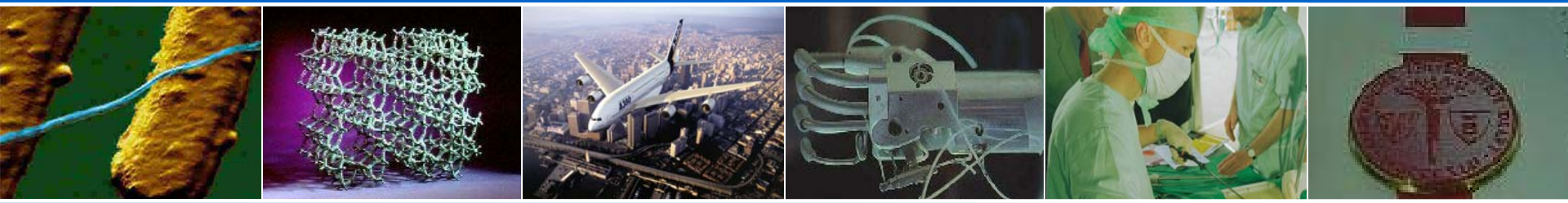
Publish or Perish



Karel Luyben

Meeting FFNT
24 april 2008

Several sheets from Ton van Raan
Calculations done by CWTS Leiden

Content

- Introduction
- Bibliometric data-analysis
- 3TU.Citation analysis
- Some results
- Conclusions

Introduction

- Publishing and being seen
- Selecting the right journals
- Quality versus Quantity
- Career development of a scientist
- Balance between research and education
- Bibliometric studies / citation analysis

Basic Concept: How do we focus on 'quality'?

Scientific performance relates to achieved quality in the contribution to the increase of our knowledge ('scientific progress')

(1) as perceived by others: peer review
(2) as measured by advanced bibliometric analysis

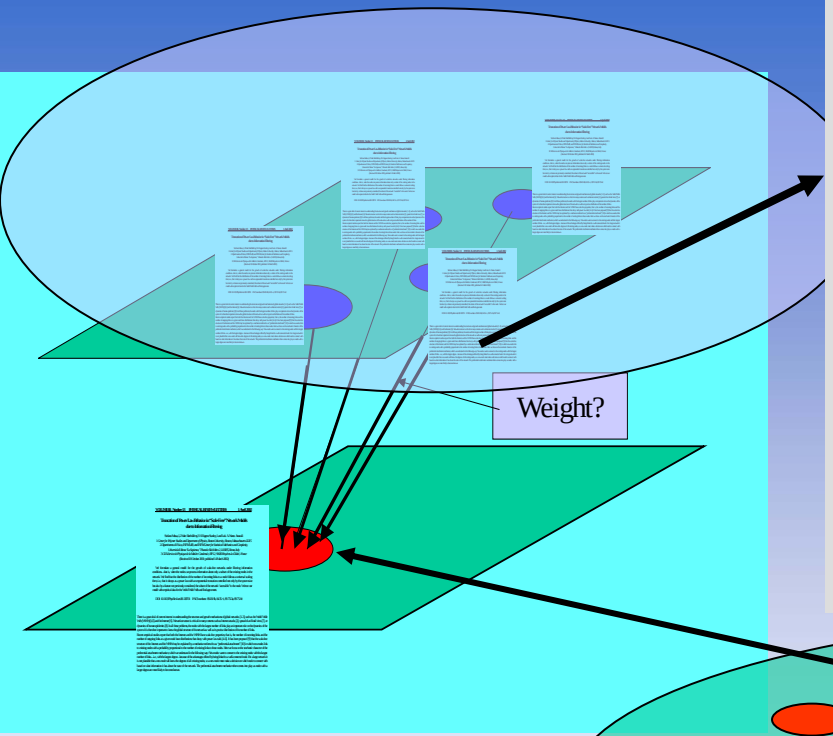
(1) and (2) correlate (very) well at group level

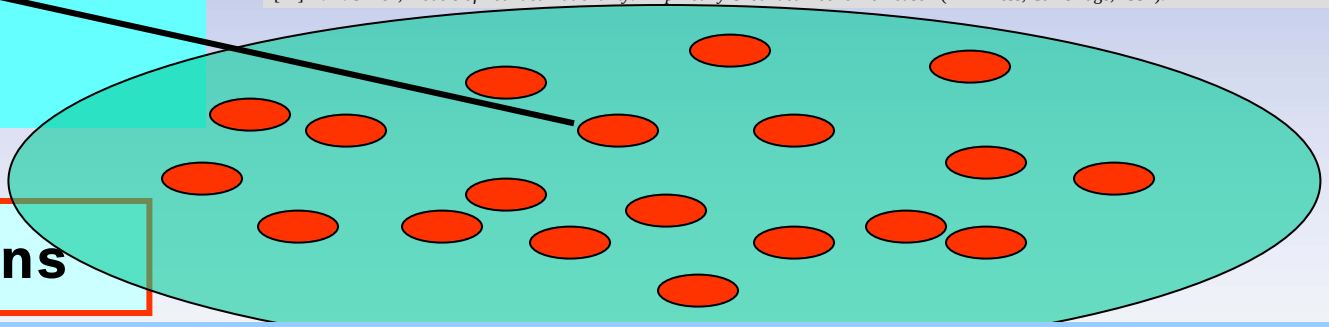
networks leading, possibly, to different dynamics, e.g., for the initiation and spread of epidemics.
In the context of network growth, the impossibility of knowing the degrees of all the nodes comprising the network due to the filtering process—and, hence, the inability to make the optimal, rational, choice—is not altogether unlike the "bounded rationality" concept of Simon [17]. Remarkably, it appears that, for the description of WWW growth, the preferential attachment mechanism, originally proposed by Simon [10], must be modified along the lines of another concept also introduced by him—bounded rationality [17].

We thank R. Albert, P. Ball, A.-L. Barabási, M. Buchanan, J. Camacho, and R. Guimerà for stimulating discussions and helpful suggestions. We are especially grateful to R. Kumar for sharing his data. We thank NIH/NCRR (P41 RR13622) and NSF for support.

[1] S. H. Strogatz, Nature (London) **410**, 268 (2001).
[2] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).
[3] S. N. Dorogovtsev and J. F. F. Mendes, Adv. Phys. (to be published).
[4] R. Albert, H. Jeong, and A.-L. Barabási, Nature (London) **401**, 130 (1999).
[5] B. A. Huberman and L. A. Adamic, Nature (London) **401**, 131 (1999); R. Kumar *et al.*, in *Proceedings of the 25th International Conference on Very Large Databases* (Morgan Kaufmann Publishers, San Francisco, 1999), p. 639; A. Broder *et al.*, Comput. Netw. **33**, 309 (2000); P. L. Krapivsky, S. Redner, and F. Leyvraz, Phys. Rev. Lett. **85**, 4629 (2000); S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, *ibid.* **85**, 4633 (2000); A. Vazquez, Europhys. Lett. **54**, 430 (2001).
[6] M. Faloutsos, P. Faloutsos, and C. Faloutsos, Comput. Commun. Rev. **29**, 251 (1999); G. Caldarelli, R. Marchetti, and L. Pietronero, Europhys. Lett. **52**, 386 (2000); A. Medina, I. Matta, and J. Byers, Comput. Commun. Rev. **30**, 18 (2000); R. Pastor-Satorras, A. Vazquez, and A. Vespignani, arXiv:cond-mat/0105161; L. A. Adamic *et al.*, Phys. Rev. E **64**, 046135 (2001).
[7] F. B. Cohen, *A Short Course on Computer Viruses* (Wiley, New York, 1994); R. Pastor-Satorras and A. Vespignagni, Phys. Rev. Lett. **86**, 3200 (2001); Phys. Rev. E **63**, 066117 (2001).
[8] F. Liljeros, C. R. Edling, L. A. Nunes Amaral, H. E. Stanley, and Y. Åberg, Nature (London) **411**, 907 (2001).
[9] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
[10] Y. Ijiri and H. A. Simon, *Skew Distributions and the Sizes of Business Firms* (North-Holland, Amsterdam, 1977).
[11] G. Bianconi and A.-L. Barabasi, Europhys. Lett. **54**, 436 (2001).
[12] A. F. J. Van Raan, Scientometrics **47**, 347 (2000).
[13] We consider a modification to the network growth rule described earlier in the paper: at each time step *t*, the new node establishes *m* new links, where *m* is drawn from a power law distribution with exponent gout.
[14] For n(I) = const, one recovers the scale-free model of Ref. [9].
[15] It is known [11] that, for an exponential or fat-tailed distribution of fitness, the structure of the network becomes much more complex; in particular, the in-degree distribution is no longer a power law. Hence, we do not consider in this manuscript other shapes of the fitness distribution.
[16] L. A. N. Amaral, A. Scala, M. Barthélémy, and H. E. Stanley, Proc. Natl. Acad. Sci. U.S.A. **97**, 11 149 (2000).
[17] H. A. Simon, *Models of Bounded Rationality: Empirically Grounded Economic Reason* (MIT Press, Cambridge, 1997).

Citing Publications





Cited Publications

All calculations are corrected for self-citations!

What do citations measure?

- Many studies showed positive correlations between citations and qualitative judgments
- In principle it is valid to interpret citations in terms of intellectual influence which is an important aspect of scientific quality
- Thus, the concepts of citation impact and scientific quality do not coincide 'automatically'

Quality judgment by peers may relate to 'tacit knowledge' such as the capability to build instruments, or to acknowledgment of 'being ahead of time' and to 'trust' that a group will come soon with breakthroughs…so peer judgments may include 'credits'….

Bibliometrics always relates to 'codified knowledge', high-value 'tacit' knowledge is immeasurable; it never works with credits; but it may concern hypes, fashions, detested by (some) peers…..

What do scientists think about important aspects of quality?

Opinions of prominent physicists in the Netherlands in discussions (1996) on the introduction of a new research council policy to allocate extra money to top-groups.

As a peer, they focus at:

First of all, quality MUST be demonstrable, no fine words. Quality of research, at least in physics, is demonstrable in the following aspects:

* Effective publication, i.e., in the best possible journals;
* Invited lectures in top-conferences of the field and/or at top-universities;
* High citation scores
* Citations in important reviews;
* Collaboration with top-groups;
* Continuation of PhD students;
* Continuation in research council grants
* Long-term financing by the government and/or the business sector

total publication universe

non-CI publ:

Books

Book chapters

Conf. proc.

Reports

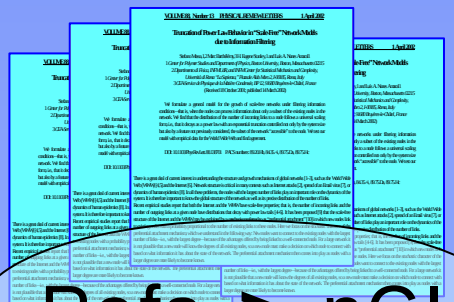
LNCS

ArXiv

Compendex

CI publ sub-universe

7,000 j; 1,000,000p/y



Refs > nCI

Medline

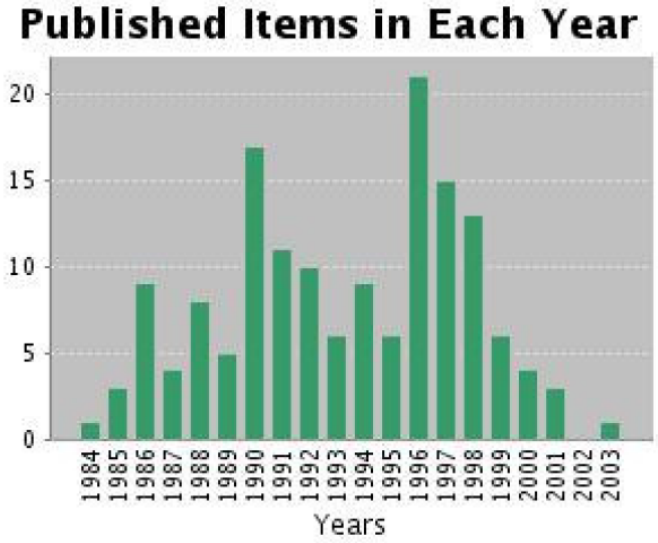
Example from WOS

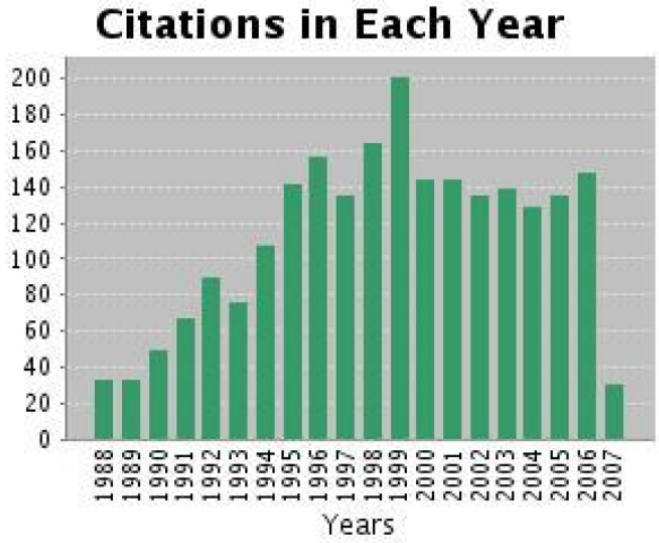
BACK

Page 1 (Records 1 -- 10):

[1]

PRINT





Results found:	**157**
Sum of the Times Cited:	**2,339**
Average Citations per Item:	**14.90**
h-index:	**27**

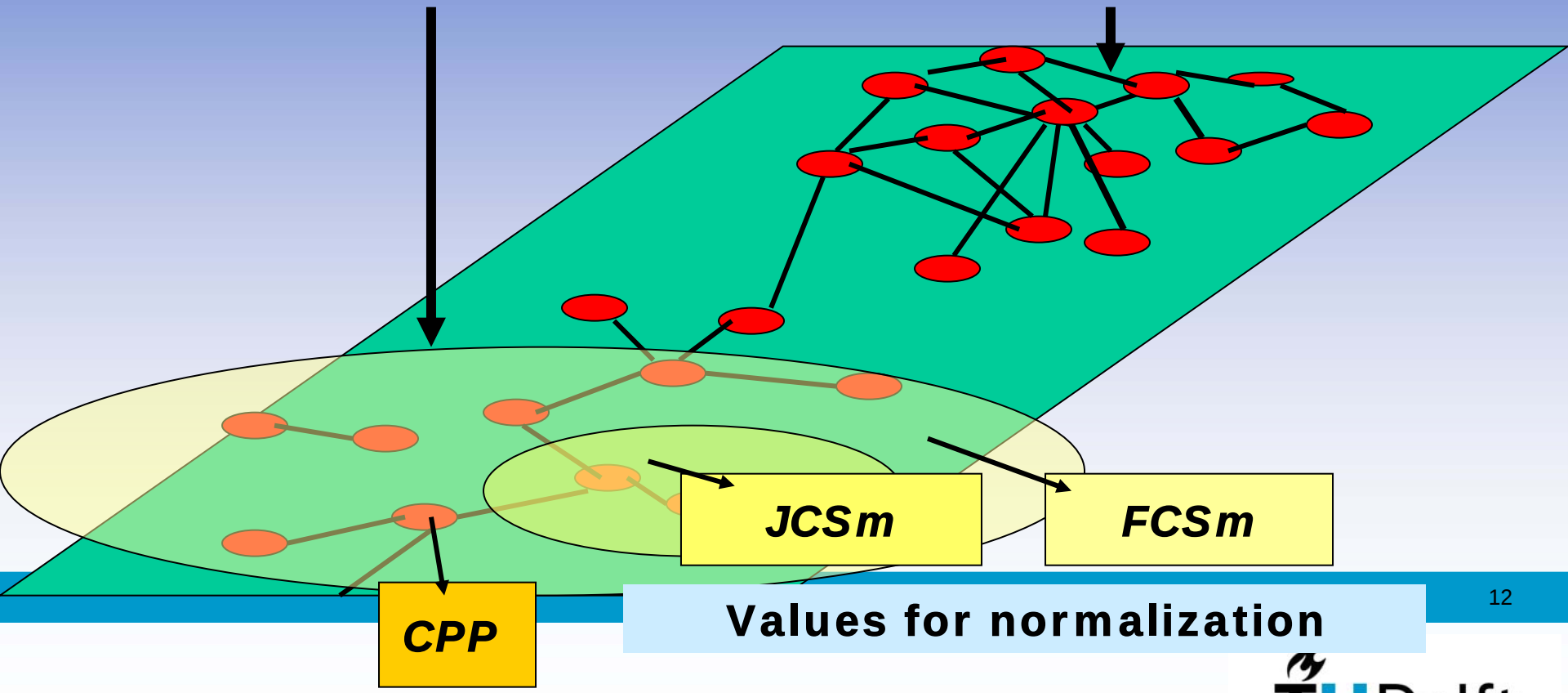
Network of publications (nodes) linked by citations (edges)

Lower citation-density

e.g., applied research, social sciences

Higher citation-density

e.g., basic natural medical research



Values for normalization

Applied Sciences

TU Delft

Calculation of *JCSm* en *FCSm (I)*

	CPP	*JCS*	*FCS*
I	17	16.9	23.7
II	4	3.1	3.0
III	6	4.8	4.1
IV	8	4.8	4.1

Calculation of *JCSm* en *FCSm (II)*

The average citation-score is determined as follows:

$$CPP = \frac{17 + 4 + 6 + 8}{1 + 1 + 1 + 1} = 8.8$$

Average journal citation-score:

$$JCSm = \frac{(1 \times 16.9) + (1 \times 3.1) + (2 \times 4.8)}{1 + 1 + 2} = 7.4$$

CPP / *JCSm*

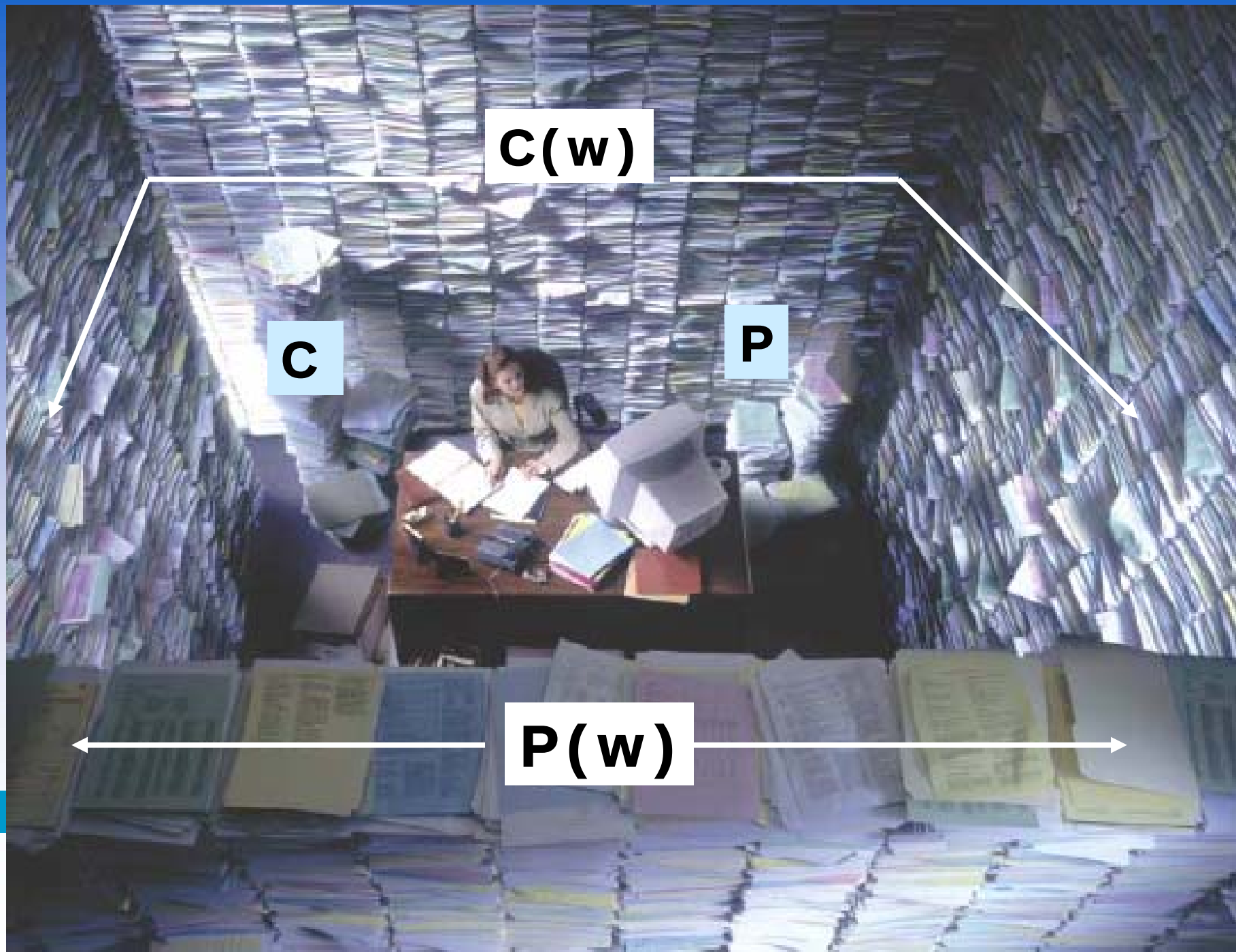
(8.8 / 7.4) = 1.19

Average field citation-score:

$$FCSm = \frac{(1 \times 23.7) + (1 \times 3.0) + (2 \times 4.1)}{1 + 1 + 2} = 8.7$$

CPP / FCSm

(8.8 / 8.7) = 1.01


C(w)
C
P
P(w)

Citing publications

Field-specific normalization

$$\frac{C(A)/P(A)}{C(f)/P(f)} = CPP/FCSm$$

cf

+ doc. type normalization
+ no self-citations, also not in C(f)!

Cited publications

CPP/FCSm $\leq$ 0.80:	performance significantly below internat. average, class A;
0.80 < *CPP/FCSm* $\leq$ 1.20:	performance about internat. average, class B;
1.20 < *CPP/FCSm* $\leq$ 2.00:	performance significantly above internat. average, class C;
2.00 < *CPP/FCSm* $\leq$ 3.00:	performance in internat. perspective is very good, class D;
CPP/FCSm > 3.00:	performance in internat. perspective is excellent, class E.

3TU.Citatie-analyse

1) Benchmarkstudie (top-down)

Een internationale vergelijking van de citatie-impact van de 3TU met een groep buitenlandse (technische) topuniversiteiten. Hiervoor hoeven de onderzoekers niet lastig gevallen te worden. De data zijn de publicaties in de ISI*-database. Op grond van de adressen van de auteurs worden de citaties toegekend aan een universiteit, zonder verdere uitsplitsing naar faculteit of ander organisatie-onderdeel. Wèl kunnen de citatiescores van een universiteit nog onderverdeeld worden naar ISI-fields (technologiegebieden die bestaan uit groepen tijdschriften).

2) Bottom-up studie

Een citatie-analyse van de publicaties over de periode 1997 – 2006 [= 10 jaar] van WP dat op 1 januari 2007 in *vaste* dienst was. De lijsten met publicaties moeten door de auteurs zelf worden gecontroleerd. Verder wordt van hen gevraagd te controleren of hun gegevens (faculteit, instituut e.d.) kloppen.

* ISI is het International Science Institute in de USA, nu Thomson geheten. Dit institute catalogiseert een zeer groot deel van alle wetenschappelijke publikaties.

Bibliometric indicators(i)

P	The number of articles (normal articles, letters, notes and reviews) published in journals processed for the Web of Science (CI) versions of the Science Citation Index, the Social Science Citation Index, the Arts and Humanities Citation Index (see Section 2.1).
C	The number of citations recorded in CI journals (as contained in Web of Science CI version) to all articles involved. Self-citations are excluded.
CPP	The average number of citations per publication. Self-citations are not included.
FCSm	Reference value. The average citation rate of all articles in the subfields in which the research unit is active. Also indicated as the world citation average in those subfields or 'world subfield average'. Subfields are defined by means of CI subject categories. Self-citations are excluded.

Bibliometric indicators(ii)

CPP/FCSm Field-normalized citation impact: The impact of a research unit's articles, compared to the world citation average in the subfields in which the research unit is active. A '+' ('-') symbol directly after the numerical value indicates that the impact of the research unit's articles is significantly above (below) world (subfield) average.

P*CPP/FCSm 'Brute force indicator': The field-normalized mean citation impact multiplied by the number of publications.

P00-03 Number of papers (normal articles and reviews) published in journals processed for the Web of Science version of Thomson-Scientific Citation Indices (CI) in the period 2000-2003.

Ptop The absolute number of papers that are among the 20%, 10%, 5%, or 1% most frequently cited of all similar papers published in the period 2000 - 2003 and cited in 2000 - 2006.

TABLE 2: BIBLIOMETRIC INDICATORS FOR ALL DISCIPLINES

University	P	Rank	C	Rank	CPP/FCSm	Rank	P* CPP/FCSm	Rank
DELFT UNIV TECHNOL	6,205	211	40,877	259	1.41	99	8,772	169
EINDHOVEN UNIV TECHNOL	4,221	321	27,887	312	1.48	77	6,230	249
UNIV TWENTE	3,159	-	21,009	-	1.42	-	4,497	-
3TU UNIVERSITIES	13,285	-	87,877	-	1.43	-	19,021	-
CALTECH	11,651	74	200,919	29	2.05	7	23,936	32
CARNEGIE MELLON UNIV	5,727	236	61,279	180	2.10	6	12,049	112
IMPERIAL COLL LONDON	18,359	22	216,764	25	1.46	85	26,737	25
INDIAN INST TECHNOL MUMBAI	1,923	-	6,392	-	0.66	-	1,274	-
MIT	16,861	27	299,718	17	2.42	1	40,835	7
NATL UNIV SINGAPORE	12,900	57	73,397	153	1.07	247	13,781	87
STANFORD UNIV	22,255	11	419,551	2	2.15	4	47,772	2
UNIV CALIF BERKELEY	19,132	20	300,134	16	2.04	9	38,978	9
UNIV CAMBRIDGE	23,194	9	311,467	12	1.67	33	38,651	11
UNIV TOKYO	33,802	2	312,658	11	1.20	161	40,726	8
AALBORG UNIV	1,681	-	8,534	-	0.94	-	1,587	-
ETH ZURICH	11,997	68	127,364	67	1.52	62	18,261	50
KATHOLIEKE UNIV LEUVEN	14,948	35	135,286	57	1.25	144	18,701	47
PARISTECH	7,457	170	48,933	226	1.22	152	9,133	159
TECH UNIV HAMBURG HARBURG	790	-	3,314	-	0.92	-	730	-
UNIV AACHEN (RWTH)	7,831	158	58,181	191	1.13	207	8,888	164
UNIV WARWICK	4,374	308	28,806	309	1.13	208	4,959	289
CHALMERS UNIV TECHNOL	4,717	286	27,210	313	1.16	189	5,481	267
EPFL LAUSANNE	5,657	239	42,313	253	1.54	59	8,690	172
GEORGIA INST TECHNOL	7,057	181	56,487	198	1.74	21	12,312	108
TECH UNIV DENMARK	4,166	327	37,393	274	1.57	48	6,556	234
TECH UNIV MUNCHEN	10,219	98	100,739	97	1.36	112	13,854	86
TSING HUA UNIV	10,904	84	32,684	289	0.62	345	6,732	232
UNIV GRONINGEN	9,374	111	95,253	106	1.26	136	11,827	118
UNIV MELBOURNE	11,765	71	104,805	87	1.19	175	13,953	85
UNIV TORONTO	26,541	3	322,587	9	1.46	84	38,781	10
UNIV UTRECHT	14,853	37	161,433	43	1.35	116	20,096	41
BENCHMARK UNIVERSITIES	308,057	-	3,309,724	-	1.48	-	456,872	-
NETHERLANDS UNIVERSITIES	87,217	-	869,900	-	1.31	-	113,934	-
NETHERLANDS	106,009	-	1,039,562	-	1.31	-	139,213	-
EUROPE	1,720,384	-	12,378,276	-	1.02	-	1,761,643	-

Table 3.1: Internal coverage percentages of the Thomson Scientific/ISI Citation Indexes

Internal Coverage Percentage			
80-100%	***60-80%***	***40-60%***	***<40%***
Biochem & Mol Biol	Appl Phys & Chem	Mathematics	Other Soc Sci
Biol Sci – Humans	Biol Sci – Anim & Plants	Economics	Humanities & Arts
Chemistry	Psychol & Psychiat	Engineering	
Clinical Medicine	Geosciences		
Phys & Astron	Soc Sci ~ Medicine		

journal coverage, in particular for engineering, social sciences, humanities

Field	1991	1996	2001	2006
MEDICAL & LIFE SCIENCES				
AGRICULTURE AND FOOD SCIENCE	66%	66%	73%	75%
BASIC LIFE SCIENCES	87%	89%	93%	93%
BASIC MEDICAL SCIENCES	76%	75%	80%	84%
BIOLOGICAL SCIENCES	72%	74%	80%	82%
BIOMEDICAL SCIENCES	86%	87%	90%	90%
CLINICAL MEDICINE	82%	82%	85%	85%
HEALTH SCIENCES	50%	47%	57%	62%
NATURAL SCIENCES				
ASTRONOMY AND ASTROPHYSICS	75%	79%	82%	86%
CHEMISTRY AND CHEMICAL ENGINEERING	77%	80%	86%	88%
COMPUTER SCIENCES	38%	37%	42%	43%
EARTH SCIENCES AND TECHNOLOGY	60%	60%	69%	74%

ENGINEERING SCIENCES

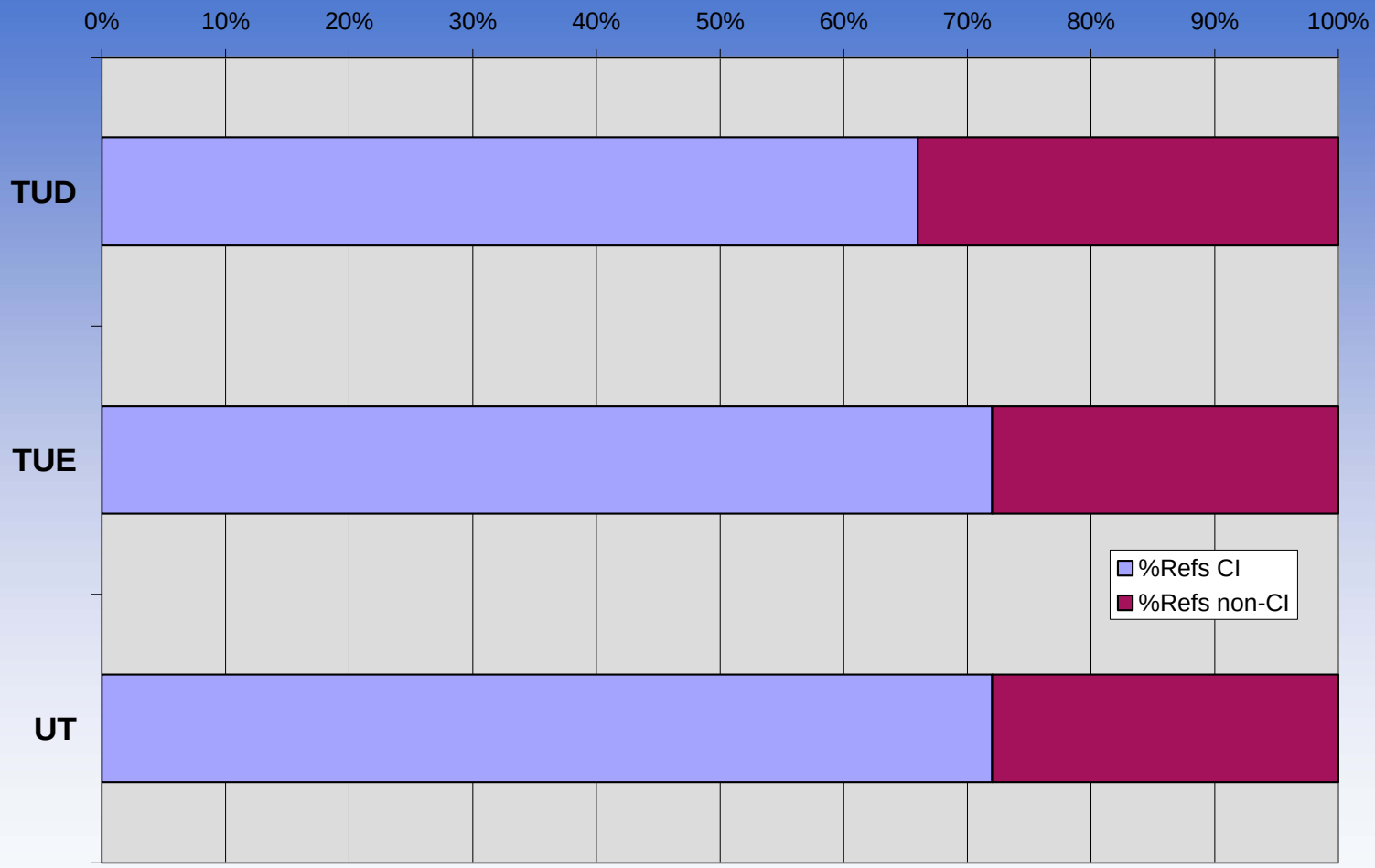
CIVIL ENG & CONSTRUCTION	**37%**	**33%**	**34%**	**45%**
ELECTRICAL ENG & TELECOMM	**54%**	**52%**	**52%**	**53%**
ENERGY SCIENCE & TECHNOLOGY	**54%**	**48%**	**53%**	**59%**
GENERAL & INDUSTRIAL ENG	**42%**	**37%**	**44%**	**54%**
INSTRUMENTS & INSTRUMENTATION	**67%**	**62%**	**71%**	**69%**
MECHANICAL ENG & AEROSPACE	**58%**	**53%**	**57%**	**64%**

ECONOMICS AND BUSINESS	35%	36%	35%	43%
EDUCATIONAL SCIENCES	27%	31%	30%	36%
MANAGEMENT AND PLANNING	23%	24%	27%	36%
POLITICAL SCIENCE AND PUBLIC ADMINISTRATION	17%	18%	20%	24%
PSYCHOLOGY	59%	59%	66%	72%
SOCIAL AND BEHAVIORAL SCIENCES, INTERDISCIPLINARY	33%	34%	36%	40%
SOCIOLOGY AND ANTHROPOLOGY	22%	27%	29%	34%
LAW, ARTS AND HUMANITIES				
CREATIVE ARTS, CULTURE AND MUSIC	17%	14%	16%	14%
HISTORY, PHILOSOPHY AND RELIGION	24%	23%	25%	27%
LAW AND CRIMINOLOGY	27%	32%	32%	31%
LITERATURE	14%	12%	11%	11%
MULTIDISCIPLINARY JOURNALS	78%	83%	87%	87%

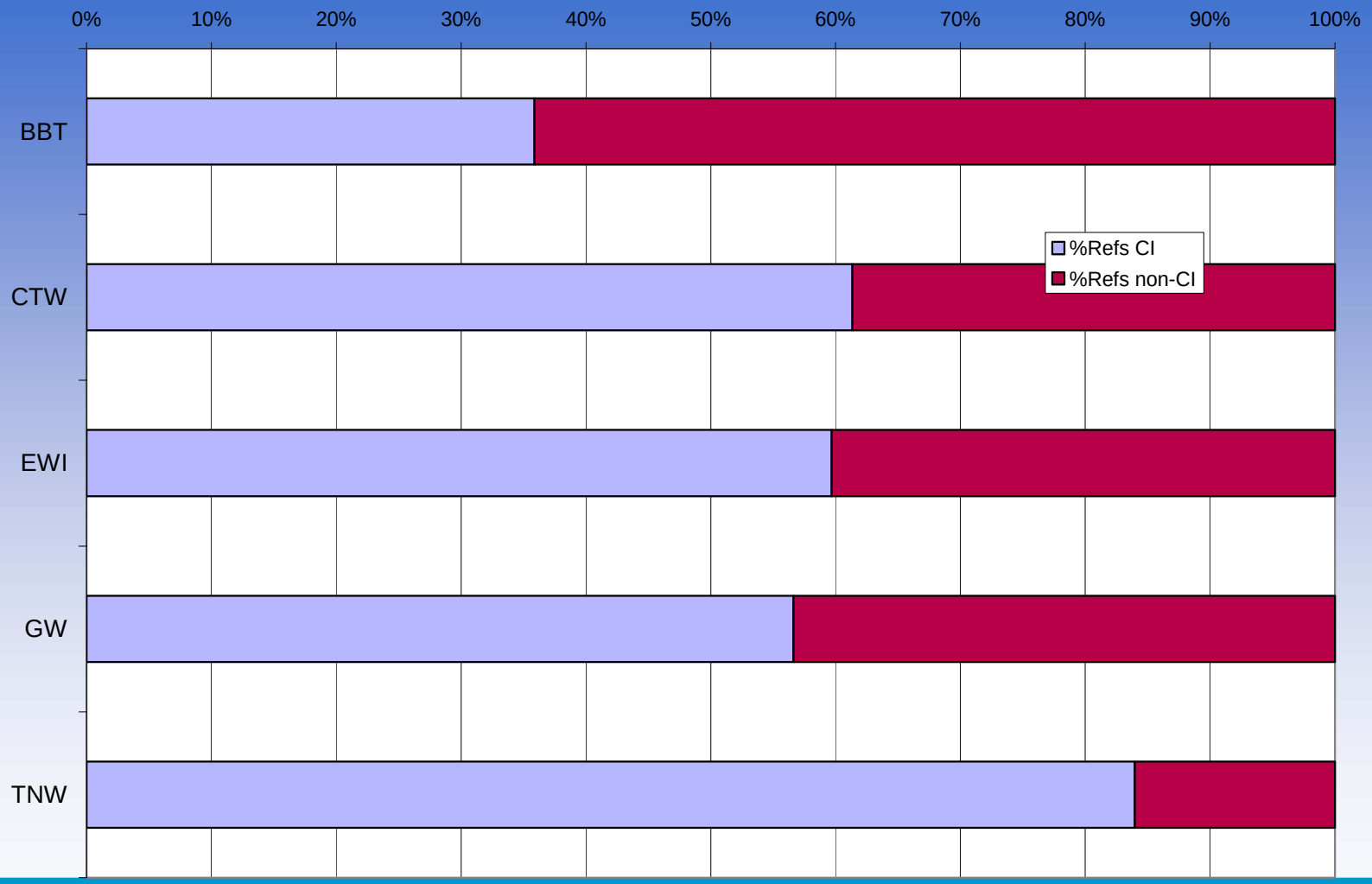
Appli

TU Delft

Coverage 3TU in Web of Science



Coverage: Schools (Faculties) UT



Output of UT-Schools (97-06)

Output of UT-Schools (97-06)

Impact of UT Schools (97-06)

Impact of UT Schools (series)

Institutes of UT: P vs. CPP/FCSm

Visibility UT in top-10%

Conclusions

- The primary tasks of universities are teaching and doing research
- Young scientists should be guided in their development:
  - from dependent to independent
  - from being coached to coaching
  - from more egoistic to altruistic
- In science quality comes for quantity
- More and more scientific result are published in journals
- More and more journals are covered in the Web of Science
- Citation studies can help in evaluating scientific quality
- Citation results should handled with care
- Especially individual results