CWTS
Meaningful metrics

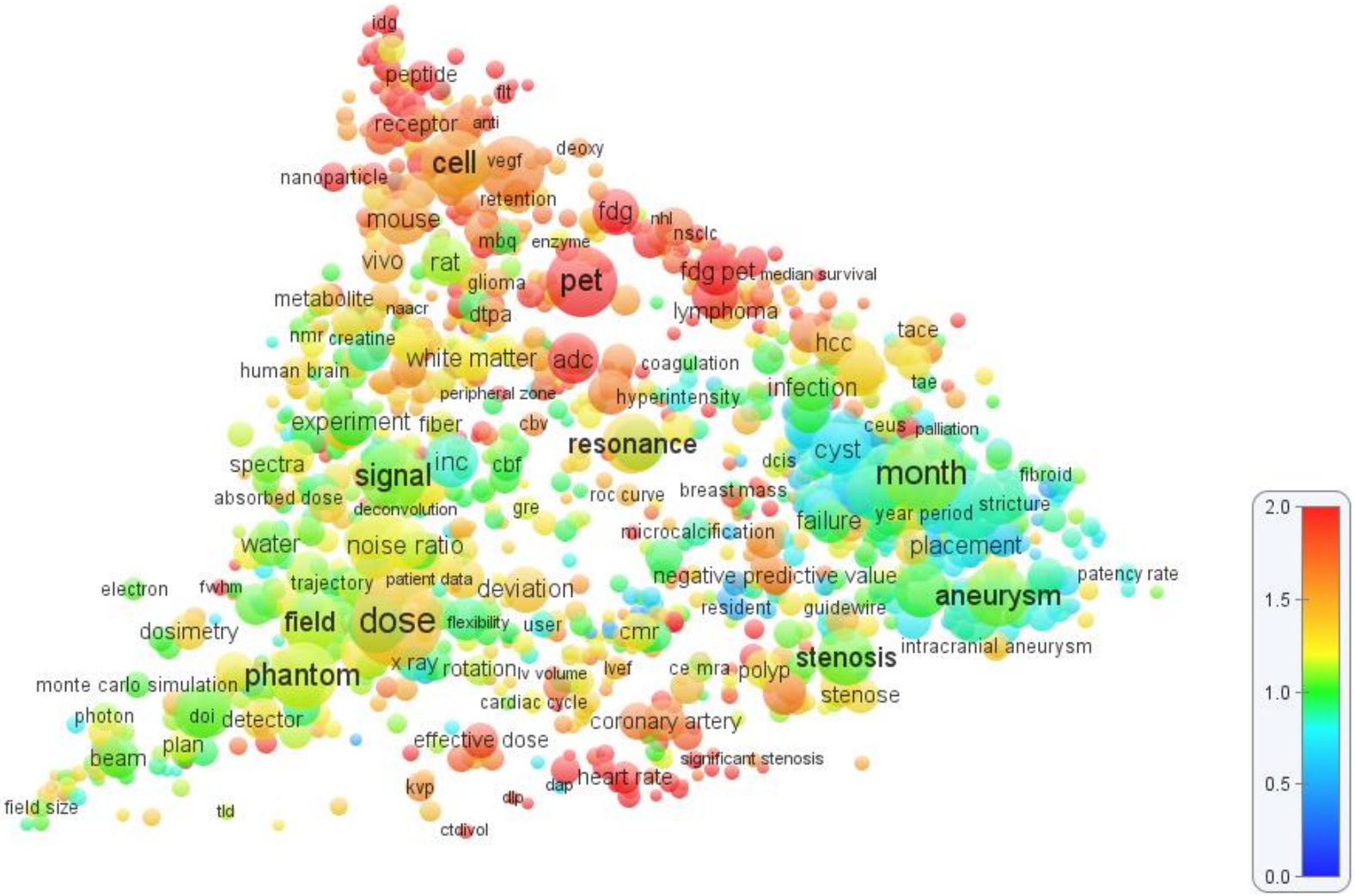
# Principles for research evaluation: altmetrics & scientometrics in context

Paul Wouters (with slides by James Wilsdon, SPRU)

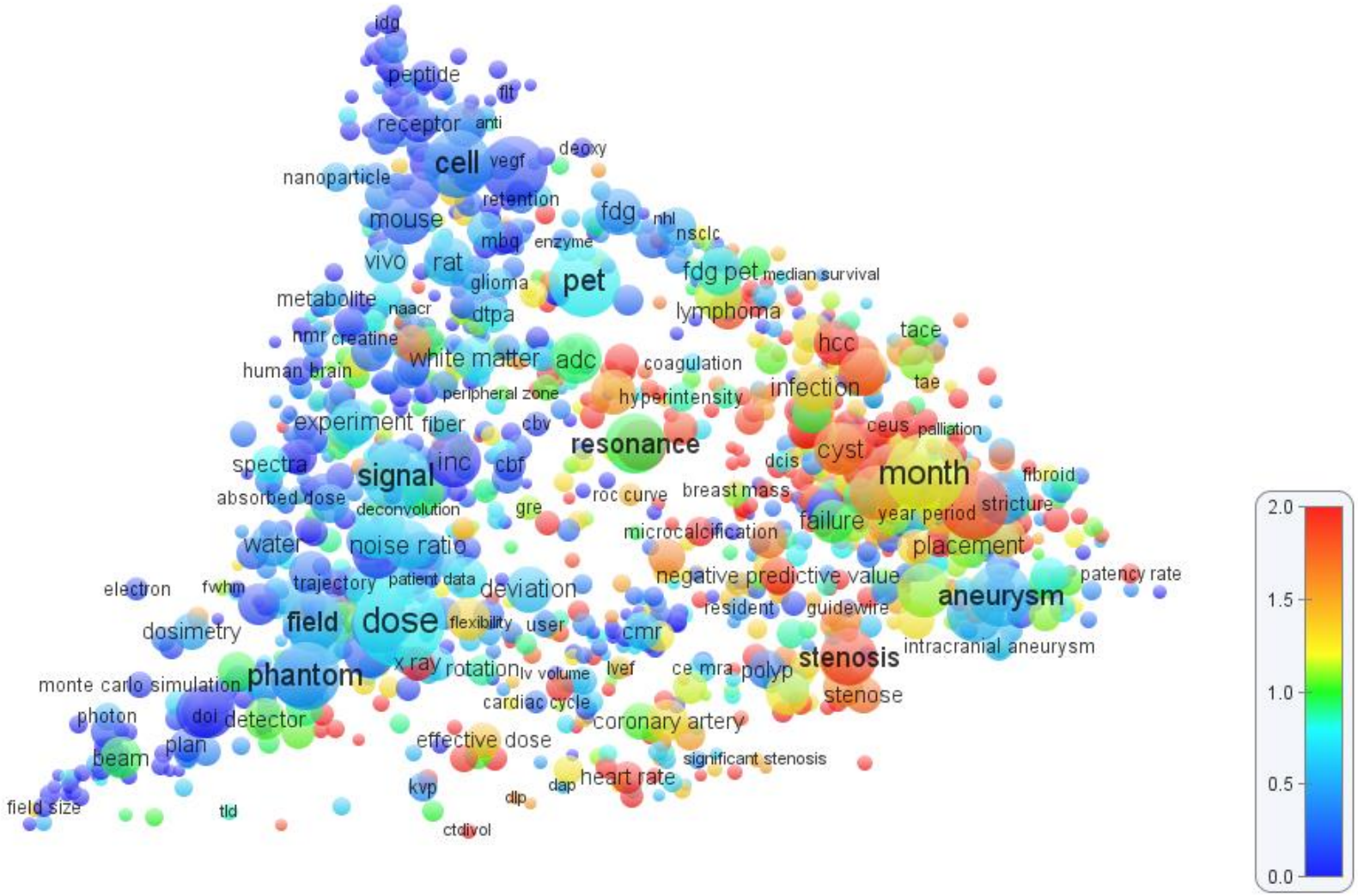
STM Innovations Seminar
London, 2 December 2015

Universiteit Leiden

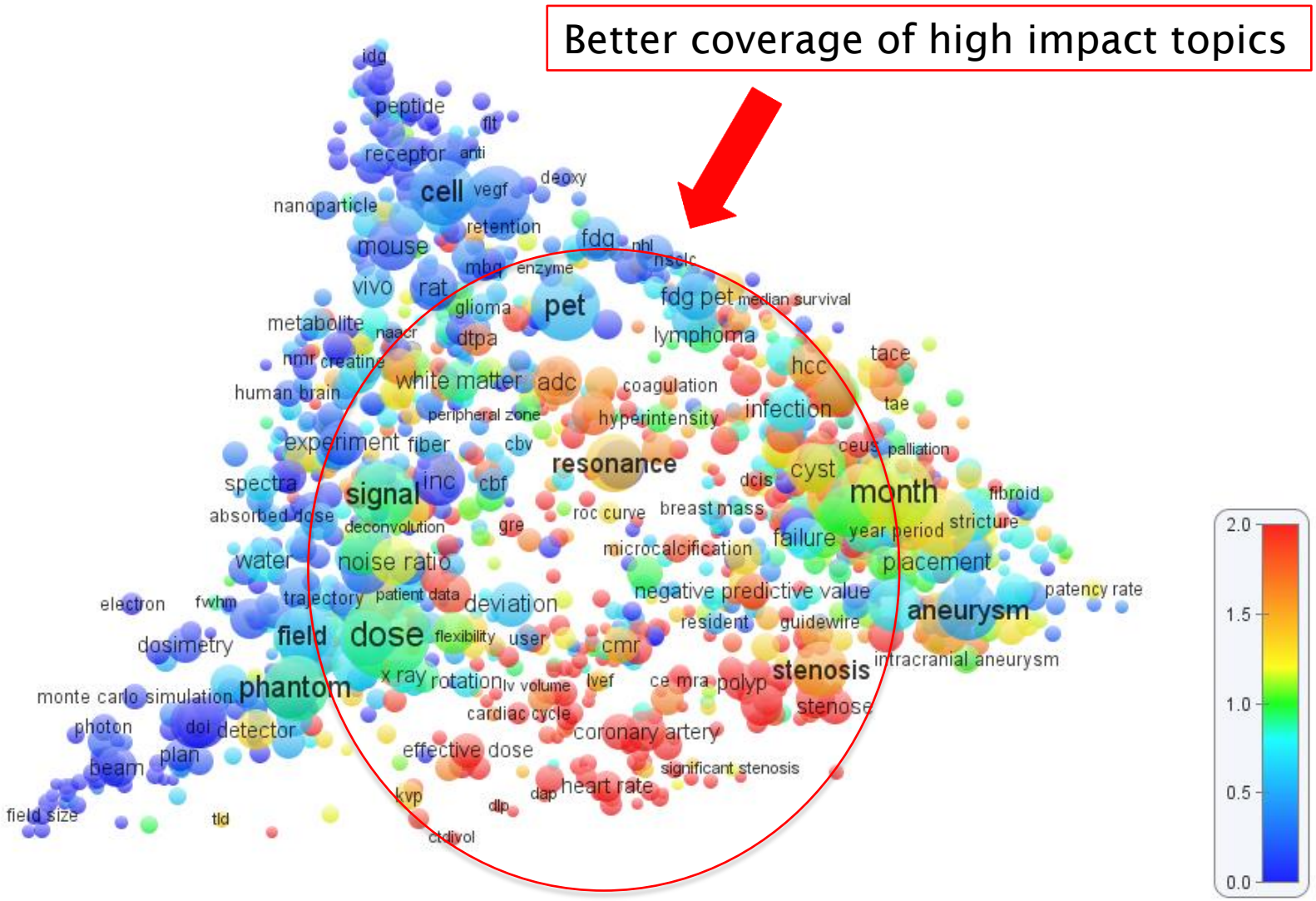
# Where is high impact in the field of Radiology?



# European Journal of Radiology (Elsevier)

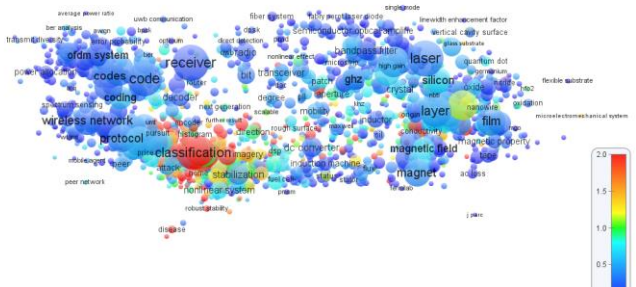


# European Radiology (Springer)



# How can we improve our position in Life Sciences

Identify: topics, authors, citing authors, etc.



# Trends in assessment

- Increased bibliometric services at university level available through databases
- Increased self-assessment via “gratis bibliometrics” on the web (h-index; publish or perish; etc.)
- Emergence of altmetrics
- Increased demand for bibliometrics at the level of the individual researcher
- Societal impact measurements required
- Career advice – where to publish?

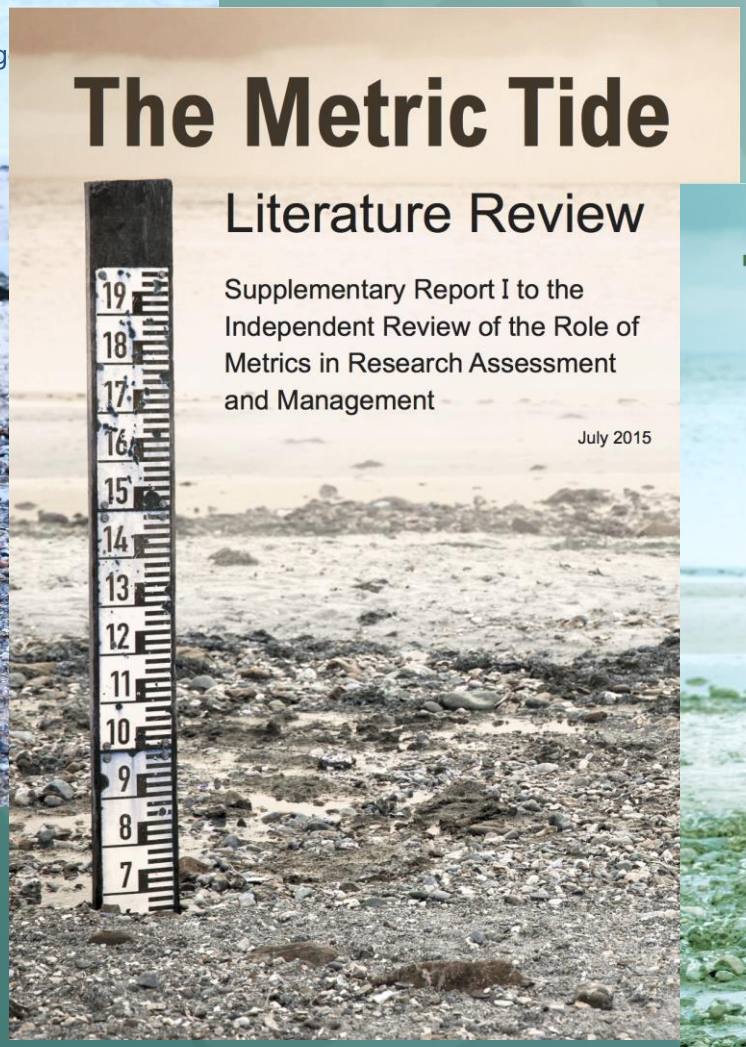
CWTS

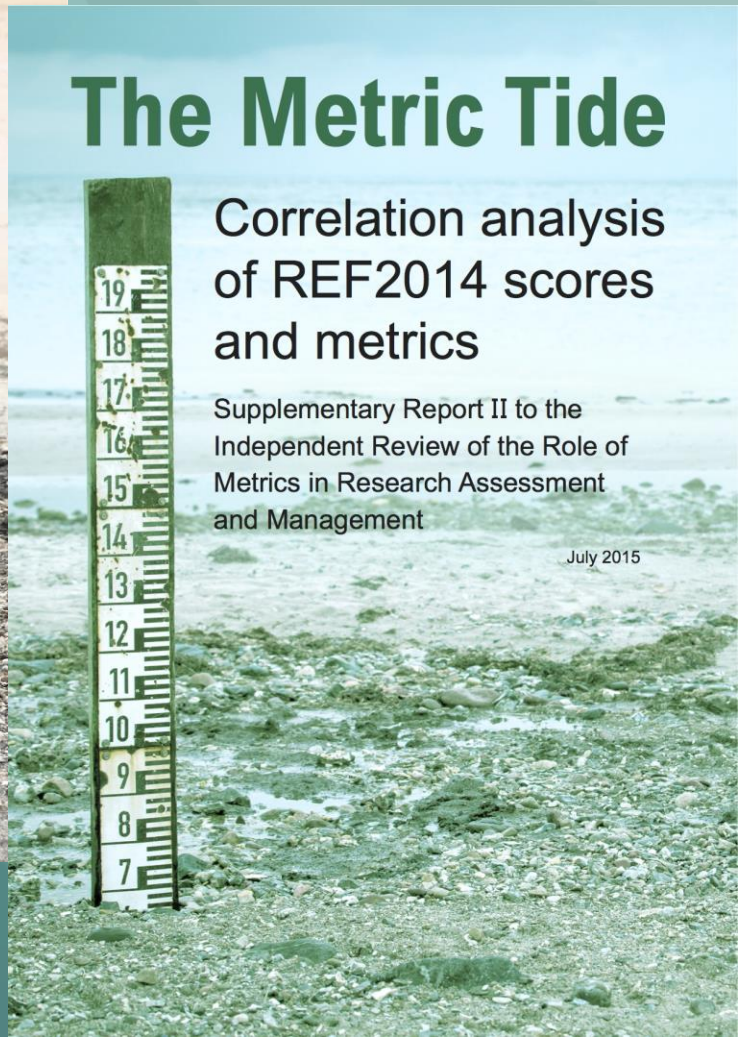
# Research leaders need more, not less, strategic intelligence

- Increasing demand for information about research:
  - hyper competition for funding
  - globalization
  - industry – academic partnerships
  - interdisciplinary research challenges
  - institutional demands on research & university management
- Increased supply of data about research:
  - web based research
  - deluge of data producing machines and sensors
  - increased social scale of research: international teams
  - large scale databases of publications, data, and applications

CWTS

http://www.hefce.ac.uk/rsrch/metrics/



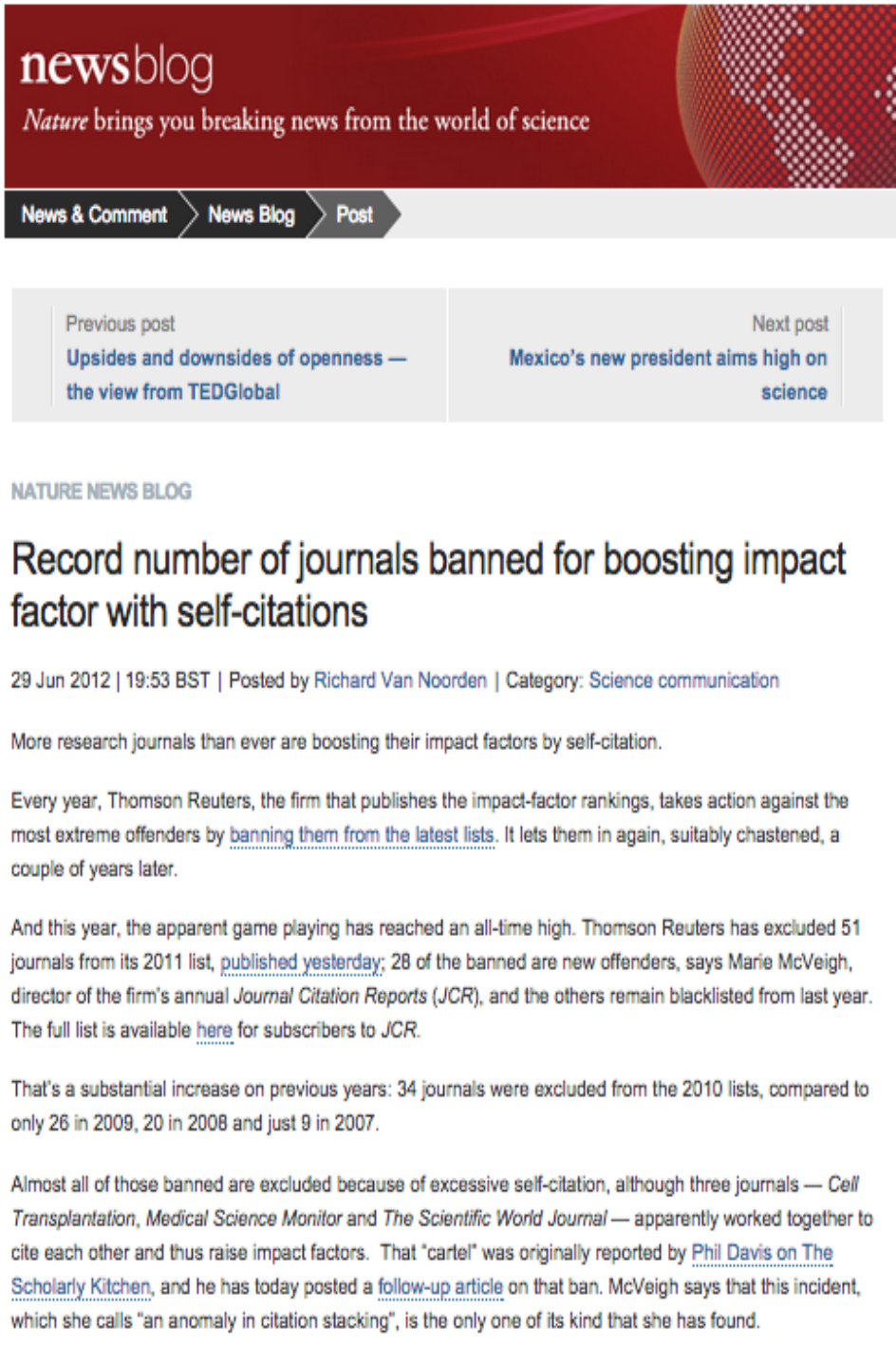

Universiteit Leiden

http://www.responsiblemetrics.org

Peer review, despite its flaws and limitations, continues to command widespread support across disciplines. Metrics should support, not supplant expert judgement.

Inappropriate indicators create perverse incentives. There is legitimate concern that some quantitative indicators can be gamed, or can lead to unintended consequences.

CWTS



newsblog

*Nature* brings you breaking news from the world of science

News & Comment | News Blog | Post

Previous post
Upsides and downsides of openness — the view from TEDGlobal

Next post
Mexico's new president aims high on science

NATURE NEWS BLOG

## Record number of journals banned for boosting impact factor with self-citations

29 Jun 2012 | 19:53 BST | Posted by Richard Van Noorden | Category: Science communication

More research journals than ever are boosting their impact factors by self-citation.

Every year, Thomson Reuters, the firm that publishes the impact-factor rankings, takes action against the most extreme offenders by banning them from the latest lists. It lets them in again, suitably chastened, a couple of years later.

And this year, the apparent game playing has reached an all-time high. Thomson Reuters has excluded 51 journals from its 2011 list, published yesterday; 28 of the banned are new offenders, says Marie McVeigh, director of the firm's annual *Journal Citation Reports* (*JCR*), and the others remain blacklisted from last year. The full list is available here for subscribers to *JCR*.

That's a substantial increase on previous years: 34 journals were excluded from the 2010 lists, compared to only 26 in 2009, 20 in 2008 and just 9 in 2007.

Almost all of those banned are excluded because of excessive self-citation, although three journals — *Cell Transplantation*, *Medical Science Monitor* and *The Scientific World Journal* — apparently worked together to cite each other and thus raise impact factors. That "cartel" was originally reported by Phil Davis on The Scholarly Kitchen, and he has today posted a follow-up article on that ban. McVeigh says that this incident, which she calls "an anomaly in citation stacking", is the only one of its kind that she has found.

Indicators can only meet their potential if they are underpinned by an open and interoperable data infrastructure.

CWTS

# Key challenges in research information system building

- Will the information infrastructure contain high quality data and indicators?
- Will it enable and support context- and mission-sensitive research assessments?
- Will it enable application of research information for primary research purposes (eg in VREs)?
- Will the public sector remain master in its own house or will it hand over control to the private sector?
- Will it be possible to truly open up the research agenda to all stakeholders – open science in a democratic society?

CWTS

# Responsible metrics

Responsible metrics can be understood in terms of:

- **Robustness**: basing metrics on the best possible data in terms of accuracy and scope;
- **Humility**: recognizing that quantitative evaluation should support – but not supplant – qualitative, expert assessment;
- **Transparency**: keeping data collection and analytical processes open and transparent, so that those being evaluated can test and verify the results;
- **Diversity**: accounting for variation by field, using a variety of indicators to reflect and support a plurality of research & researcher career paths;
- **Reflexivity**: recognizing the potential & systemic effects of indicators and updating them in response.

CWTS

Across the research community, the description, production and consumption of ‘metrics’ remains contested and open to misunderstandings.

CWTS

COMMENT

**SUSTAINABILITY** Data needed to drive UN development goals **p.432**

**CONSERVATION** Economics and environmental catastrophe **p.434**

**GEOLOGY** Questions raised over proposed Anthropocene dates **p.436**

**HISTORY** Music inspired Newton to add more colours to the rainbow **p.436**

# The Leiden Manifesto for research metrics

Use these ten principles to guide research evaluation, urge **Diana Hicks**, **Paul Wouters** and colleagues.

Data are increasingly used to govern science. Research evaluations that were once bespoke and performed by peers are now routine and reliant on metrics[1]. The problem is that evaluation is now led by the data rather than by judgement. Metrics have proliferated: usually well intentioned, not always well informed, often ill applied. We risk damaging the system with the very tools designed to improve it, as evaluation is increasingly implemented by organizations without knowledge of, or advice on, good practice and interpretation.

Before 2000, there was the Science Citation Index on CD-ROM from the Institute for Scientific Information (ISI), used by experts for specialist analyses. In 2002, Thomson Reuters launched an integrated web platform, making the Web of Science database widely accessible. Competing citation indices were created: Elsevier’s Scopus (released in 2004) and Google Scholar (beta version released in 2004). Web-based tools to easily compare institutional research productivity and impact were introduced, such as InCites (using the Web of Science) and SciVal (using Scopus), as well as software to analyse individual citation profiles using Google Scholar (Publish or Perish, released in 2007).

In 2005, Jorge Hirsch, a physicist at the University of California, San Diego, proposed the *h*-index, popularizing citation counting for individual researchers. Interest in the journal impact factor grew steadily after 1995 (see ‘Impact-factor obsession’).

Lately, metrics related to social usage ▸

ILLUSTRATION BY DAVID PARKINS

# The Leiden Manifesto

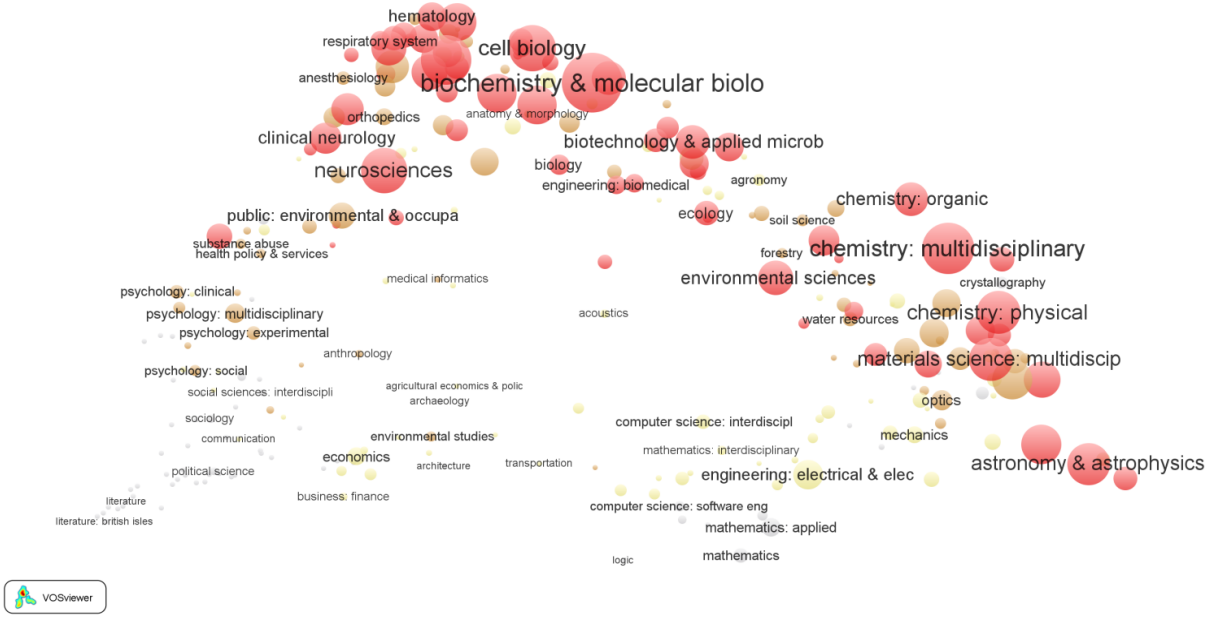
- Quantitative evaluation should support expert assessment.
- Measure performance in accordance with the research mission.
- Protect excellence in locally relevant research
- Keep data collection and analytical processes open, transparent and simple.
- Allow for data verification
- Account for variation by field in publication and citation practices
- Data should be interpreted taking into account the difficulty of credit assignment in the case of multi-authored publications.
- Base assessment of individual researchers on *qualitative* judgment.
- False precision should be avoided (eg. the JIF).
- Systemic effects of the assessment and the indicators should be taken into account and indicators should be updated regularly

Diana Hicks (Georgia Tech), Paul Wouters (CWTS), Ismael Rafols (SPRU/Ingenio), Sarah de Rijcke and Ludo Waltman (CWTS)

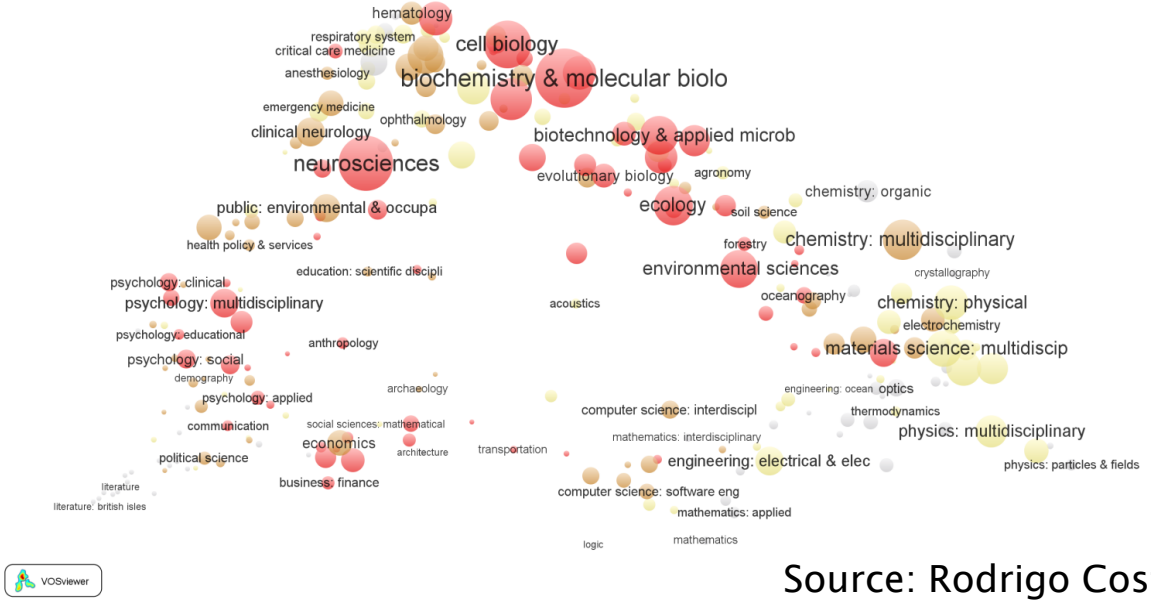
CWTS

# Coverage by fields

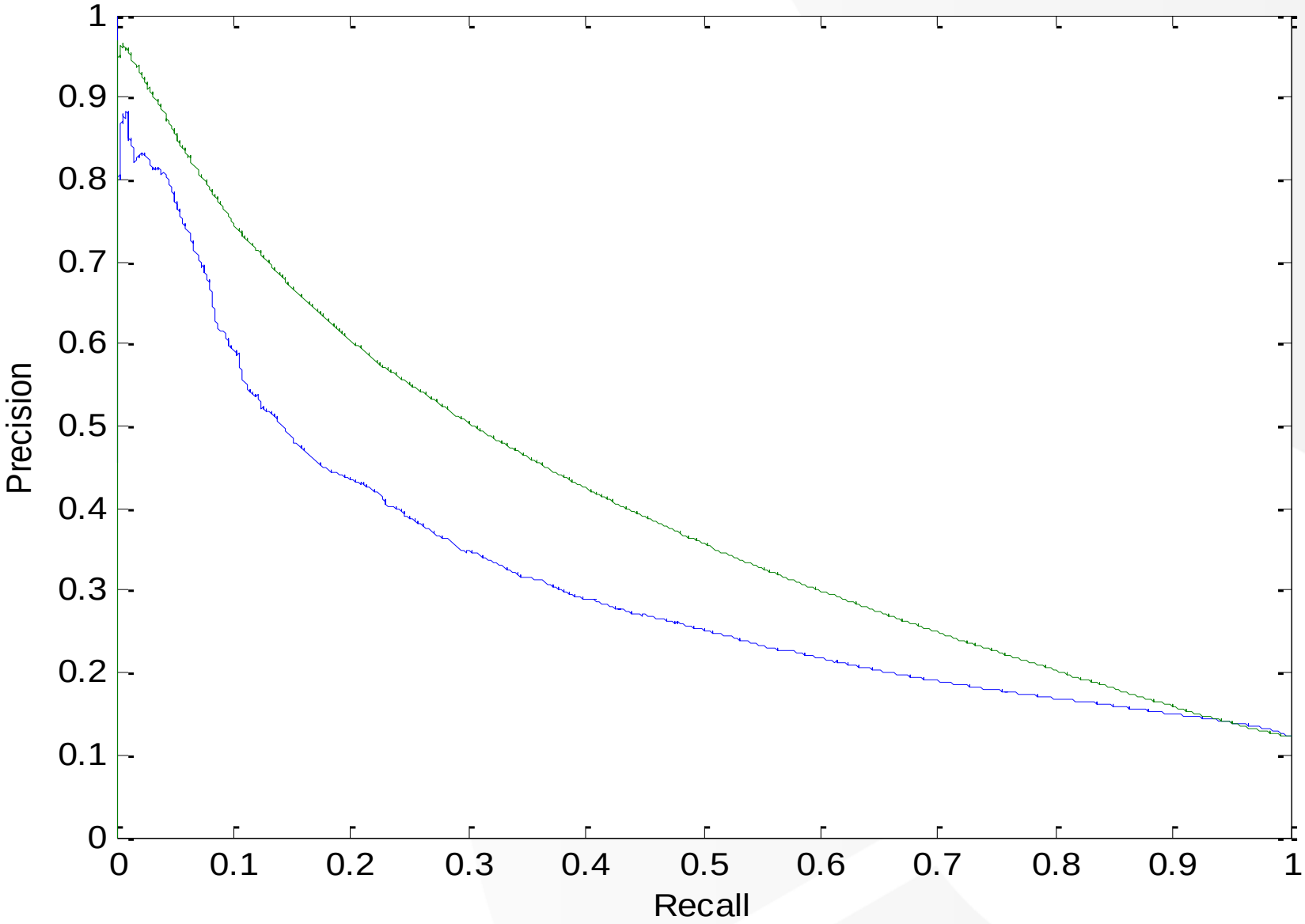
Citations




Mendeley




Source: Rodrigo Costas

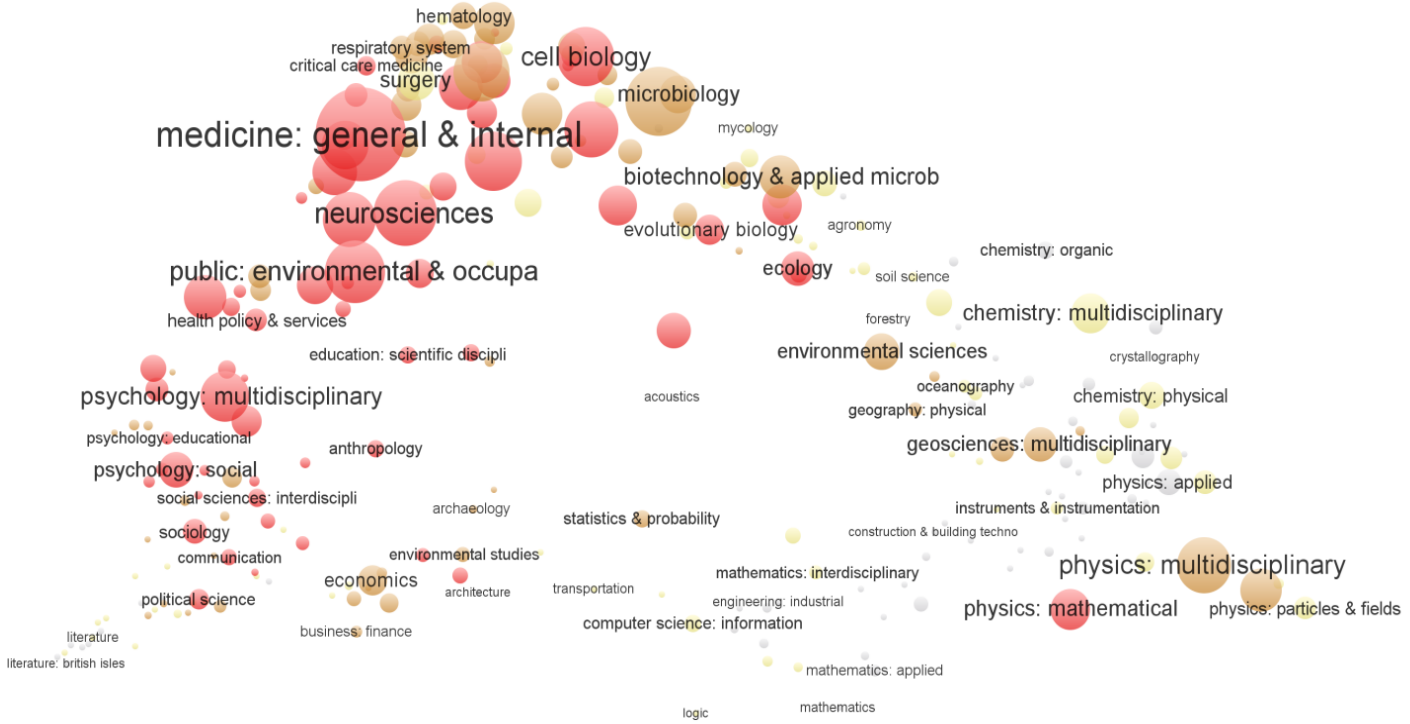



**General Precision-recall curves for JCS (blue line) and total readerships (green line) for identifying PPtop10% most highly cited publications**

Source: Zohreh Zahedi

# Coverage by fields

- Other relevant patterns:
  - Twitter: **stronger in Social Sciences and General medicine, weaker in Natural Sciences and Humanities**



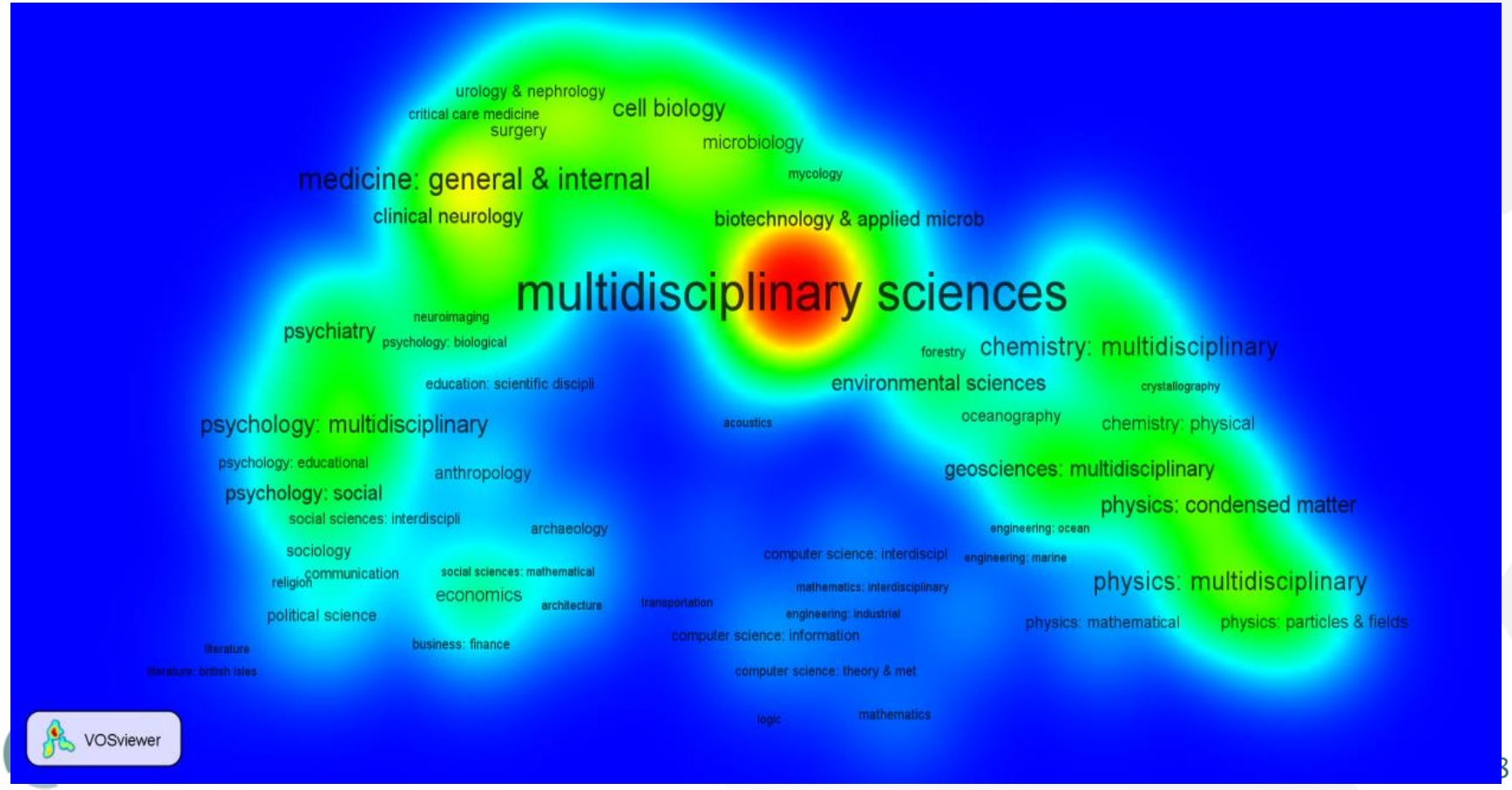

VOSviewer



Source: Rodrigo Costas

# Coverage by fields

- Blogs and news media have a strong **focus on multidisciplinary journals!**




Source: Rodrigo Costas

# Context Counts!

# cwts.nl