruy

- Introduction of the term “bibliometrics” by Pritchard, 1969:

  “the application of mathematical and statistical methods to books and other media of communication”

- Introduction of the term “scientometrics by Nalimov & Mulchenko, 1969:

  “the application of those quantitative methods which are dealing with the analysis of science viewed as an information process”

esss european summer school for scientometrics

Introduction

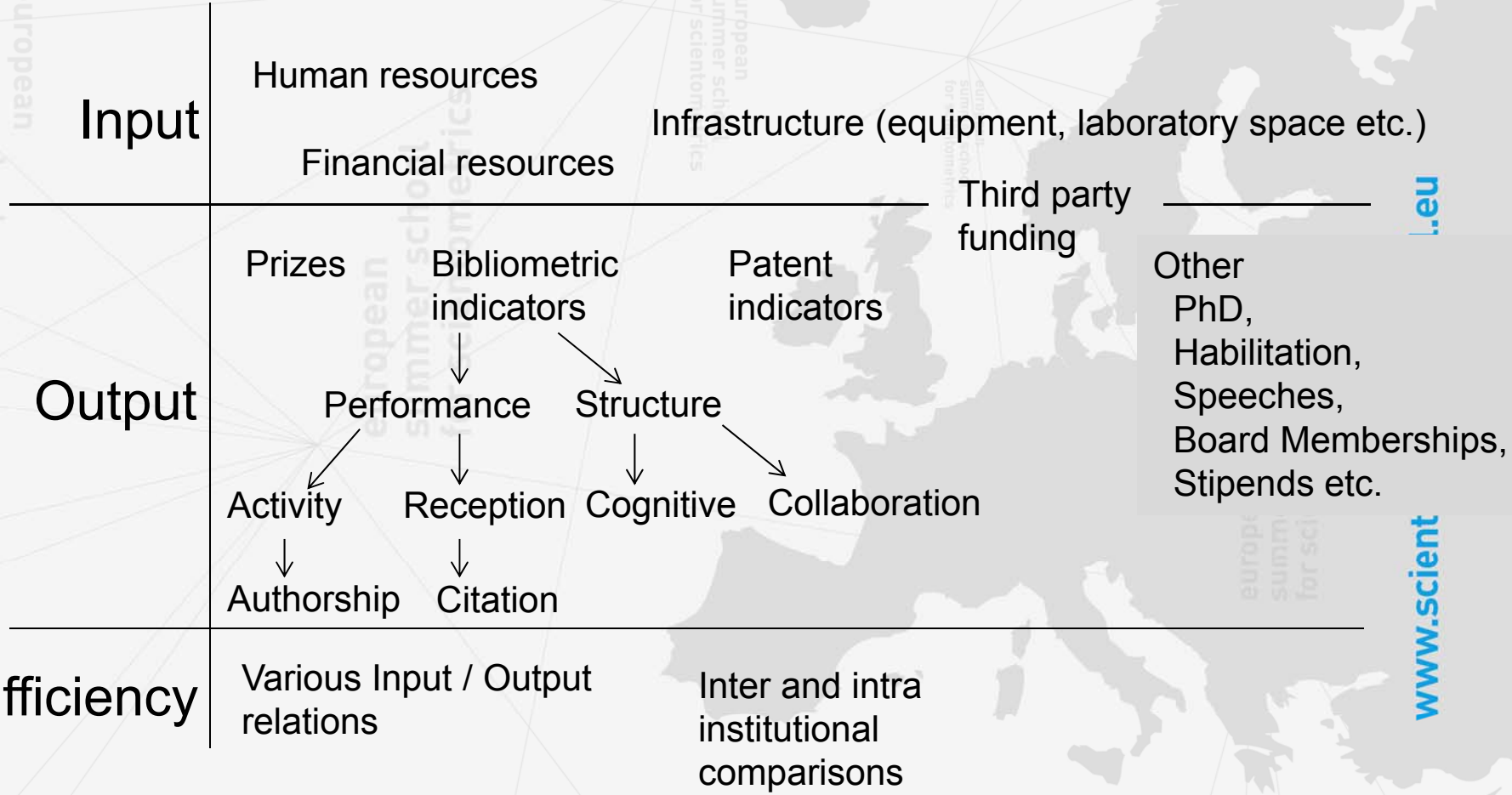
Scientometrics / Bibliometrics

- depicts essential aspects of scientific activities by *quantitative* and *statistical* methods, and its output proved to be a valuable supplement to qualitative methods such as peer reviews

- has developed tools to quantify that part of research output, which is documented in the framework of scholarly communication

esss european summer school for scientometrics

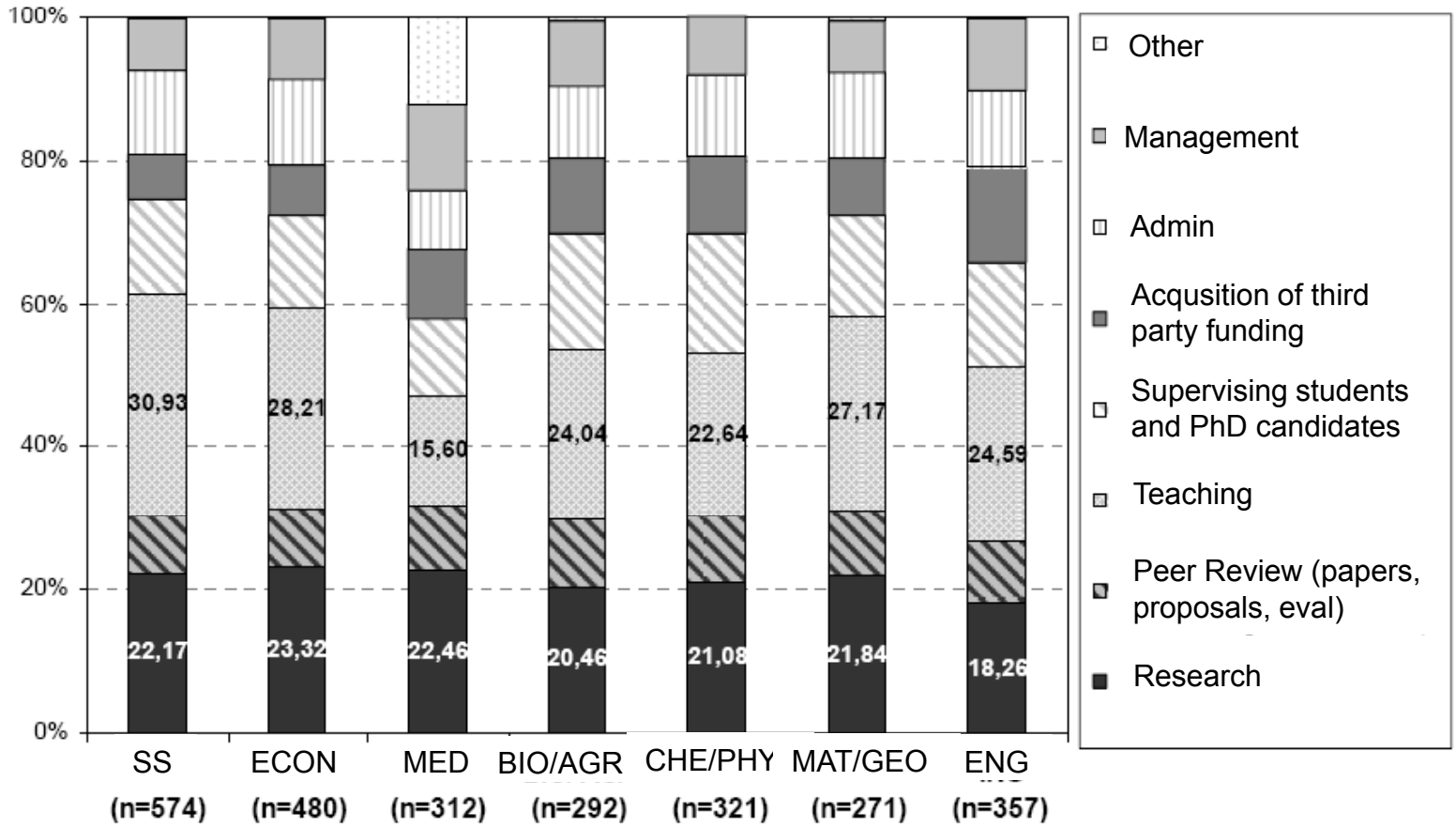
Types of indicators



Based upon Hornbostel 1999, p. 59

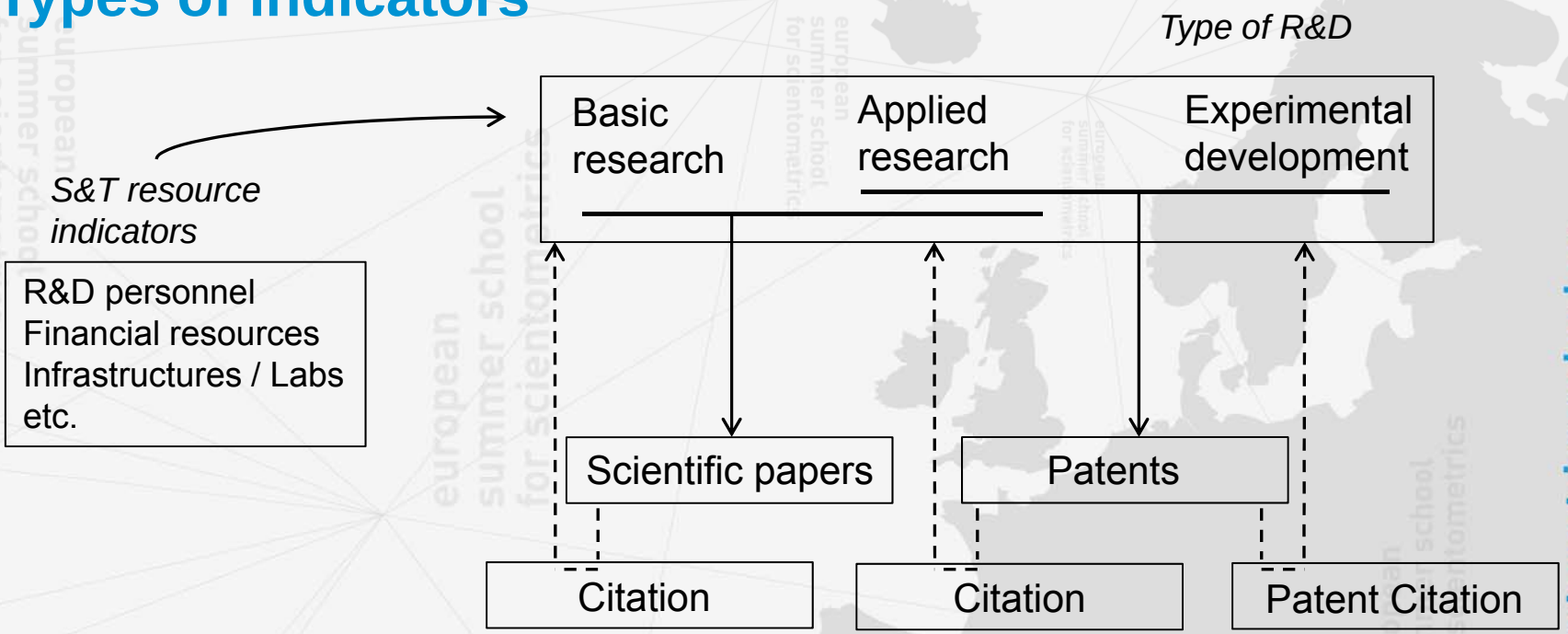
esss european summer school for scientometrics

Types of indicators



Böhmer et al. 2011, p. 129

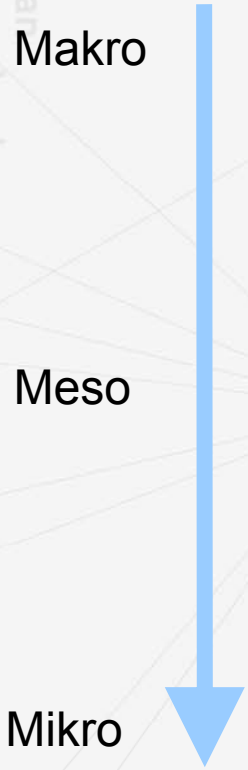
Types of indicators



Based upon Schmoch et al. 1996, p.118

esss european summer school for scientometrics

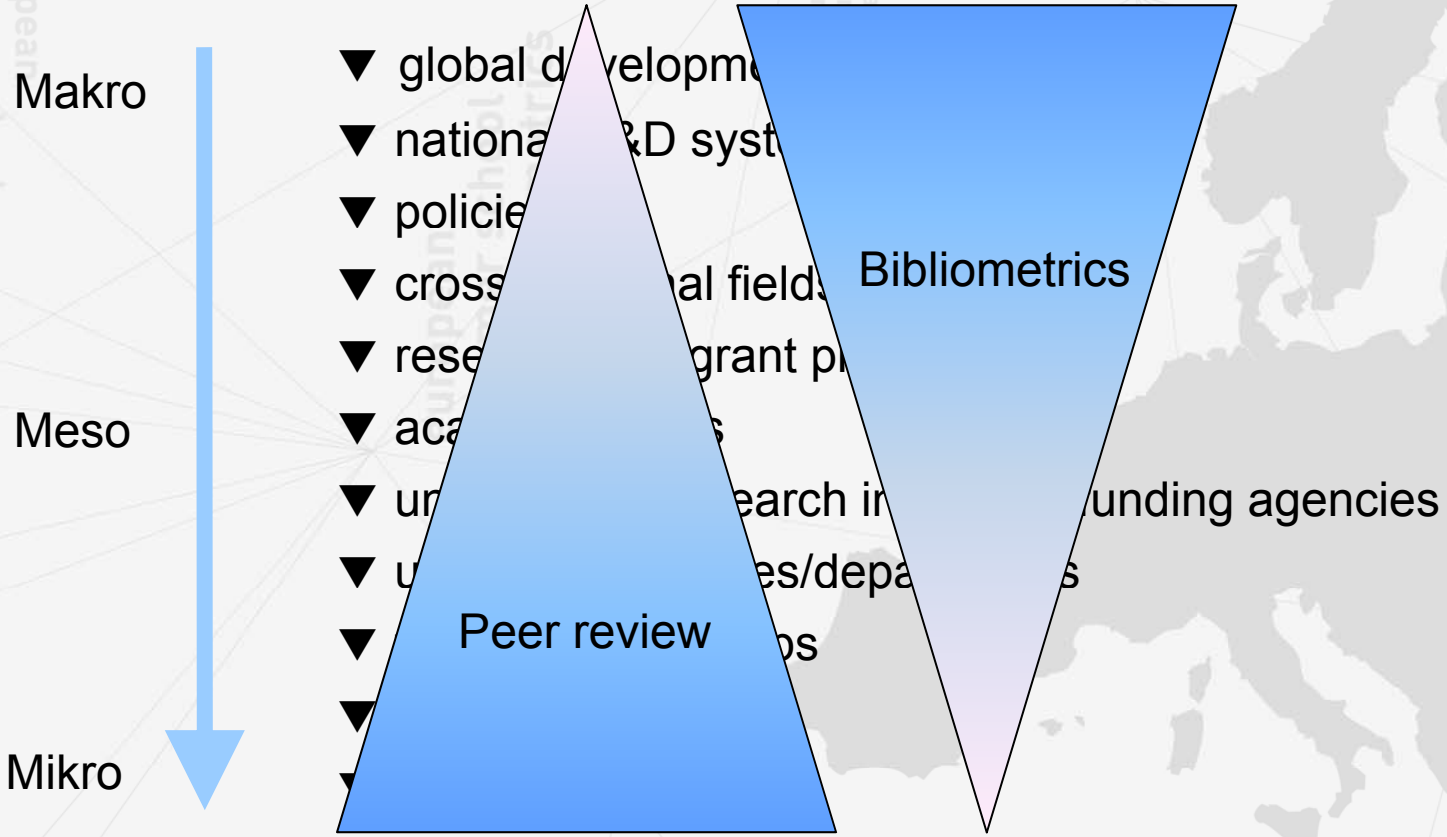
Types of indicators



- ▼ global developments
- ▼ national R&D systems
- ▼ policies
- ▼ cross-sectional fields
- ▼ research and grant programs
- ▼ academic fields
- ▼ universities, research institutes, funding agencies
- ▼ university institutes/departments
- ▼ target/status groups
- ▼ research groups
- ▼ individuals

esss european summer school for scientometrics

Types of indicators



www.scientometrics-school.eu

esss european summer school for scientometrics

Bibliometric indicators

- Productivity / Activity: publication output
- Collaboration: co-authorship
- Reception / Impact: citation rates
- Cognitive structures: co-occurrences of words / classifications / citations

or the combination thereof.

Most indicators are derived from simple counts of items extracted from various bibliographies and databases. Advanced measures are “network indicators” derived from the analysis of co-authorship-, citation-, co-word- etc. networks.

esss european summer school for scientometrics

Bibliometric indicators – Document types

Data sources: usually papers published in periodicals and serials.

Only conveyors of original scientific information are included.
These are considered as citable items.

Citable items = *research articles*, *short communications* and *notes*, *letters*, *reviews*, and *proceedings papers*.

Book reviews, editorials, corrections/errata, meeting abstracts and reprints are not considered original research output.

Bibliometric indicators

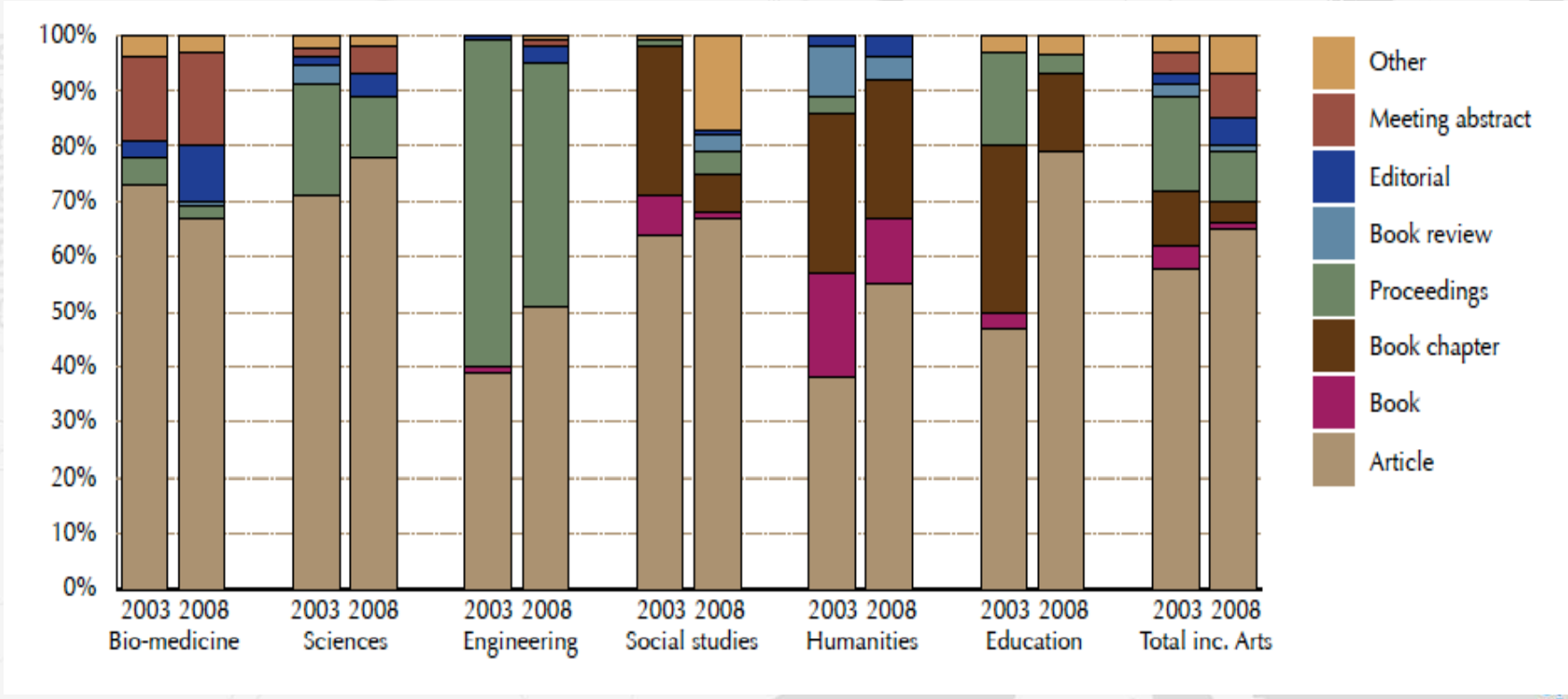
Disciplinary coverage in the ISI citation indices

EXCELLENT (> 80%)	**Good (60-80%)**	**Good (40-60%)**	**MODERATE (<40 %)**
Molecular biology & biochemistry	Applied physics & chemistry	Mathematics	Other social sciences
Biological sciences primarily related to humans	Biological sciences primarily related to animals and plants	Economics	Humanities and arts
Chemistry	Psychology and psychiatry	Engineering	
Clinical medicine	Geosciences		
Physics & astronomy	Other social sciences primarily related to medicine and health		

Source: Moed, H., Evaluation of Research Performance and Funding Programme in Social Sciences. At: Norface Workshop on Research Programme Development and Management. 7.Feb.2006, Bonn

esss european summer school for scientometrics

Bibliometric indicators



Source: Fry et al. 2009, Communicating knowledge: how and why researchers publish and disseminate their findings. https://dspace.lboro.ac.uk/dspace-jspui/bitstream/2134/5465/1/Communicating-knowledge-report.pdfb

esss european summer school for scientometrics

Bibliometric indicators – Document types

Data sources: usually papers published in periodicals and serials.

Only conveyors of original scientific information are included.
These are considered as citable items:

= *research articles*, *short communications* and *notes*, *letters*, *reviews*, and *proceedings papers*.

Book reviews, editorials, corrections/errata, meeting abstracts and reprints are not considered original research output.

esss european summer school for scientometrics

Bibliometric indicators

National publication counts and percentages by document type – Web of Science (2007)

Country	All papers	Article	Letter	Review	Book review	Editorial	Meeting abstract	Rest
USA	392.488	66,5%	2,4%	5,0%	0,6%	5,4%	19,4%	0,7%
UK	104.561	65,9%	4,6%	5,9%	1,2%	5,5%	16,2%	0,6%
Germany	95892	72,3%	1,7%	4,6%	0,1%	3,2%	17,4%	0,6%
China PR	95231	92,0%	0,6%	1,5%	0,0%	0,7%	4,8%	0,3%
Japan	89575	78,8%	1,5%	2,6%	0,0%	1,2%	15,4%	0,4%
France	63656	77,6%	2,0%	4,4%	0,1%	2,8%	12,5%	0,6%
Canada	57500	71,7%	2,1%	4,9%	0,5%	3,9%	16,2%	0,6%
Italy	55223	72,7%	3,5%	4,6%	0,1%	2,6%	16,0%	0,5%
Spain	41274	75,9%	3,2%	4,0%	0,1%	2,5%	13,8%	0,5%
Australia	35327	72,4%	3,3%	5,9%	0,6%	4,0%	13,3%	0,5%
India	32842	86,4%	3,5%	2,6%	0,0%	2,1%	4,7%	0,6%
World total	1299678	68,6%	2,8%	3,7%	0,5%	4,8%	16,6%	2,9%

Source: Zhang et al. 2011

esss european summer school for scientometrics

Counting schemes

= method according to which publications are to be assigned to contributing units

- The *fractional counting scheme*:

  if *n* units (authors, institutions, countries, etc.) have contributed to the paper in question, each contributing unit takes the value 1/*n* for this paper (partially additive)

- The *first address count*:

  a paper is assigned to one unit only, on the basis of the first address in the address list of a paper (additive)

- The *full or integer counting scheme:*

  assigns a co-publication fully to each contributing unit (non-additive)

esss european summer school for scientometrics

Counting schemes

Nowadays only *fractional and full counting* are used while the first-address count is obsolete; in the past it was used due to the coverage of bibliographic databases, which as a rule recorded only one address.

Databases providing incomplete author address information are not appropriate for bibliometric analyses.

Counting schemes - Example

SCI CDE with Abstracts (Jan 93 - Jul 93) (D4.0)

Authors: Prassides-K Kroto-HW Taylor-R Walton-DRM David-WIF Tomkinson-J Haddon-RC Rosseinsky-MJ Murphy-DW
Title: Fullerenes and Fullerides in the Solid-State - Neutron-Scattering Studies
Source: CARBON 1992, Vol 30, Iss 8, pp 1277-1286
Address: UNIV-SUSSEX, SCH CHEM & MOLEC SCI, BRIGHTON BN1-9QJ, E-SUSSEX, ENGLAND
RUTHERFORD-APPLETON-LAB, DIDCOT OX11-0QX, OXON, ENGLAND
AT&T-BELL-LABS, MURRAY-HILL, NJ07974, USA

Source: Glänzel, *Bibliometrics as a Research Field*, 2003

esss european summer school for scientometrics

Example: Counting options by level of aggregation

- 9 co-authors, 3 institutions, 2 different countries

	Full Count	Fractional Count
Prassides K	1	0,111
Kroto HW	1	0,111
…	…	…
Murphy DW	1	0,111
UNIV SUSSEX	1	0,333
RUTHERFORD APPLETON LAB	1	0,333
AT&T BELL LAB	1	0,333
United Kingdom	1	0,5
USA	1	0,5

- Apply fractional counting only within the same level of aggregation, otherwise inconsistencies will arise e.g. national versus supra-national level.

Source: Glänzel 2011

esss european summer school for scientometrics

Example: Counting options by level of aggregation

Example: national versus supra-national level analysis

- 3 addresses, 3 countries: DE / FR / USA

  National level = 1/3 for each country

  Supra-national level (EU versus USA)

  2/3 : 1/3 (based on sum of individual country contributions)

  1/2 : 1/2 (delimitation of the regions)

Source: Glänzel 2011

esss european summer school for scientometrics

Counting schemes

Shares of publications by country and counting method

Country	Full count	Fractional count
USA	28.0 %	23.9 %
Germany	7.3 %	5.4 %
United Kingdom	7.8 %	4.9 %
France	5.2 %	3.9 %
China	11.0 %	9.9 %
Other	36.1 %	27.3 %
Total	126.3 %	100.0 %

Data source Web of Science, calculations by Fraunhofer ISI 2011

esss european summer school for scientometrics

Subject classification

- Disciplines / fields analysis based on classification schemes
  - Specialized databases often offer hierarchical subject classification at the document level e.g. Medline (MeSH – Medical Subject Headings), Mathematical Reviews (MSC - Mathematics Classification System)
  - Multidisciplinary databases often assign whole journals to their subject classification scheme

Due to multiple assignments of papers / journals to classification categories indicators are not additive over subject categories, sub-fields, fields etc.

esss european summer school for scientometrics

Bibliometric indicators

Their use of bibliometric data to inform science policy

- National and international reporting
- Research Evaluation and Monitoring
- Identifying research priorities

www.scientometrics-school.eu

National and international reporting


National Science Board

SCIENCE AND

Contents

Acronyms and Abbreviations x

About Science and Engineering Indicators xii
SEI's Different Parts xii
Presentation xiii

Overview O-1
Introduction O-3
A Bird's Eye View of the World's Changing S&T Picture O-3

Undergraduate Education, Enrollment, and Degrees in the United States 2-11
Graduate Education, Enrollment, and Degrees in the United States 2-17
Postdoctoral Education 2-30
International S&E Higher Education 2-31
Conclusion 2-37
Notes 2-38
Glossary 2-38
References 2-39

• vii

Research Outputs: Journal Articles and Patents

Expanding International Research Collaborations

New Research Patterns Reflected in World's Citations Base

[4] STC key figures report 2008/2009

Table of contents

Chapter 3. The scientific and technological outputs of R&D activities and their high-tech outcomes

3.1 Has the EU increased its efficiency in producing scientific publications since 2000?

3.2 Has the EU's inventiveness, as measured by patent applications, improved since 2000?

3.3 Has the EU moved towards a more knowledge-intensive economy since 2000?

National and international reporting

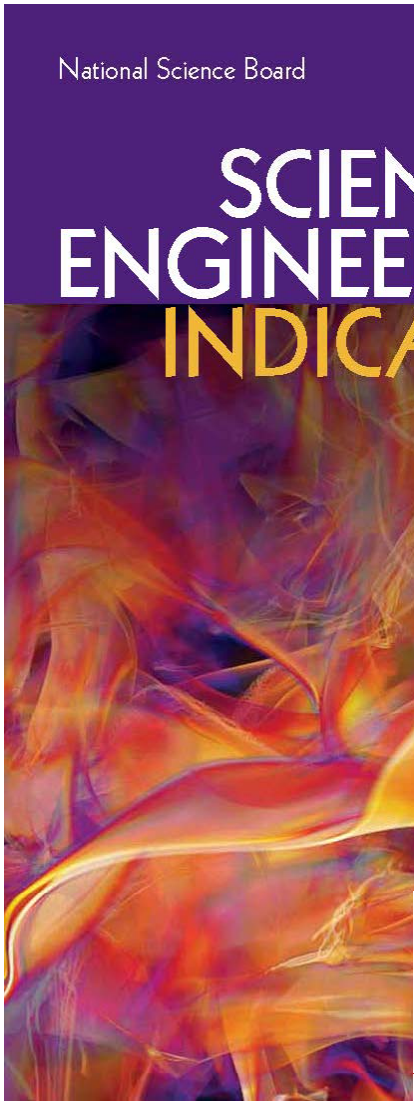
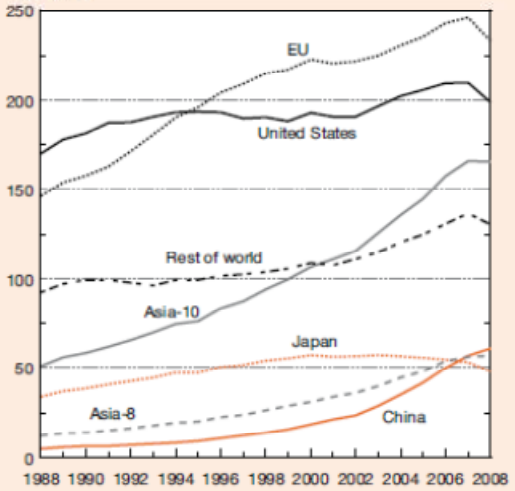


Figure O-13
S&E journal articles produced by selected regions/ countries: 1988–2008



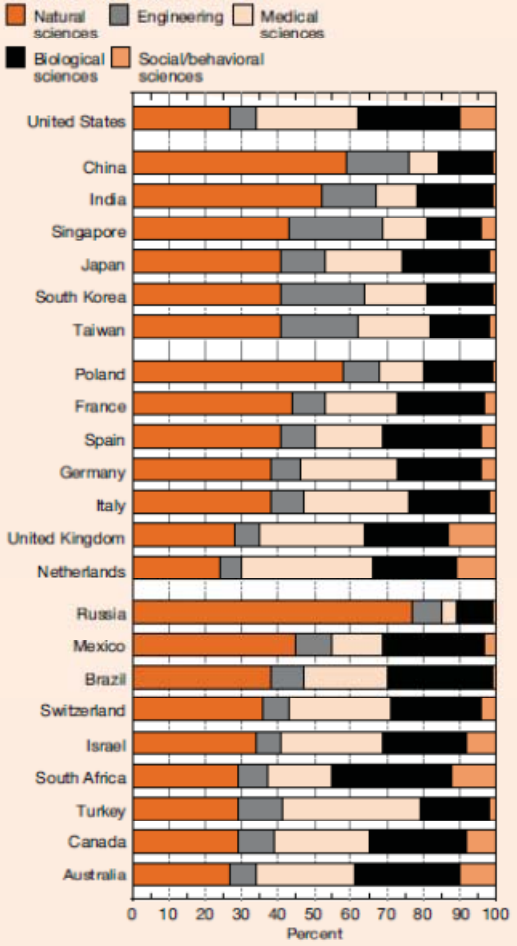
EU = European Union

NOTES: See glossary for countries included in Asia-8 and Asia-10. EU includes all 27 member states. Articles classified by year of publication and assigned to region/country on basis of authors' institutional address(es). For articles with collaborating institutions from multiple countries/ economies, each country/economy receives fractional credit on basis of proportion of its participating institutions. Counts for 2008 are incomplete.

SOURCES: Thomson Reuters, Science Citation Index and Social Sciences Citation Index, http://thomsonreuters.com/products_services/science/; The Patent Board™; and National Science Foundation, Division of Science Resources Statistics, special tabulations.

Science and Engineering Indicators 2010

Figure O-14
Field shares of research articles for selected countries/economies: 2007



NOTE: Natural sciences include astronomy, chemistry, physics, geosciences, mathematics, and computer sciences.

SOURCES: Thomson Reuters, Science Citation Index and Social Sciences Citation Index, http://thomsonreuters.com/products_services/science/; The Patent Board™; and National Science Foundation, Division of Science Resources Statistics, special tabulations.

Science and Engineering Indicators 2010

Source: NSF Science and Engineering Indicators 2010

National and international reporting

28 A 7 ESTABLISHING EXTENSIVE EVALUATION RESEARCH

30 B CORE TOPICS 2010

30 B 1 THE GERMAN R&I-SYSTEM IN AN INTERNATIONAL COMPARISON

46 B 2 THE BOLOGNA PROCESS – AN INTERIM ASSESSMENT

54 B 3 RESEARCH AND INNOVATION IN EASTERN GERMANY

65 B 4 ELECTROMOBILITY

76 B 5 CURRENT DESIGN OF THE PATENT SYSTEM

84 C STRUCTURE AND TRENDS

86 C1 EDUCATION AND QUALIFICATIONS

93 C 2 RESEARCH AND DEVELOPMENT

101 C 3 INNOVATION BEHAVIOUR IN THE GERMAN PRIVATE SECTOR

107 C 4 NEW ENTERPRISES

111 C 5 PATENTS IN INTERNATIONAL COMPETITION

116 C 6 CIENTIFIC PUBLICATIONS AND PERFORMANCE

120 C 7 PRODUCTION, VALUE CREATION AND EMPLOYMENT

Schmoch/Mallig/Michels/Neuhäusler/Schulze: Performance and Structures of the German Science System in an International Comparison 2010 with a Special Analysis of Public Non-university Research Institutions.Studien zum deutschen Innovationssystem Nr. 8-2011

Productivity and developmental dynamics

Publication shares of Countries and Regions 2000 to 2009 (Index 2000 = 100)

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
US	100	99	99	98	98	96	94	92	89	87
JP	100	99	99	98	94	90	86	81	76	73
DE	100	100	100	97	96	95	93	91	89	90
GB	100	97	9	93	91	90	90	89	85	84
FR	100	99	98	96	93	93	92	89	90	90
CH	100	97	97	99	102	101	104	103	102	105
CA	100	99	100	103	104	108	109	108	108	108
SE	100	103	102	98	97	96	64	92	88	89
IT	100	104	105	109	110	110	109	112	111	112
NL	100	100	103	103	103	106	105	104	104	107
FI	100	104	101	100	99	95	97	94	92	91
KR	100	117	126	142	159	165	170	167	182	194
CN	100	118	129	150	176	203	233	247	264	296
EU15	100	100	99	98	97	97	96	95	93	93
EU27	100	100	100	99	98	97	97	97	96	96
World	100	100	100	100	100	100	100	100	100	100

Data web of science, Source: Schmoch et al.: Performance and Structures of the German Science System in an International Comparison 2010 with a Special Analysis of Public Non-university Research Institutions. Studien zum deutschen Innovationssystem Nr. 8-2011, p. 13

esss european summer school for scientometrics

Productivity and developmental dynamics

Sharpe Ratio – Adjusted Growth Rate

— growth indicator borrowed from financial economics/stock market analysis

$$BW = (W_F - W_G) / S_{WF}$$

W_F = Growth of a field

W_G = Growth of all fields

S_{WF} = Standard deviation of growth of a field

Based on Fischer 2001, p. 271

esss european summer school for scientometrics

Specialization / Profiles

Activity Index - AI

$$AI = \left[\left(P_{ij} / \sum_i P_{ij} \right) / \left(\sum_j P_{ij} / \sum_{ij} P_{ij} \right) \right]$$

Revealed Literature Advantage - RLA

$$RLA = 100 \tanh \ln \left[\left(P_{ij} / \sum_i P_{ij} \right) / \left(\sum_j P_{ij} / \sum_{ij} P_{ij} \right) \right]$$

P_{ij}= No. of Publications of a country i in field j

$\Sigma_i P_{ij}$= No. of Publications of all countries in field j

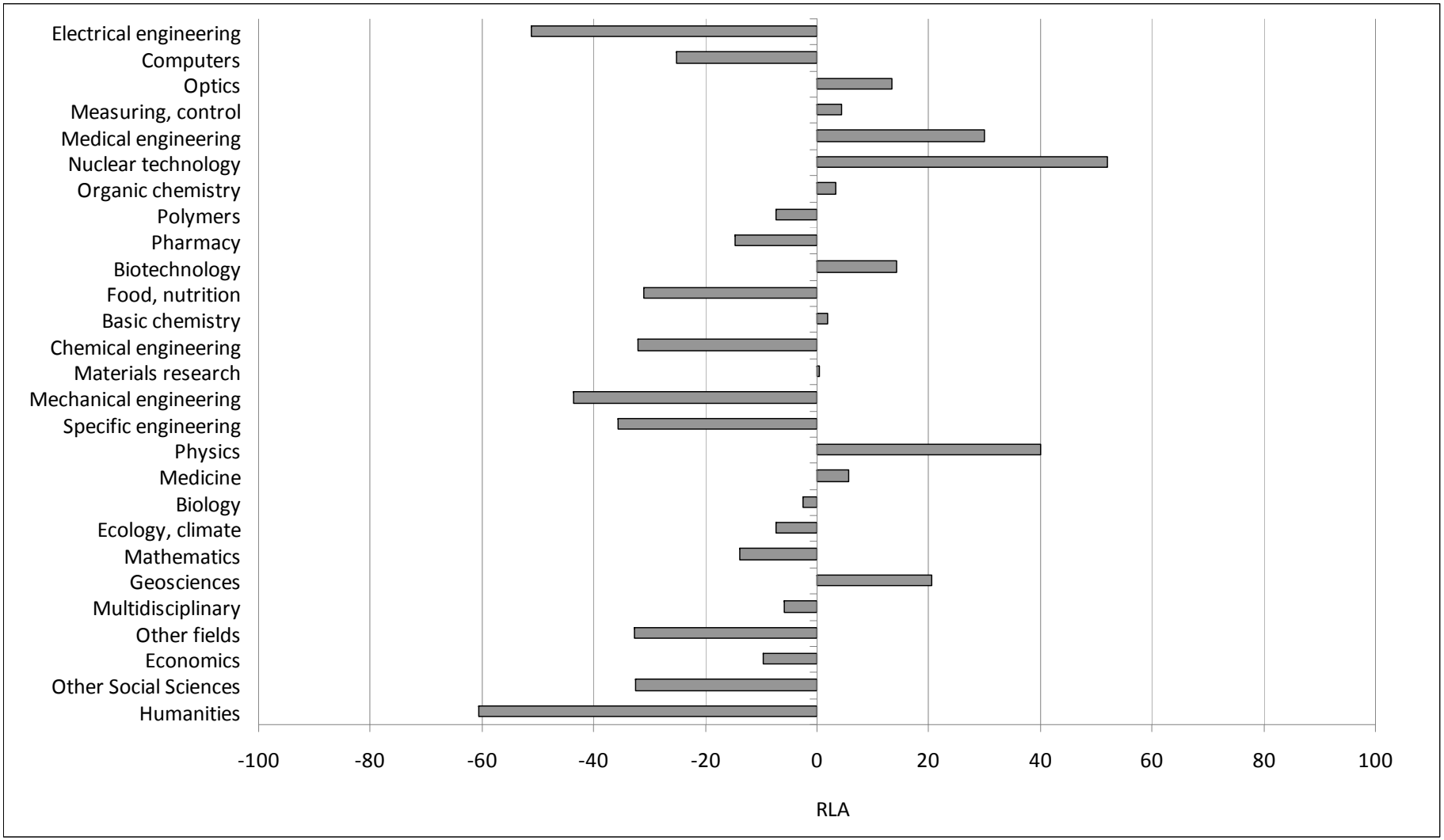
$\Sigma_j P_{ij}$= No. of all Publications of country i

$\Sigma_{ij} P_{ij}$= No. of all Publications of all countries

Specialization / Profiles

Revealed Literature Advantage - RLA

Publications in WoS differentiated by Science Fields, 2009



Source: Schmoch et al.: Performance and Structures of the German Science System in an International Comparison 2010 with a Special Analysis of Public Non-university Research Institutions. Studien zum deutschen Innovationssystem Nr. 8-2011, p. 13

esss european summer school for scientometrics

Collaboration

- **Co-authorships used as proxi**
  - Partial indicator due to the fact that neither all collaboration leed to a common publication nor all co-authored papers are based on collaboration (see Katz & Martin 1997)
  - But, the higher the level of collaboration the better the approximation by „Co-publication“ (see Glänzel & Schubert 2004)

esss european summer school for scientometrics

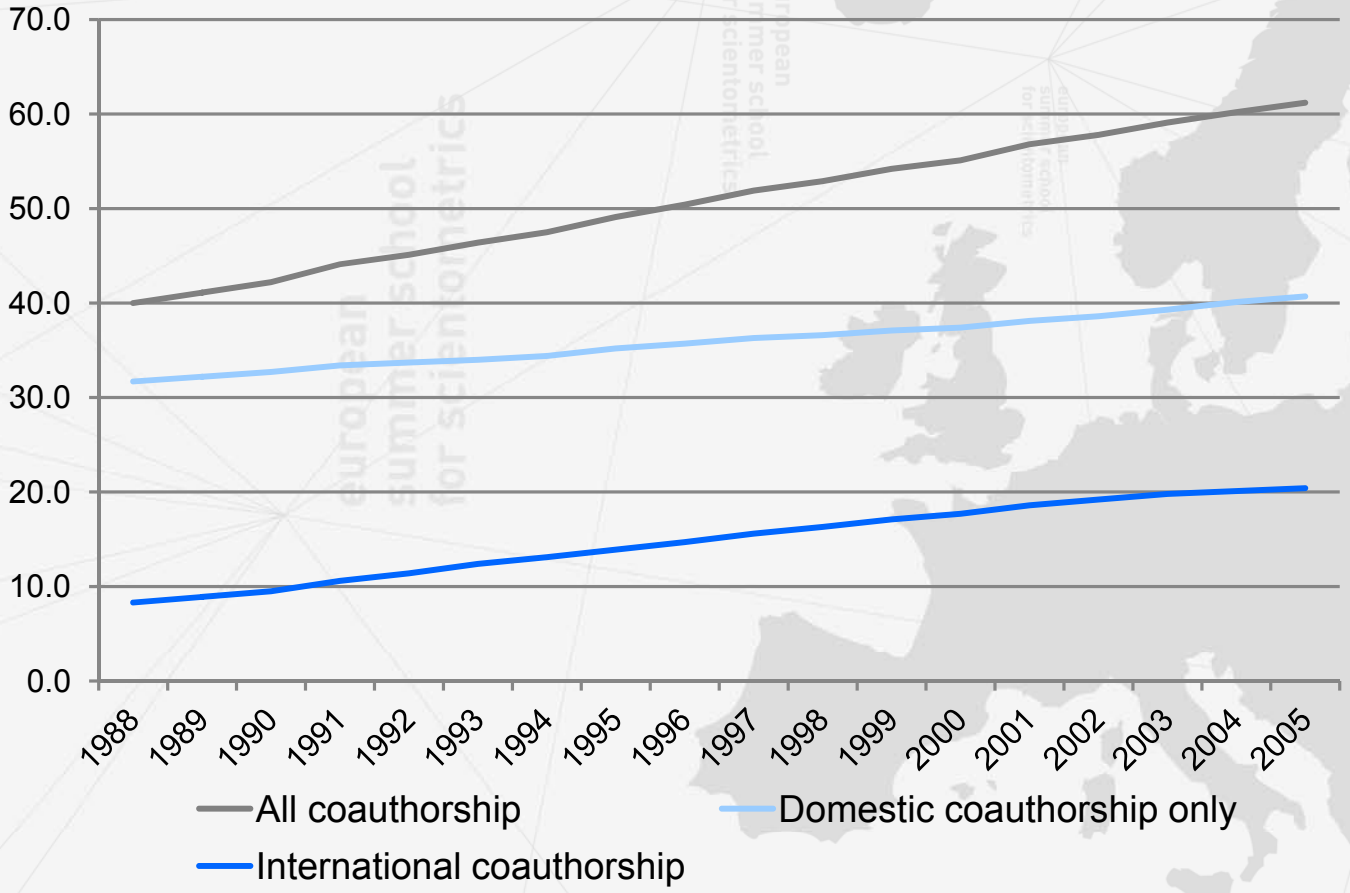
Collaboration

- **Aggregation levels used in co-authorship analysis**
  - Individual authors
  - Institutions - domestic
  - International collaboration – institutions and countries
  - Collaboration between sectors

esss european summer school for scientometrics

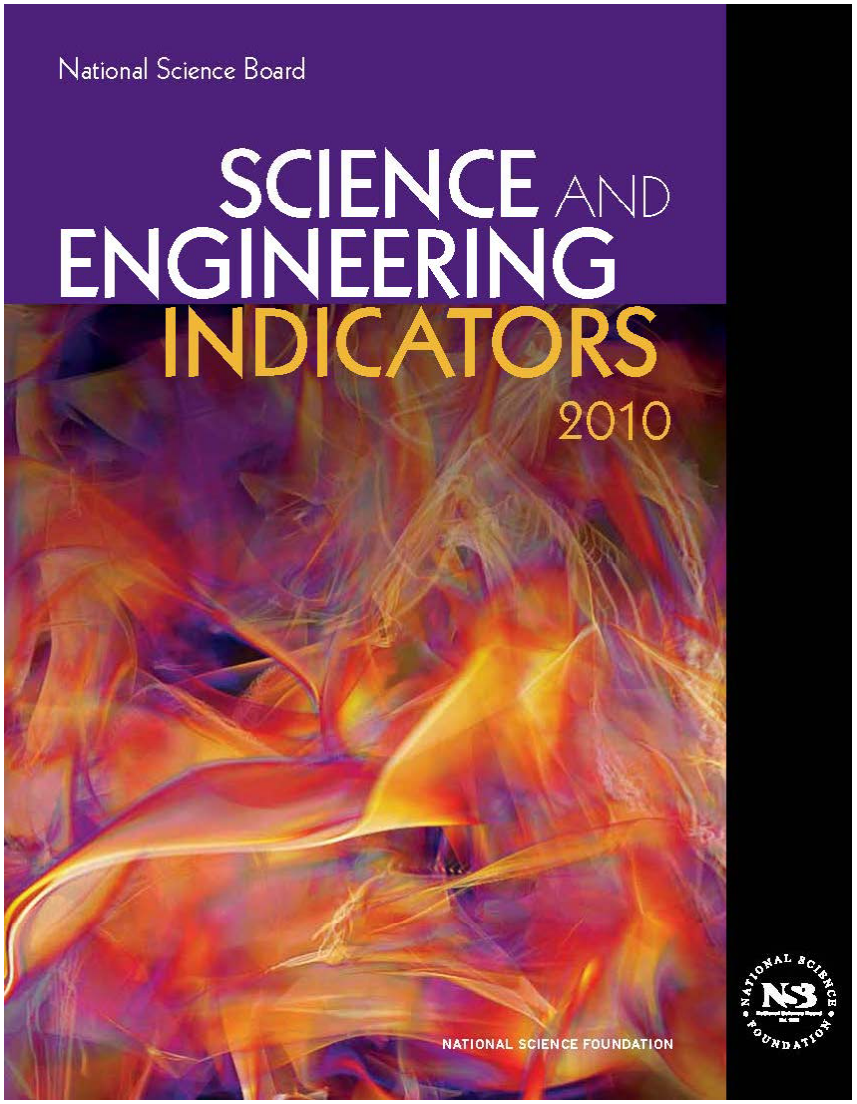
Collaboration

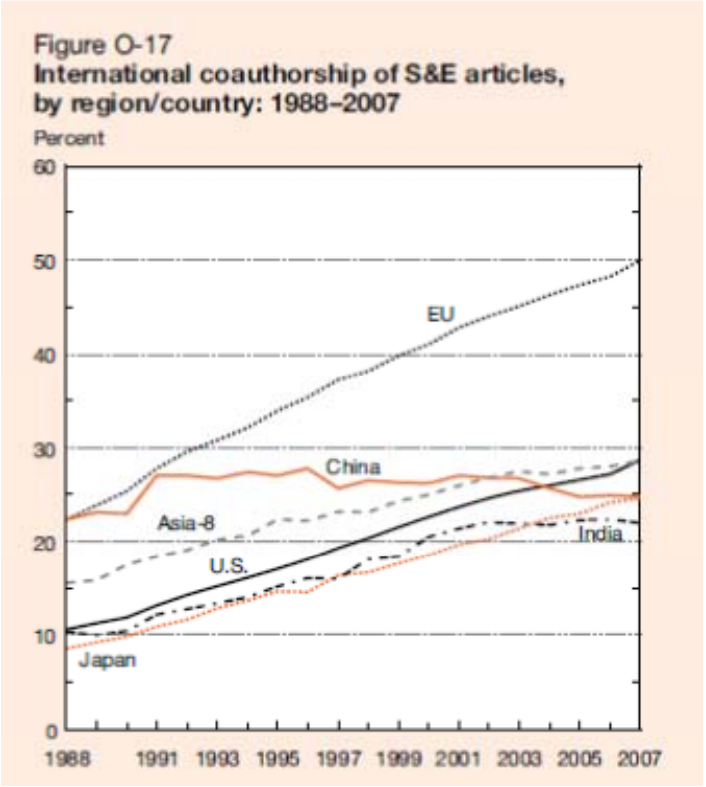
Share of worldwide S&E articles coauthored domestically and internationally



Source: NSF Science and Engineering Indicators 2008

Collaboration



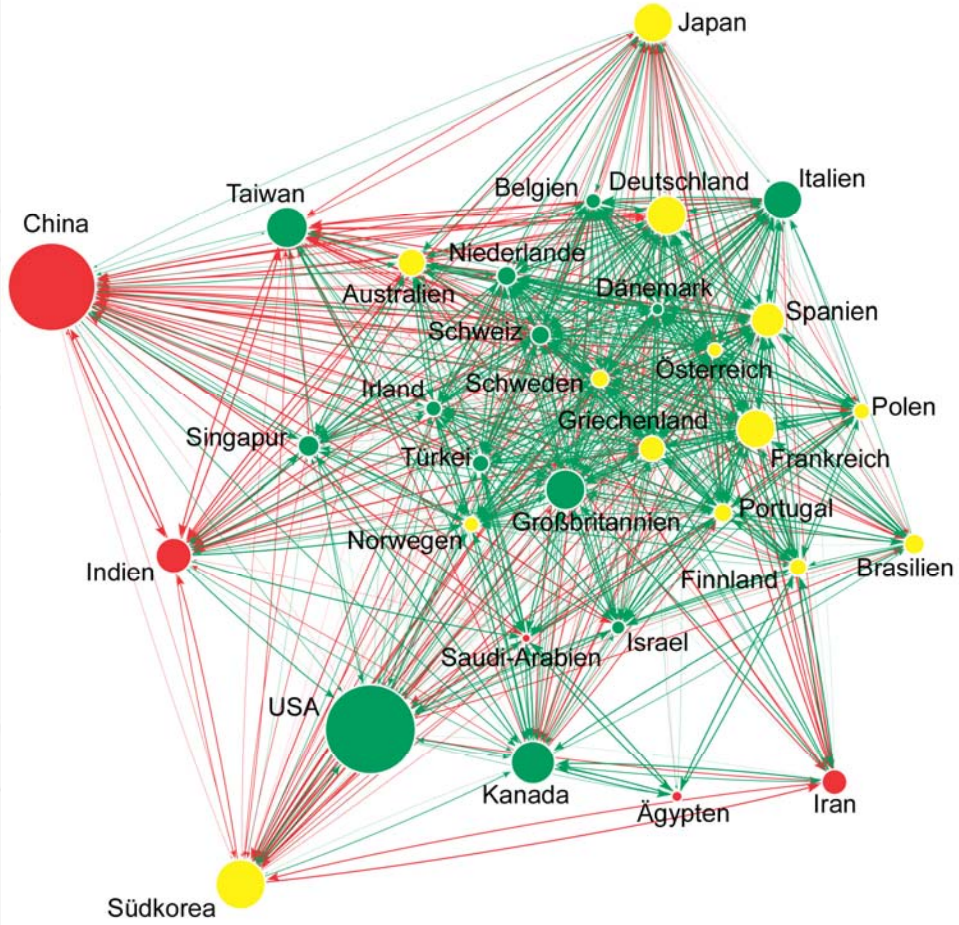


Source: NSF Science and Engineering Indicators 2010

esss european summer school for scientometrics

Collaboration

Research collaboration in the field future internet – 2008-2010



Data source: TR WoS, Source: iFQ/IWT, 2012

www.scientometrics-school.eu

esss european summer school for scientometrics

Citation based indicators

- Citation used as proxi to reflect reception or impact

Basically two notions of citations have become prevalent in bibliometrics,

- the information science related and
- the sociological approach.

According to the first notion, citation is "*one important form of use of scientific information within the framework of documented science communication*". (Glänzel & Schöpflin, *Information Processing & Management*, 1999)

Sociology of science considers citations part of the reward system in science, atoms of peer recognition. (Merton, Science, 1968)

Holmes & Oppenheim found that citations are not primarily a measure of quality, though they significantly correlate with other quality measures. (Holmes & Oppenheim, *(Information Research*, 2001)

esss european summer school for scientometrics

Citation based indicators

- 15 reasons to cite other‘s work

1. Paying homage to pioneers
2. Giving credit for related work (homage to peer)
3. Identifying methodology, equipment, etc.
4. Providing background reading
5. Correcting one’s own work
6. Correcting the work of others
7. Criticising previous work
8. Substantiating claims
9. Alerting to forthcoming work
10. Providing leads to poorly disseminated, poorly indexed, or uncited work
11. Authenticating data and classes of facts – physical constants, etc.
12. Identifying original publications in which an idea or concept was discussed
13. Identifying original publications or other work describing an eponymic concept or term
14. Disclaiming work or ideas of others (negative claim)
15. Disputing priority claims of others (negative homage)

Source: Garfield, Current Contents, 1970

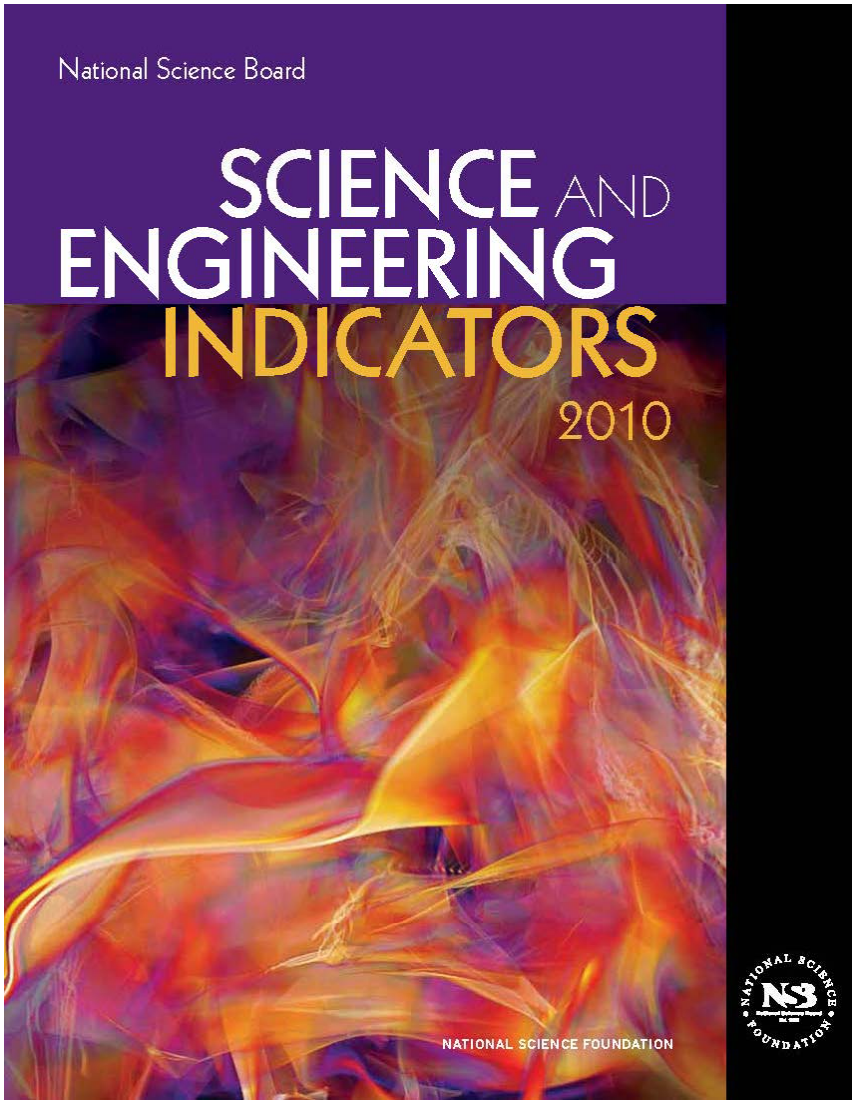
esss european summer school for scientometrics

Citation based indicators

A bibliometricians' view

"if a paper receives 5 or 10 citations a year throughout several years after its publication, it is very likely that its content will become integrated into the body of knowledge of the respective subject field; if, on the other hand, no reference is made at all to the paper during 5 to 10 years after publication, it is likely that the results involved do not contribute essentially to the contemporary scientific paradigm system of the subject field in question."

Source: Braun et al., Scientometric Indicators, 1985



Source: NSF Science and Engineering Indicators 2010

Figure O-18
Citations in U.S. S&E articles to non-U.S. publications: 1992–2007

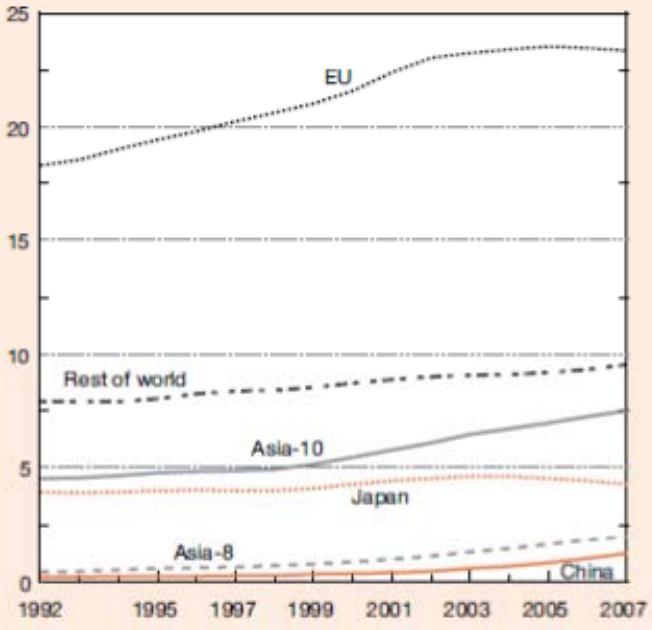
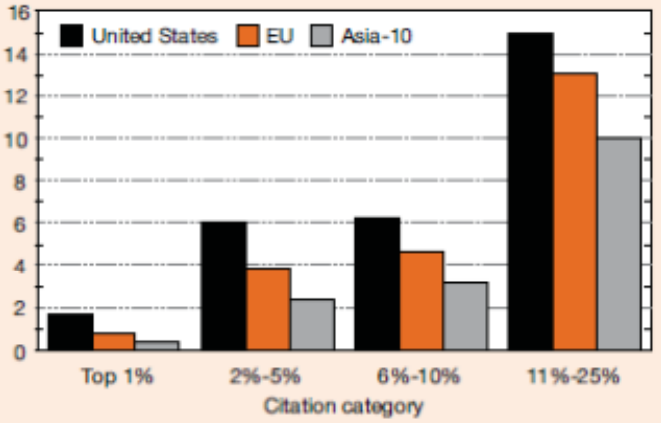


Figure O-21
Share of region's/country's papers among world's most cited S&E articles: 2007



esss european summer school for scientometrics

Citation based indicators

- Self citations
  - Author self-citation

    Inevitable part of scholarly communication

  - Journal self-citation

    Large share of journal self-citations hints to the fact that a journal might be highly specialized while a low share of self citations is, for example, characteristic for review journals (Schubert & Braun, 1993)

    Increasing extent of journal self-citations has been reported in the context of possible manipulation of the Journal Impact Factor (Smith, BMJ, 1997; Weingart, Scientometrics, 2005)

esss european summer school for scientometrics

Citation based indicators

- Factors influencing citation impact
  - Subject area or discipline
  - Age of the paper
  - „Social status“ of the paper (through co-author(s) and the journal)
  - Document type
  - Observation period

esss european summer school for scientometrics

Citation based indicators

- Citation window

  Rule of thumb: the larger the citation window the more reliable the results.

  But: Science policy is interested in timely analysis and reporting

  As a compromise usually citation windows between three and five years are applied.

(see also Moed, Scientometrics 1996; Glänzel Scientometrics, 1997; van Raan, JASIST, 2006)

esss european summer school for scientometrics

Standard citation indicators

The following notations are used:

- c_i number of citations to paper I
- n number of publications
- x_i impact of journal J_i, where the paper is published
- F_i impact of the subject F_i the paper belongs to

see also Braun et al., Scientometric Indicators, 1985; Braun & Glänzel, Scientometrics, 1990; Moed et al, Scientometrics, 1995

esss european summer school for scientometrics

Standard citation indicators

- Observed citations
  - Total citations (within a defined citation window)
  - Share of uncited papers

  - Mean Observed Citation Rate (MOCR)

$$\text{MOCR} = \frac{\sum_{i=1}^{n} c_i}{n}$$

esss european summer school for scientometrics

Standard citation indicators

— Expected citation rates

— Mean Expected Citation Rate (MECR)

$$MECR = \frac{\sum_{i=1}^{n} x_i}{n}$$

— Field Expected Citation Rate (FECR)

$$FECR = \frac{\sum_{i=1}^{n} f_i}{n}$$

The ratio of these two indicators MECR / FECR reflects whether a unit publishes in higher / lower impact journals than it would be expected taken the field the unit is active in.

esss european summer school for scientometrics

Relative citation indicators

— Normalized citation rate

$$NMCR = MOCR / FECR$$

Indicates whether a paper is cited above / below average compared to the field it is assigned to.

— Relative citation rate

$$RCR = MOCR / MECR$$

Indicates whether a paper is cited above / below average compared to the journal it appeared in.

esss european summer school for scientometrics

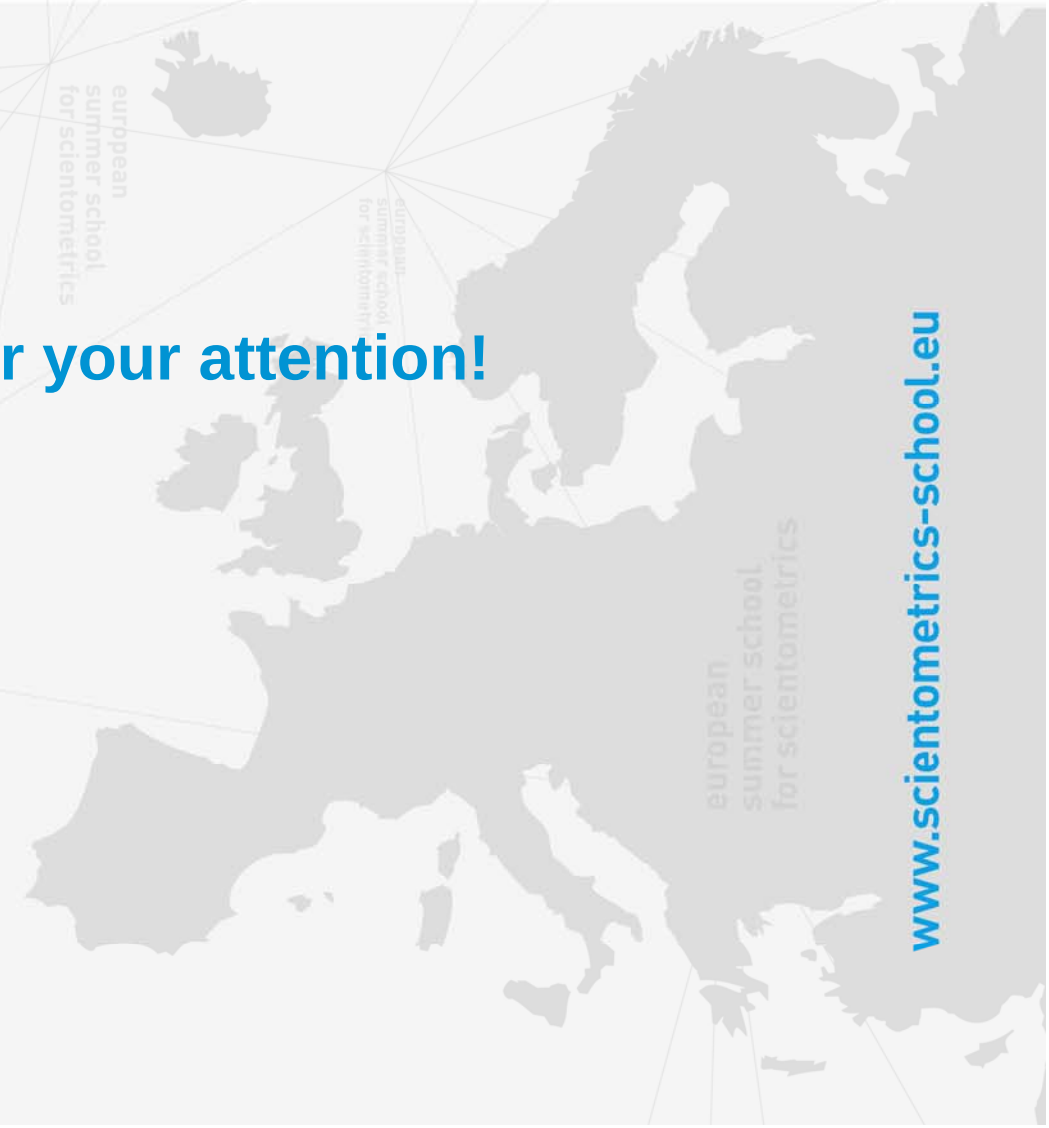
References

- A. Pritchard (1969): Statistical bibliography or bibliometrics. In: Journal of Documentation. 1969, 25 (4): 348-349)
- V.V. Nalimov & Z.M. Mulchenko (1969): Naukometriya. Izuchenie Razvitiya Nauki kak Informatsionnogo Protsessa. [Scientometrics. Study of the Development of Science as an Information Process], Nauka, Moscow, (English translation: 1971. Washington, D.C.: Foreign Technology Division. U.S. Air Force Systems Command, Wright-Patterson AFB, Ohio. (NTIS Report No.AD735-634).
- A.F.J. van Raan (1988): Introduction to the Handbook. In: A.F.J. van Raan (ed.): Handbook of Quantitative Studies of Science and Technology. Amsterdam : North Holland, Elsevier Science Publishers, 1988. p. 1-8.
- S. Hornbostel (1999): Welche Indikatoren zu welchem Zweck, Input, Throughput, Output, in: *Röbbecke, Martina / Simon, Dagmar (Ed.)*, Qualitätsförderung durch Evaluation? Ziele, Aufgaben und Verfahren im Wandel. Berlin: Wissenschaftszentrum Berlin für Sozialforschung, 99-003.
- H. Moed (2005): Citation Analysis in Research Evaluation. Springer. Dorderecht.
- L. Zhang, R. Rousseau, W. Glänzel (2011): Document-type country profiles. Journal of the American Society for Information Science and Technology, 62(7), 1403-1411.et al
- B.R. Fischer (2001): Performance Analyse in der Praxis. Performanzmaße, Attributionsanalyse, DFVA-Performance Presentation Standards. München und Wien: Oldenburg Verlag
- W. Glänzel, A. Schubert (2004): Analyzing scientific networks through co-authorship. In: Moed & Glänzel & Schmoch (ed.) Handbook of Quantitative Studies of Science and Technology Research. Dordrecht, Boston, London: Kluwer, 2004. 257-276
- J. S. Katz, B. R. Martin (1997): 1997. What is research collaboration? Research Policy 26, 1–18.

www.scientometrics-school.eu

esss european summer school for scientometrics

Thank you for your attention!



www.scientometrics-school.eu