CWTS
Meaningful metrics

Identification of organisations in bibliometric data

Martijn Visser and Clara Calero

RISIS COURSE – Universities, research organizations and rankings

28 June 2017
CWTS, Leiden

Universiteit Leiden

Research organisations

- Higher education institutions
- Research organisations
- Hospitals
- Companies
- Governmental departments
- ….

CWTS

RISIS
Research infrastructure for research and innovation policy studies

Hierarchical Structures

Collegiate Universities / Federal Universities

Universities

Campuses

CWTS

RISIS
Research infrastructure for research and innovation policy studies

Public Research Organisations: definition

- Performs research activities as *a core activity*
- Is *independent* in the conduct of research
- Its research activities are *not for-profit* oriented

CWTS

RISIS
Research infrastructure for research and innovation policy studies

Public Research Organisations: classification

- Mission Oriented Centres (MOC) – organizations that carry out research in specific topics, owned or run by government department or ministries, also to support to policy making (included National Agencies and Federal Offices);
- Independent Research Institutes (IRI) – Centres carrying out basic and applied research focused on issues, at the boundaries between public and private sector;
- Public Research Centres and Councils (PRC) – Overarching institutions of considerable size performing research in several fields (ex. CNR, CNRS, Academies of Sciences);
- Research Technologies Organizations (RTO) – Industrial research institutes, semi-public or private not for profit, for the development and transfer of S&T to the private sector and society (ex. TNO, SINTEF).

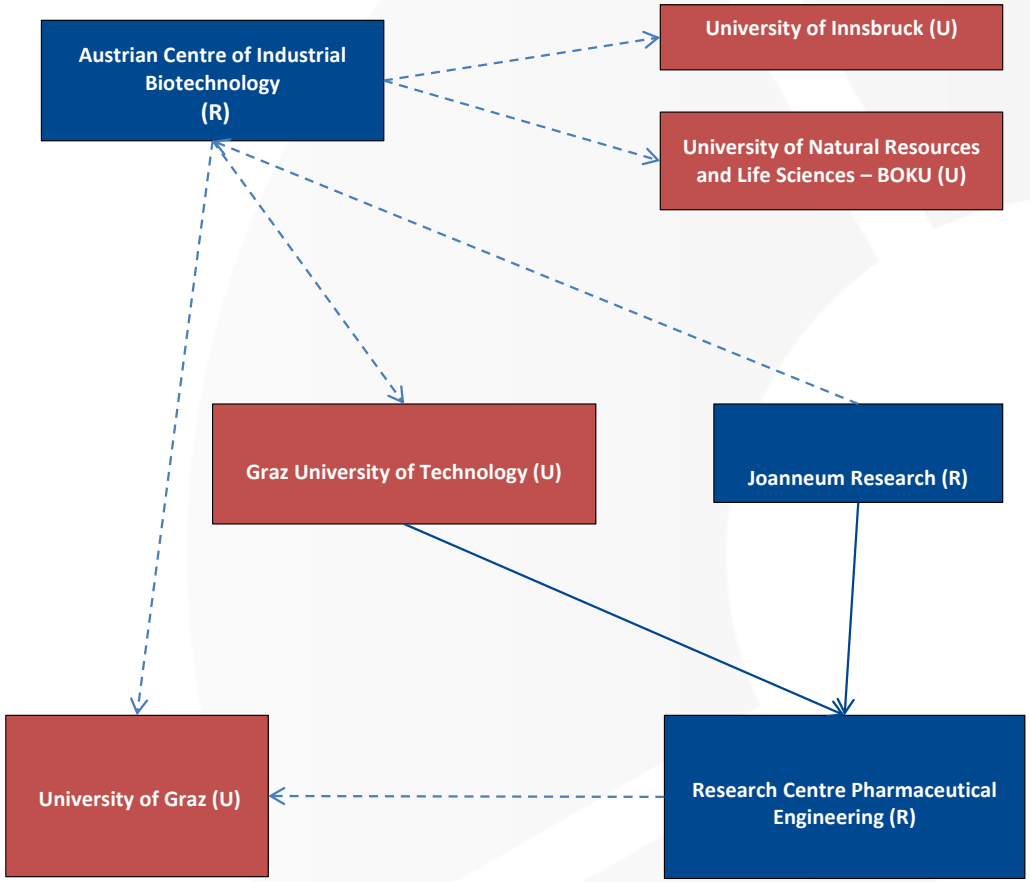
CWTS

RISIS
Research infrastructure for research and innovation policy studies

	number institutes	annual turnover	funding	indus. ties	Univ. ties	Class OECD	Freq. pub.
Max Planck Society	82	1.3 b. €	40% federal 40% länder 20% third party	?	staff overlap	PRC*	>123,000
Fraunhofer Society	80 +	1.65 b. €	85% contract 15% government	?	?	RTO*	> 13,000
Helmholtz Association	17	3 b. €	63% federal 30% third party 7% länder	?	?	IRI	> 13,000
Leibnitz Association	87, independent!	1.3 b. €	33% federal 33% länder 33% third party	?	no	IRI	> 13,000
Academies of Science	?	?	government	?	staff overlap	PRC*	> 500
Government Institutes	100+ federal 38 länder	?	government	no	?	MOC	?

CWTS

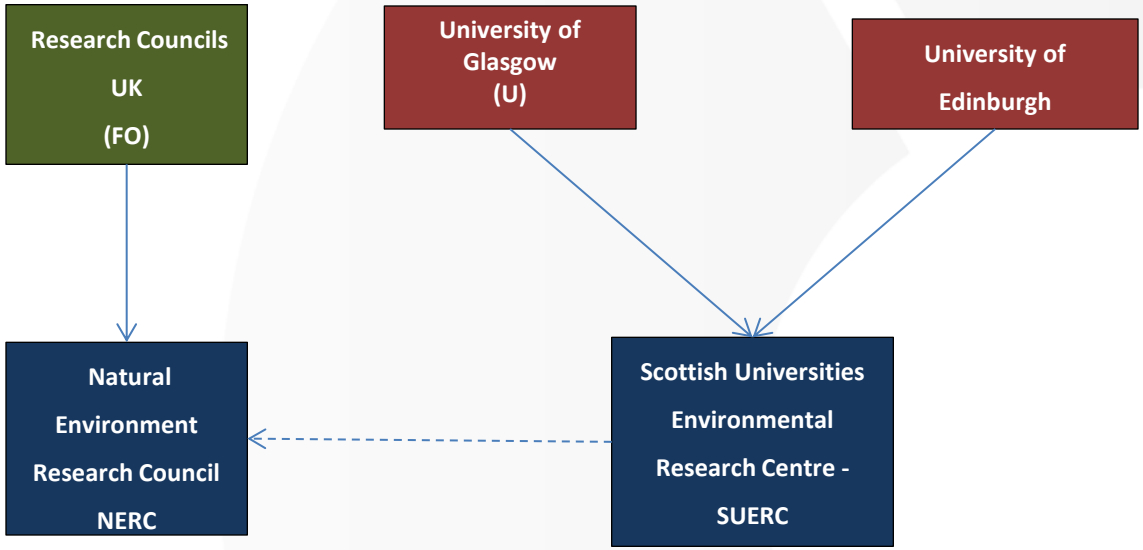
Public Research Organisations: an example of relationships



CWTS

RISIS
Research infrastructure for research and innovation policy studies

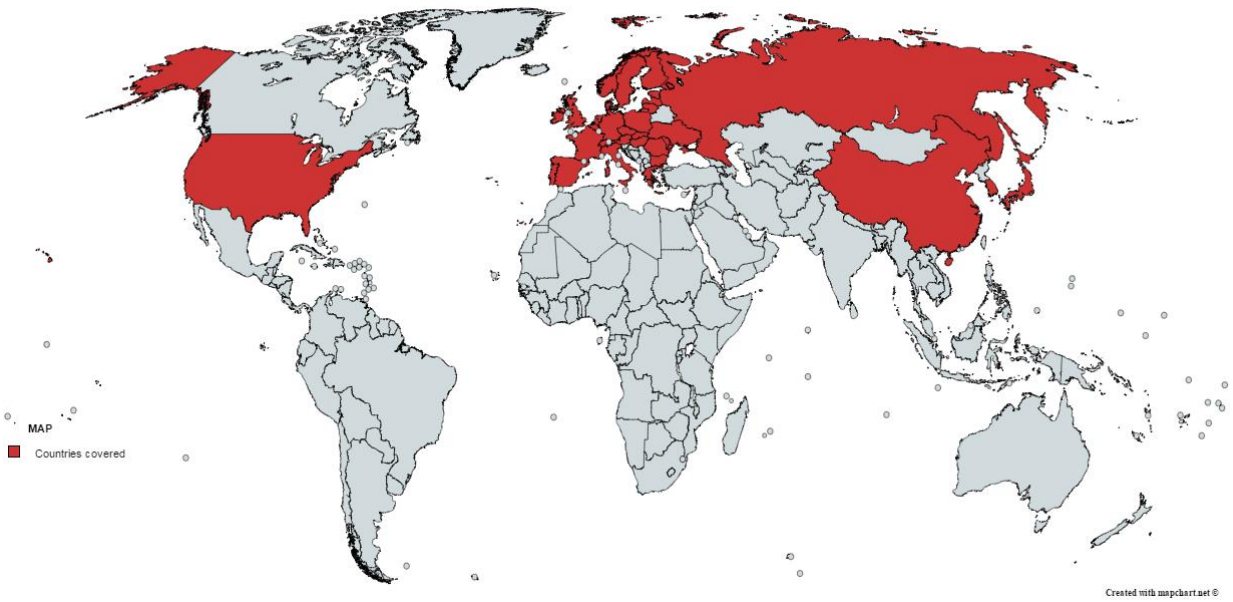
Public Research Organisations: example



CWTS

RISIS
Research infrastructure for research and innovation policy studies

Public Research Organisations



CWTS

RISIS
Research infrastructure for research and innovation policy studies

Important differences between national research systems

- Institutional relationships between hospitals and medical faculties
- The role of public and private research institutes
- Distinction between higher education institutions

CWTS

Research organisations

- Name variations
  - Components (institutes, departments)
  - Mergers, splits, take-overs, split-offs
  - Affiliated institutes

Example: Aarhus University

8452	AARHUS	AARHUS UNIV
4886	AARHUS	AARHUS UNIV HOSP
4496	AARHUS	UNIV AARHUS
1421	TJELE	DANISH INST AGR SCI
896	AALBORG	AALBORG HOSP
765	ROSKILDE	NATL ENVIRONM RES INST
431	SILKEBORG	NATL ENVIRONM RES INST
316	AARHUS	AARHUS KOMMUNE HOSP
273	AARHUS	AARHUS SCH BUSINESS
257	AALBORG	AARHUS UNIV HOSP
242	RISSKOV	UNIV AARHUS
238	SLAGELSE	DANISH INST AGR SCI
227	AARSLEV	DANISH INST AGR SCI
214	TJELE	RES CTR FOULUM
198	RONDE	NATL ENVIRONM RES INST
159	AARHUS	SKEJBY UNIV HOSP
155	AALBORG	AALBORG UNIV HOSP
130	RISSKOV	AARHUS UNIV HOSP
113	AARHUS	SKEJBY HOSP

- Ca. 700 different name variations
- More than half of these name variations just appear once
- Searching for“AARHUS UNIV” or “UNIV AARHUS” captures ca. 70% of all papers

CWTS

Important differences between WoS and Scopus

- Scopus performs less preprocessing of data extracted from publications
- Scopus provides an ‘as-is’ representation of author affiliations
- Scopus does not distinguish between reprint address and affiliation

CWTS

PLOS ONE | Subject Areas | For Authors | About Us | Search | advanced search

OPEN ACCESS PEER-REVIEWED
RESEARCH ARTICLE

2,091 VIEWS | 11 CITATIONS | 30 SAVES | 8 SHARES

Citation Analysis May Severely Underestimate the Impact of Clinical Research as Compared to Basic Research

Nees Jan van Eck, Ludo Waltman, Anthony F. J. van Raan, Robert J. M. Klautz, Wilco C. Peul

Published: April 24, 2013 • DOI: 10.1371/journal.pone.0062395

Article | About the Authors | Metrics | Comments | Related Content

Download PDF | Print | Share

CrossMark

About the Authors

Nees Jan van Eck, Ludo Waltman, Anthony F. J. van Raan
Centre for Science and Technology Studies, Leiden University, Leiden, The Netherlands

Robert J. M. Klautz
Department of Cardiothoracic Surgery, Leiden University Medical Center, Leiden, The Netherlands

Wilco C. Peul
Department of Neurosurgery, Leiden University Medical Center, Leiden, The Netherlands

Subject Areas

Bibliometrics
Cardiac surgery
Cardiology
Medicine and health ...
Orthopedic surgery

Scopus

van Eck, N.J.[a], Waltman, L.[a], van Raan, A.F.J.[a], Klautz, R.J.M.[b], Peul, W.C.[c]

[a] Centre for Science and Technology Studies, Leiden University, Leiden, Netherlands

[b] Department of Cardiothoracic Surgery, Leiden University Medical Center, Leiden, Netherlands

[c] Department of Neurosurgery, Leiden University Medical Center, Leiden, Netherlands

Web of Science

Author Information

Reprint Address: van Eck, NJ (reprint author)

+ Leiden Univ, Ctr Sci & Technol Studies, Leiden, Netherlands.

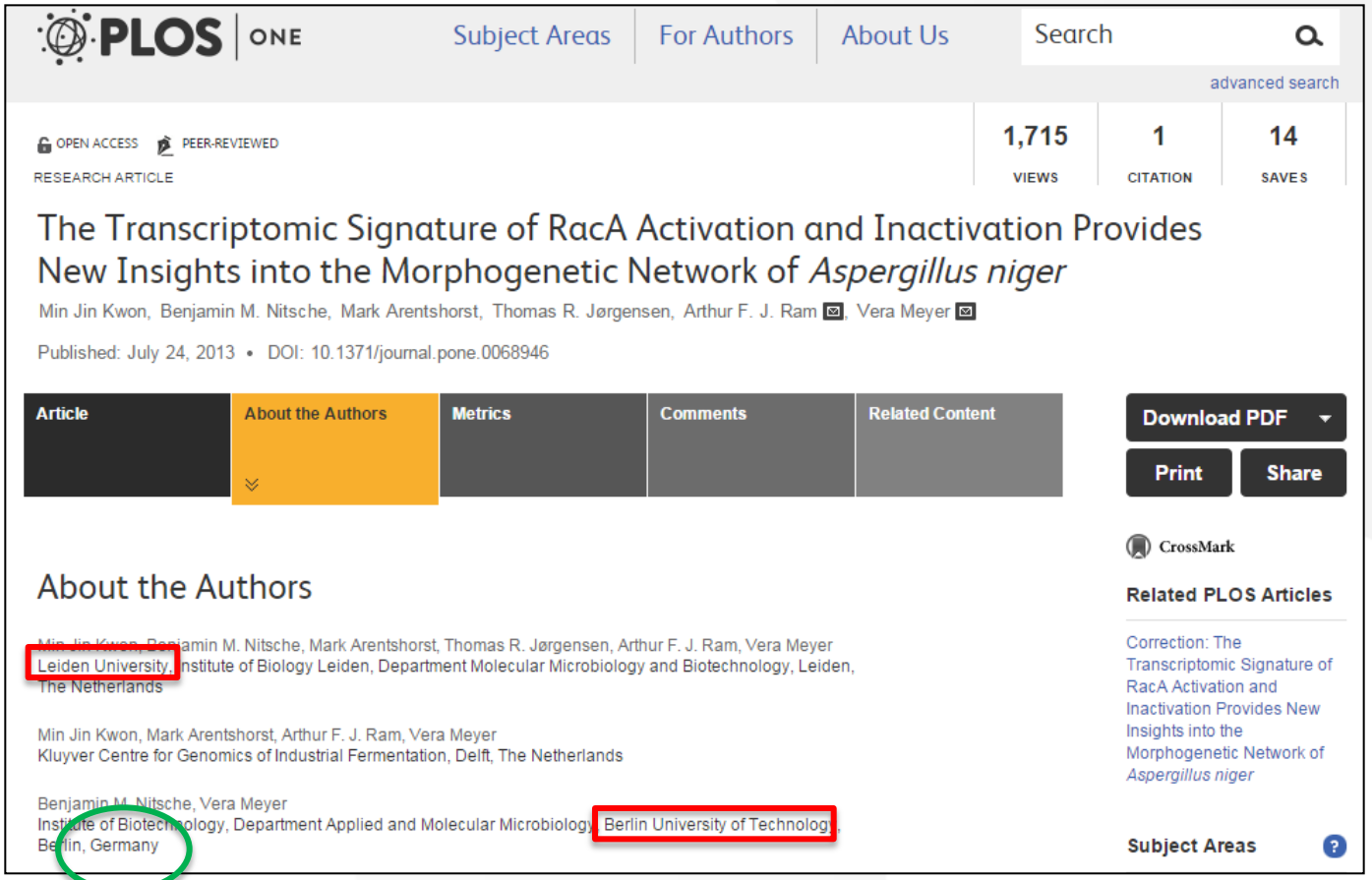
Addresses:

+ [1] Leiden Univ, Ctr Sci & Technol Studies, Leiden, Netherlands

+ [2] Leiden Univ, Med Ctr, Dept Cardiothorac Surg, Leiden, Netherlands

+ [3] Leiden Univ, Med Ctr, Dept Neurosurg, Leiden, Netherlands

CWTS



PLOS ONE | Subject Areas | For Authors | About Us | Search | advanced search

OPEN ACCESS PEER-REVIEWED
RESEARCH ARTICLE

1,715 VIEWS | 1 CITATION | 14 SAVES

The Transcriptomic Signature of RacA Activation and Inactivation Provides New Insights into the Morphogenetic Network of *Aspergillus niger*

Min Jin Kwon, Benjamin M. Nitsche, Mark Arentshorst, Thomas R. Jørgensen, Arthur F. J. Ram, Vera Meyer

Published: July 24, 2013 • DOI: 10.1371/journal.pone.0068946

Article | About the Authors | Metrics | Comments | Related Content

Download PDF | Print | Share

CrossMark

About the Authors

Min Jin Kwon, Benjamin M. Nitsche, Mark Arentshorst, Thomas R. Jørgensen, Arthur F. J. Ram, Vera Meyer
Leiden University, Institute of Biology Leiden, Department Molecular Microbiology and Biotechnology, Leiden, The Netherlands

Min Jin Kwon, Mark Arentshorst, Arthur F. J. Ram, Vera Meyer
Kluyver Centre for Genomics of Industrial Fermentation, Delft, The Netherlands

Benjamin M. Nitsche, Vera Meyer
Institute of Biotechnology, Department Applied and Molecular Microbiology, Berlin University of Technology, Berlin, Germany

Related PLOS Articles

Correction: The Transcriptomic Signature of RacA Activation and Inactivation Provides New Insights into the Morphogenetic Network of *Aspergillus niger*

Subject Areas

Scopus

Kwon, M.J.[ab], Nitsche, B.M.[ac], Arentshorst, M.[ab], Jørgensen, T.R.[a], Ram, A.F.J.[ab], Meyer, V.[abc]

[a] Leiden University, Institute of Biology Leiden, Department Molecular Microbiology and Biotechnology, Leiden, Netherlands

[b] Kluyver Centre for Genomics of Industrial Fermentation, Delft, Netherlands

[c] Institute of Biotechnology, Department Applied and Molecular Microbiology, Berlin University of Technology, Berlin, Netherlands

Web of Science

Author Information

Reprint Address: Ram, AFJ (reprint author)

+ Leiden Univ, Inst Biol Leiden, Dept Mol Microbiol & Biotechnol, Leiden, Netherlands.

Addresses:

+ [1] Leiden Univ, Inst Biol Leiden, Dept Mol Microbiol & Biotechnol, Leiden, Netherlands

[2] Kluyver Ctr Genom Ind Fermentat, Delft, Netherlands

+ [3] Berlin Univ Technol, Dept Appl & Mol Microbiol, Inst Biotechnol, Berlin, Germany

CWTS

WEB OF SCIENCE™

Organizations - Enhanced List

*** Use this list to find the preferred name for an organization and the variants we have identified and associated with it. Note: Not all organizations have beer*

Use the Browse and Find features to locate organizations to add to your query.

Click on a letter or number to browse organizations alphabetically by title

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z 0 1 2 3 4 5 6 7 8 9

Enter text to find organizations containing or related to the text.

Example: PRAGUE to find ACAD OF FINE ARTS PRAGUE and CHARLES UNIV PRAGUE ACAD SCI CZECH REPUBL

aarhus [Find]

DETAILS

KEY: [Add] = add to query

Organization Name:	Add	**AARHUS UNIVERSITY**
Other Names:		AARHUS UNIVERSITET; AARHUS UNIVERSITY; AARHUS UNIV
Address:		NORDRE RINGGADE 1 ,AARHUS, DENMARK ,DK-8000
Website:		http://www.au.dk/en
Name Variants:	Add	AAARHUS UNIV
	Add	AAHUS UNIV HOSP
	Add	AAKH
	Add	AALBERG HOSP N
	Add	AALBORG AARHUS UNIV HOSP

Scopus

Scopus SciVal | Register Login Help

Search | Alerts | Lists | My Scopus

The Scopus Affiliation Identifier assigns a unique number to groups of documents affiliated with an organization via an algorithm that matches affiliation names based on certain criteria.

Affiliation **"aarhus"** Edit

20 affiliation results About Scopus Affiliation Identifier

Sort on: **Document Count** ↓ Affiliation (A-Z) ...

Show documents | Give feedback

Refine

Limit to Exclude

City

- Arhus (10)
- Aalborg (2)
- Risskov (2)
- Aarhus (1)
- Copenhagen (1)

Country/Territory

- Denmark (20)

Limit to Exclude

Export refine

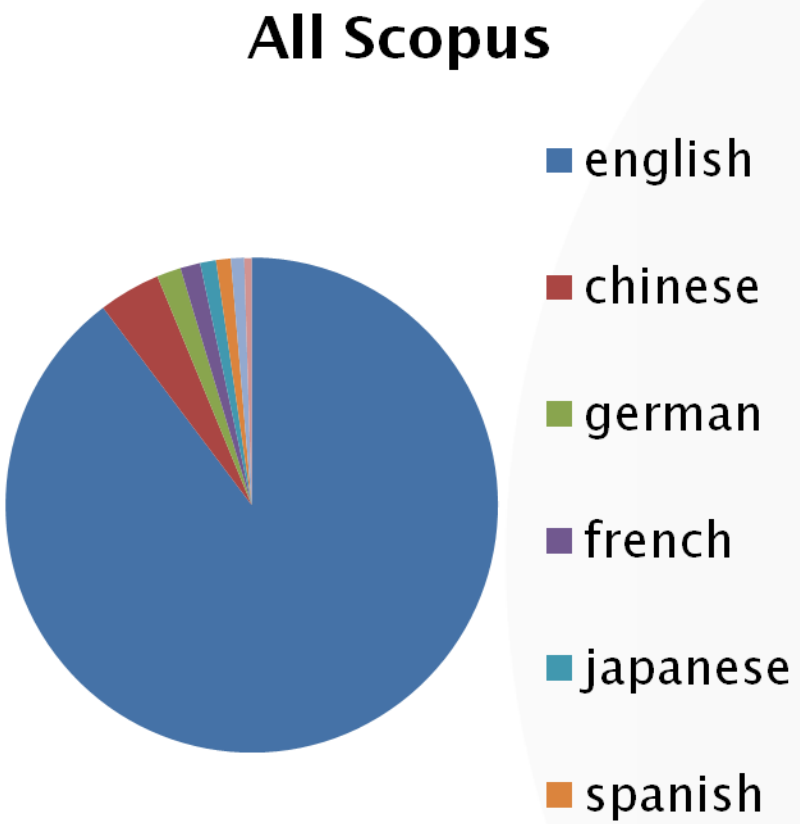
1	Aarhus Universitet Aarhus University University of Aarhus	52689	Aarhus	Denmark
2	Arhus Universitetshospital Aarhus University Hospital	15776	Arhus	Denmark
3	Aalborg Sygehus Aalborg Sygehus Aarhus University Hospital	6238	Aalborg	Denmark
4	Aarhus Amtssygehus, Aarhus University Hospital Aarhus University Hospital	5292	Arhus	Denmark
5	Danmarks Miljoundersogelser Natl. Environ. Research Institute Aarhus University	3400	Roskilde	Denmark
6	Skejby Sygehus, Aarhus University Hospital Skejby Sygehus Aarhus University Hospital	2711	Arhus	Denmark
7	Det Jordbrugsvidenskabelige Fakultet, Aarhus Universitet Aarhus University University of Aarhus	2448 documents	Tjele	Denmark
8	Aarhus Universitetshospital Psychiatric Hospital in Aarhus Psychiatric Hospital	1328	Risskov	Denmark
	Handelshøjskolen, Aarhus Universitet	1226	Arhus	Denmark

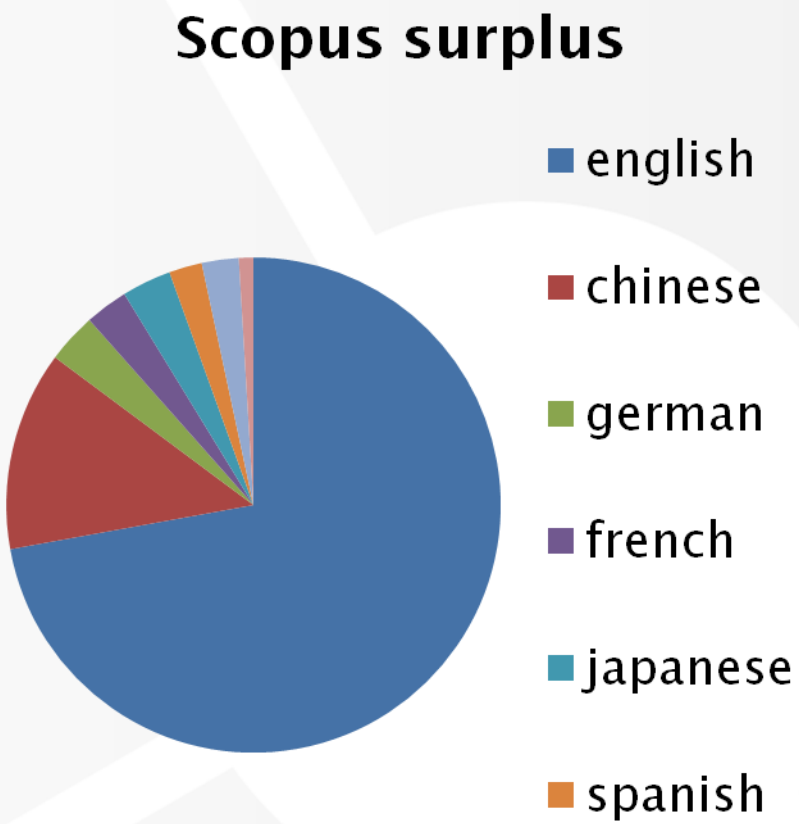
Relative Share of countries in Scopus Surplus

Country	% Affiliations in Scopus	% Affiliations in Scopus-Surplus
China	8.8%	18.6%
Malaysia	0.3%	0.7%
Pakistan	0.2%	0.3%
Ukraine	0.3%	0.4%
India	2.3%	3.9%
Croatia	0.2%	0.3%
Iran	0.8%	1.3%
Egypt	0.3%	0.4%
Thailand	0.3%	0.4%
Czech Republic	0.5%	0.7%
Poland	1.0%	1.4%
Romania	0.3%	0.4%
Russia	1.5%	1.9%
Brazil	1.9%	2.3%
Hungary	0.3%	0.4%
Mexico	0.6%	0.6%

CWTS

Language distribution





CWTS

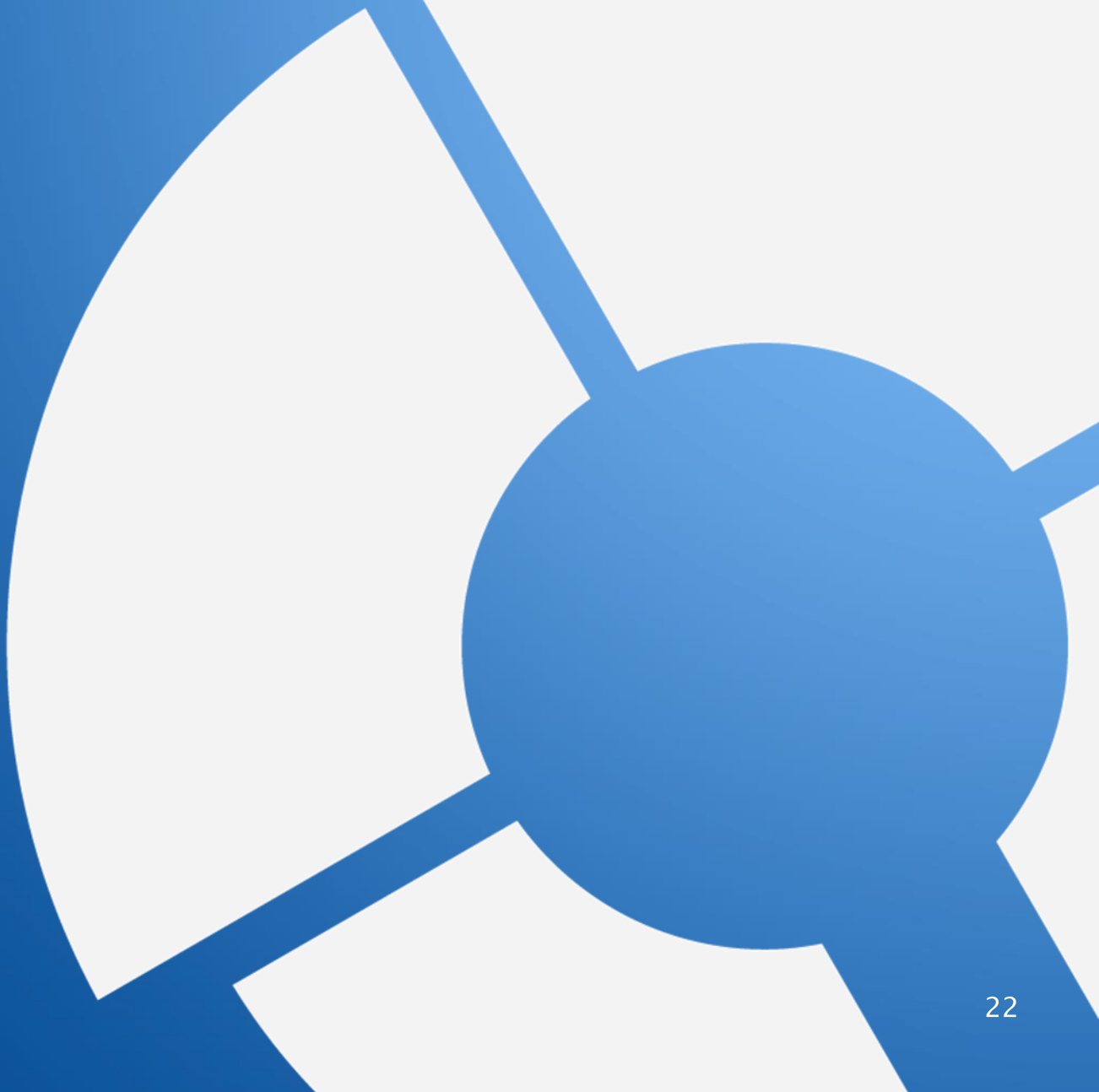
Number of References and Citations
(articles, reviews and conference papers)

	Scopus	Scopus Surplus	%
Nr References	25.7	16.3	63%
Nr Citations (MCS)	14.7	2.1	13%

CWTS

Assignments object identification

Counting methods



Counting methods

- Full (or whole) counting:
  - A publication is fully assigned to each co-author
  - Example: A publication co-authored by three universities is assigned to each university with a weight of 1
- Fractional counting:
  - A publication is fractionally assigned to each co-author
  - Example: A publication co-authored by three authors from three different universities is assigned to each university with a weight of 1/3 = 0.33

CWTS

Field normalization at the publication level

- By definition, the mean normalized citation score of all publications in a field equals 1, i.e., MNCS = 1
- Also, provided that an appropriate approach for handling ties is used, the proportion of all publications in a field that belong to the top 10% equals 10%, i.e., $PP_{top\ 10\%} = 10\%$

CWTS

Field normalization at the country/institution level

- Consider a field and take all countries (or institutions) active in the field
- Assign each country a weight equal to its number of publications in the field
- The weighted average of countries' MNCS values should be 1
- The weighted average of countries' $PP_{top\ 10\%}$ values should be 10%

CWTS

Example: Full counting

	Authors	Citations	Norm. citations
Publication 1	Country A	3	0.6
Publication 2	Country B	4	0.8
Publication 3	Countries A and B	8	1.6

- MNCS country A: (0.6 + 1.6) / 2 = 1.1
- MNCS country B: (0.8 + 1.6) / 2 = 1.2
- Average MNCS value: (1.1 + 1.2) / 2 = 1.15
- All countries perform above the world average!

CWTS

Example: Fractional counting

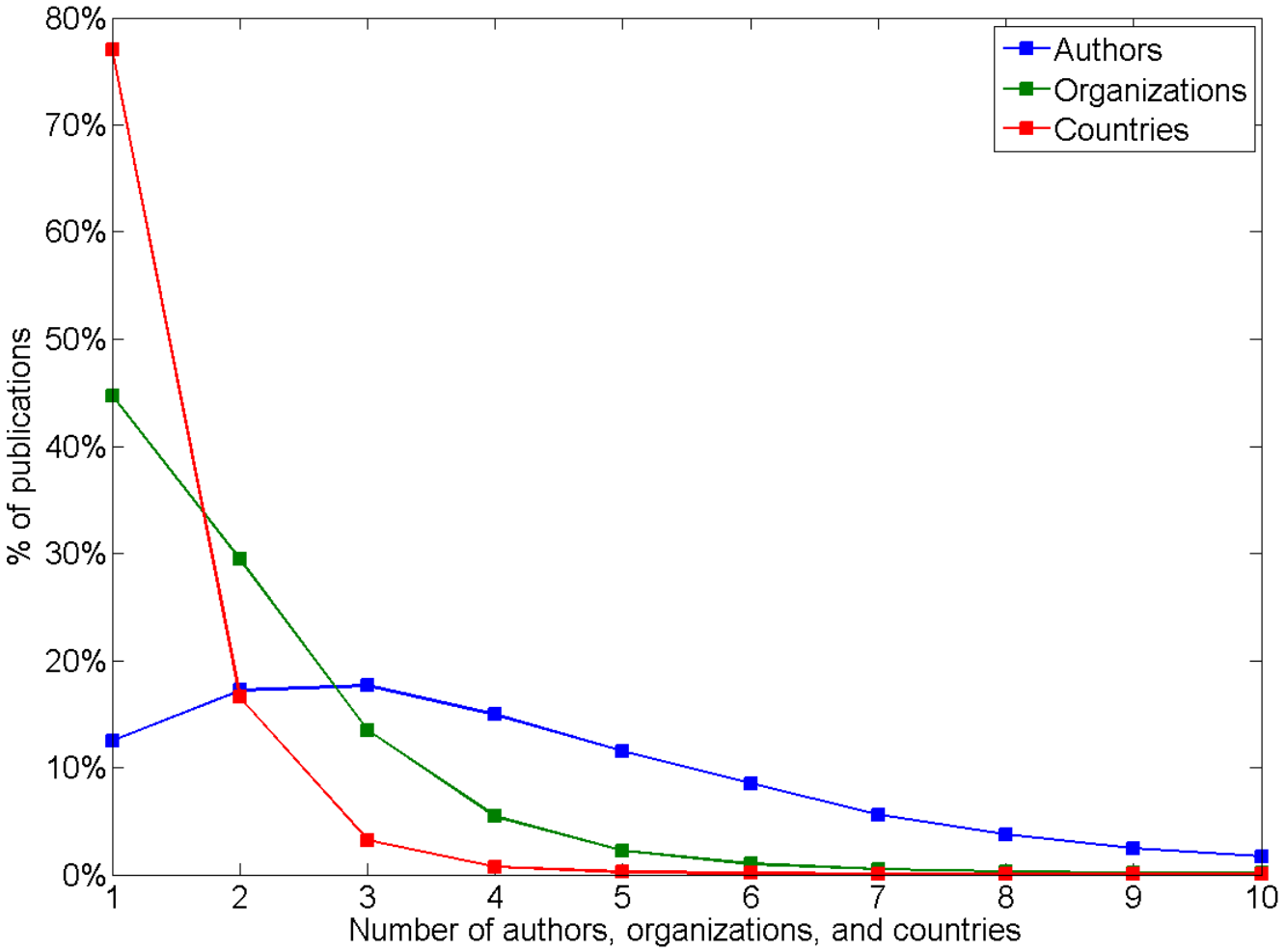
	Authors	Citations	Norm. citations
Publication 1	Country A	3	0.6
Publication 2	Country B	4	0.8
Publication 3	Countries A and B	8	1.6

- MNCS country A: $(1 \times 0.6 + 0.5 \times 1.6) / (1 + 0.5) = 0.93$
- MNCS country B: $(1 \times 0.8 + 0.5 \times 1.6) / (1 + 0.5) = 1.07$
- Average MNCS value: $(0.93 + 1.07) / 2 = 1$

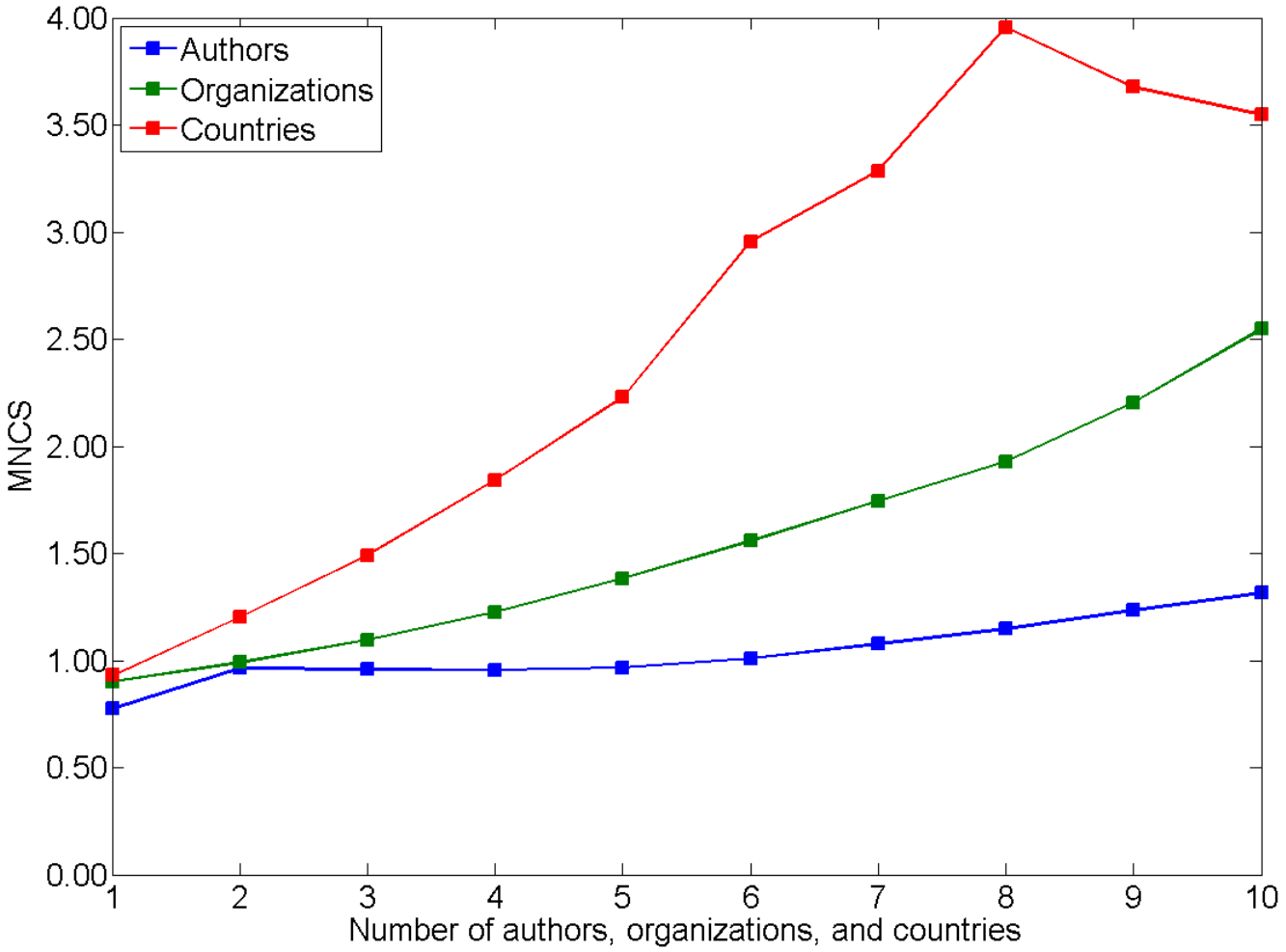
CWTS

Assignments counting methods

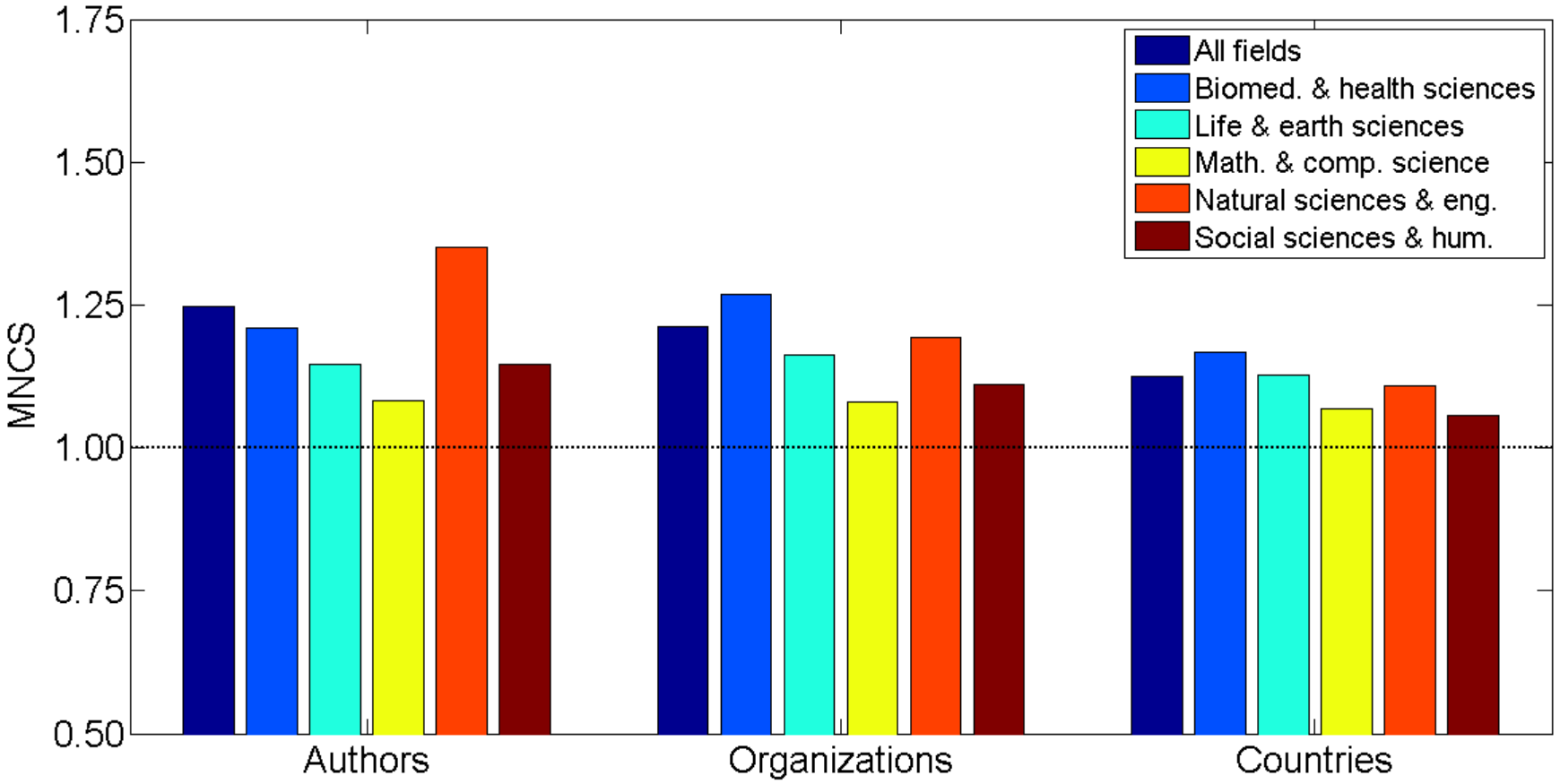
Intensity of collaboration



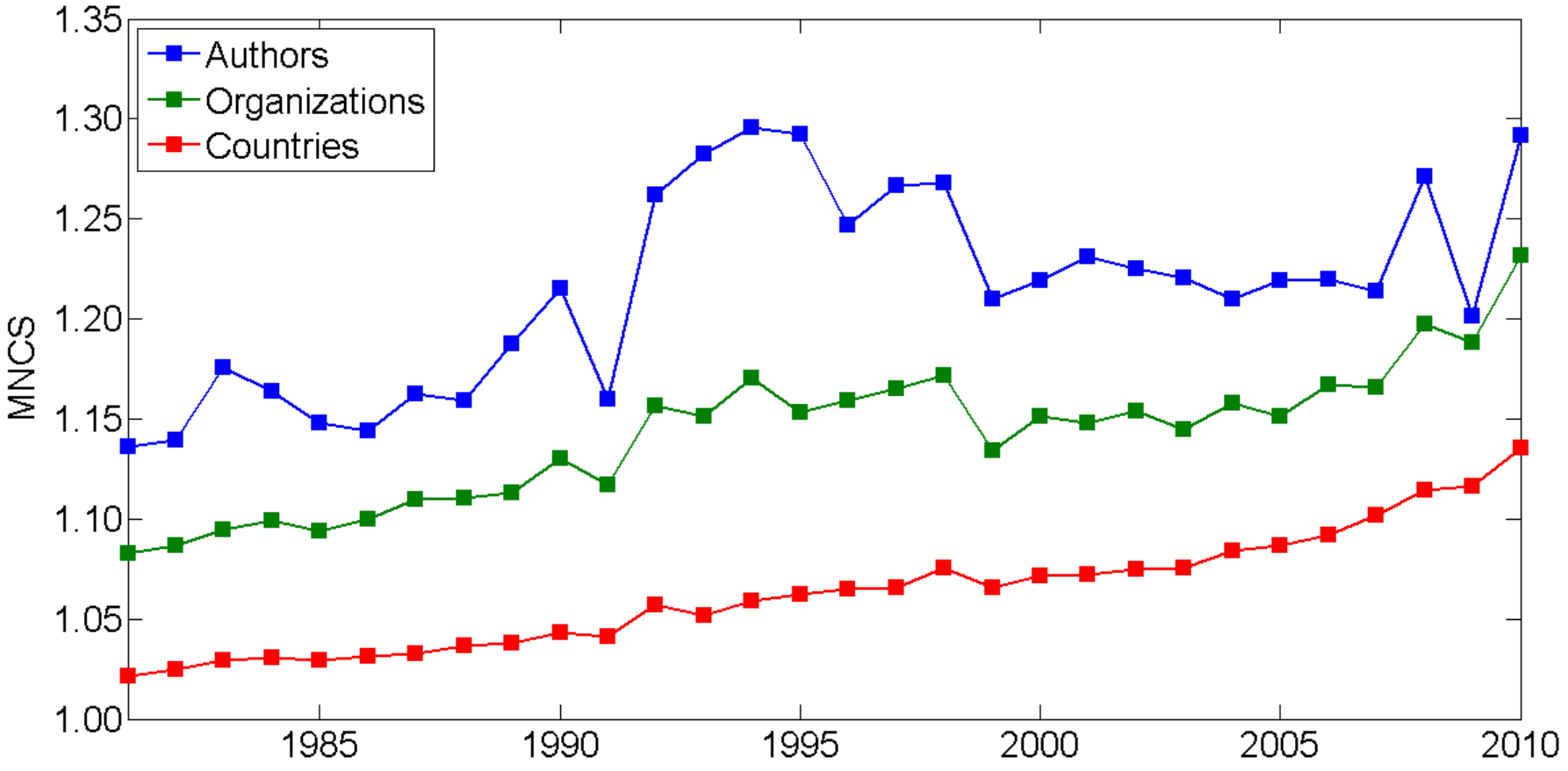
Citation advantage of collaborative publications



Full counting bonus per field of science



Full counting bonus time trend



Fractional counting in the CWTS Leiden Ranking

- Leiden Ranking uses author-based affiliation fractional counting

BMJ Open 2014;**4**:e004468 doi:10.1136/bmjopen-2013-004468

Research methods

Mapping patient safety: a large-scale literature review using bibliometric visualisation techniques

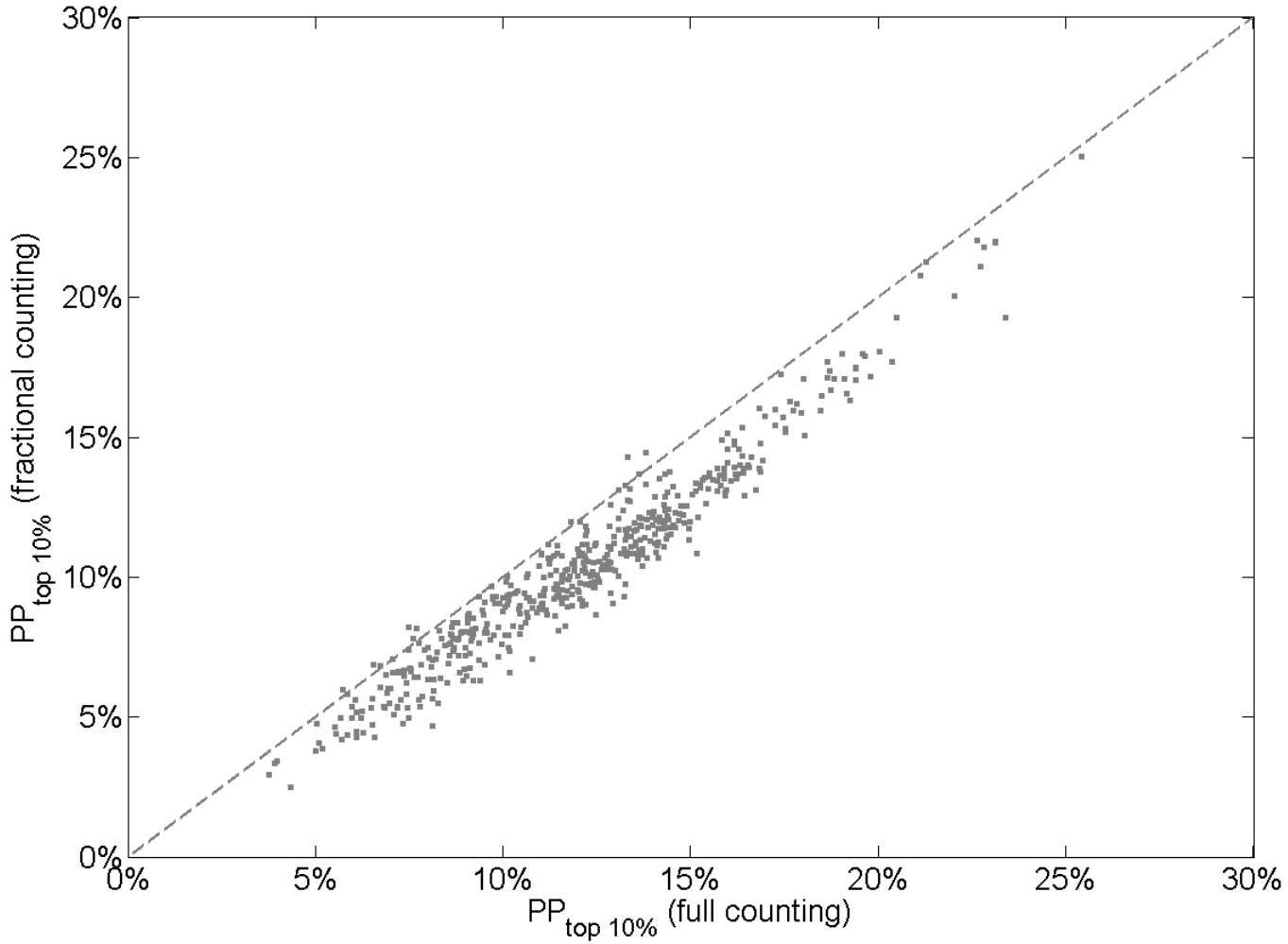
S P Rodrigues[1], N J van Eck[2], L Waltman[2], F W Jansen[1,3]

Author Affiliations

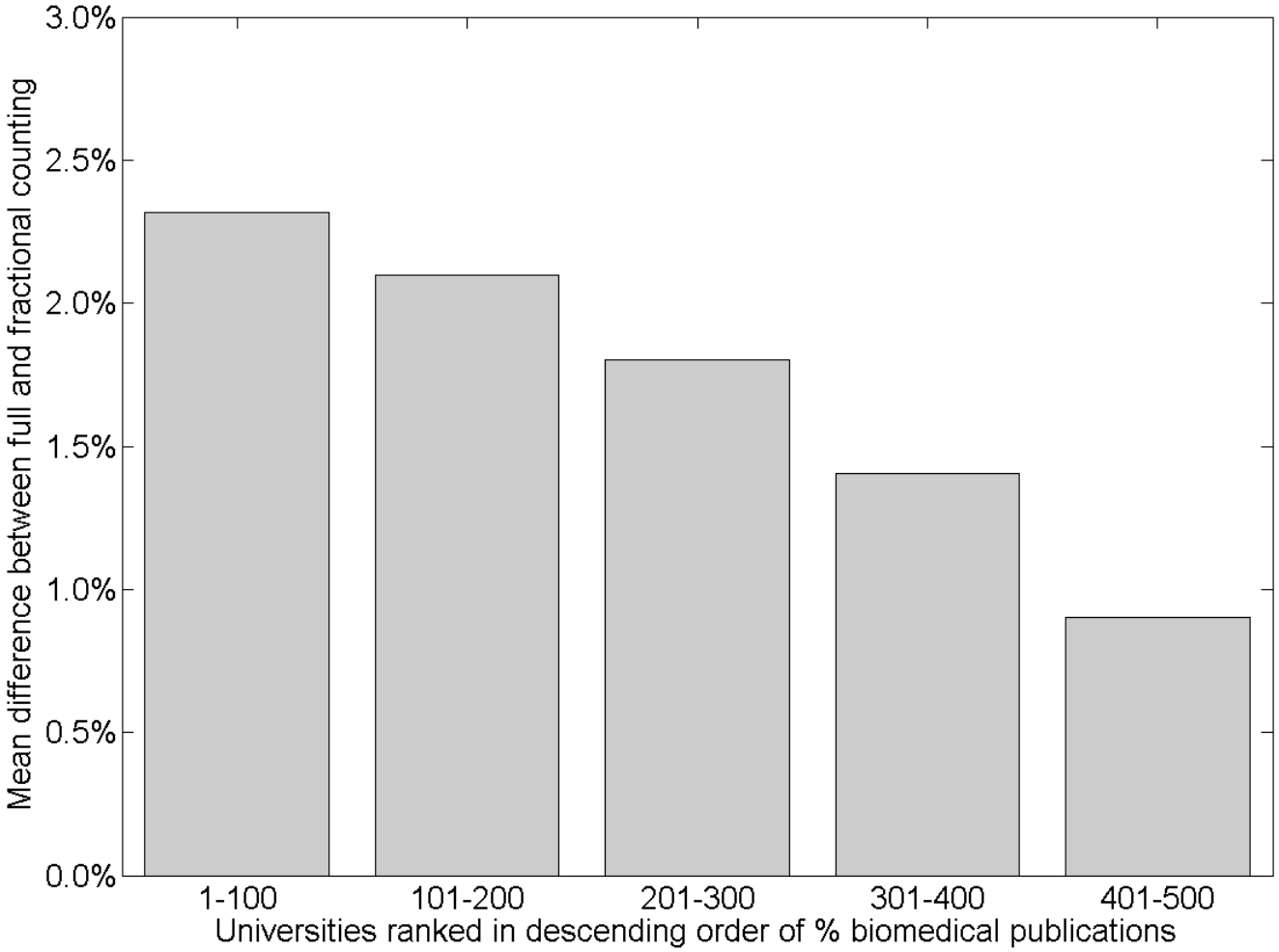
[1]Department of Gynecology, Leiden University Medical Center, Leiden, The Netherlands
[2]Centre for Science and Technology Studies, Leiden University, Leiden, The Netherlands
[3]Department of BioMechanical Engineering, Delft University of Technology, Delft, The Netherlands

- The above publication is assigned to Leiden Univ. with a weight of 7/8 and to Delft Univ. of Technology with a weight of 1/8

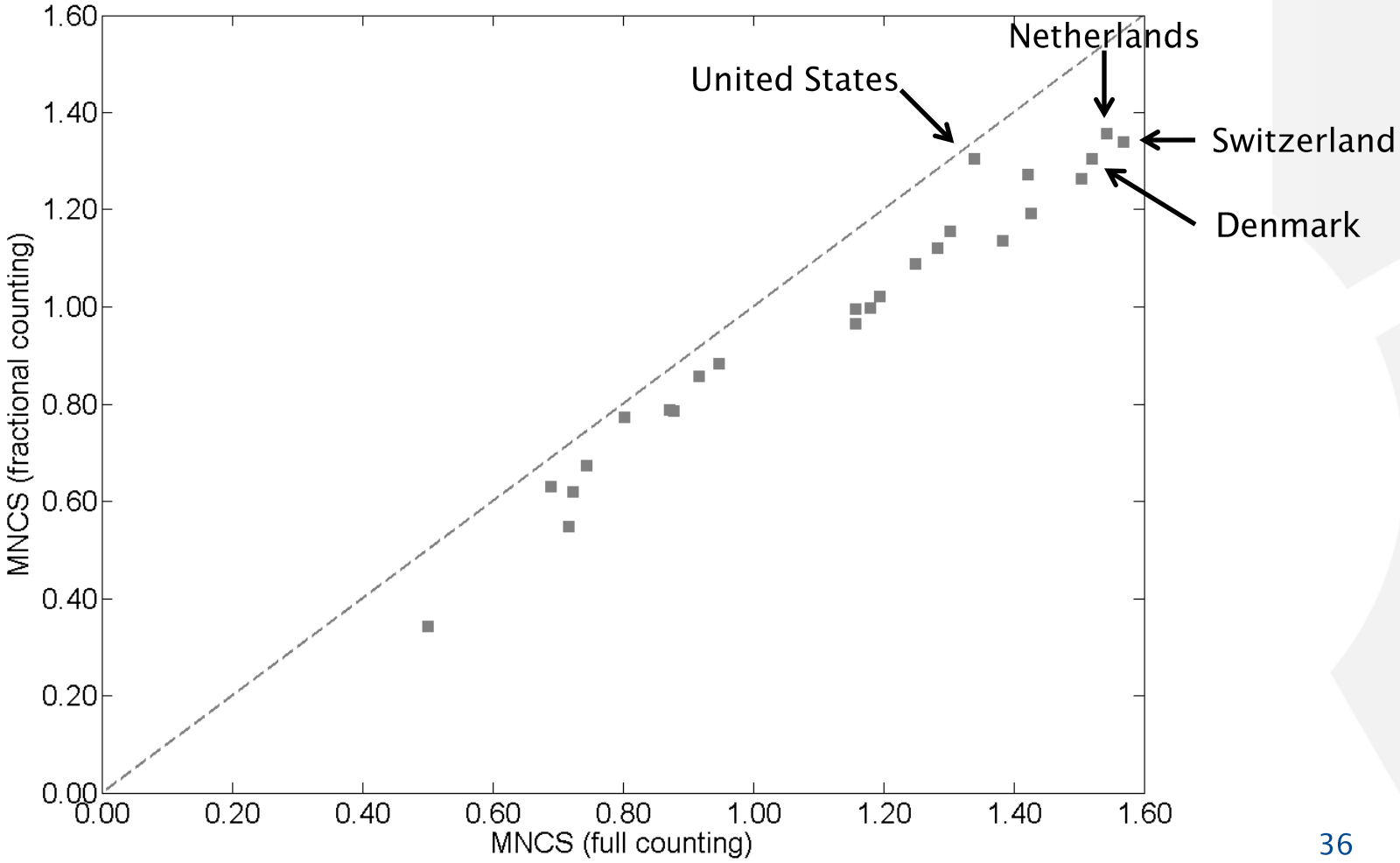
Full counting vs. fractional counting at the university level



Full counting is biased in favor of biomedical research



Full counting vs. fractional counting at the country level



Summary

- Collaborative publications are cited more frequently than non-collaborative publications
- Full counting therefore yields higher citation scores than fractional counting
- Full counting is biased in favor of certain fields of science, in particular biomedicine
- Use fractional counting in analyses at the level of institutions or countries

CWTS