

The Economist
 OCTOBER 19TH - 25TH 2013
 Economist.com

Britain's angry white men
 How to do a nuclear deal with Iran
 Investment tips from Nobel economists
 Junk bonds are back
 The meaning of Sachin Tendulkar

HOW SCIENCE GOES WRONG.
 Einsteinium

OCTOBER 19TH - 25TH 2013
 Worldwide con

ESA European Sociological Association

Call to Action
 Sign the petition addressed to European Commission for more support of Social Sciences and Humanities in Europe
 Click here to sign the petition

- 1
- 2
- 3
- 4
- 5
- 6

the guardian

News Sport Comment Culture Business Money Life & Style Travel Environment
 Others Jobs

News Science People in science

Nobel winner declares boycott of top science journals
 Randy Schickman says his lab will no longer send papers for Nature, Cell and Science as they distort scientific process

By Sam Kean, science correspondent
 The Guardian, Monday 5 December 2011 11:42 GMT

Share 144
 Tweet 327
 Like 45
 Email 23

Print History

COMMENT

The Leiden Manifesto for research metrics

Download the full report at www.leidenmanifesto.nl

D

San Francisco
DORA
 Declaration on Research Assessment

The Metric Tide

Report of the Independent Review of the Role of Metrics in Research Assessment and Management

July 2015

Uses of indicators: Pressing demands of research management and evaluation --- Can indicators help?

Yes, indicators can help make decisions...

- Reduce time and costs
- Increase transparency and sense of objectivity
- Reduce complexity, accessible to managers

but do they lead to the “right” decisions?

Evaluation gap (Wouters):

“discrepancy between evaluation criteria [implicit in indicators] and the social and economic functions of science”

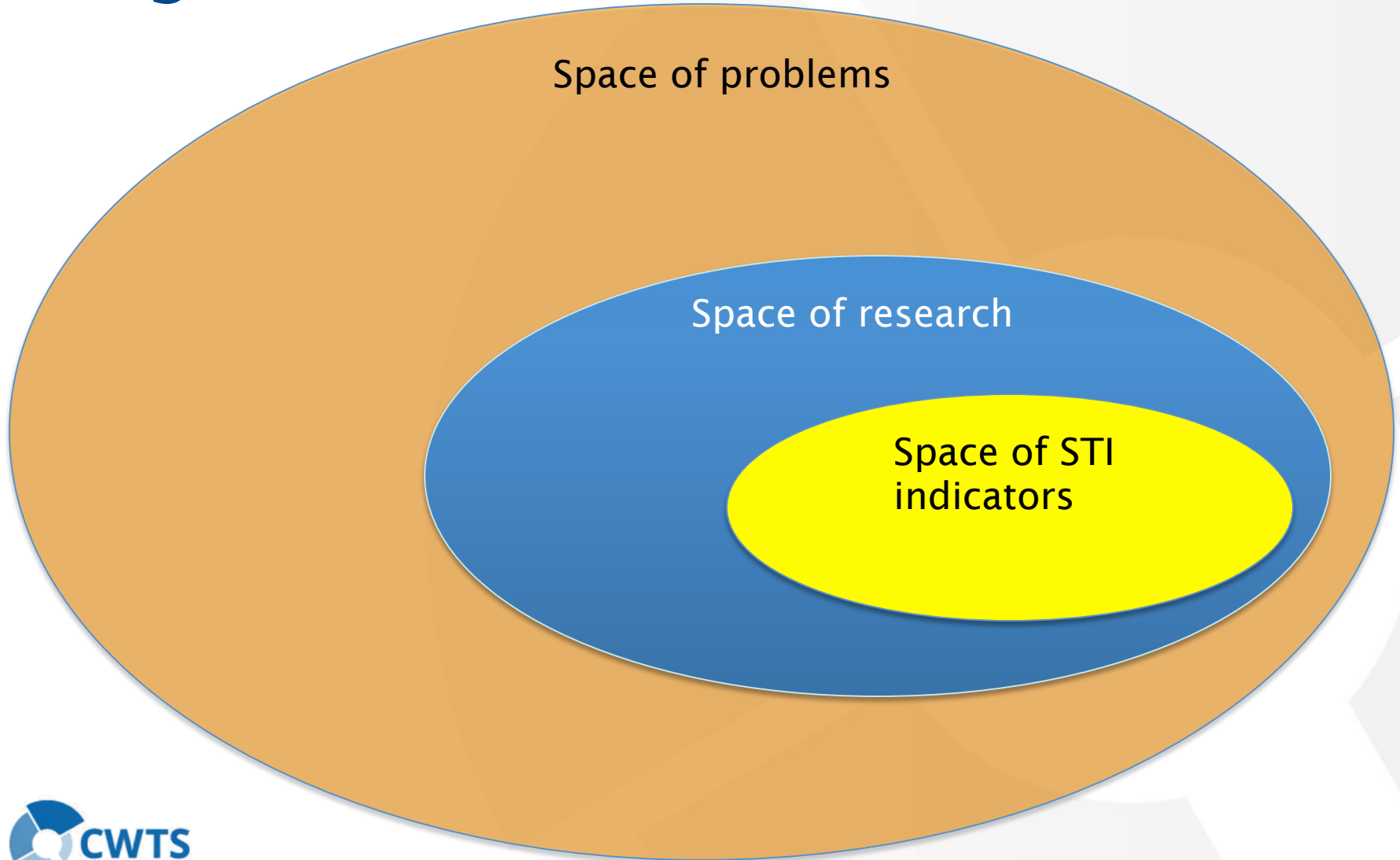
***Academia – “excellence” *Innovation – economic “growth”**

Missions *not well covered*: agriculture, public health, defence, development, social inclusion,...

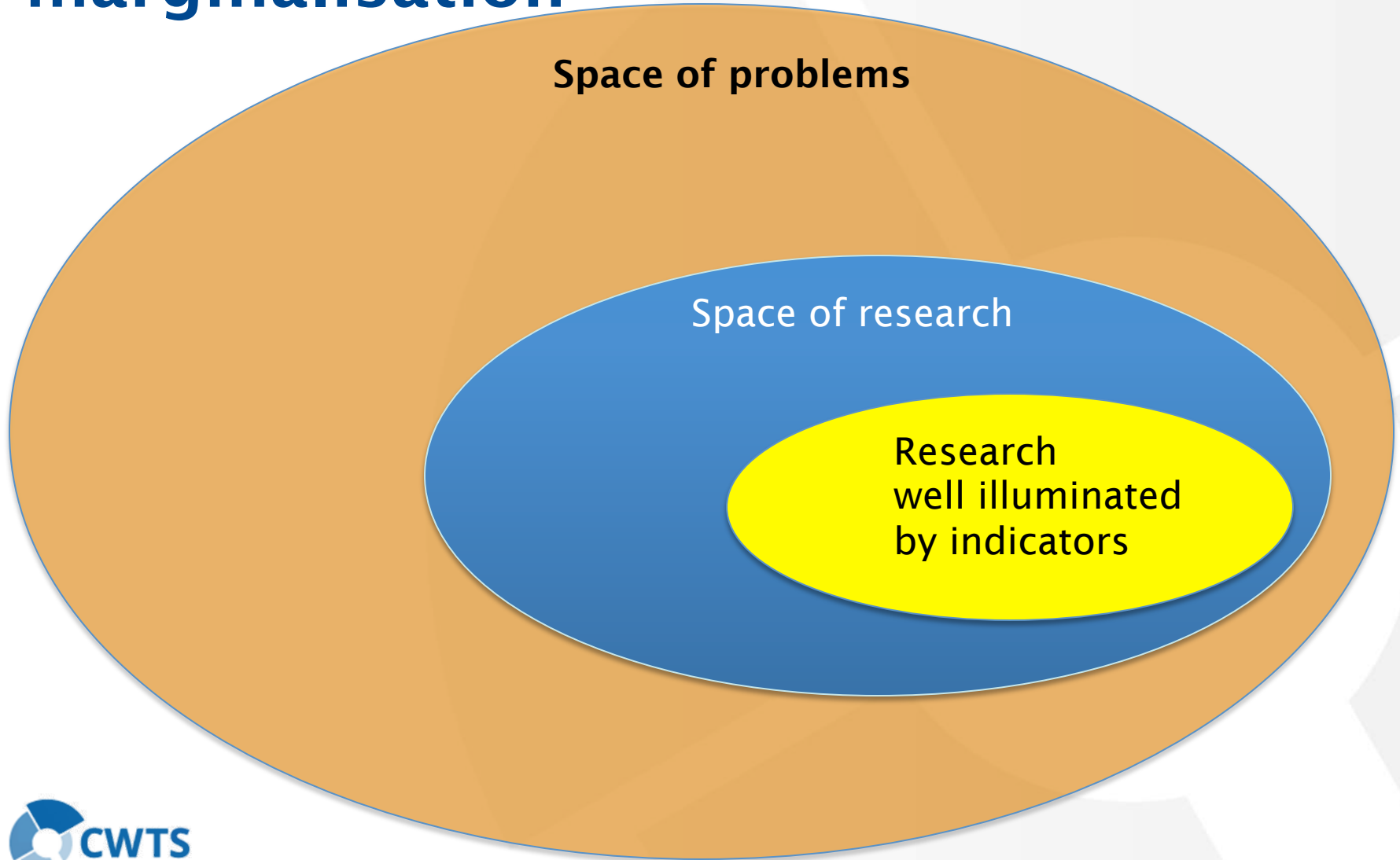
Often related to marginalised / “neglected” populations?



Problems, research, indicators and marginalisation



Problems, research, indicators and marginalisation



Problems, research, indicators and marginalisation

Space of problems



STI Peripheries:
research spaces not
well captured
by indicators

Research
well illuminated
by indicators

Problems, research, indicators and marginalisation

Multiple types of space:

Geographical: regional, “South”

Cognitive: SSH,
engineering

Linguistic:
non-English

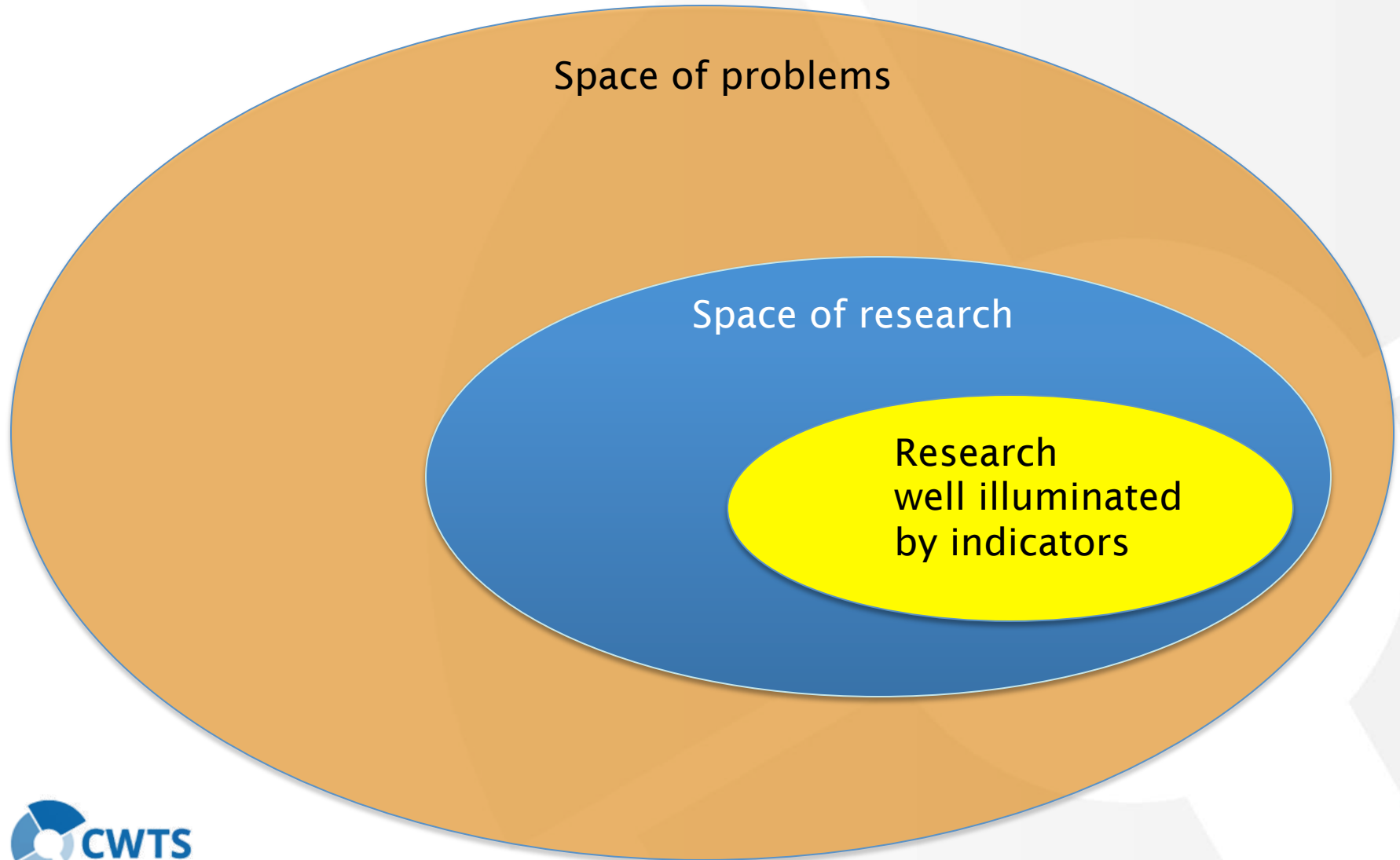
Sectoral: low-tech,
agriculture, creative ind.

Social: gender,
minorities

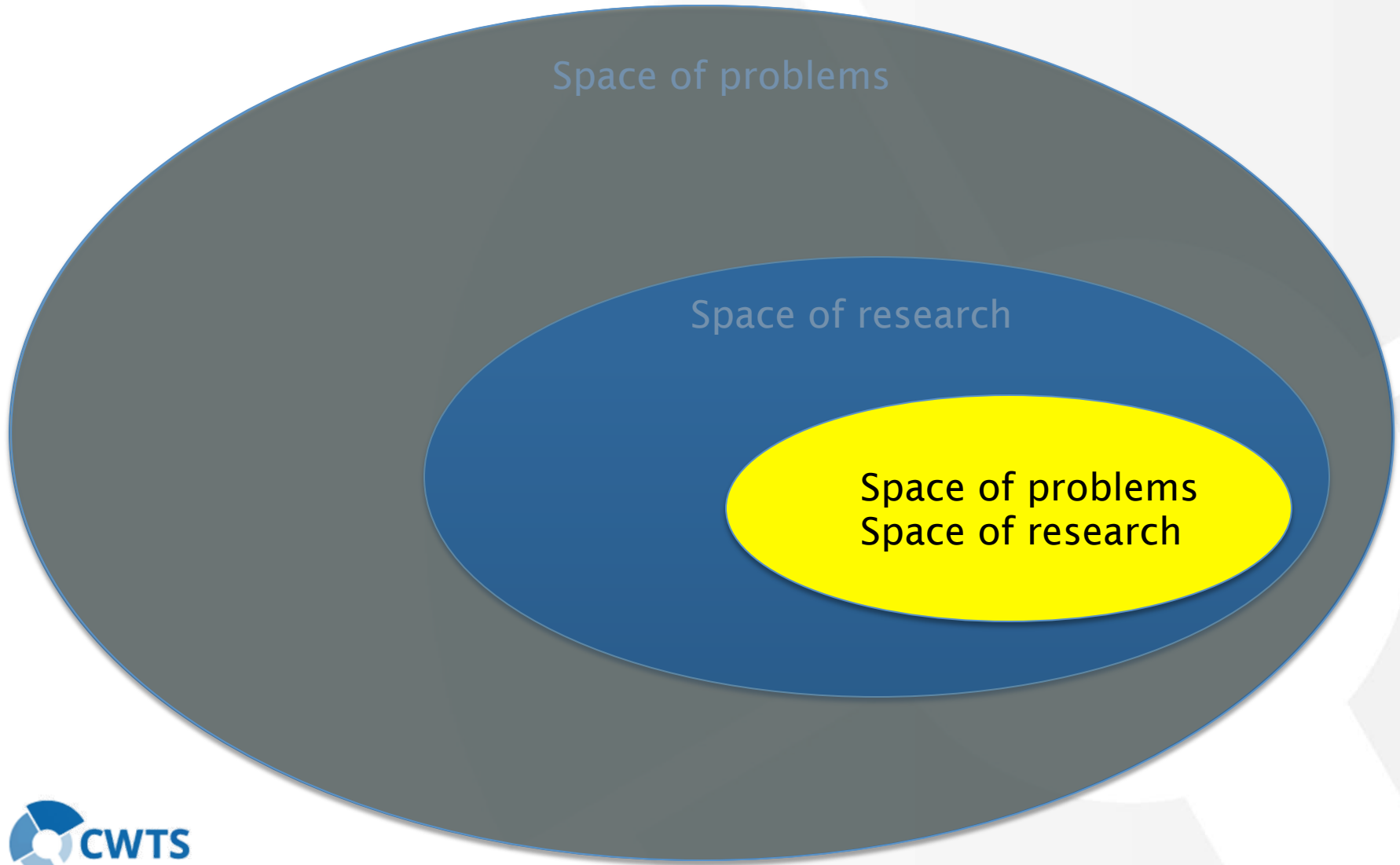
STI Peripheries:
research spaces not
well captured
by indicators

Research
well illuminated
by indicators

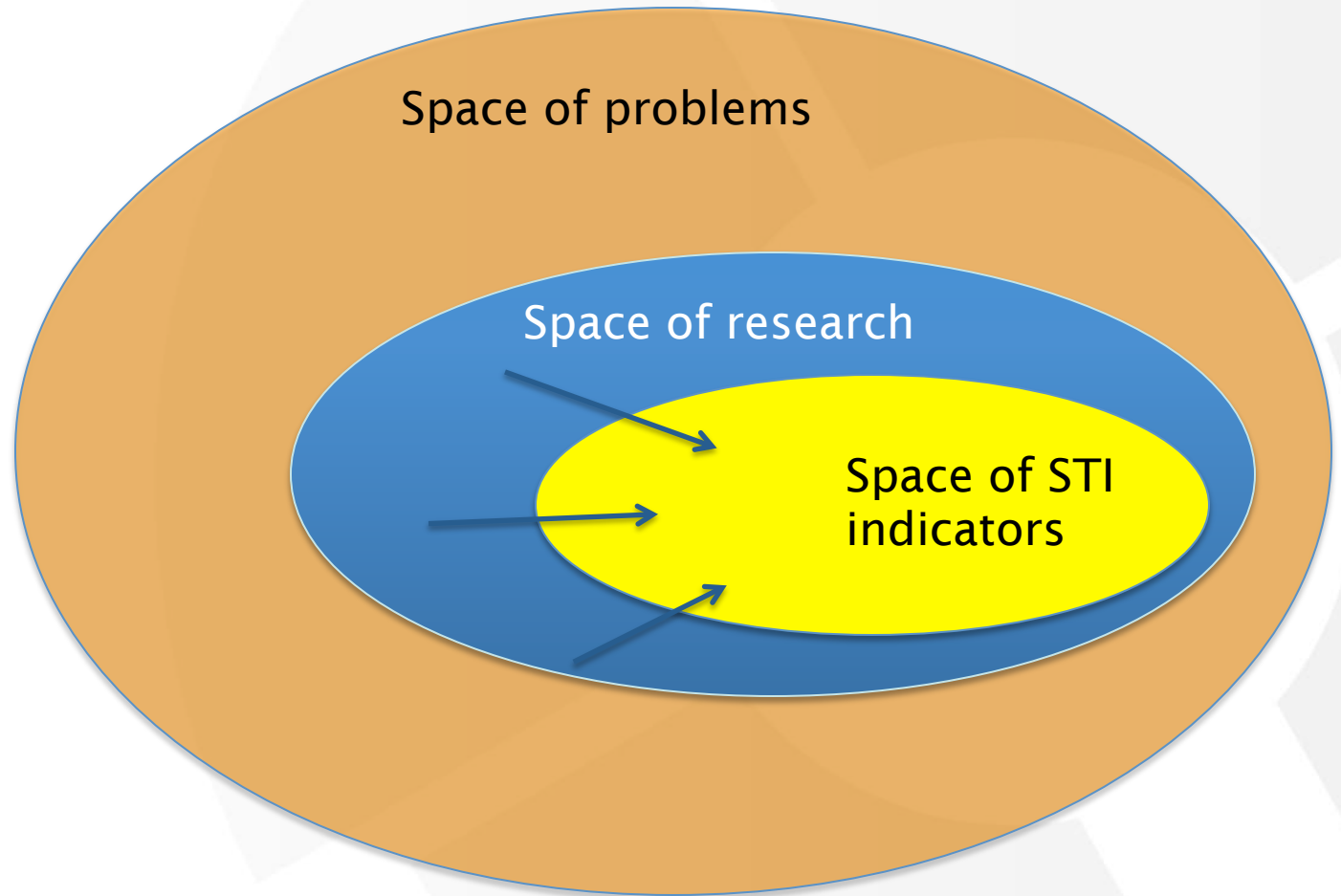
Streetlight effect in indicators: mistaking light with “problems”



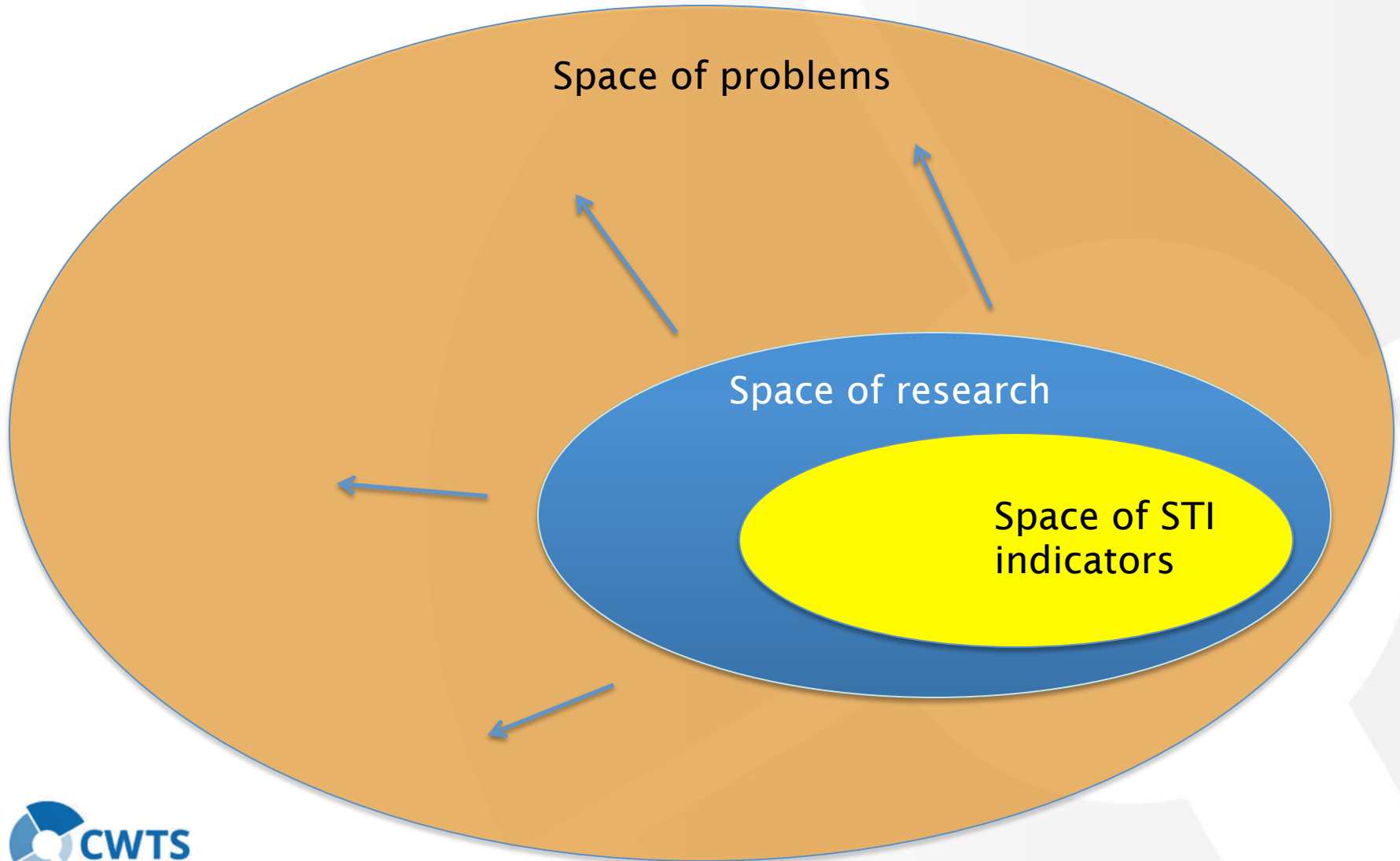
Streetlight effect in indicators: mistaking light with “problems”



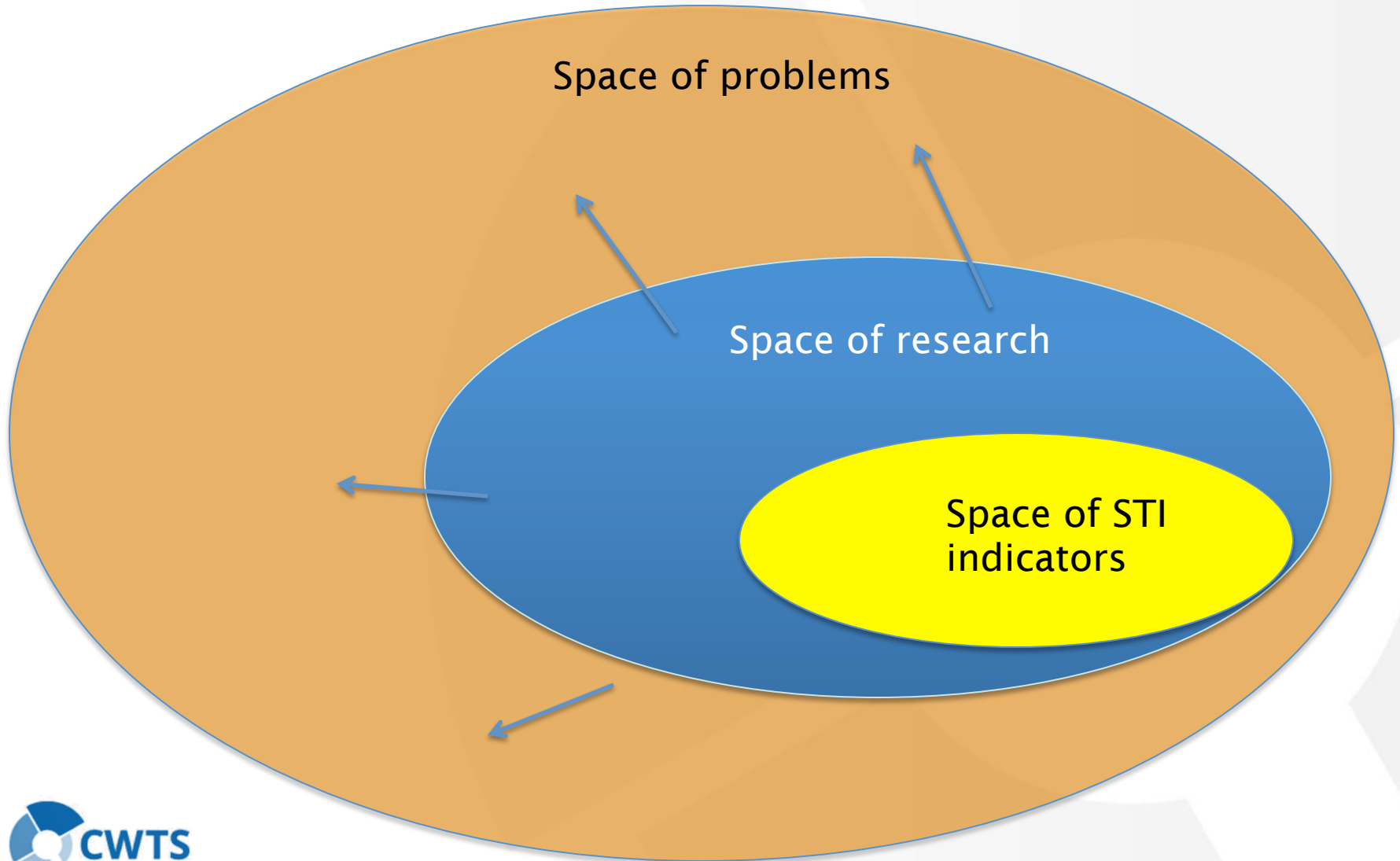
Hypothesis: reduced indicator coverage may contract research space



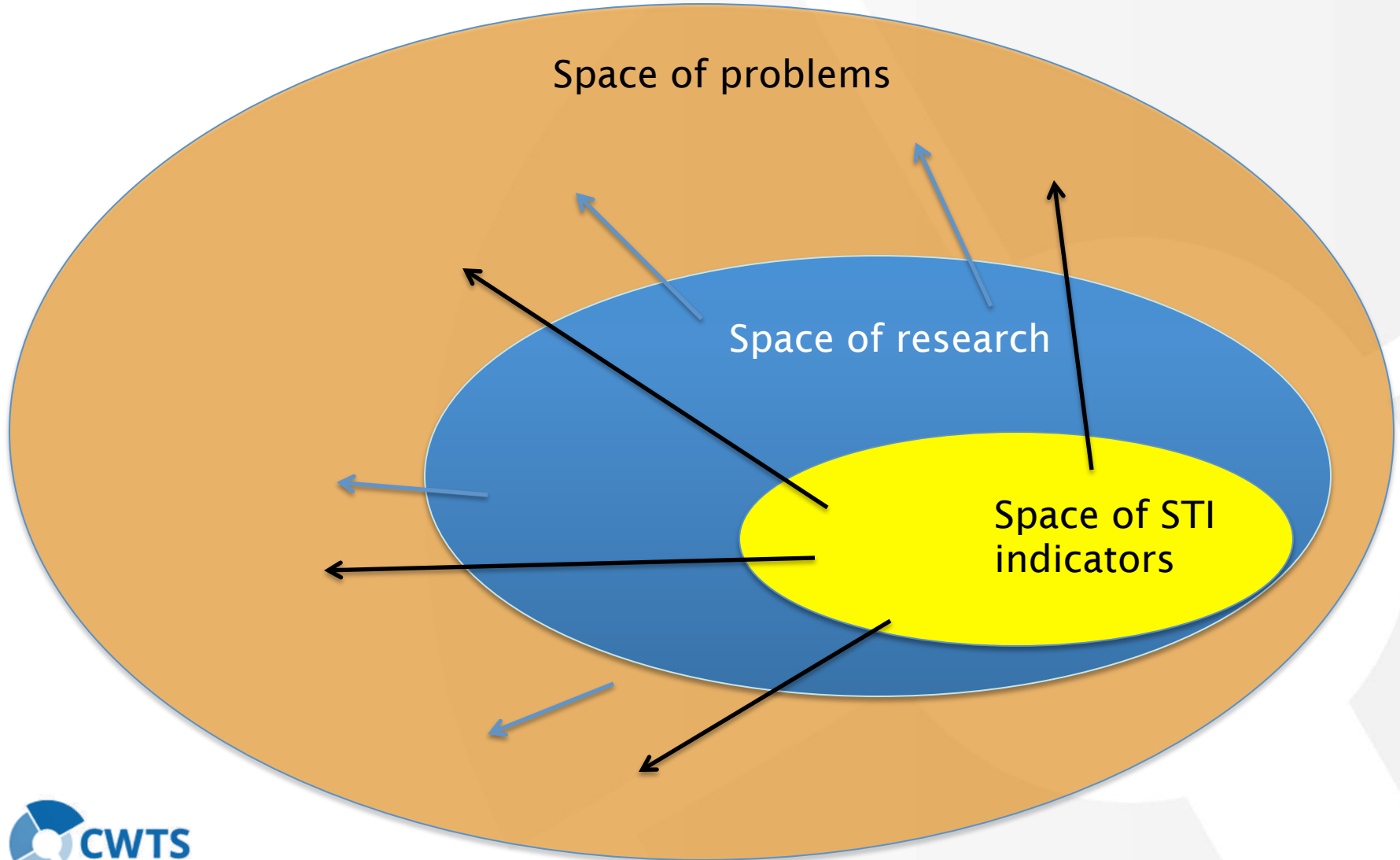
Demands for expanding role of science in society...




Demands for expanding role of science in society...



...may require an expanded set of indicators: MORE



Principles for responsible metrics

An abstract graphic composed of several overlapping blue shapes. On the left, there is a large blue circle. To its right, a smaller blue circle is partially overlapping it. Several thick blue lines radiate from the right side of the smaller circle, extending towards the right edge of the frame. The background is a light blue gradient.

COMMENT

SUSTAINABILITY Data needed to drive UN development goals p.432

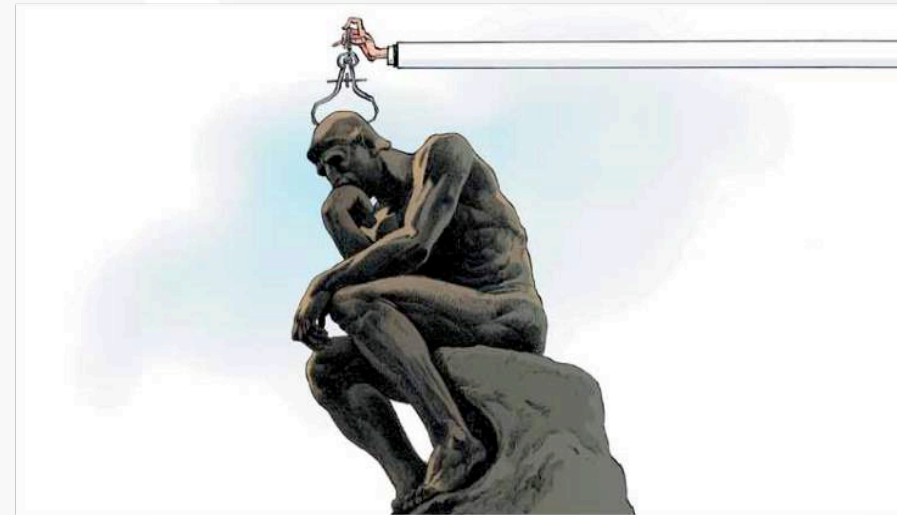


CONSERVATION Economics and environmental catastrophe p.434

GEOLOGY Questions raised over proposed Anthropocene dates p.436

HISTORY Music inspired Newton to add more colours to the rainbow p.438

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



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'. 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

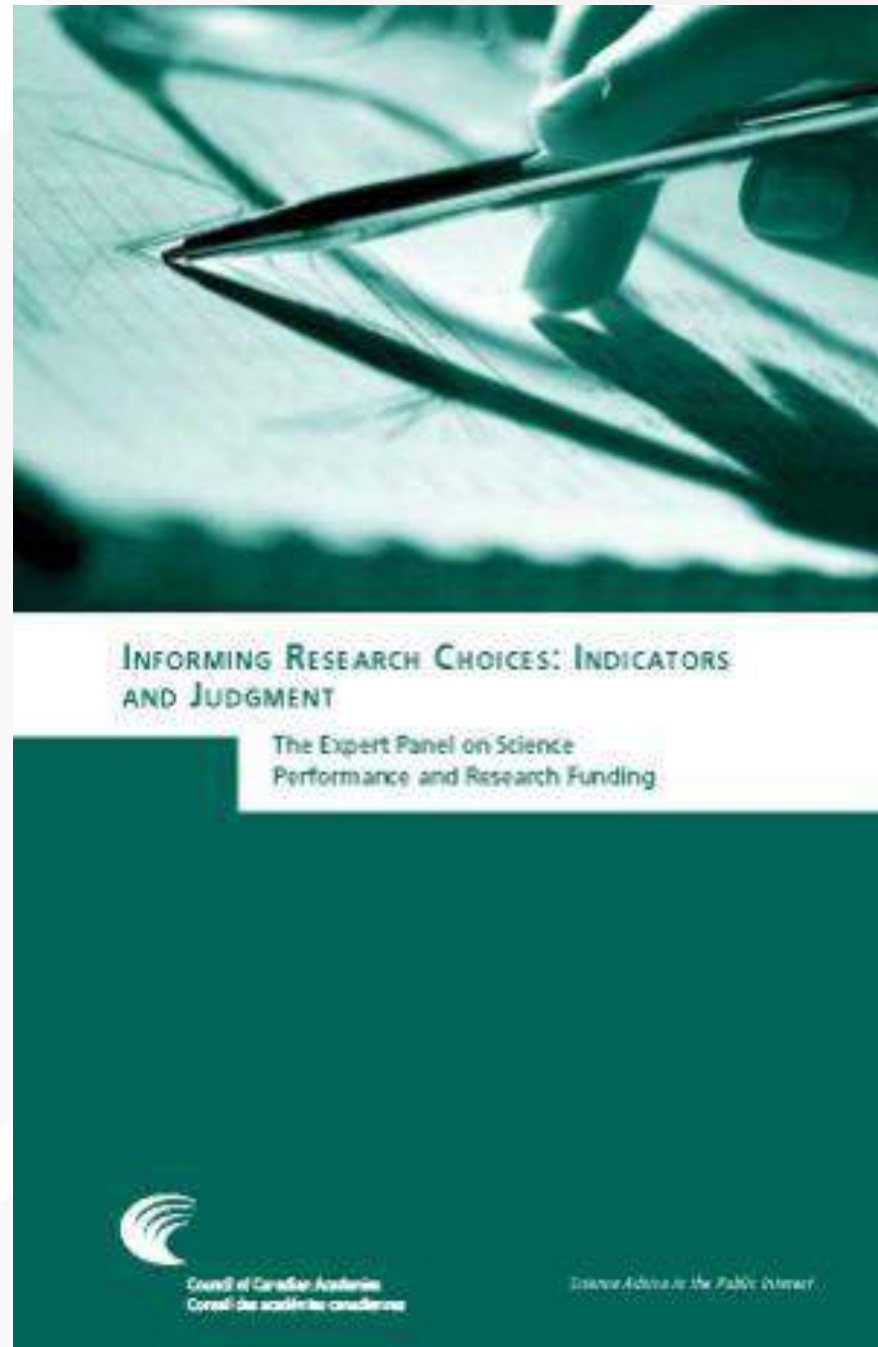
advice on, good practice and interpretation. Before 2000, there was the Science Citation Index on CD-ROM from the Institute for Scientific Information (ISI), used by experts for specialist analyses. In 2002, Thomson Reuters launched an integrated web platform, making the Web of Science database widely 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 ▶

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

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.



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.

The logo for ORCID, with 'ORCID' in a grey sans-serif font and 'iD' in a green sans-serif font.

Responsible metrics

- **Robustness**
- **Humility**
- **Transparency**
- **Diversity**
- **Reflexivity**



Best practice: a portfolio – Step 1. Broaden out

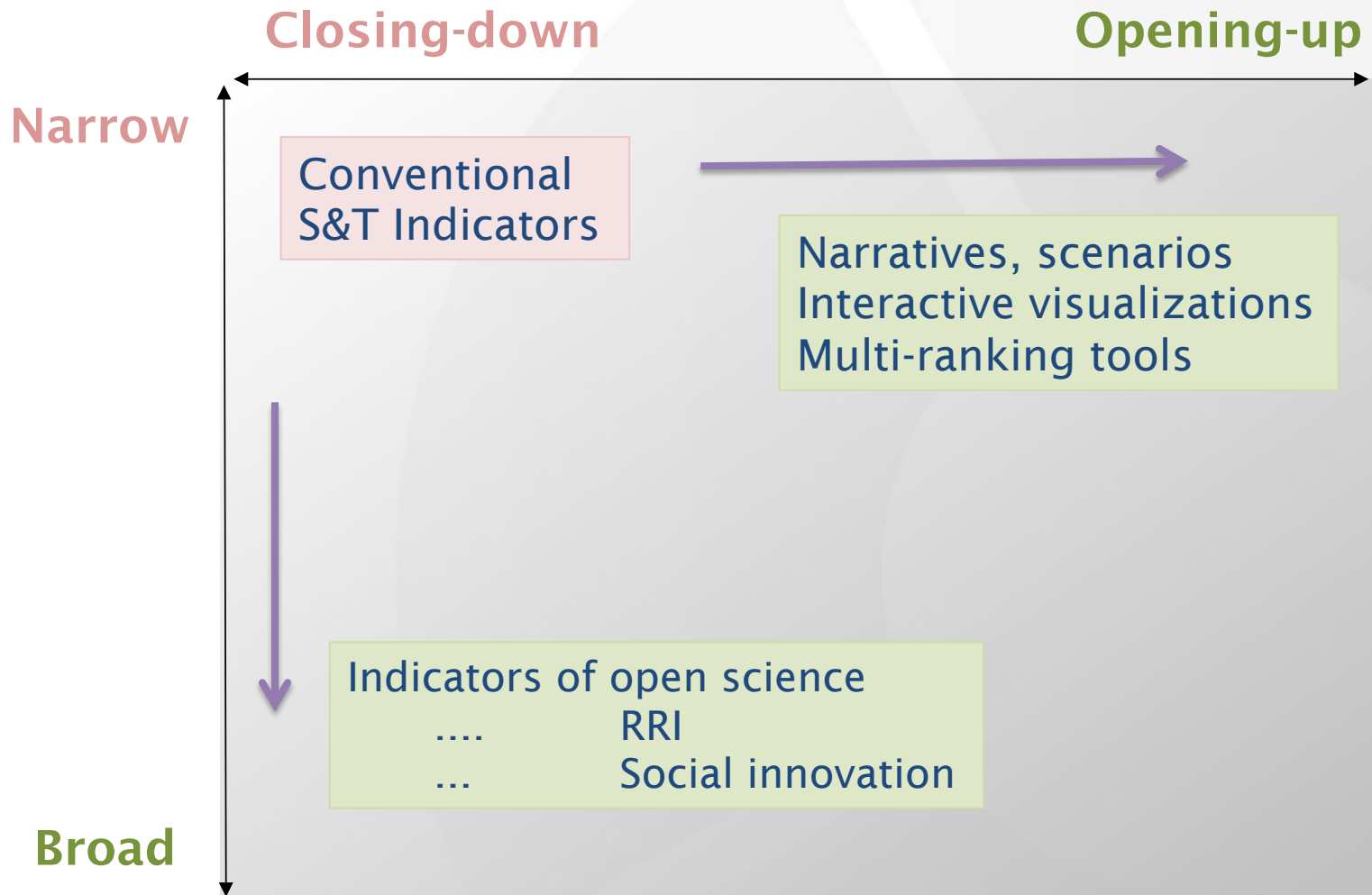
- A suite of indicators that include conventional outcome measurements, evaluations of academic citizenship across communities, and assessments of structure and process
- Move towards including advanced and next generation metrics (indicators for open science, RRI, social innovation)

Best practice: a portfolio

Step 2. Open up

- Focus on content and quality – with view on responsible research assessment
- Bibliometrics can be reductive
- BUT: some may lack vocabulary to
 - discuss less-quantifiable dimensions -> ‘bibliometric creep’
 - use metrics in interesting ways -> indicators as ‘tin openers’
- By formally requiring qualitative indicators and a descriptive portfolio, and new metrics where possible, we open up what can be discussed in an assessment

Opening up S&T Indicators



Back to the Netherlands

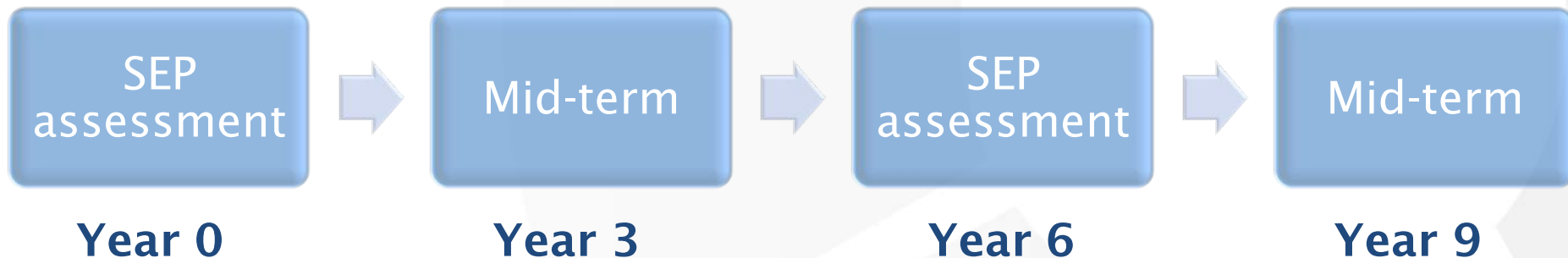
The background features a large, abstract graphic composed of thick blue lines and shapes. A prominent element is a large white circle on the left side, partially enclosed by a blue arc. To the right, a solid blue circle is connected to the main structure by several thick blue lines that radiate outwards, creating a sense of movement and connectivity. The overall aesthetic is clean, modern, and minimalist.

Purposes of evaluation practice (Molas-Gallart 2012)

3 main purposes of evaluation practice:

- distributive, improvement, controlling use

Steps in the evaluation process



Content self-assessment report (i)

- Description of unit's organisational structure + financing
- Strategy past 6 years
- Targets past 6 years (research, societal relevance)
- Strategy and targets next 5-10 years
- Most important (and relevant) performance indicators
- Results research and societal relevance past 6 years (latter in a narrative)
+ link results to SEP criteria (quality, relevance, viability)

Content self-assessment report (ii)

- Relevant environmental factors/developments past six years
- Forecast of trends and developments in the coming years
- SWOT analysis and benchmarking
- PhD Programs
- Research Integrity

Our consultancy:

- Key steps:
 - *information gathering process for “self-evaluation,” part of the Dutch Standard Evaluation Protocol*
 - *Inventory of how central goals and core values are operationalized, resulting in outputs for academic and societal networks*
 - *Make explicit academic and societal quality and relevance, by collecting and processing information from researchers as well as societal stakeholders*
 - *Inventory of impact pathways can serve as a framework to efficiently outline policy for conducting research in the next years*

Four phases

1. Exploratory phase

- Articulation questions and issues

2. Data gathering

- *Document analysis*
- *Quantitative methods (e.g. Scientometrics, Contextual Respons Analysis, ABC, proximities)*
- *Interviews*

3. Workshop

4. Data analysis and reporting

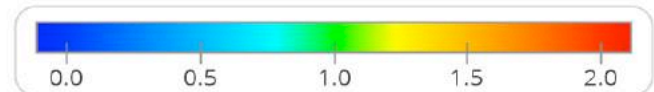
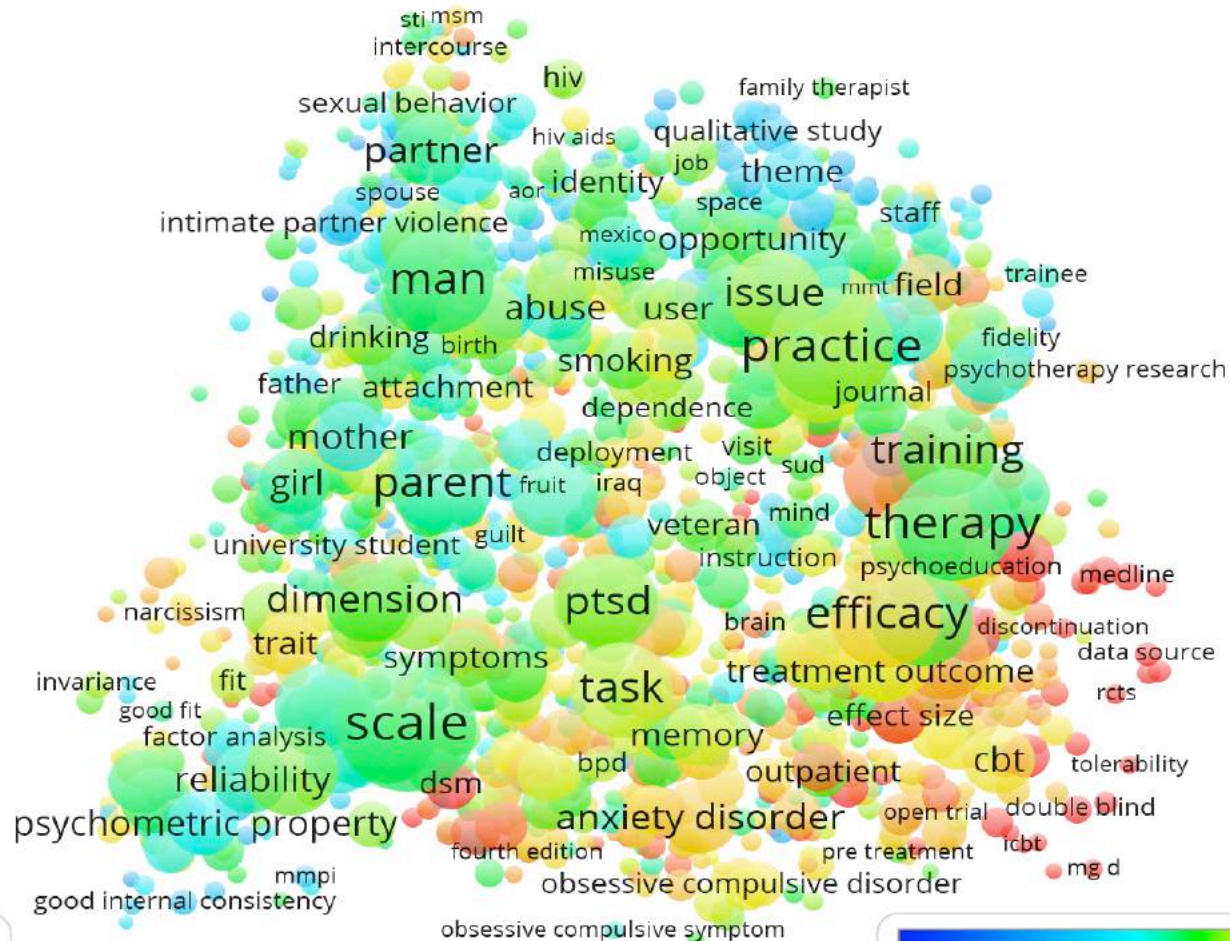


Quantitative methods

An abstract graphic composed of several overlapping blue shapes. On the left, there is a large blue arc. To its right is a large blue circle. Several thick blue lines radiate from the center of the circle towards the top and bottom edges of the frame, creating a stylized, geometric design.

- Descriptive scientometrics – cartography of field national – international perspective and break-down to institutional level
- Performance analyses at institutional level
- Research profile analysis
- Collaboration profile analysis
- Contextualisation of citation-impact (Knowledge user analysis)

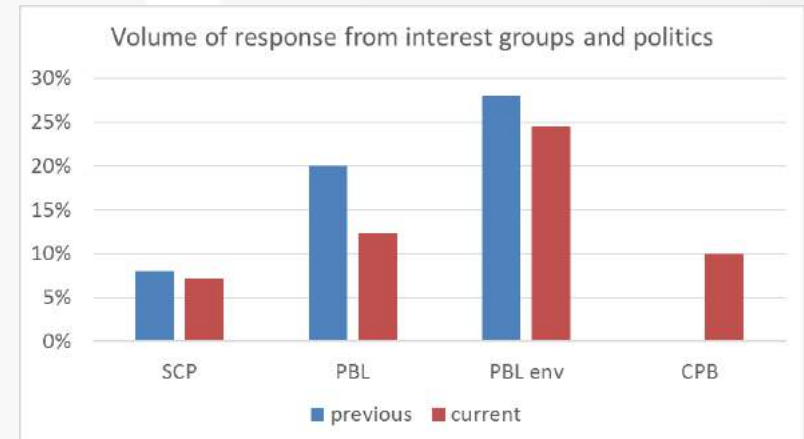
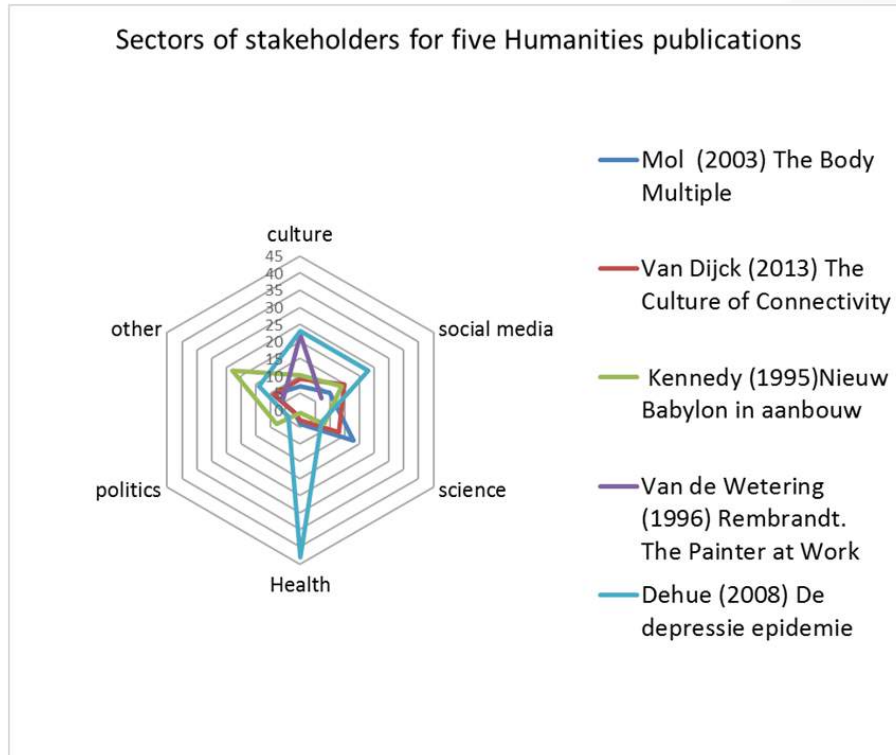
2: international literature psychology



Context Response Analysis (CRA)

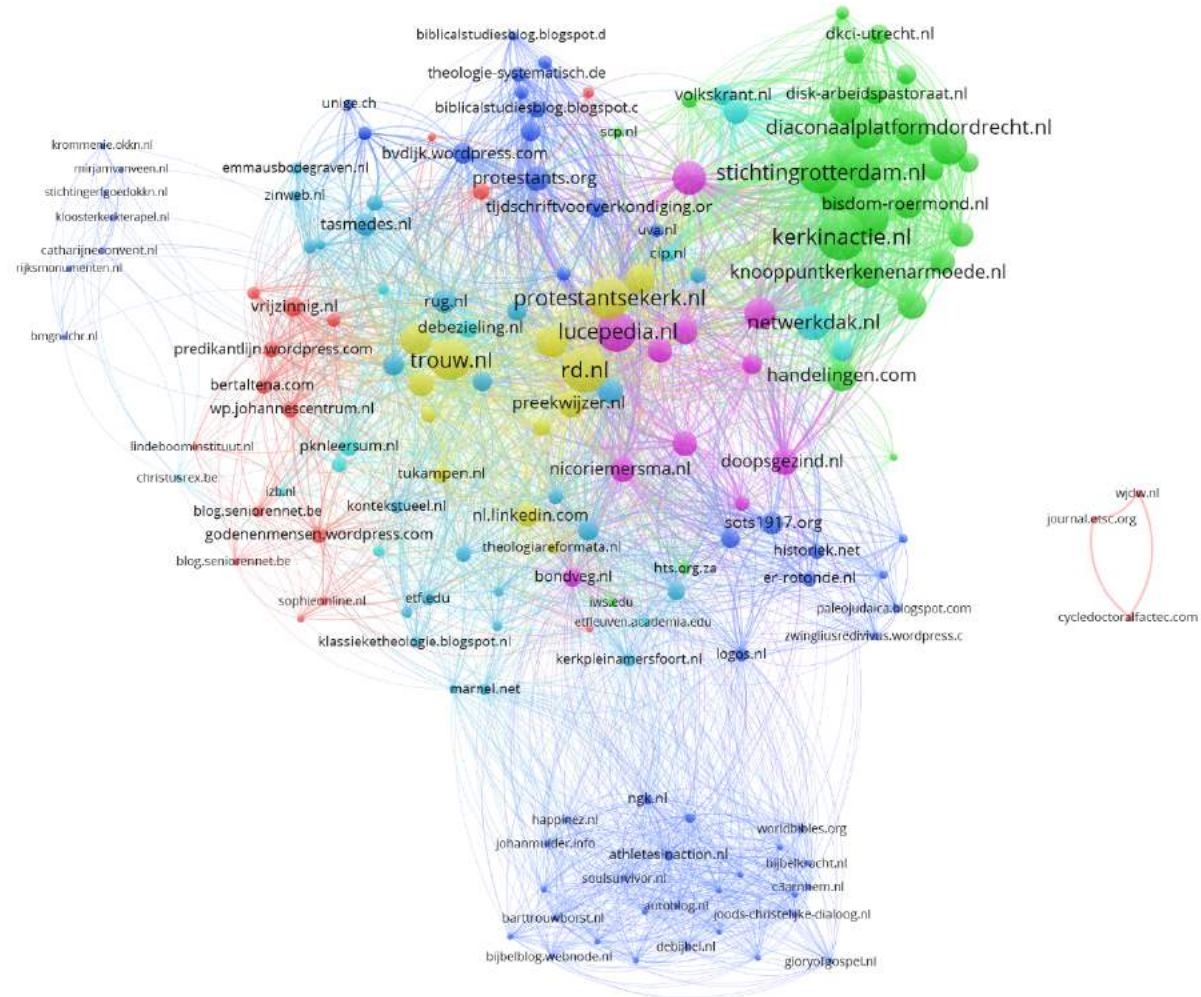
- Tracing use, classifying the user in society, science, news and politics
 - Goal: Enabling evaluation in terms of learning from past performance
 - Developed in concurrence with SIAMPI, ERiC
 - Method: formalized searches in Parliament, LexisNexis, Google & Bing, GS
 - Result: profiles of research units, identified stakeholders
- Flexibility
 - Operationalize the method to address issues and questions in the evaluation
 - Identifying meaningful interactions or groups of stakeholders

CRA types of outcome



Hybrid outcomes of research: combined academic and societal interest

Unit's stakeholders



Four phases

1. Exploratory phase

- Articulation questions and issues

2. Data gathering

- *Document analysis*
- *Quantitative methods (e.g. Scientometrics, Contextual Respons Analysis, ABC, proximities)*
- *Interviews*

3. Workshop

4. Data analysis and reporting



Results in

- Detailed analysis of:
 - Evolution and transformation of research topics
 - Possibility to diversify and protect local excellence
 - Translation knowledge in outputs, outcomes, impacts
 - Distinction between different phases of generating impact (scientific, societal)
 - Involved networks of actors and types of resources
 - Influence research on academic and societal networks and fields

- **Characteristics**
 - tailor made and modular
 - content oriented, learning capacity central
 - mixed methods approach
 - Indicators as ‘proxies’ for narratives
- **Makes visible:**
 - mission and research topics
 - communication and collaboration patterns
 - all types of output and results
 - conditions for research and infrastructure
 - process determinants (eg open science, gender diversity)

What can units do with it (I)

- Research organizations grapple with changing societal, economic and political contexts and expectations
- Give an overview of goals and missions and the ways these are embedded within the organization (goal > mobilization > output > reach)
- By using multiple methods

What can units do with it (II)

- Can serve as a starting point to develop or refine the missions of the organization
- Based on the views and experiences of researchers and users (bottom up)
- Articulating what is already going on
- And identifying new possibilities
 - New audiences, existing ones
 - Ways of communication next to books and articles
 - A clearer structure of the organization, in terms of programs, centres and projects

Context counts

- Metrics are not a 'holy grail'
- Need for narratives, context and broader perspectives

One size does not fit all!

