New bibliometric mapping methods to identify structures and new developments in science

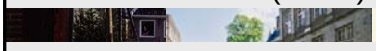
CRIS 2010, Aalborg University
June 3, 2010

Ton van Raan
Center for Science and Technology Studies (CWTS)
Leiden University

Leiden University,

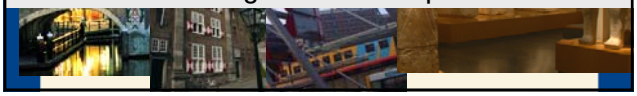
oldest in the Netherlands, 1575,

European League of Research Universities (LERU)



Leiden, historic city (2th, 11th C.), strong cultural (arts, painting) & scientific tradition

one of the largest science parks in EU



Basics of bibliometric analysis:

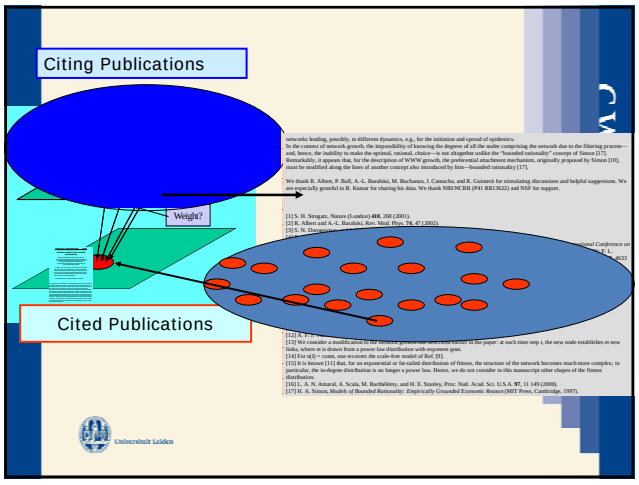
science as a system structured by communication

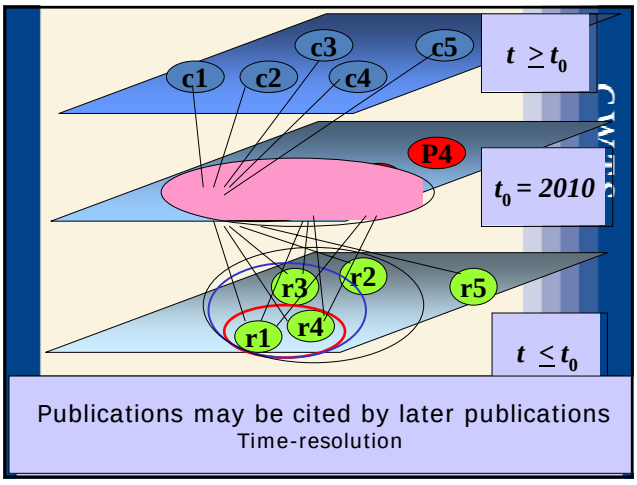
not just simple counting

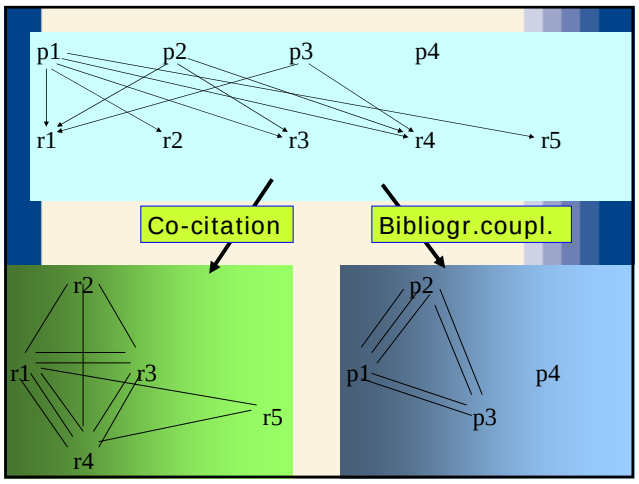
discovery of patterns

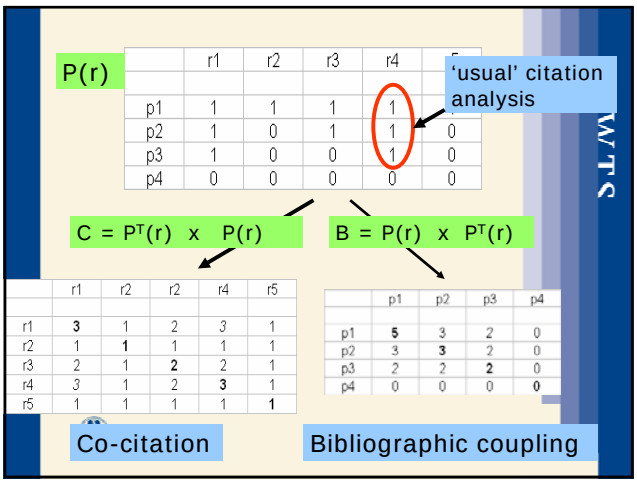
mathematical foundations

CWTS


Citing Publications
Weight?
Cited Publications
Universiteit Leiden


c1
c2
c3
c4
c5
t ≥ t0
P4
t0 = 2010
r1
r2
r3
r4
r5
t ≤ t0
Publications may be cited by later publications
Time-resolution


p1
p2
p3
p4
r1
r2
r3
r4
r5
Co-citation
Bibliogr.coupl.


P(r)
'usual' citation analysis
C = PT(r) x P(r)
B = P(r) x PT(r)
Co-citation
Bibliographic coupling

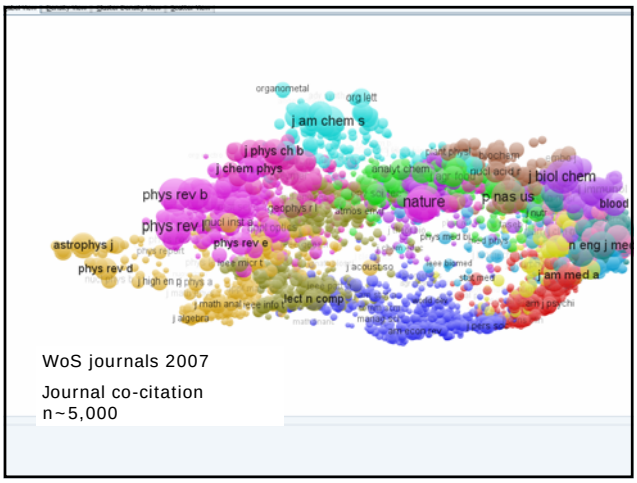
Structures on different levels of publication-aggregation:

Publications
Journals (sets of publications)
Fields (set of journals, WoS)

CWTS

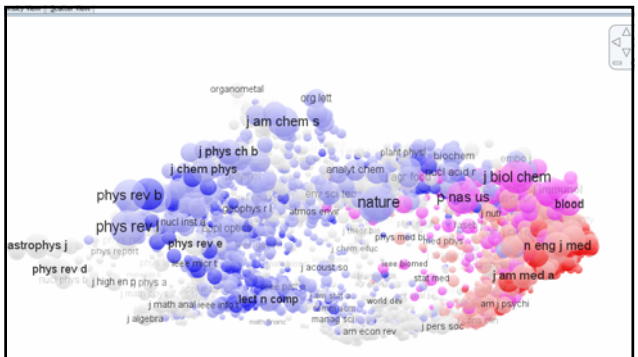
Universiteit Leiden

9

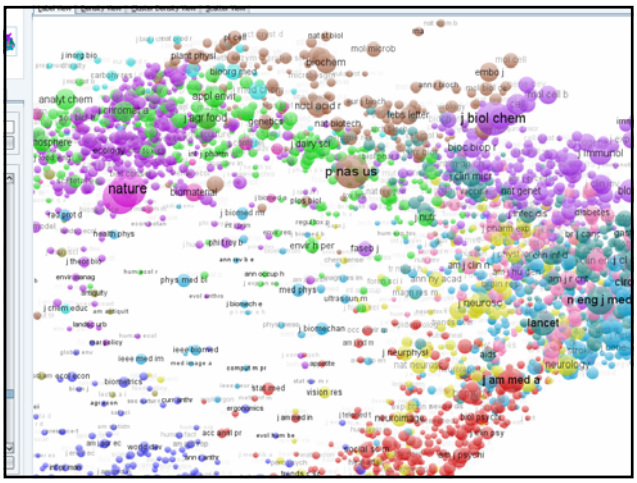


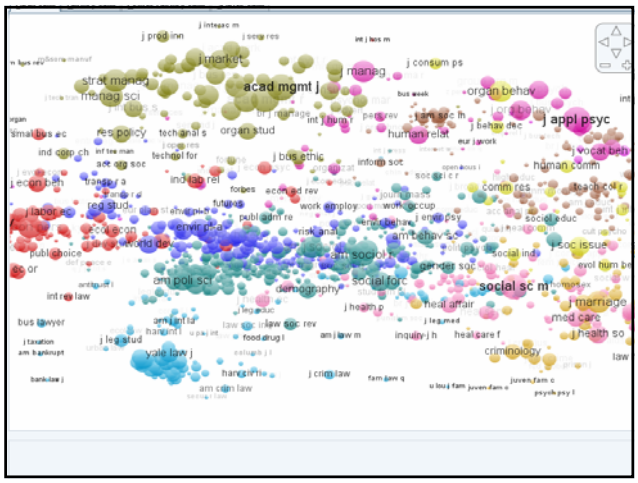
WoS journals 2007

Journal co-citation
n~5,000



Same, now highlighting areas of university-industry collaboration

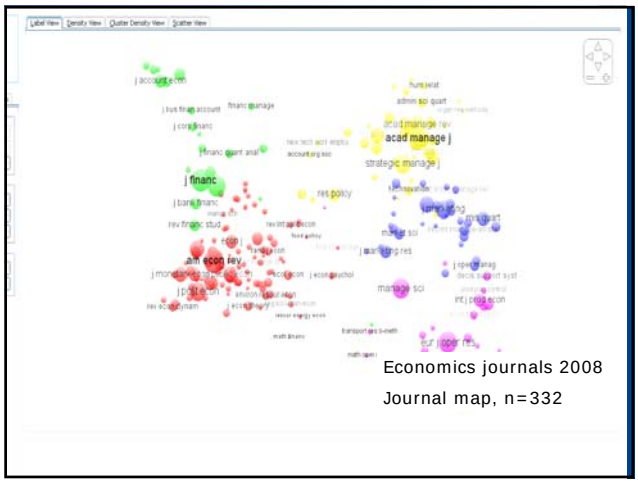




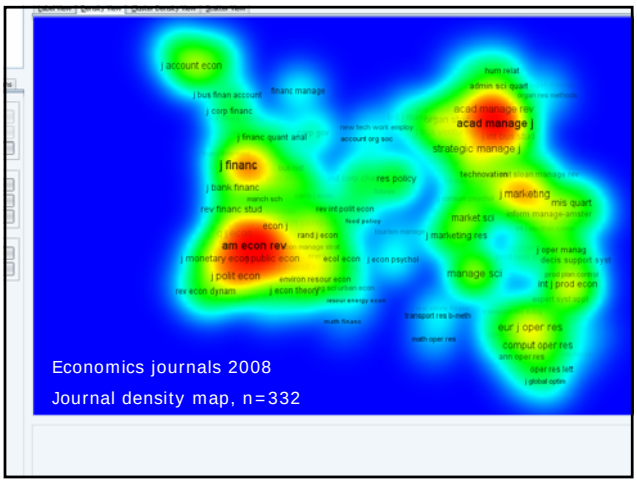
Journal Classification by j-to-j citations is severely hampered by 'everything relates to everything':

Need for consistent 'local' clustering

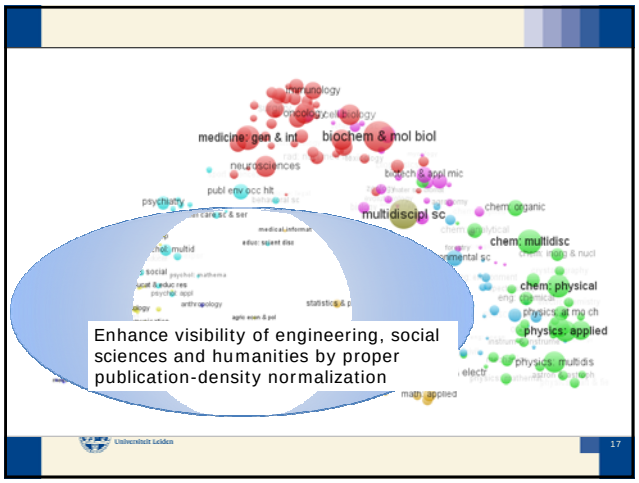
CWTS is developing novel mathematical methods to unify clustering and mapping ('Lay-out')

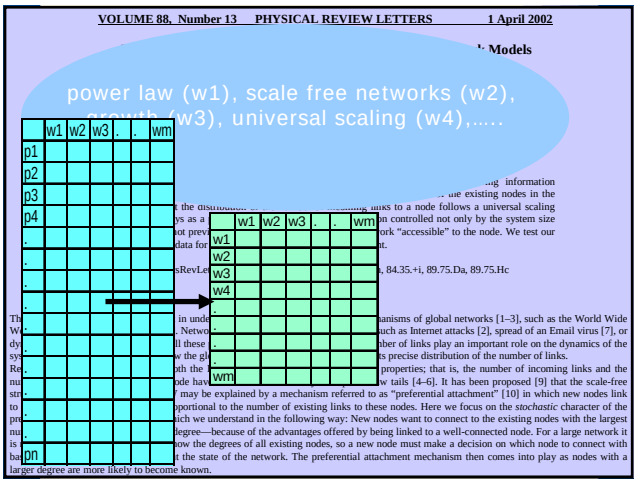


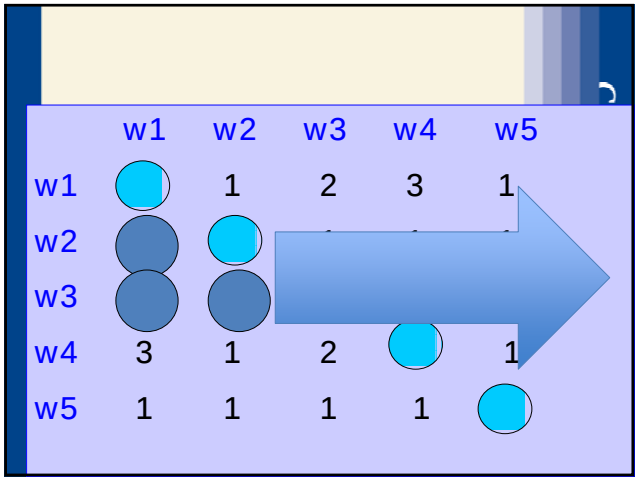
Economics journals 2008
Journal map, n=332

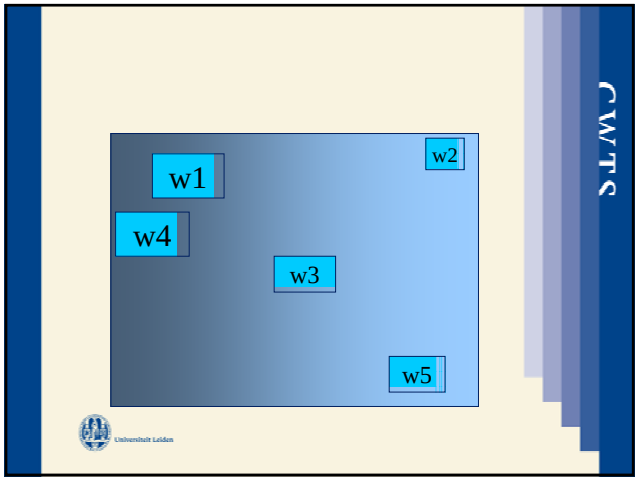


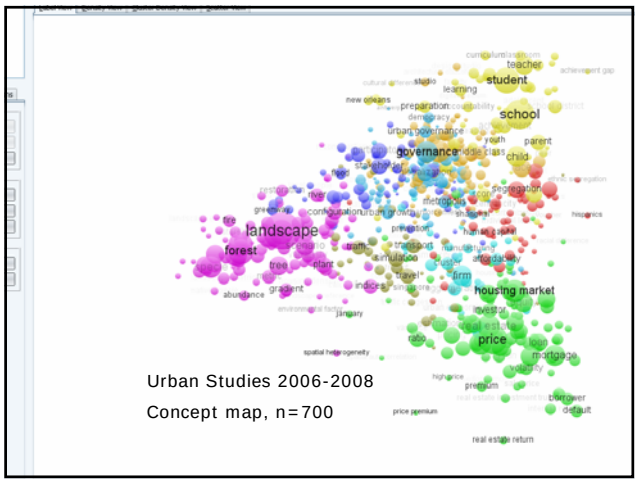
Economics journals 2008
Journal density map, n=332

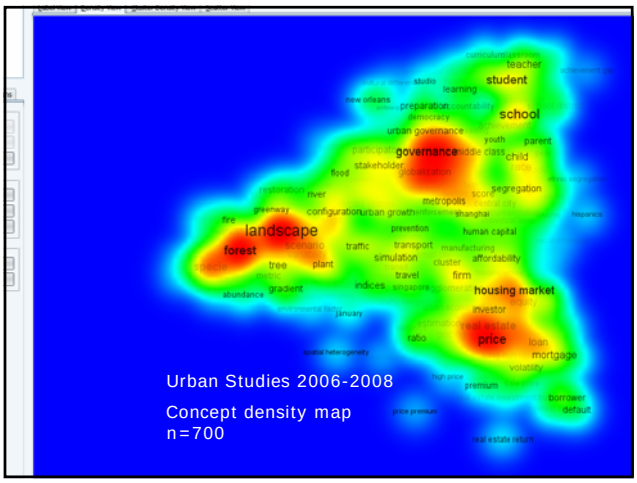

immunology
oncology
cell biology
medicine: gen & int
biochem & mol biol
neurosciences
biotech & appl mic
publ env occ hlt
psychiatry
multidiscipl sc
chem: organic
chem: multidisc
chem: physical
physics: applied
physics: multidis
math: applied
Enhance visibility of engineering, social sciences and humanities by proper publication-density normalization
Universiteit Leiden
17

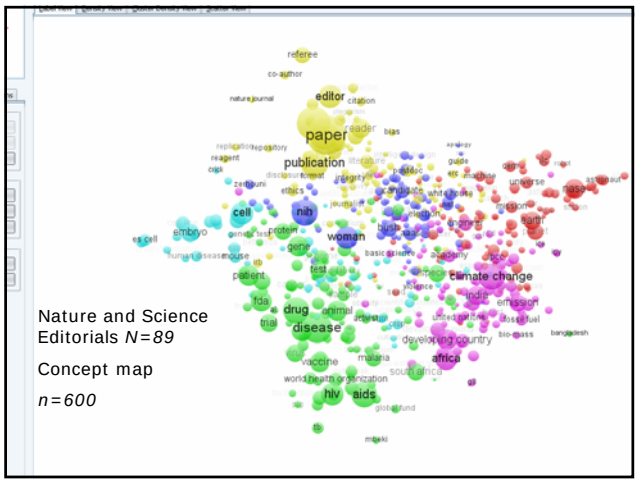

VOLUME 88, Number 13 PHYSICAL REVIEW LETTERS 1 April 2002
Models
power law (w1), scale free networks (w2),
growth (w3), universal scaling (w4),.....
w1 w2 w3 . . wm
p1
p2
p3
p4
wm
information
the existing nodes in the
links to a node follows a universal scaling
on controlled not only by the system size
ork "accessible" to the node. We test our
nt.
84.35.+i, 89.75.Da, 89.75.Hc
anisms of global networks [1–3], such as the World Wide
such as Internet attacks [2], spread of an Email virus [7], or
ber of links play an important role on the dynamics of the
ts precise distribution of the number of links.
properties; that is, the number of incoming links and the
w tails [4–6]. It has been proposed [9] that the scale-free
may be explained by a mechanism referred to as "preferential attachment" [10] in which new nodes link
portional to the number of existing links to these nodes. Here we focus on the stochastic character of the
ich we understand in the following way: New nodes want to connect to the existing nodes with the largest
legree—because of the advantages offered by being linked to a well-connected node. For a large network it
ow the degrees of all existing nodes, so a new node must make a decision on which node to connect with
t the state of the network. The preferential attachment mechanism then comes into play as nodes with a
larger degree are more likely to become known.

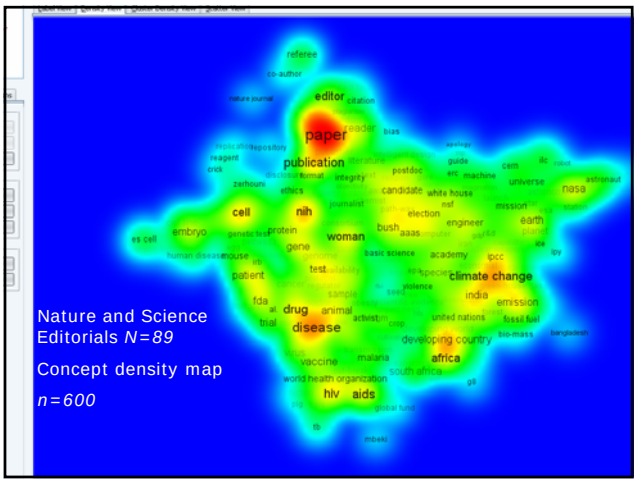

w1 w2 w3 w4 w5
w1 1 2 3 1
w2
w3
w4 3 1 2 1
w5 1 1 1 1


w1
w2
w3
w4
w5
Universiteit Leiden


Urban Studies 2006-2008
Concept map, n=700

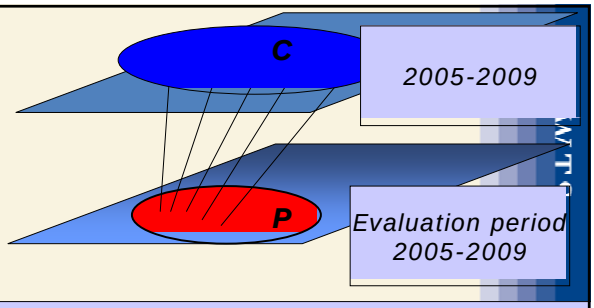

Urban Studies 2006-2008
Concept density map
n=700


Nature and Science
Editorials N=89
Concept map
n=600


Nature and Science
Editorials N=89
Concept density map
n=600

Mapping is crucial for proper citation impact normalization in the context of evaluation studies

CWTS



In most cases we work with aggregates of publications:
groups, institutes, fields, universities, countries….
as well as aggregates over time

CWTS

From other disciplines

From emerging fields

From research devoted to societal, economical and technological problems

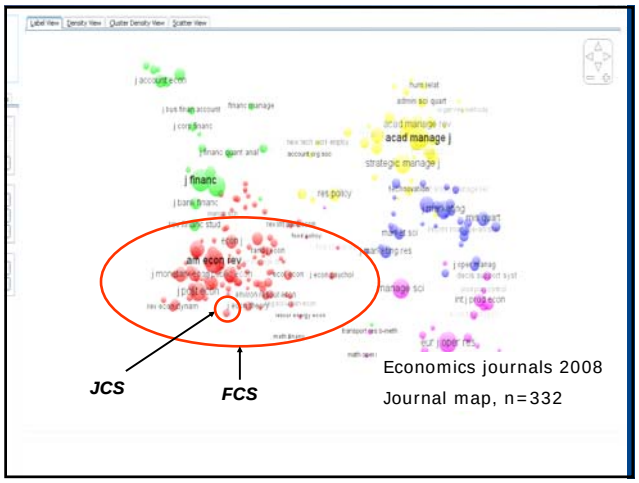
From industry

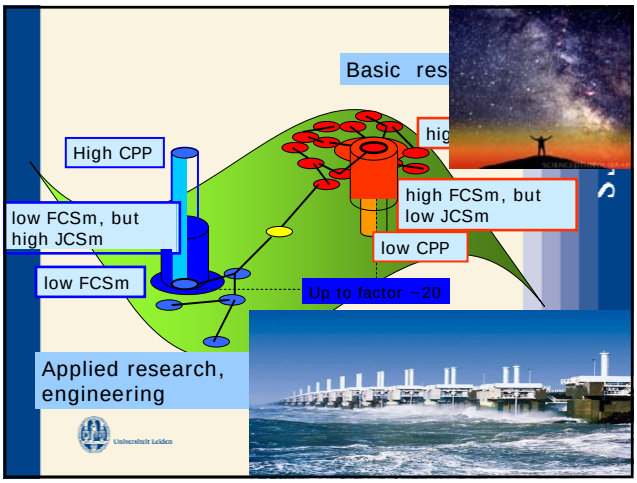
From international top-groups

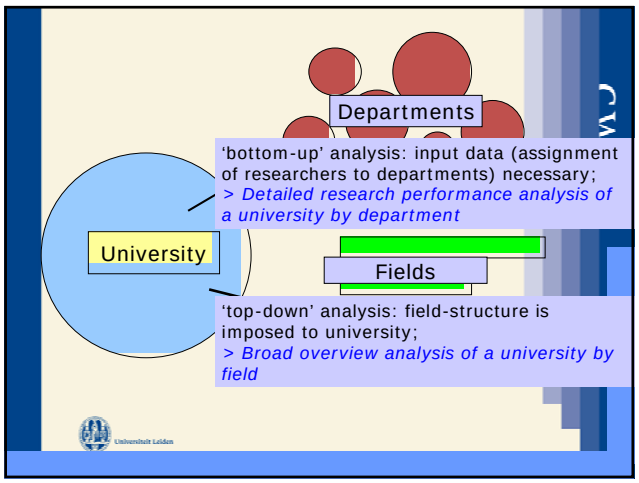
These all f(t)!! > Sleeping Beauties

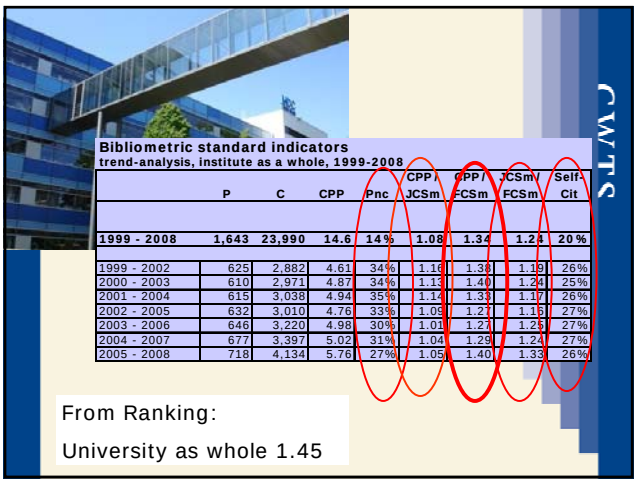
A single publication remains a 'building block' of the whole system, it has its own international environment, namely the journal and the field.

For *each* publication we *MUST* normalize its citation characteristics to the overall citation characteristics of the journal and/or the field







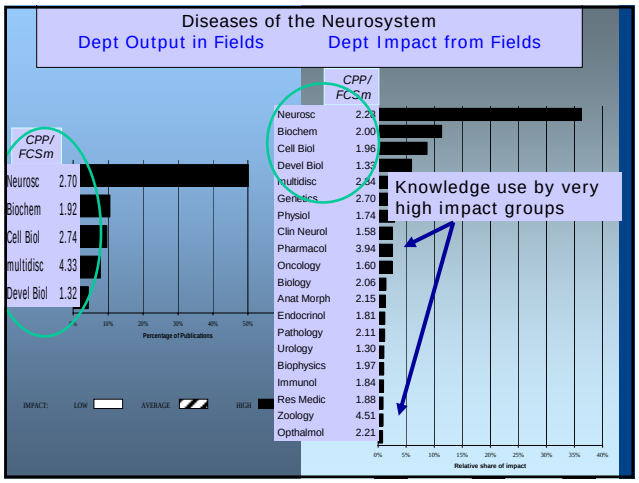


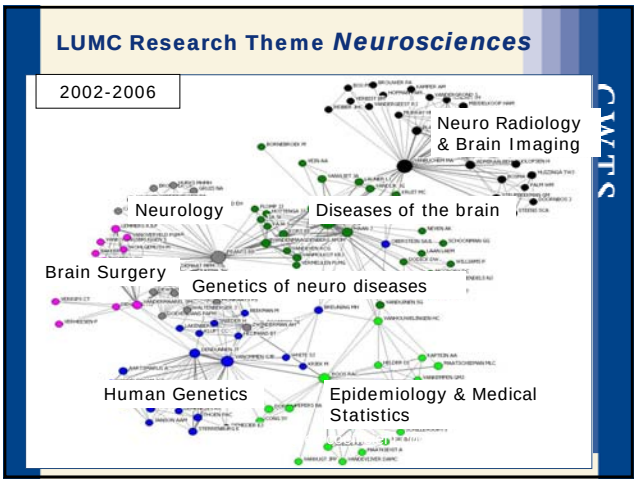
Bibliometric standard indicators
trend-analysis, institute as a whole, 1999-2008

	P	C	CPP	Pnc	CPP / JCSm	CPP / FCSm	JCSm / FCSm	Self-Cit
1999 - 2008	**1,643**	**23,990**	**14.6**	**14%**	**1.08**	**1.34**	**1.24**	**20%**
1999 - 2002	625	2,882	4.61	34%	1.16	1.38	1.19	26%
2000 - 2003	610	2,971	4.87	34%	1.13	1.40	1.24	25%
2001 - 2004	615	3,038	4.94	35%	1.14	1.33	1.17	26%
2002 - 2005	632	3,010	4.76	33%	1.09	1.27	1.16	27%
2003 - 2006	646	3,220	4.98	30%	1.01	1.27	1.25	27%
2004 - 2007	677	3,397	5.02	31%	1.04	1.29	1.24	27%
2005 - 2008	718	4,134	5.76	27%	1.05	1.40	1.33	26%

From Ranking:

University as whole 1.45


Diseases of the Neurosystem
Dept Output in Fields
Dept Impact from Fields
CPP/FCSm
Neurosc 2.70
Biochem 1.92
Cell Biol 2.74
multidisc 4.33
Devel Biol 1.32
Percentage of Publications
IMPACT: LOW AVERAGE HIGH
CPP/FCSm
Neurosc 2.23
Biochem 2.00
Cell Biol 1.96
Devel Biol 1.33
multidisc 2.34
Genetics 2.70
Physiol 1.74
Clin Neurol 1.58
Pharmacol 3.94
Oncology 1.60
Biology 2.06
Anat Morph 2.15
Endocrinol 1.81
Pathology 2.11
Urology 1.30
Biophysics 1.97
Immunol 1.84
Res Medic 1.88
Zoology 4.51
Opthalmol 2.21
Knowledge use by very high impact groups
Relative share of impact


LUMC Research Theme Neurosciences
2002-2006
CWTS
Neuro Radiology & Brain Imaging
Neurology
Diseases of the brain
Brain Surgery
Genetics of neuro diseases
Human Genetics
Epidemiology & Medical Statistics

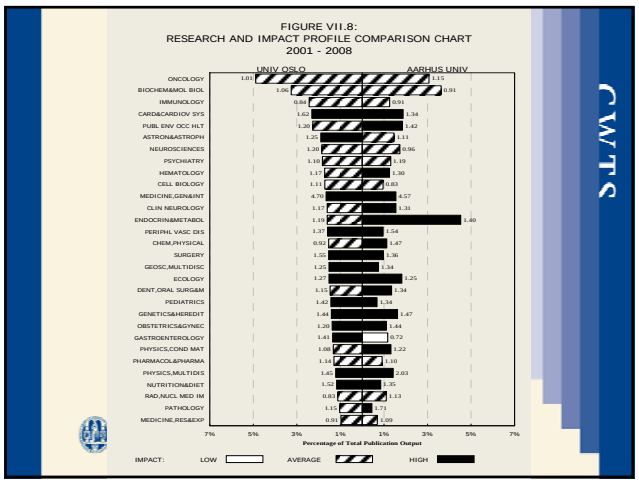


FIGURE VII.8:
RESEARCH AND IMPACT PROFILE COMPARISON CHART
2001 - 2008
UNIV OSLO
AARHUS UNIV
CWTS
Percentage of Total Publication Output
IMPACT: LOW AVERAGE HIGH

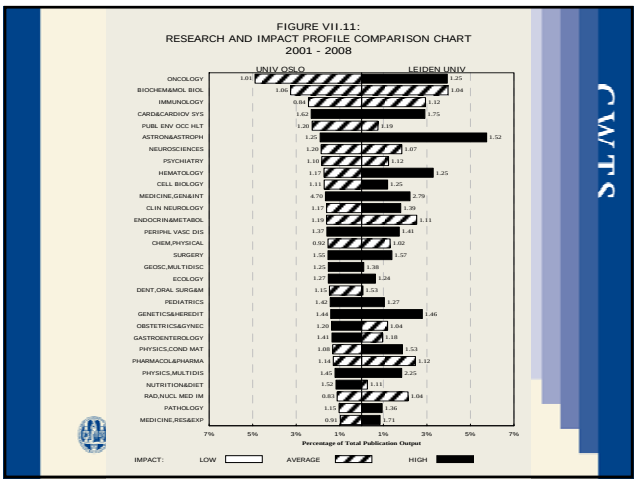
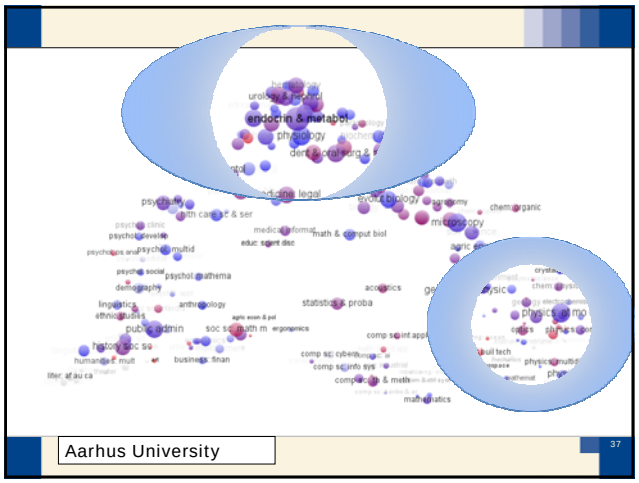
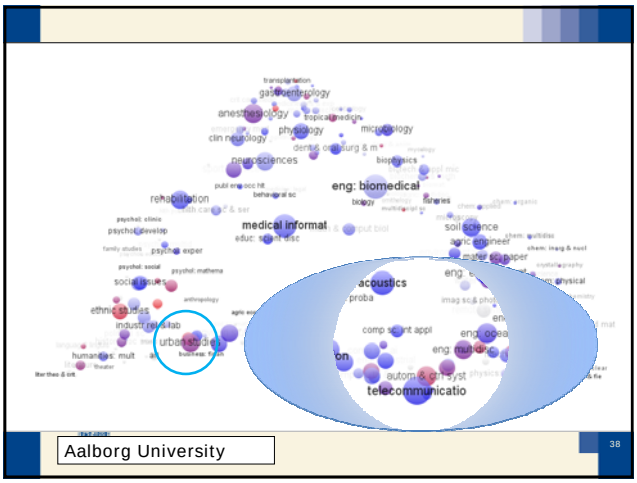


FIGURE VII.11:
RESEARCH AND IMPACT PROFILE COMPARISON CHART
2001 - 2008
UNIV OSLO
LEIDEN UNIV
CWTS
Percentage of Total Publication Output
IMPACT: LOW AVERAGE HIGH



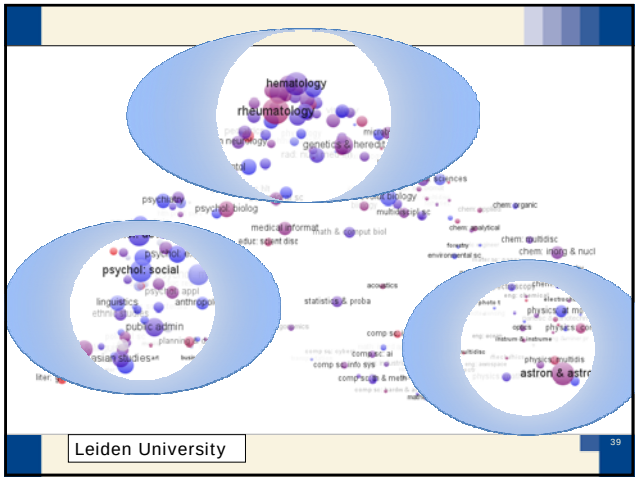
Aarhus University

37



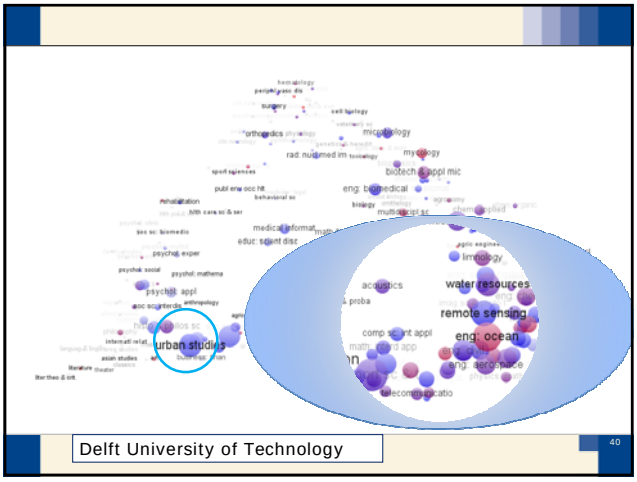
Aalborg University

38



Leiden University

39

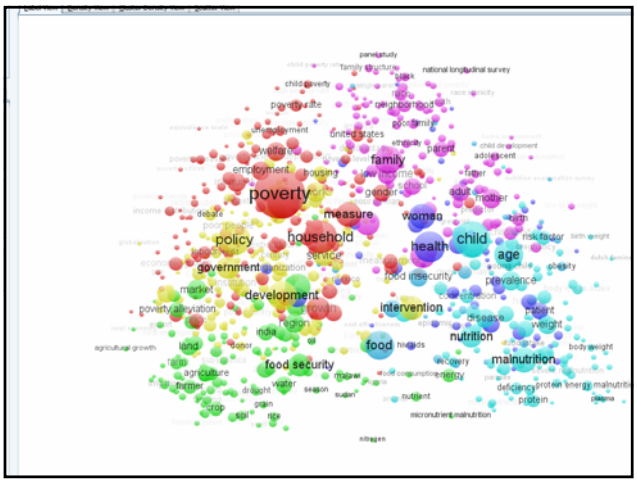


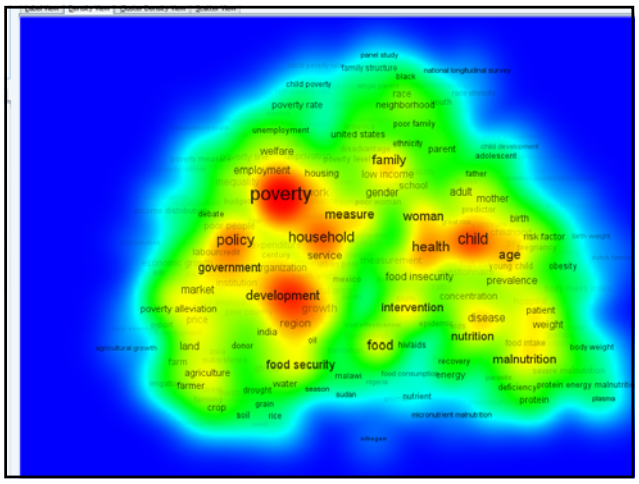
Delft University of Technology

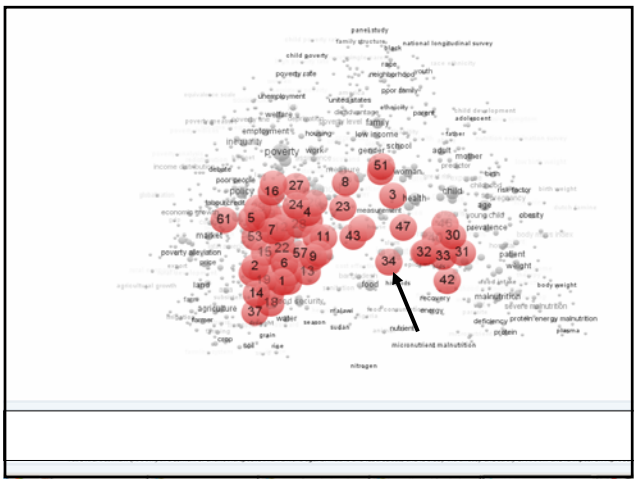
40

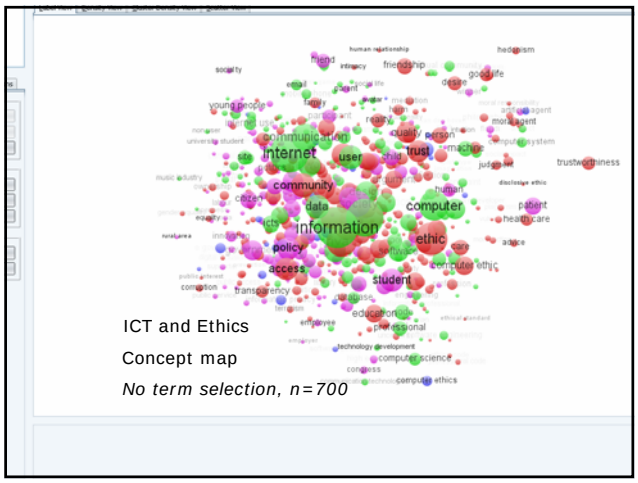
Mapping research on socio-economic problems

Universiteit Leiden





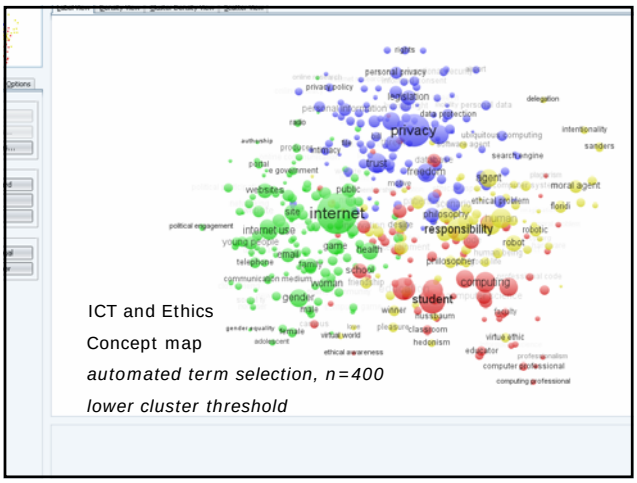




ICT and Ethics

Concept map

No term selection, n=700

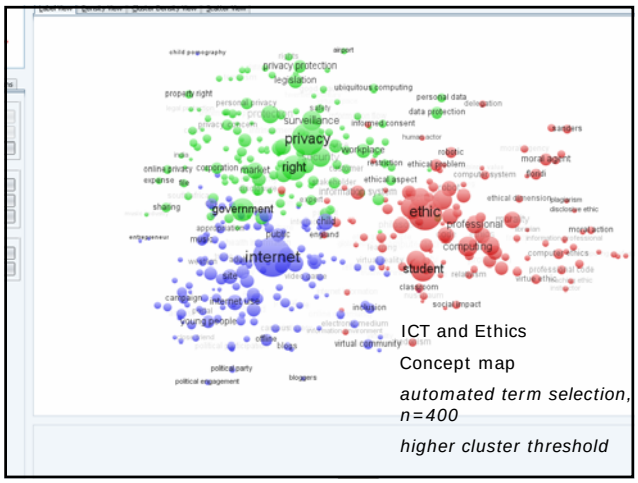


ICT and Ethics

Concept map

automated term selection, n=400

lower cluster threshold

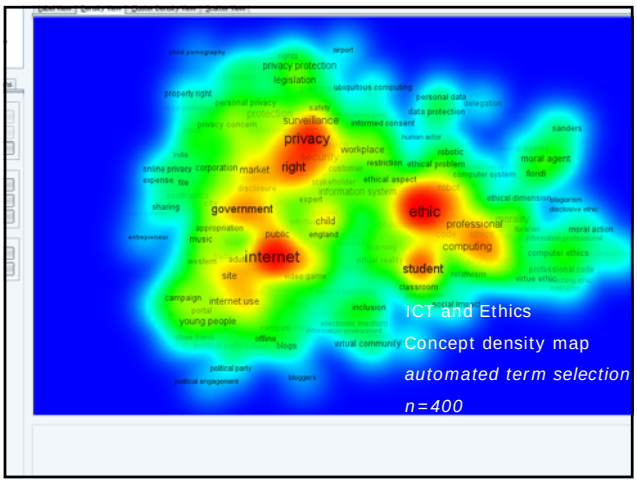


ICT and Ethics

Concept map

automated term selection, n=400

higher cluster threshold

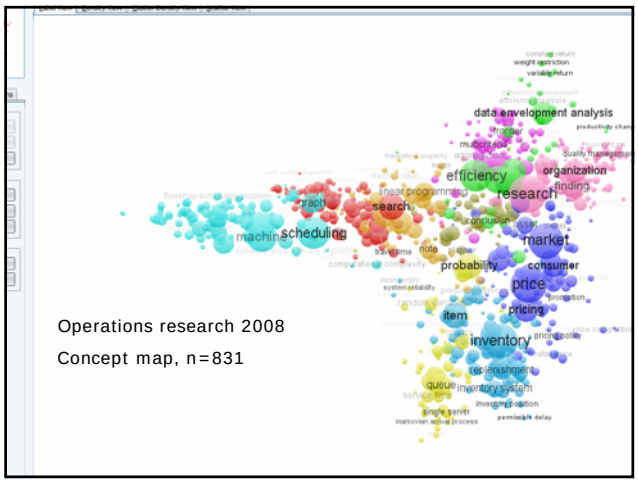


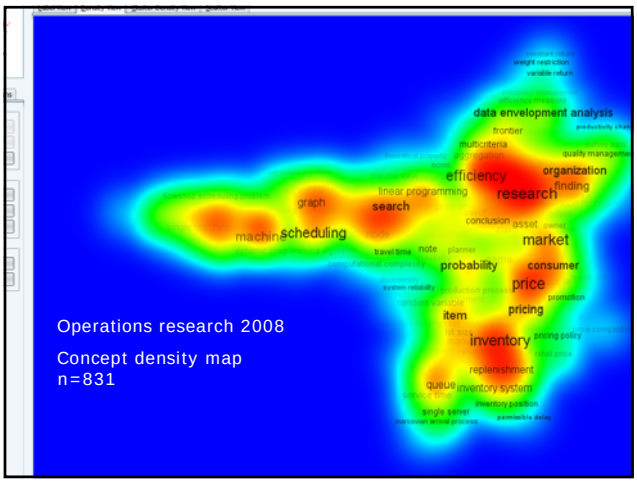
ICT and Ethics

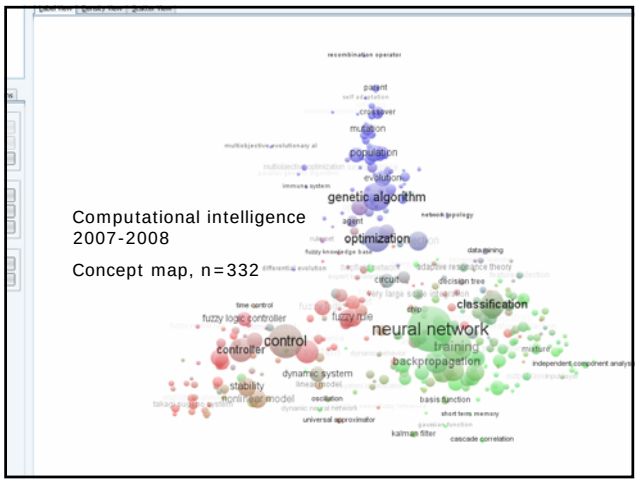
Concept density map

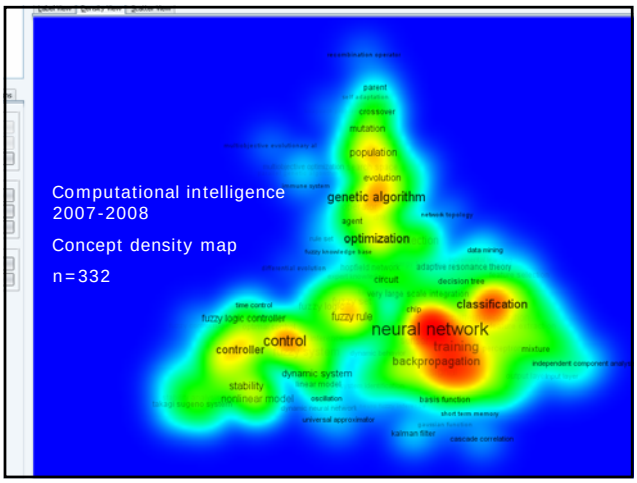
automated term selection

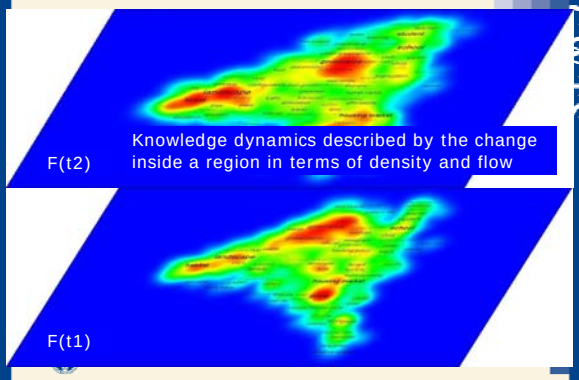
n=400


data envelopment analysis
efficiency
organization
research
search
graph
scheduling
market
probability
consumer
price
pricing
item
inventory
queue
Operations research 2008
Concept map, n=831


data envelopment analysis
efficiency
organization
research
search
graph
scheduling
market
probability
consumer
price
pricing
item
inventory
queue
Operations research 2008
Concept density map
n=831


genetic algorithm
optimization
classification
neural network
control
fuzzy logic controller
backpropagation
dynamic system
Computational intelligence
2007-2008
Concept map, n=332


genetic algorithm
optimization
classification
neural network
control
fuzzy logic controller
backpropagation
dynamic system
Computational intelligence
2007-2008
Concept density map
n=332



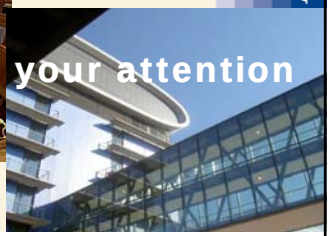
Conclusion

Advanced bibliometric analysis is a powerful method to

(1) Assess the international influence (impact, collaboration, profile) of scientific work in a reliable, transparent and objective way, particularly in the natural science and medical fields, and in several of the engineering and social science fields.

(2) Discover patterns in the structure of science which enables us to identify interdisciplinarity, knowledge flows and knowledge diffusion, and maps of research around socio-economic themes.

Thank you for your attention



more information: www.cwts.leidenuniv.nl