Monitoring the Evolution and Benefits of Responsible Research and Innovation - MoRRI

Presentations MoRRI final event (D15) PART 3

Day 2 – Discussions on policy aspects

Final Event – Discussion on policy aspects

Date: 7 March 2018

Location: Fondation Universitaire – rue d'Egmont 11, Brussels

Presentations

- Responsible metrics for RRI – Ismael Rafols
- Limits in measurement – Ingeborg Meijer

Brussels, MoRRI, March 2018

RRI^RRI (RRI square)?

Responsible (use of) metrics of responsible research & innovation

Ismael Ràfols

Ingenio (CSIC-UPV), Universitat Politècnica de València

CWTS (Centre for Science and Technology Studies), University of Leiden
SPRU (Science Policy Research Unit), University of Sussex

The argument

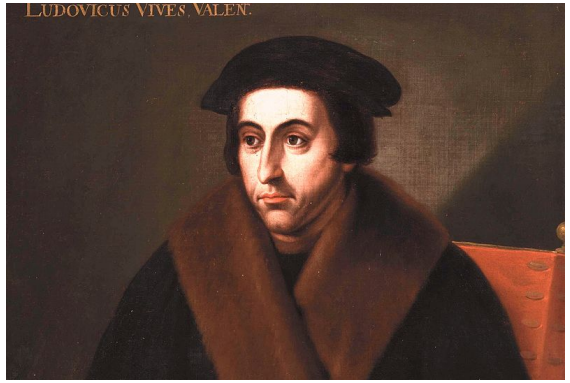
- Indicators are performative
  - Stakeholders in the STI sytems will respond to monitoring

- Indicators may mean different things in different contexts
  - Dangerous sense of objectivity and precision
  - **Broadening out** – taking into account different knowledge sources
  - **Opening up** –consider different perspectives
  - Pluralistic use of indicators in **deliberative processes**

- General prescriptions on the use of indicators
  - **Humility, diversity (context) and reflexivity**

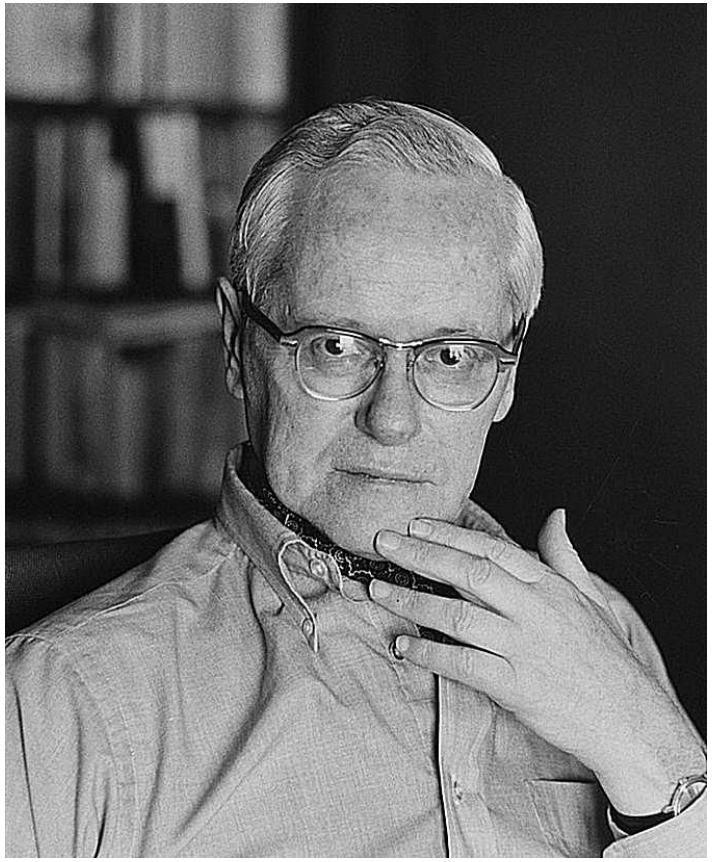
Governance beyond science: music and humanism

Joan Lluís Vives, humanist
(1493 - València, 1540 - Bruges)



Pau Casals playing a Catalan folksong in exile since 1939

Robert K. Merton (1942)
The normative structure of science



Science blossoms in
pluralistic, democratic societies

PERFORMATIVE EFFECTS OF INDICATORS

The parable of Prussian scientific forestry (Seeing like a state, *J. Scott*)

Forests in Old Prussia

- Wild
- Uncontrolled
- Unpredictable
- Inefficient

Enlightenment and Scientific forestry:

- Cut the wild forest
- Plant Norway spruce –reduce diversity
- Increase yield and predictability
- Loss of forest activities for peasants:

(fruits, hunting, medicinal herbs, etc.)

Now…Predictive analytics?

The parable of Prussian scientific forestry (Seeing like a state, *J. Scott*)

Monocultures and Forest death

- Nutrient depletion leading to 20-30% production loss in 2nd generation
- Storm felling
- Pests due to loss of ‘services’ of insects, birds and animals.

Restoration forestry or forest hygiene:

- Artificial ant colonies & spiders
- Wooden boxes to provide bird nests
- The dangers of dismembering a complex set of relations and processes to isolate a single element of instrumental value

The parable of Prussian scientific forestry (Seeing like a state, *J. Scott*)

Task reduction
“Exaggerating only slightly, one might say that the crown's interest in forests was resolved through its fiscal lens into **a single number: the revenue yield of the timber** that might be extracted annually.. ”

Performativity
“Backed by state power through records, courts, and ultimately coercion, **these state fictions transformed the reality they presumed to observe**, although never so thoroughly as to precisely fit the grid.”

The challenge: indicators that reflect goals

"Our statistics and accounts reflect our aspirations, the values that we assign to things. They are inseparable from our vision of the world and the economy, of society, and our conception of human being and our interactions.

Treating these as objective data, as if they are external to us, beyond question or dispute, is undoubtedly reassuring and comfortable, but it's dangerous.

It's dangerous because we get to the point where we stop asking ourselves about the purpose of what we are doing; what we are actually measuring, and what lessons we need to draw."

Nicolas Sarkozy in prologue to

"Mismeasuring our lives" by Stiglitz, Sen and Fitoussi (2010)

Indicators of RRI

What can be the effects of RRI indicators use?

1. Restoration of the destruction by conventional indicators (IF, h-factor, based on productivist narrative) ? (Producing complexity)

2. Can RRI indicators become a goal of their own?
   - Goal displacement & task reduction? (Reduction of complexity)
     - Particularly dangerous in regulation in bureaucratic traditions:
       - e.g. Open Access (via gold –reducing resources?)
       - e.g. Gender equality ‘plans’

PLURALISM AND INCLUSIVENESS
CLOSING DOWN VS. OPENING UP

Policy use of S&T indicators: Appraisal

Appraisal:

'the ensemble of processes through which knowledges are gathered and produced in order to inform decision-making and wider institutional commitments' Leach et al. (2010)

Example:

Allocation of resources based on research "excellence"

Breadth: extent to which appraisal covers diverse dimensions of knowledge

Narrow: citations/paper

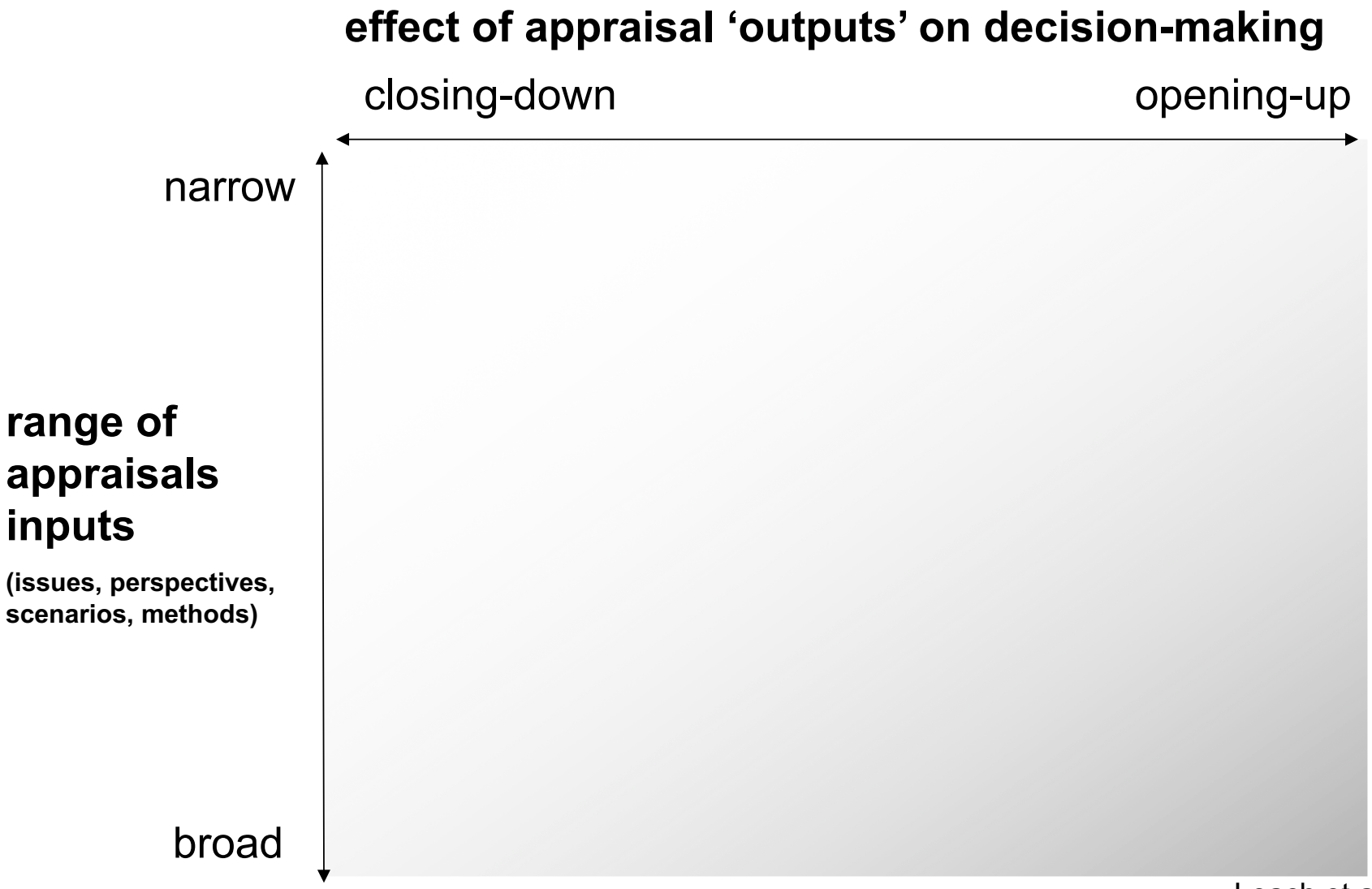
Broad: citations, peer interview, stakeholder view, media coverage, altmetrics

Openness: degree to which outputs provide an array of options for policies.

Closed: fixed composite measure of variables → unitary and prescriptive

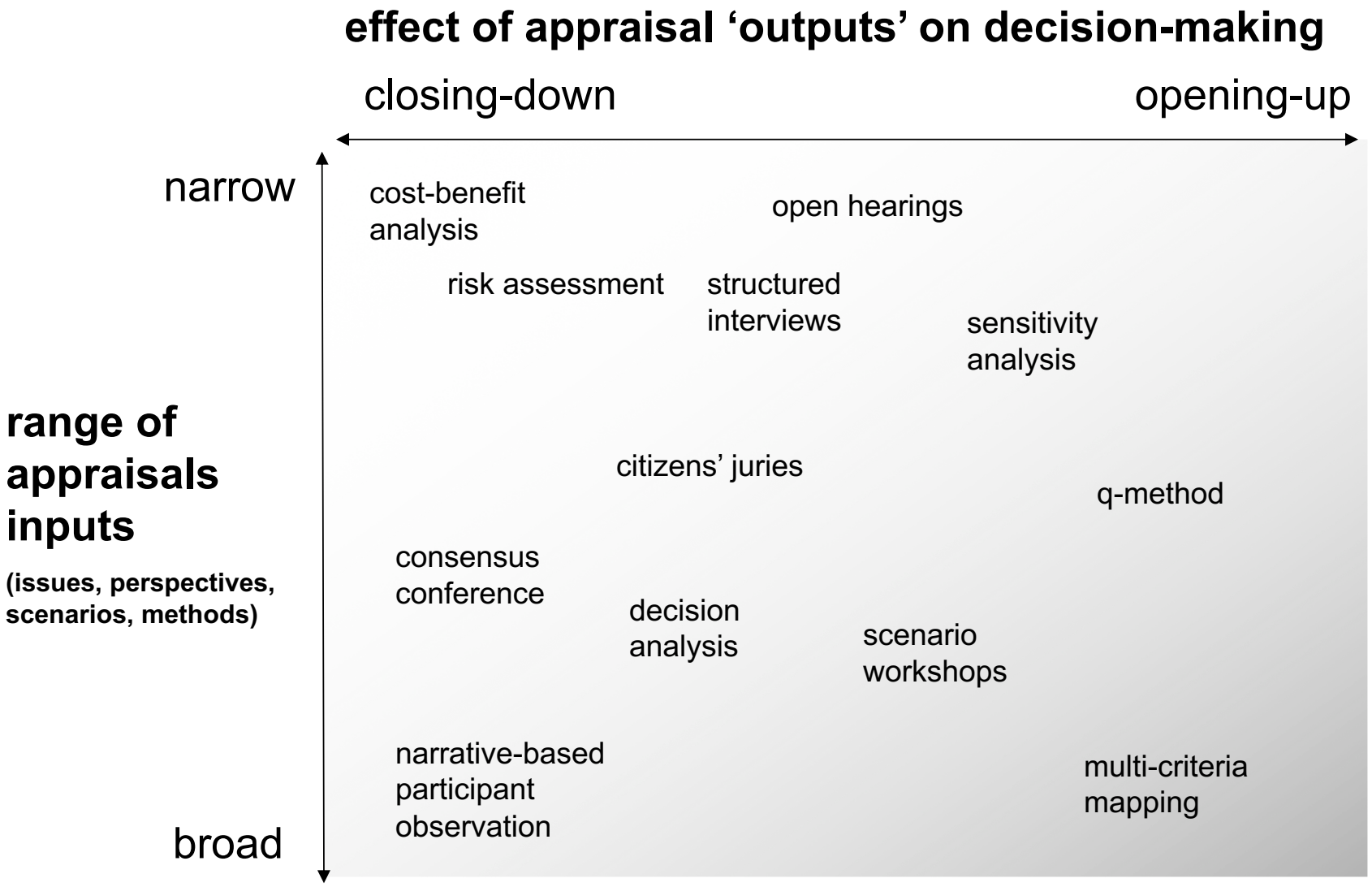
Open: consideration of various dimensions → plural and conditional

Appraisal methods: broad vs. narrow & closing vs. opening



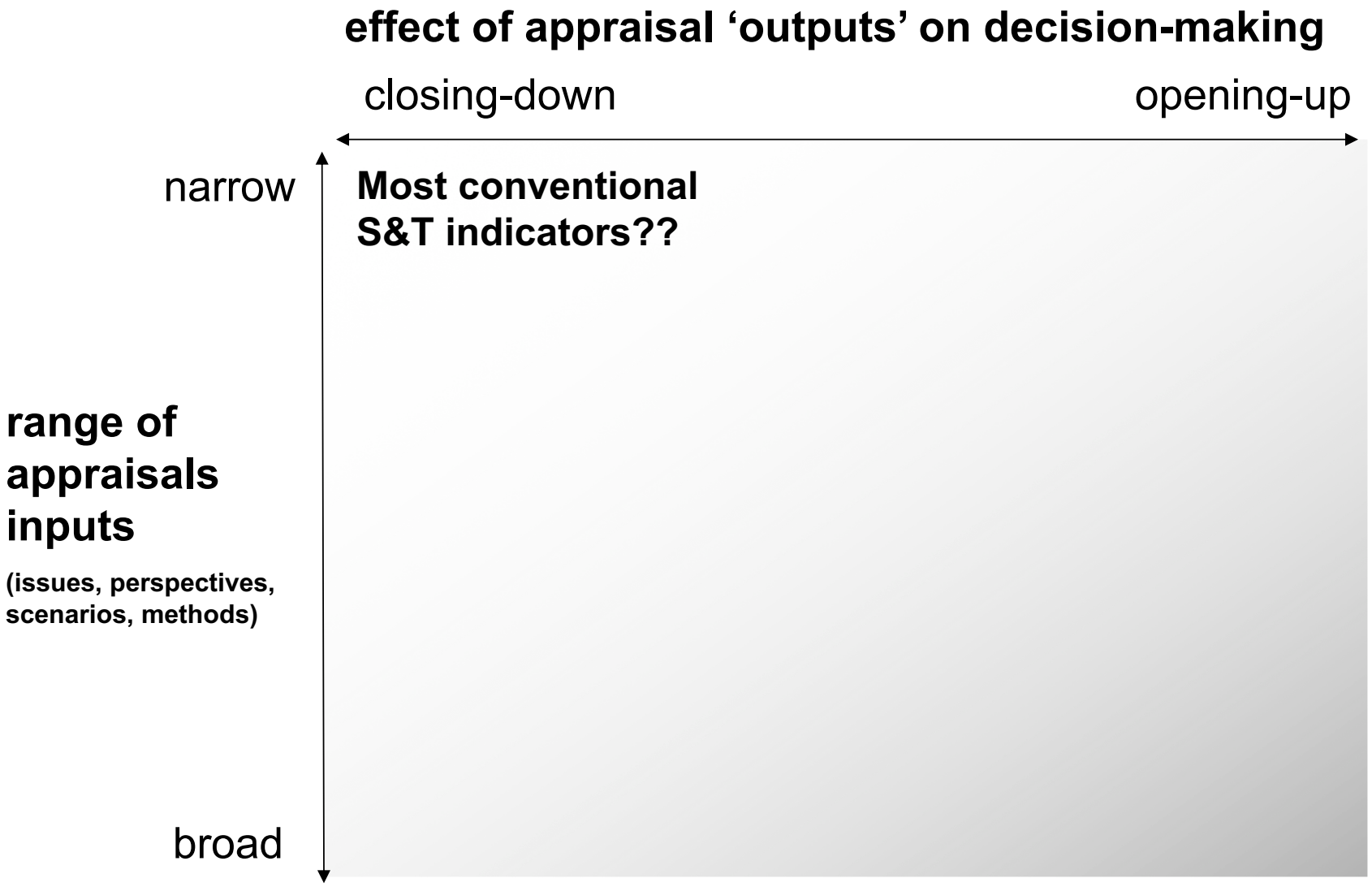
Leach et al. 2010

Appraisal methods: broad vs. narrow & close vs. open

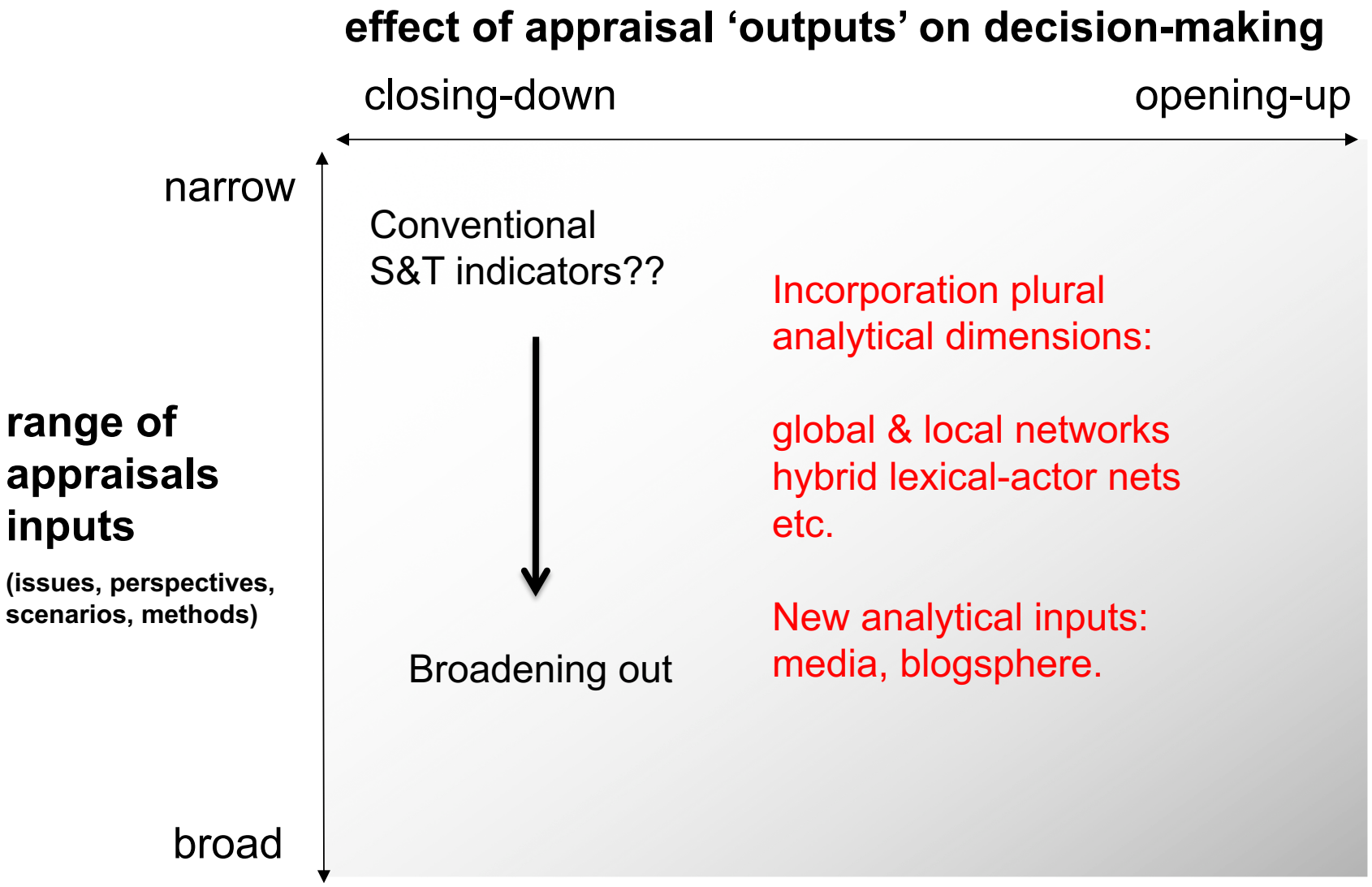


Stirling et al. (2007)

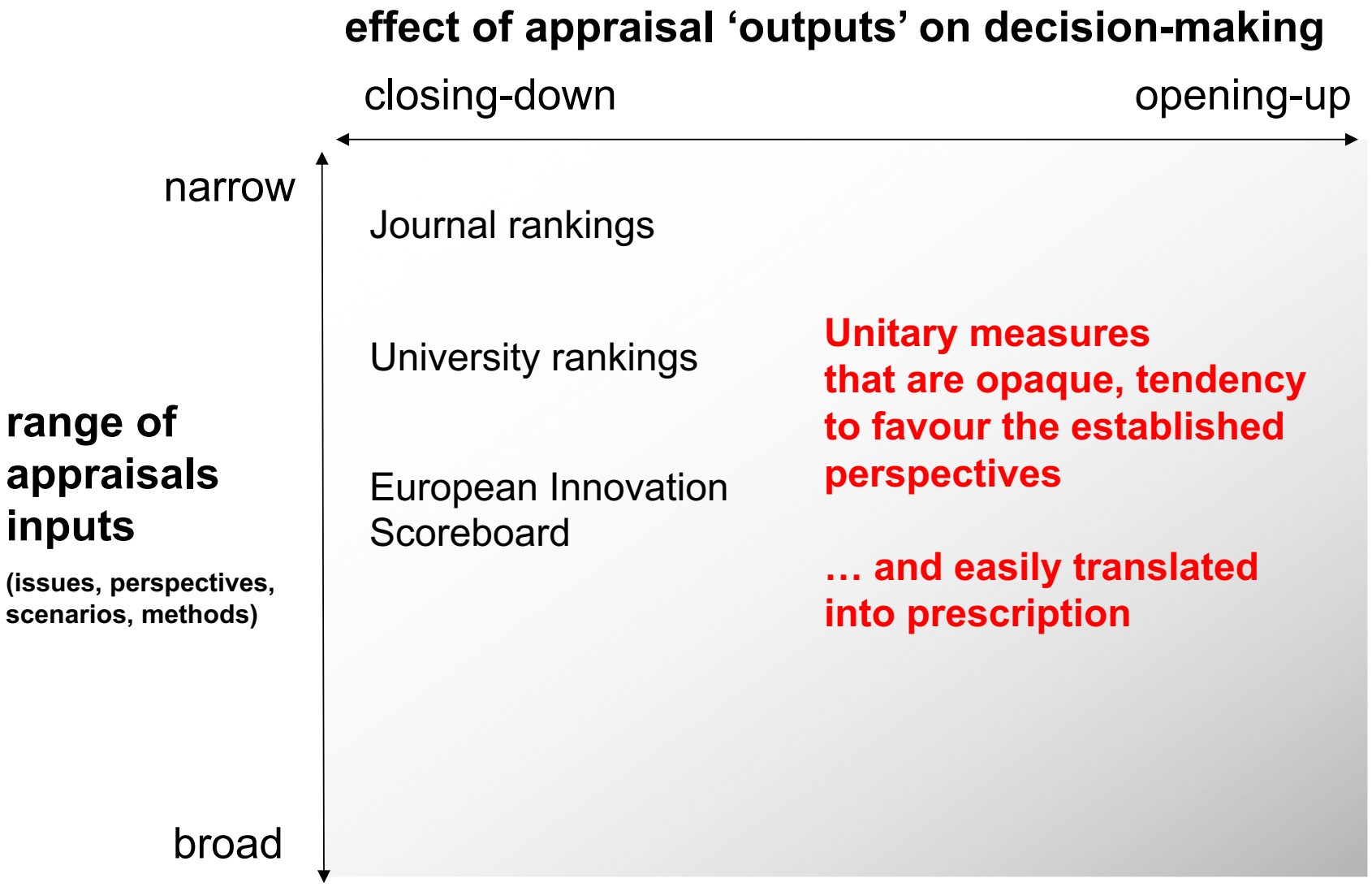
Appraisal methods: broad vs. narrow & closing vs. opening



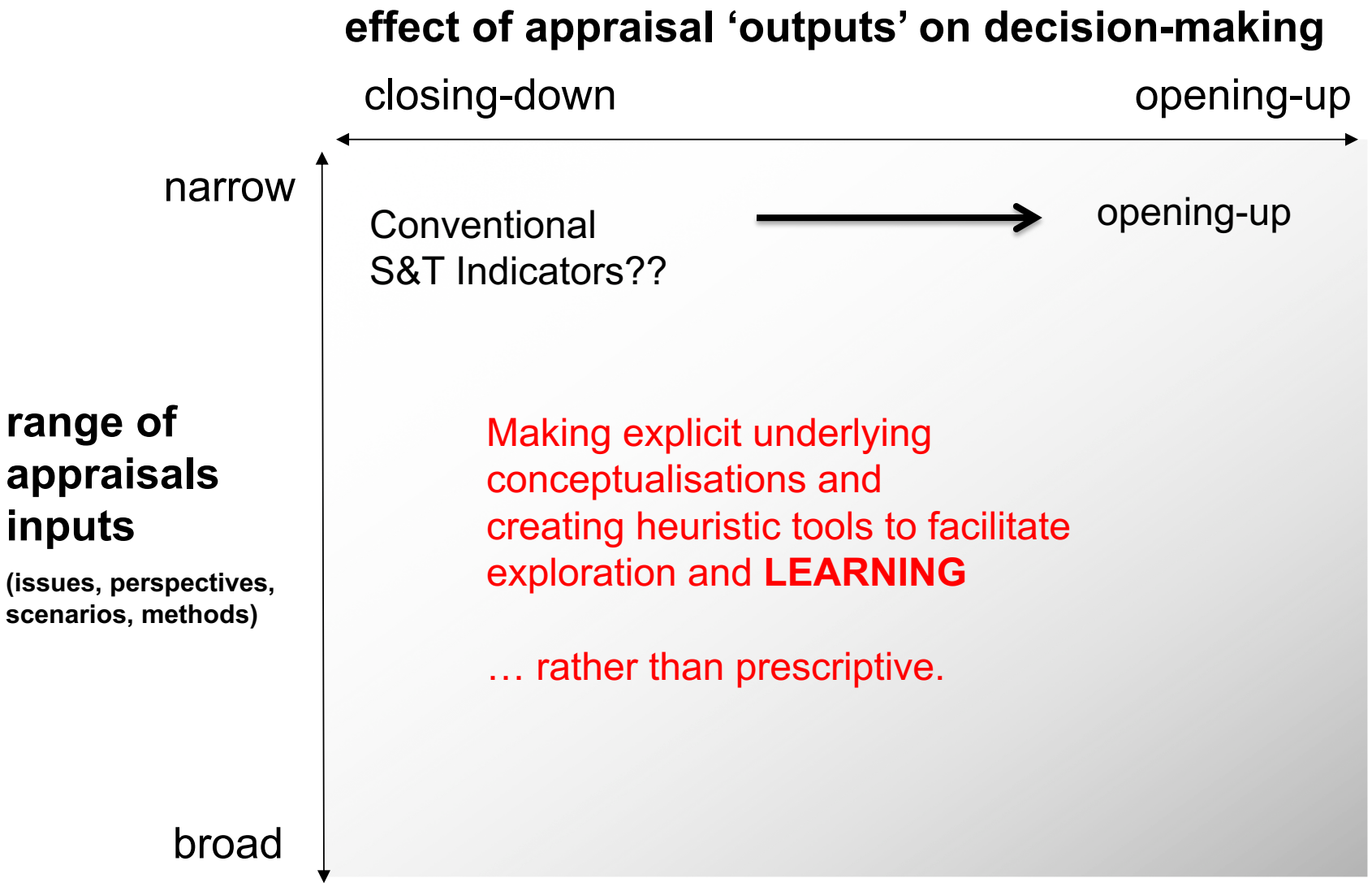
Broadening out S&T Indicators



Appraisal methods: broad vs. narrow & closing vs. opening



Opening up S&T Indicators



Composite Innovation Indicators (25-30 indicators)

European (Union) Innovation Scoreboard

Grupp and Schubert (2010) show that order is highly dependent on indicators weightings.

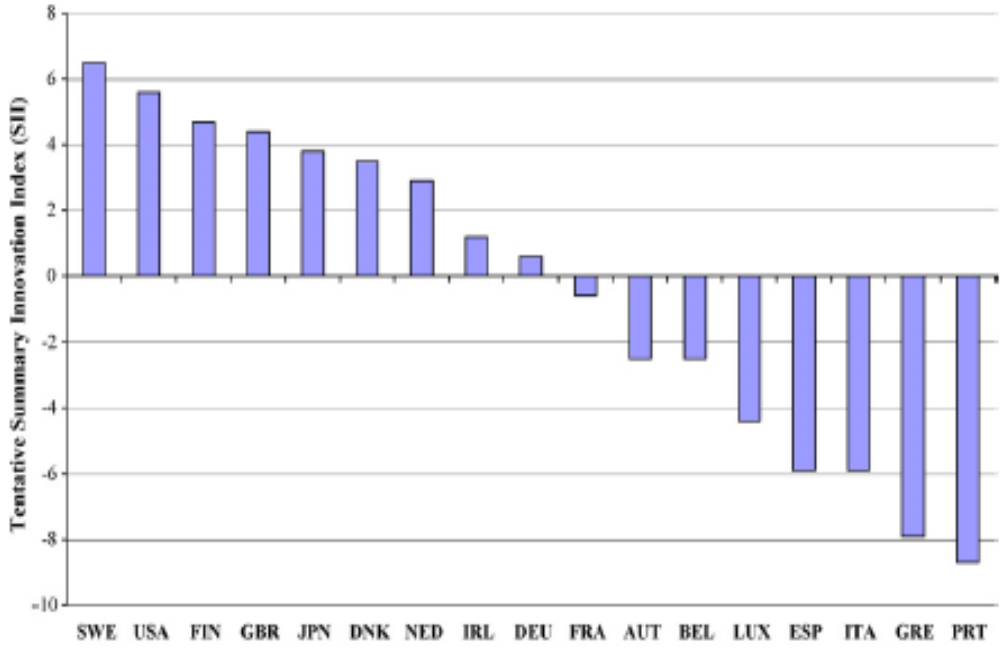


Fig. 1. Summary Innovation Index of the European Innovation Scoreboard 2001 (modified graph from European Commission, 2001, p. 12).

Sensitivity analysis

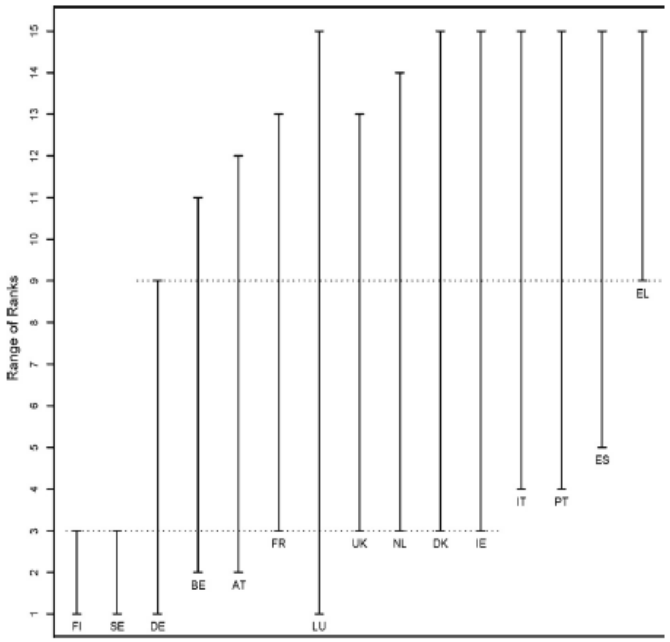
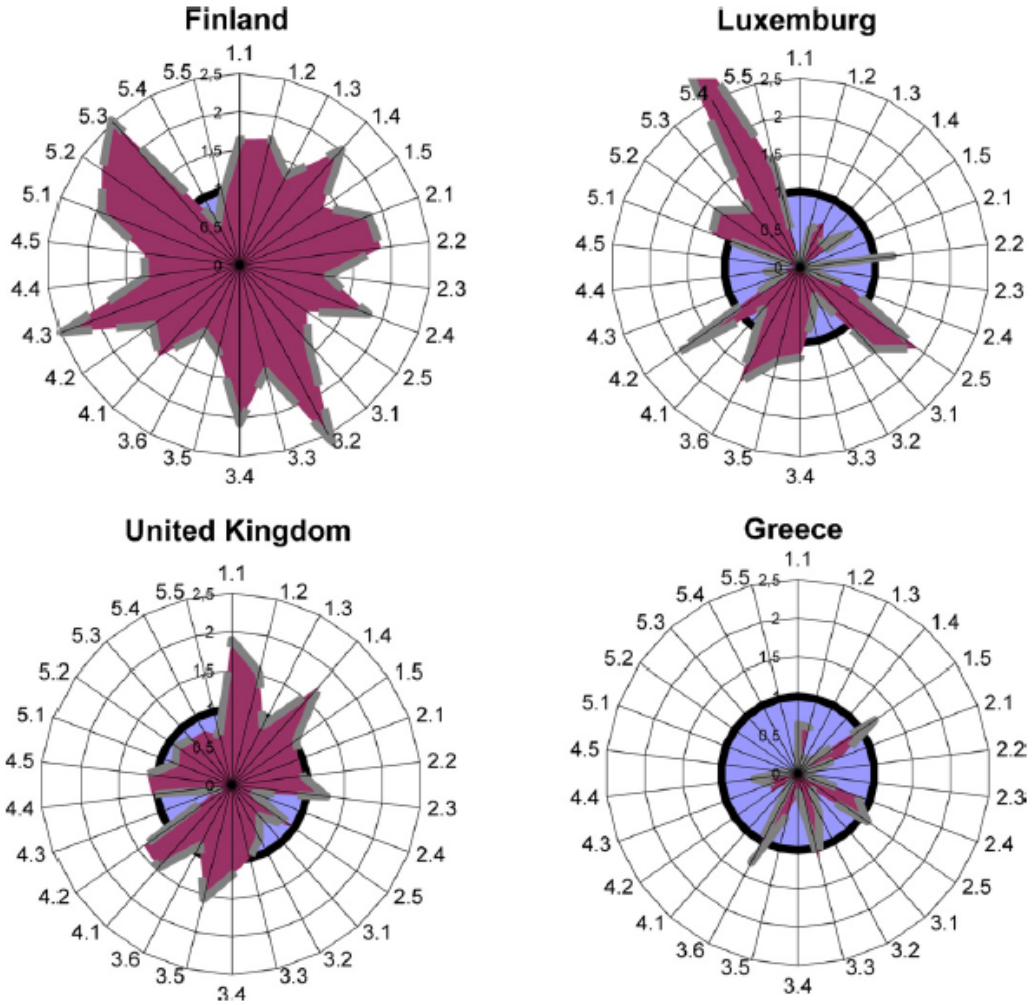


Fig. 2. Sensitivity analyses for the European Innovation Scoreboard 2005 (sources: original data European Commission, 2005).

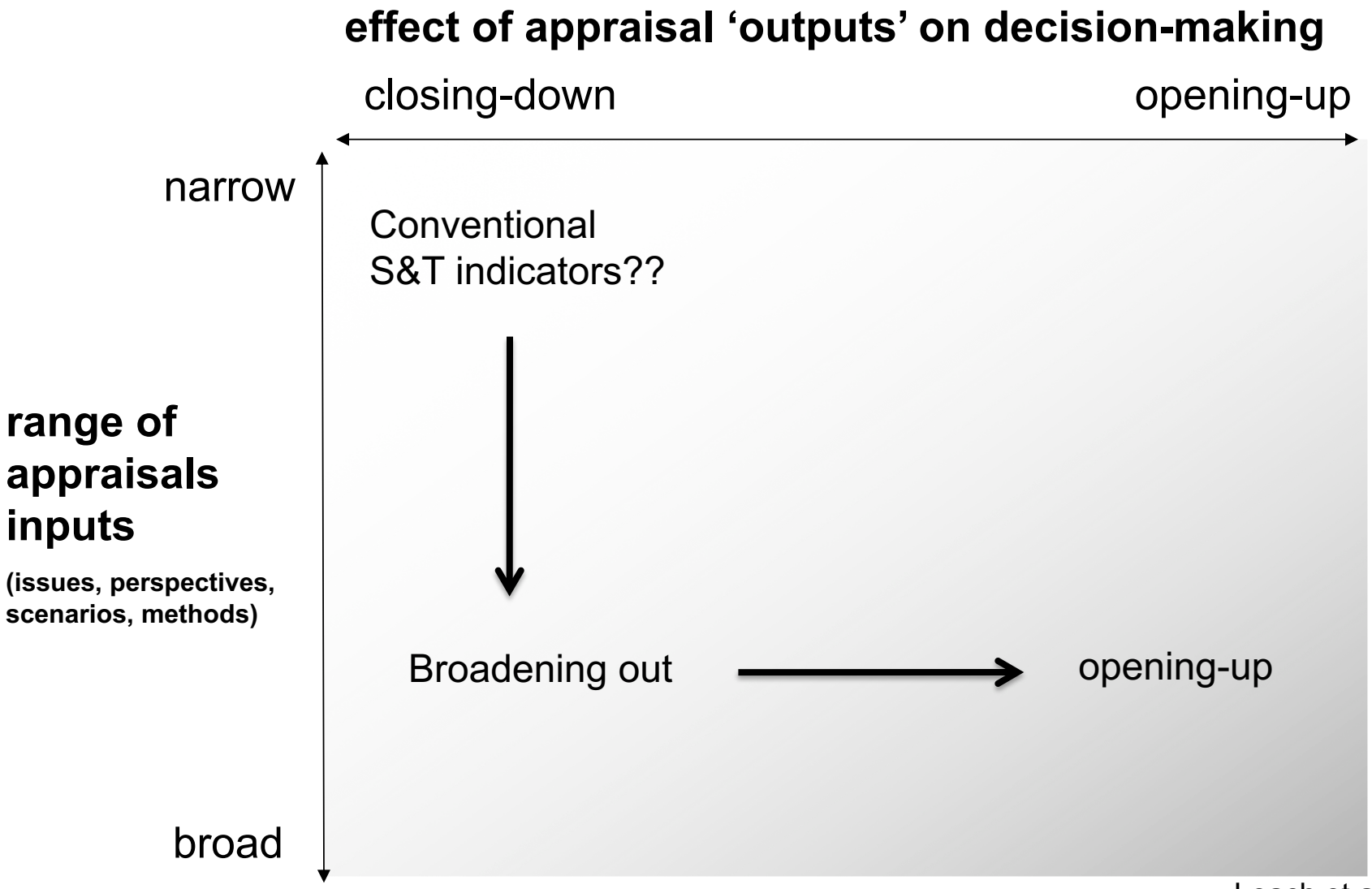
Solution: representing multiple dimensions
(critique by Grupp and Schubert, 2010)



Use of spider diagrams allows comparing like with like

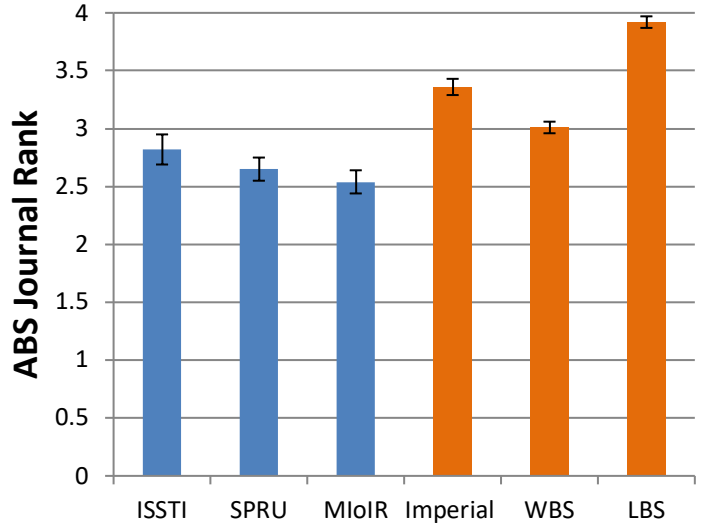
Charles Edquist:
At least differentiate Input from Output Indicators

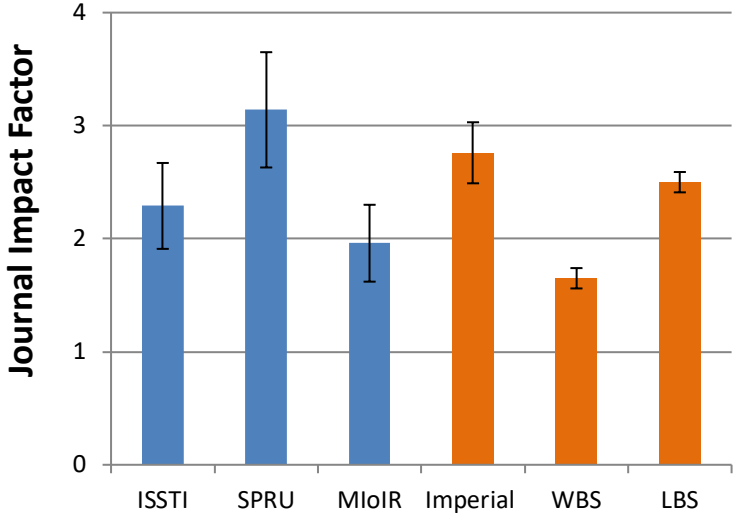
1. Preserving multiple dimensions in broad appraisals

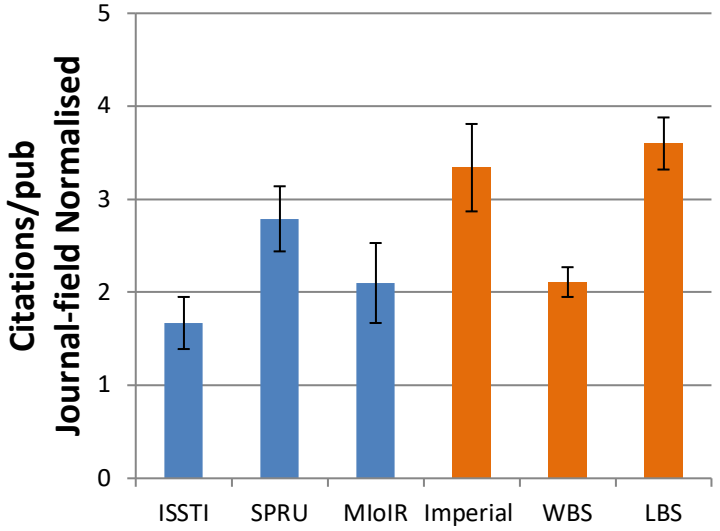


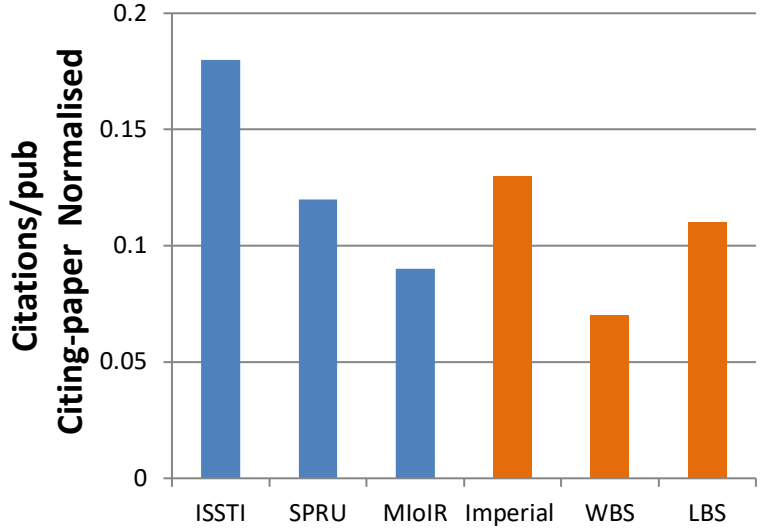
Leach et al. 2010

Multiple & non-convergent indicators of “excellence”









Which one is more meaningful??

Rafols et al. (2012, Research Policy)

What uses of indicators?

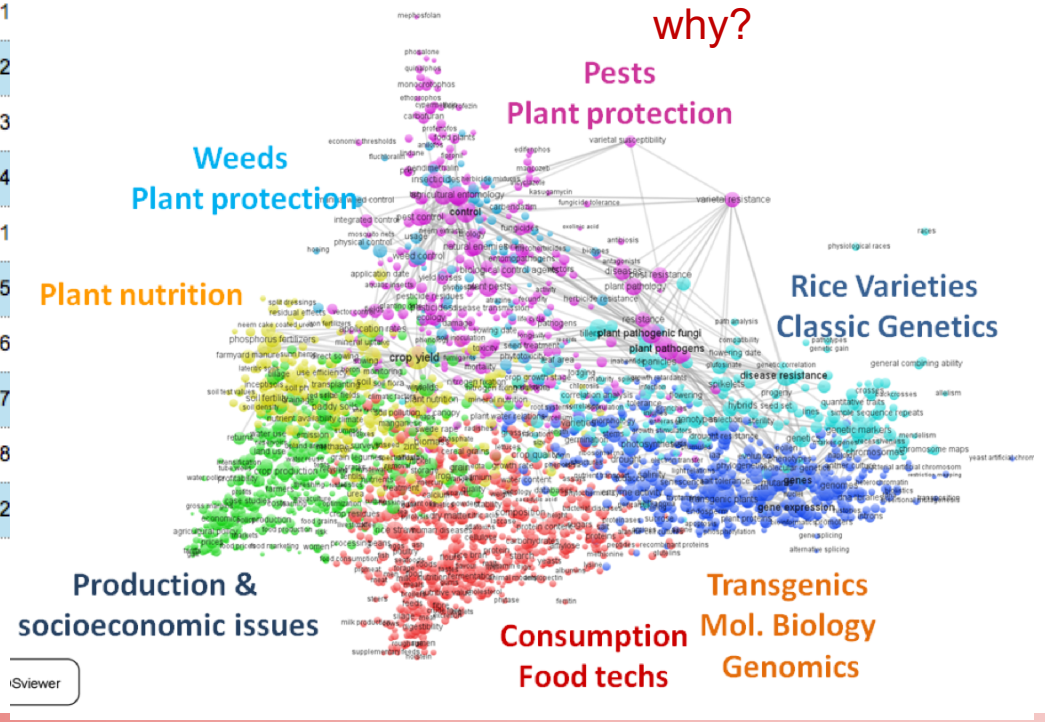
Model 1: Unique and prescriptive
“Best choices” for perfornmace
Rankings -- ranking list of preferences

Academic Ranking of World Universities - 2011

World Rank	Institution	Country	National Rank	Total Score
1	Harvard University		1	
2	Stanford University		2	
3	Massachusetts Institute of Technology (MIT)		3	
4	University of California, Berkeley		4	
5	University of Cambridge		1	
6	California Institute of Technology		5	
7	Princeton University		6	
8	Columbia University		7	
9	University of Chicago		8	
10	University of Oxford		2	

Model 2: Plural and conditional
Exploring choices and strategies
Learning about patterns and trajectories

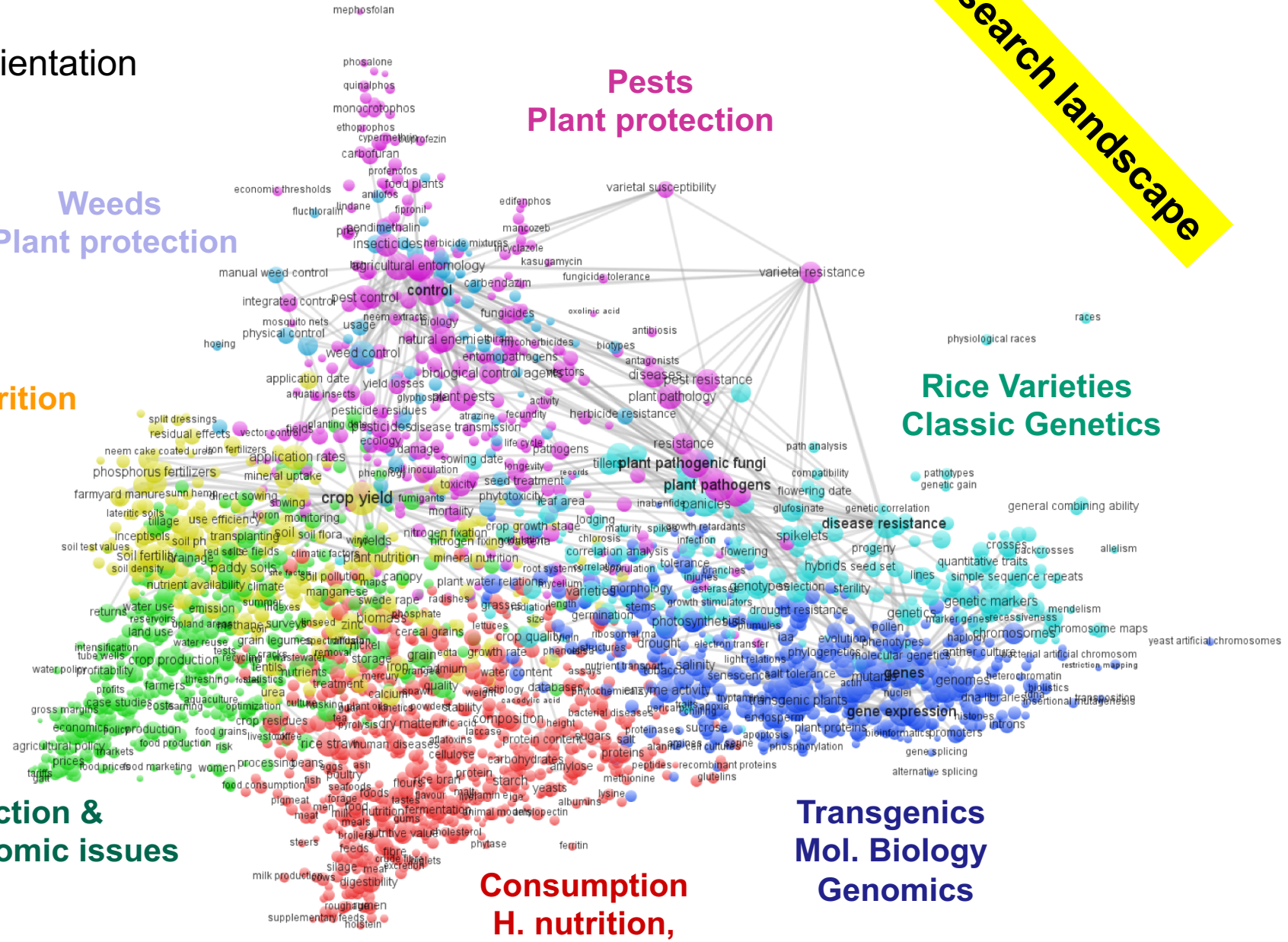
which way?
what alternatives?
why?



What are the “options” in rice research?

Research landscape

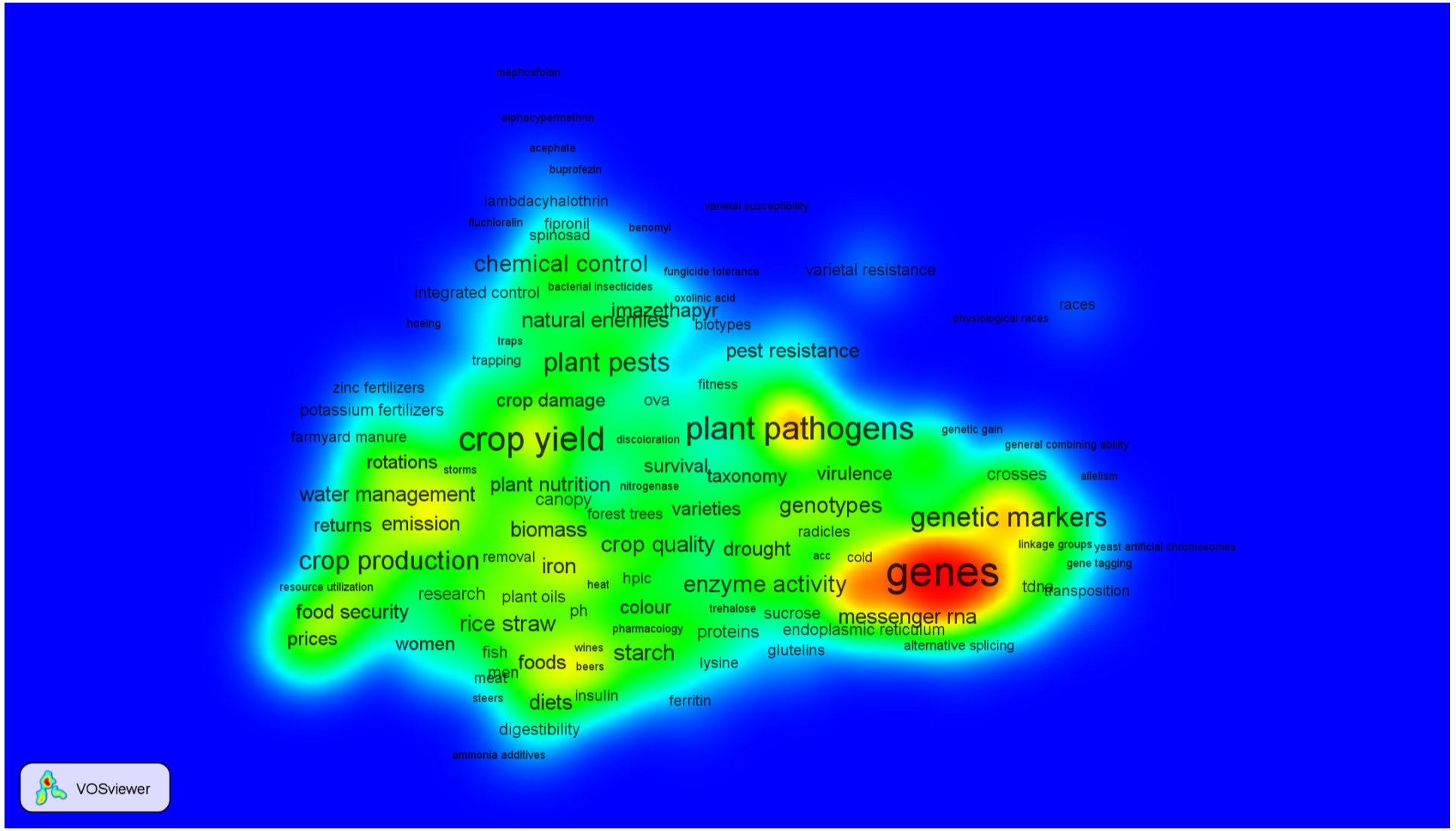
Mission-orientation



VOSviewer

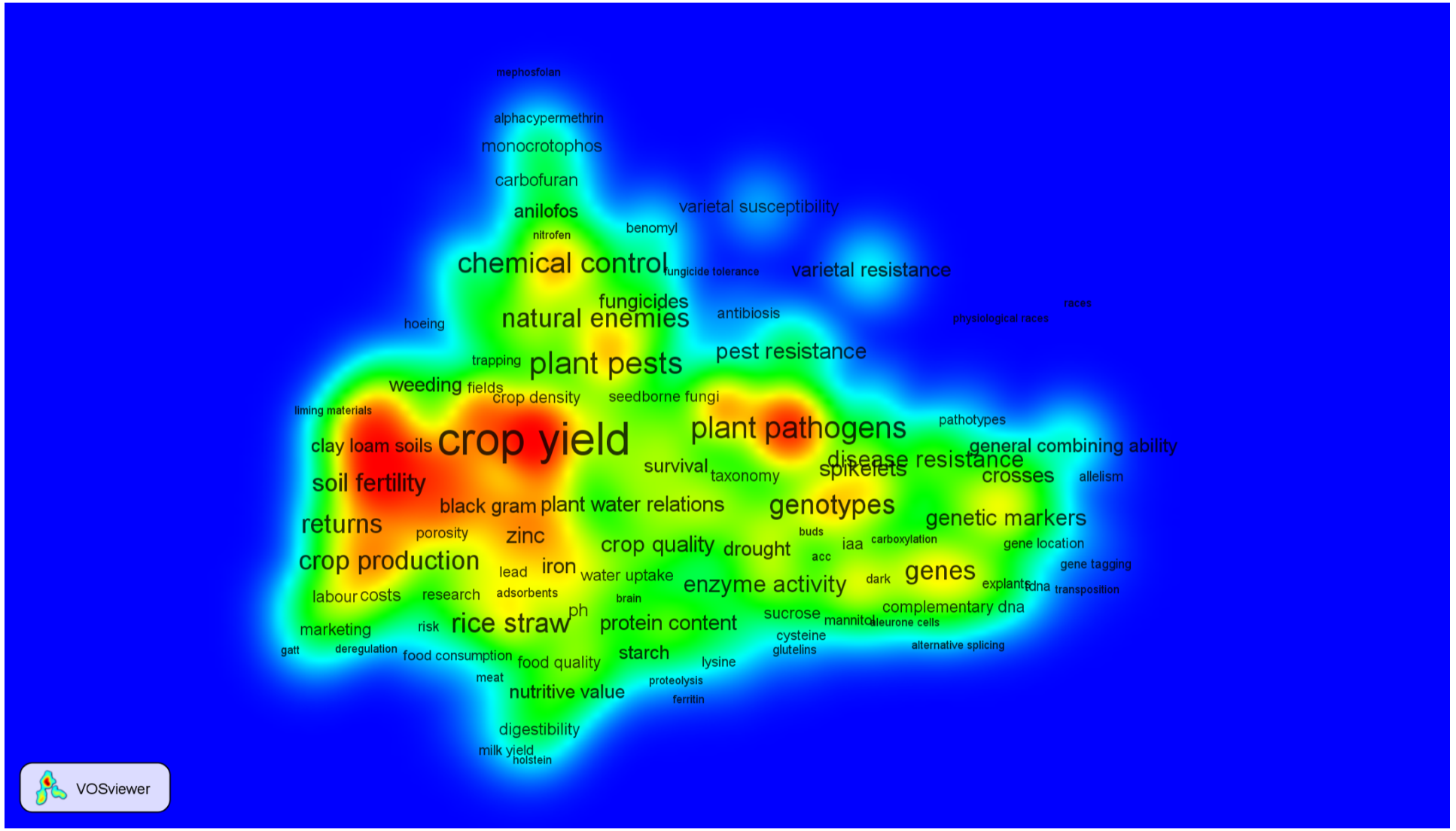
Not a fined-grain perspective

How research priorities differ by country



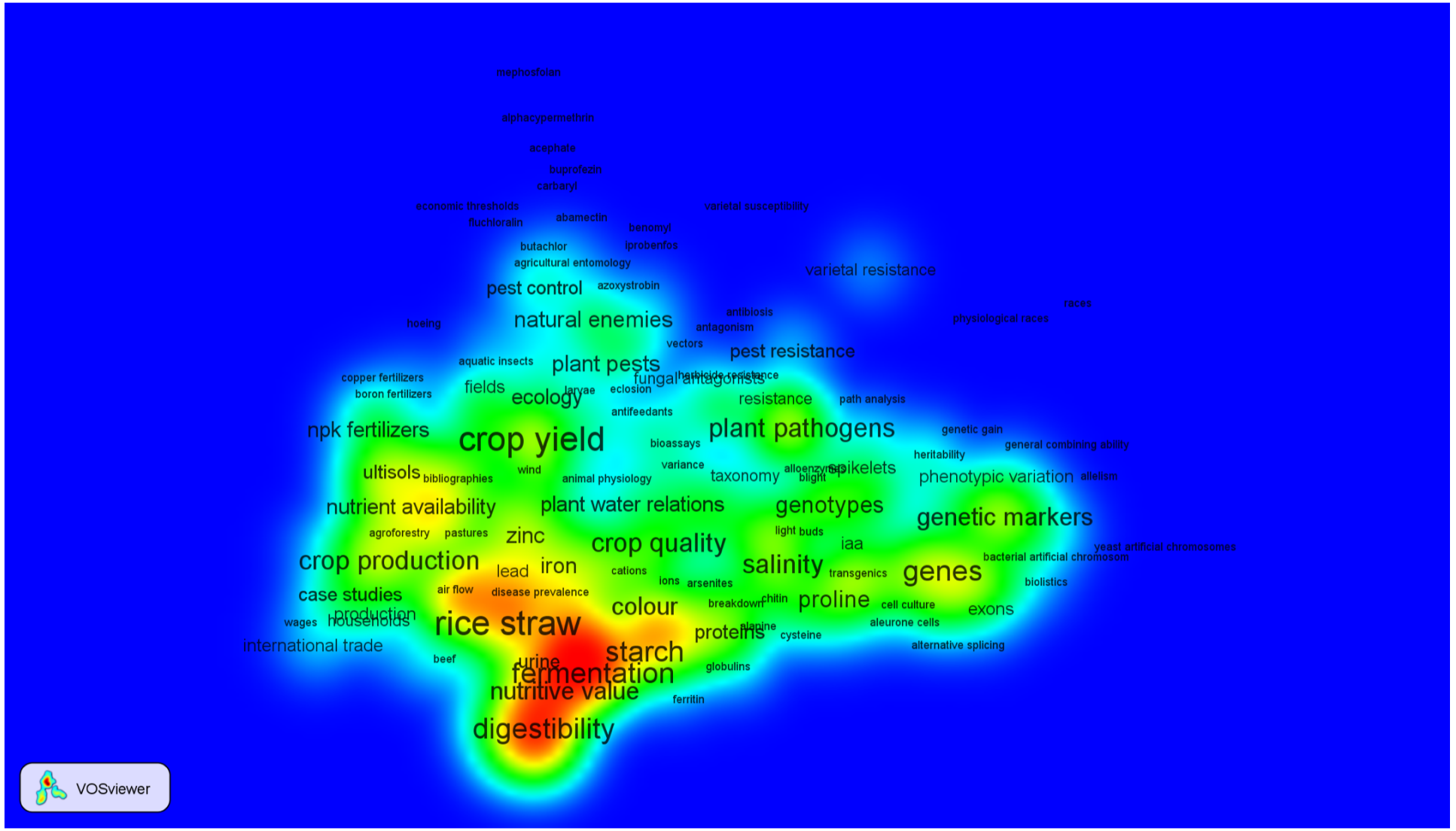
US, 2000-12

How research priorities differ by country



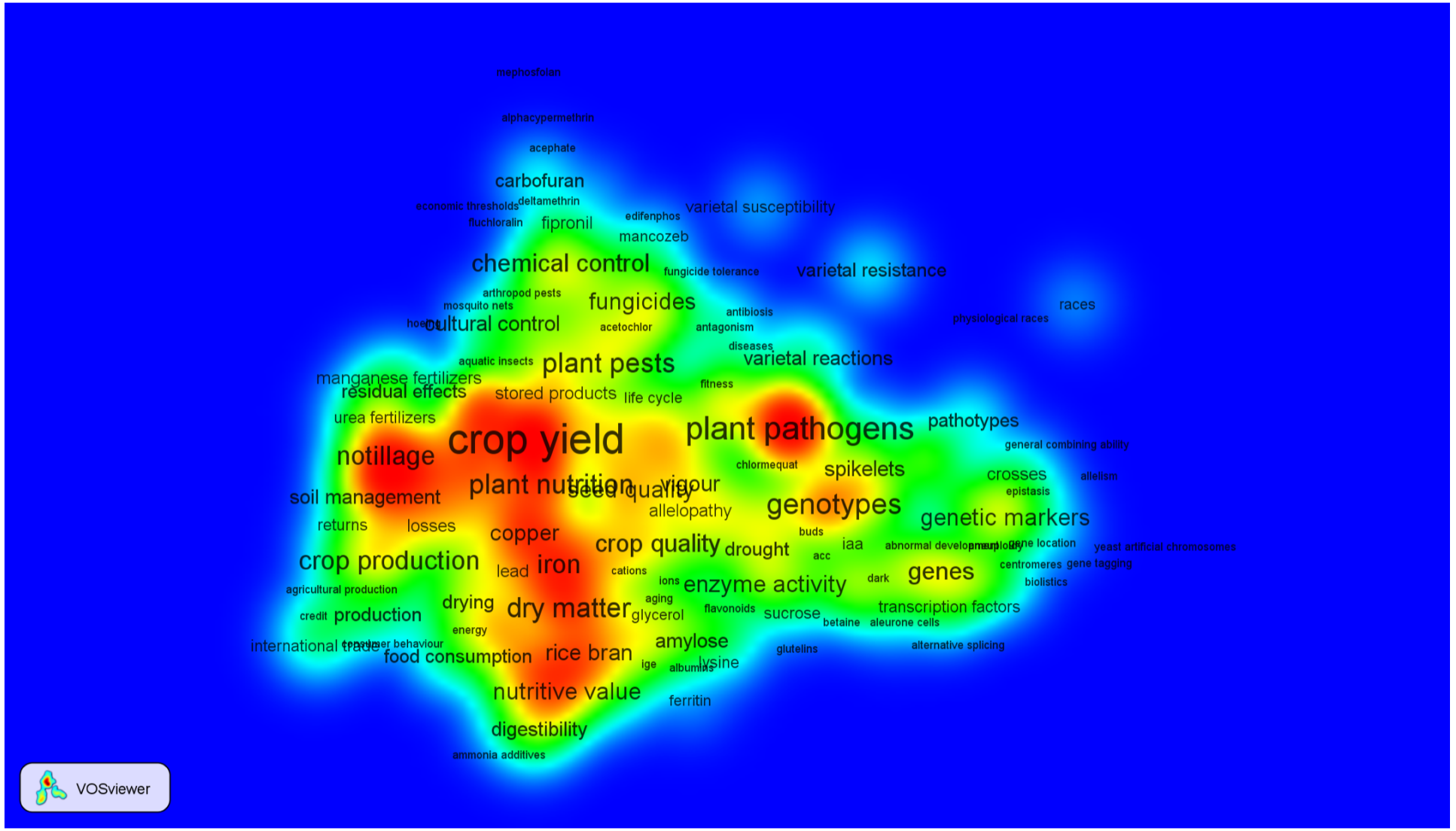
India 2000-12

How research priorities differ by country



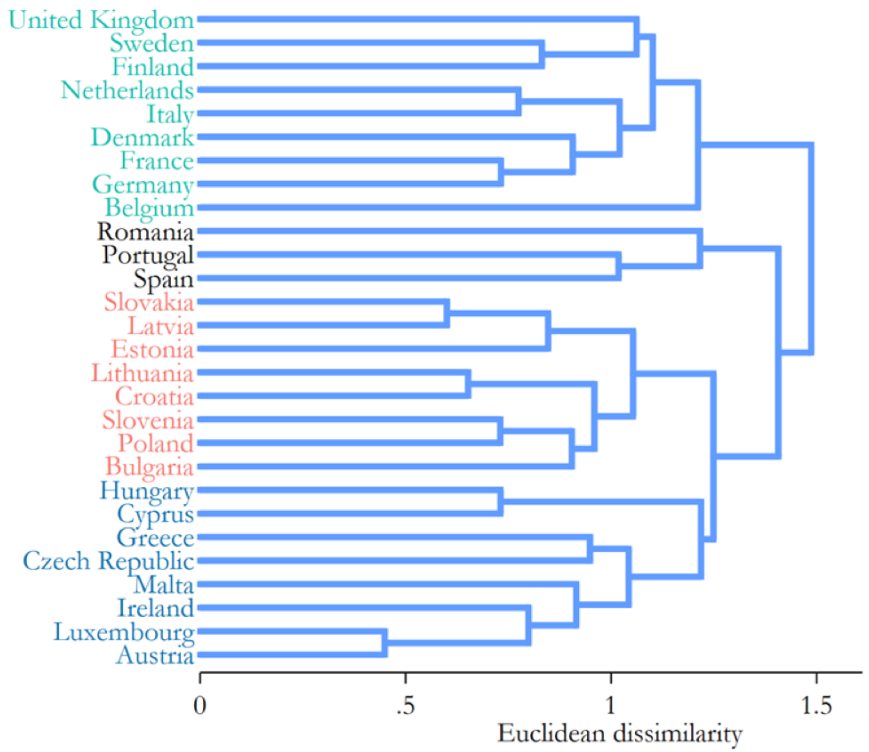
Thailand 2000-12

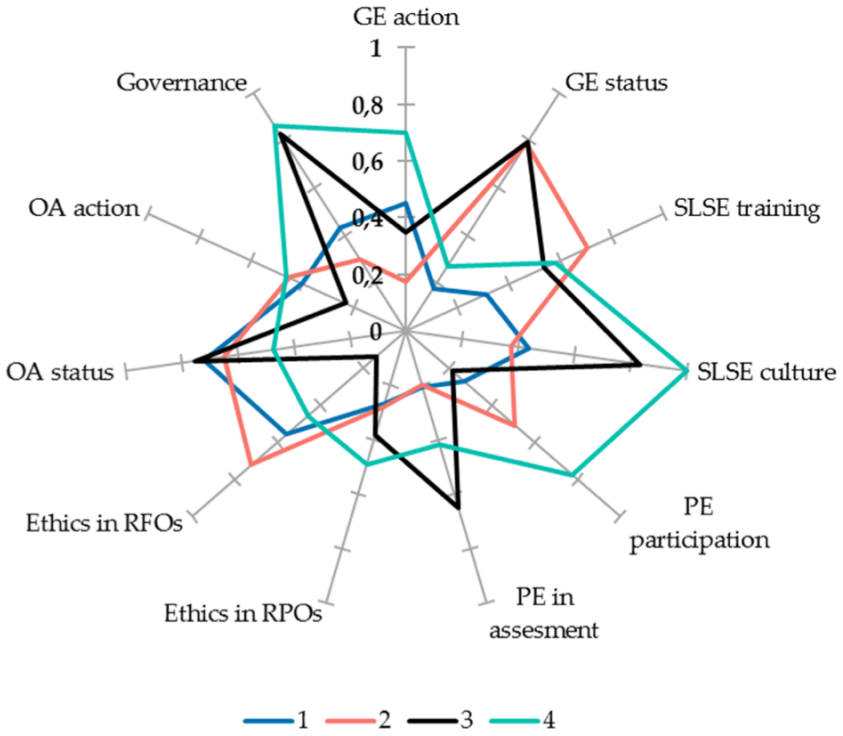
How research priorities differ by country



Brazil 2000-12

MoRRI – not rankings, but patterns





MoRRI Final Report

S&T indicator as a tools in policy deliberation

- ‘Conventional’ use of indicators (‘Science Arbiter’--Pielke)
  - Purely analytical character (i.e. free of normative assumptions)
  - Seeking convergence (partial converging indicators, Martin and Irvine, 1983)
  - Aimed at justifying ‘best-choices’ (e.g. excellence)

  → **Unitary and prescriptive advice**

- ‘Opening up’ indicators(‘Honest broker’ --Pielke)
  - Aimed at locating the actors in their context and dynamics

    → Not predictive, or explanatory, but exploratory
  - Construction of indicators is based on choice of perspectives

    → Make explicit the possible choices on what matters
  - Supporting debate

    → Making science policy more ‘socially robust’

  → **Plural and conditional advice**

Barré (2001, 2004, 2010), Stirling (2008)

NAVIGATING WITH THE METRIC TIDE

Responsible Metrics: “The Metric” Tide on uses of indicators

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

Towards a more ‘inclusive’ use of indicators?

- **Inclusiveness in the inputs (multiple perspectives)**
  - *Broadening out:* Create more diverse indicators
    - Improve representation of SSH, non-English, the “South”,…
    - Indicators of open science, RRI, hidden, social innovation

- **Inclusiveness in the outputs (multiple goals)**
  - *Opening up:* develop toolkits that allow exploration of directions. New ways of presenting indicators.
    - Multi-criteria tools
    - Interactive visualisations

- **Inclusiveness in the monitoring process**
  - Develop new social processes in which:
    - indicators are tools for interpretation and deliberation (Barré, 2010)
    - broader participation of ’experts’ (including social stakeholders)

Principles of the “The Leiden Manifesto”

1. Quantitative evaluation should support qualitative, expert assessment.

2. **Measure performance against the research missions of the institution, group or researcher.**
3. **Protect excellence in locally relevant research.**

4. Keep data collection and analytical processes open, transparent and simple.
5. Allow those evaluated to verify data and analysis.

6. **Account for variation by field in publication and citation practices.**
7. Base assessment of individual researchers on a qualitative judgement of their portfolio.
8. Avoid misplaced concreteness and false precision.

9. Recognize the systemic effects of and indicators.
10. Scrutinize indicators regularly and update them.

Judgement

Sensitive to Contexts

Transparency

Pluralism

Reflexifity

Hicks, Wouters, Waltman, de Rijcke and Rafols (Nature, 2015)

Misalignment between research and societal needs

Perceived mismatch between discourses (or expectations) of research and actual outcomes.

Societal challenges:

Energy, environment, health.

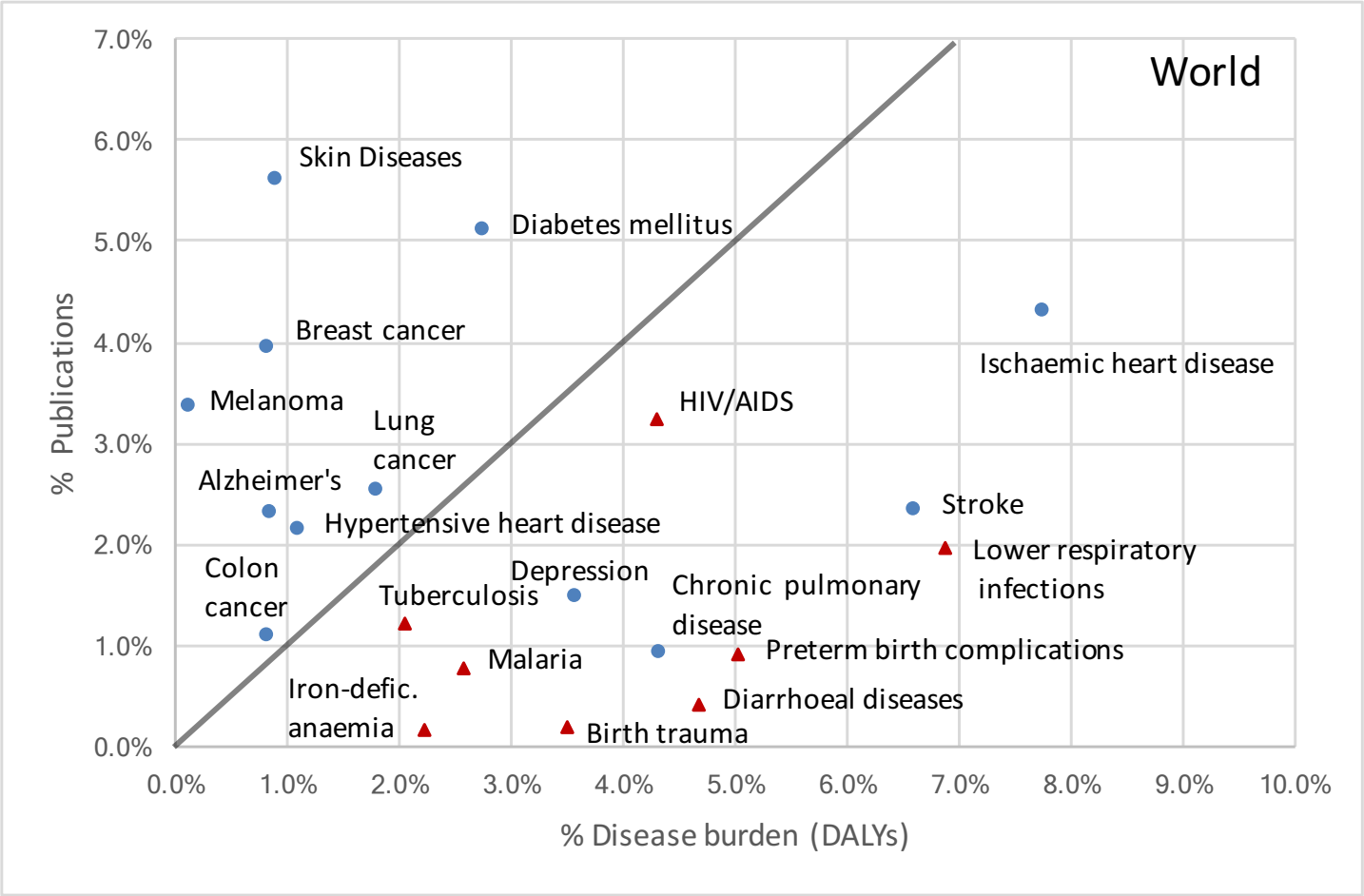
More research does not mean better societal outcomes

Monitoring tools and incentives (S&T indicators!) are part of the problem.

Source: Daniel Sarewitz – Saving Science – New Atlantis August, 2016

S&T indicator as a tools in policy deliberation

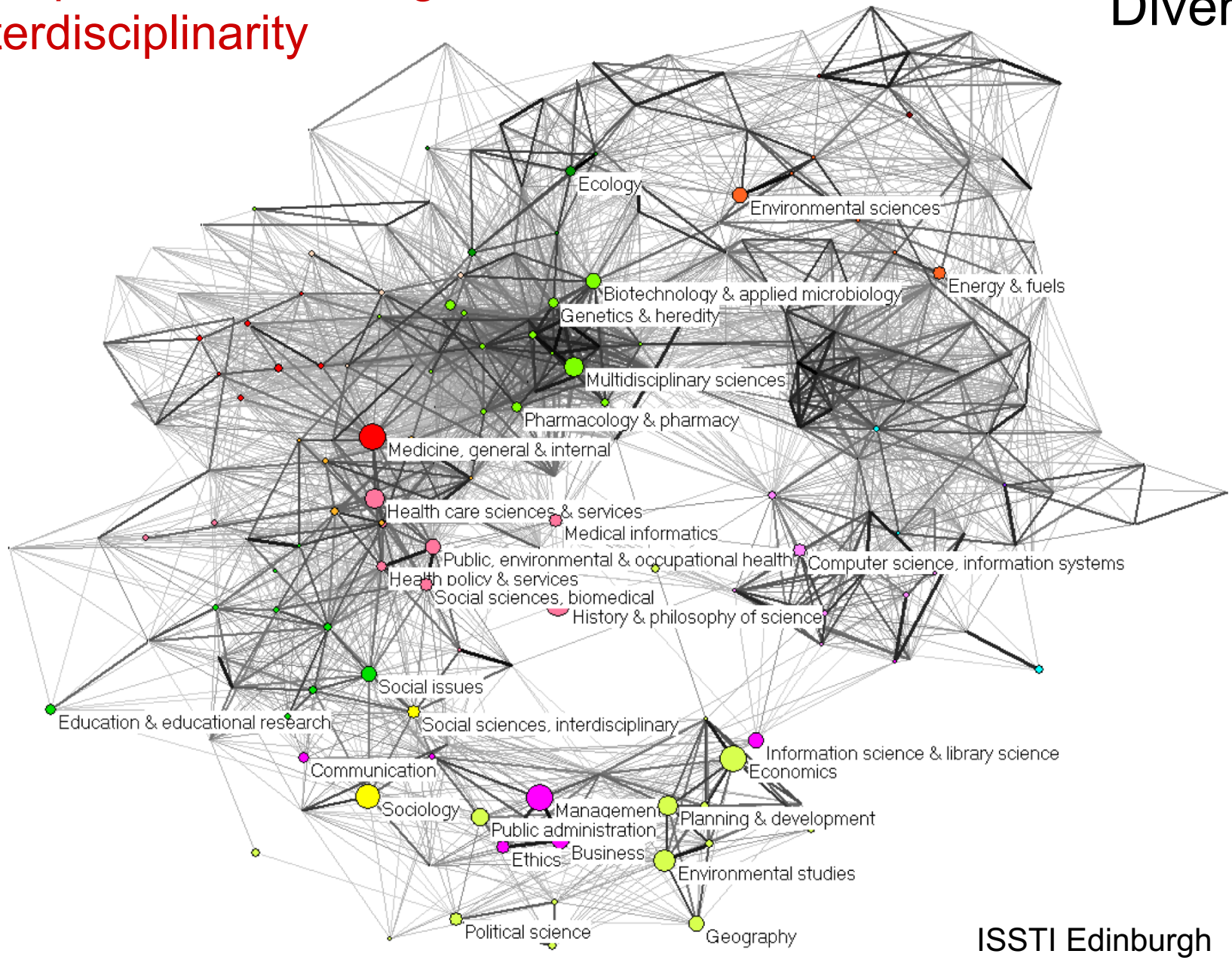
Not necessarily complex data – e.g. contrasting data that needs experts interpretation.



Research Portfolios on Diseases. Yegros et al. (2017)

Multiple understandings of interdisciplinarity

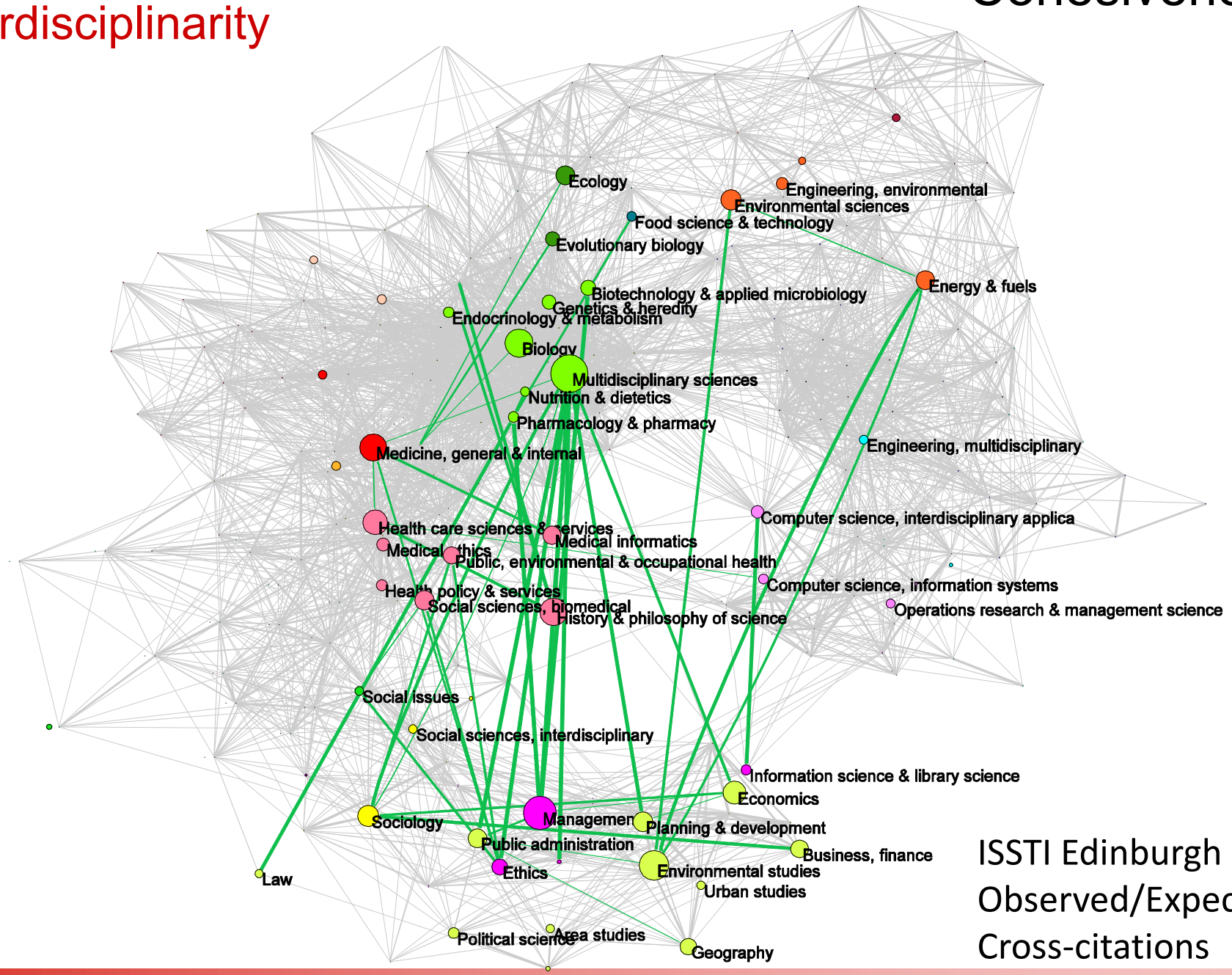
Diversity



ISSTI Edinburgh
WoS Cats of references

Multiple understandings of interdisciplinarity

Cohesiveness



ISSTI Edinburgh
Observed/Expected
Cross-citations

Mixed methods in assessment

Hybrid forums are spaces "in which (…) uncertainties predominate, and everyone contributes information and knowledge that enrich the discussion." Callon et al. 2001

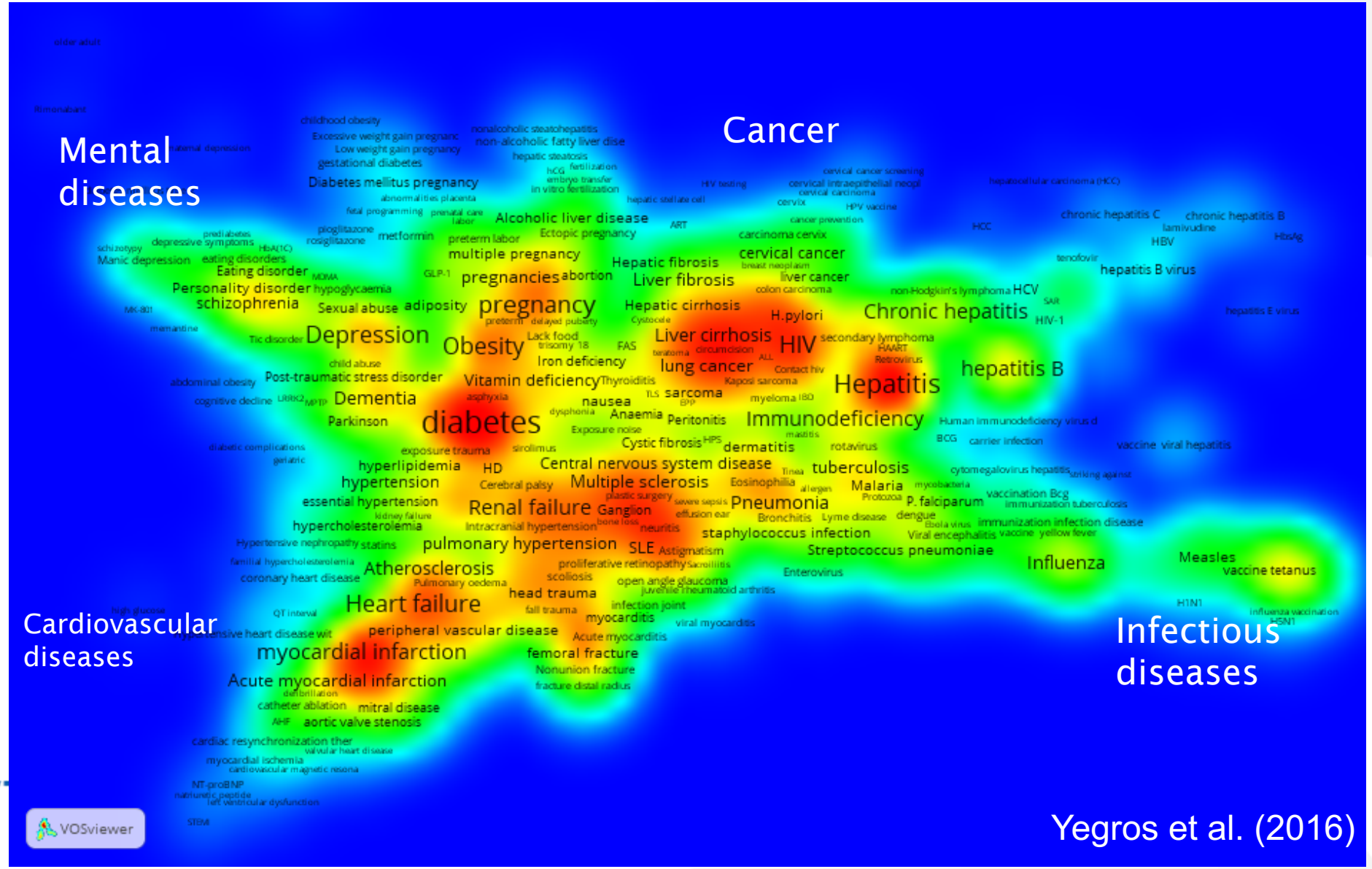
Plural and conditional **quantitative and qualitative evidence**

- **Multicriteria** Decision Analysis Methods (MCM, Stirling)
- **Participatory** Impact Pathway Analysis (PIPA)

Evaluative inquiry: De Rijcke and Wouters:
Makes visible:

-- mission and research topics
-- communication and collaboration patterns
-- all types of output and results
-- conditions for research and infrastructure
-- process determinants (eg gender diversity)

Global map of research on diseases

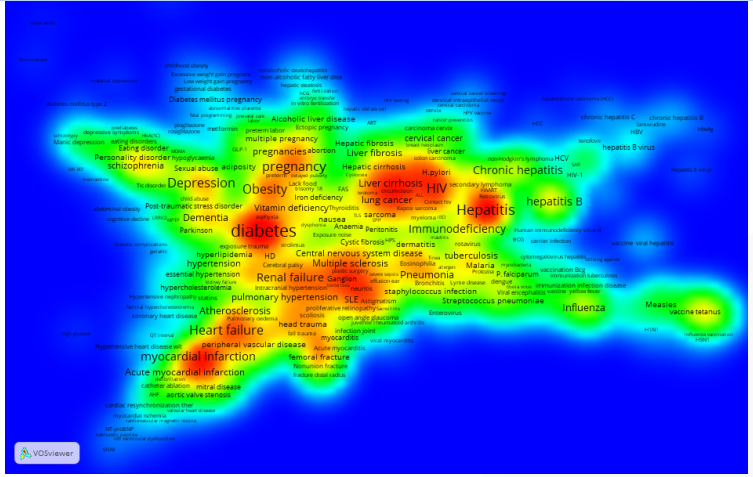


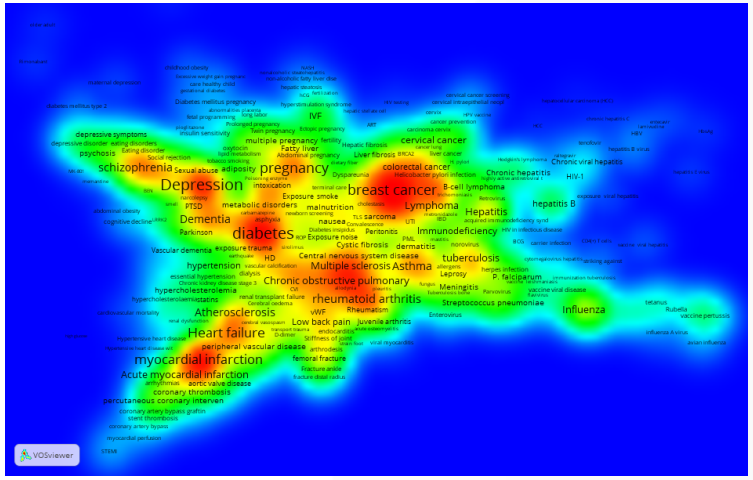
Yegros et al. (2016)

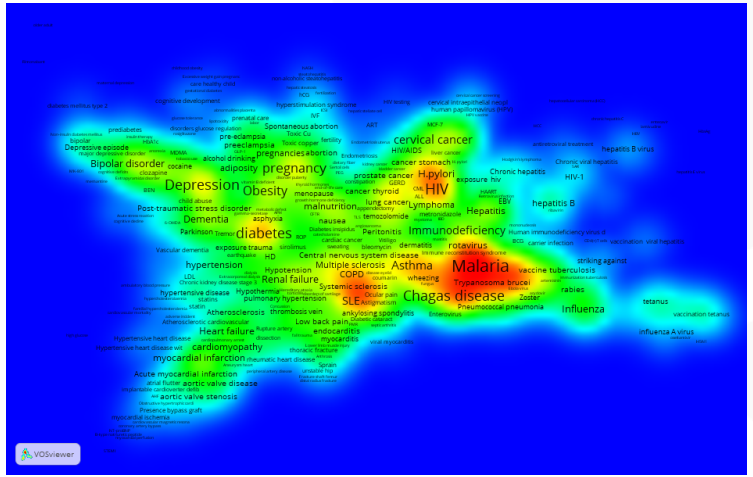
World

Netherlands

Colombia

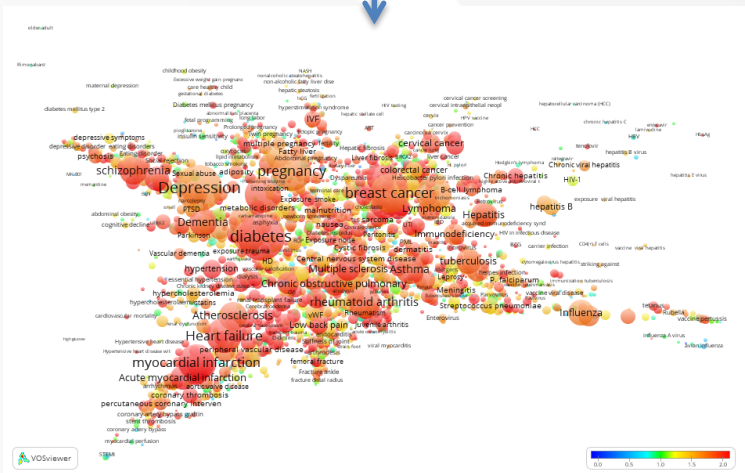


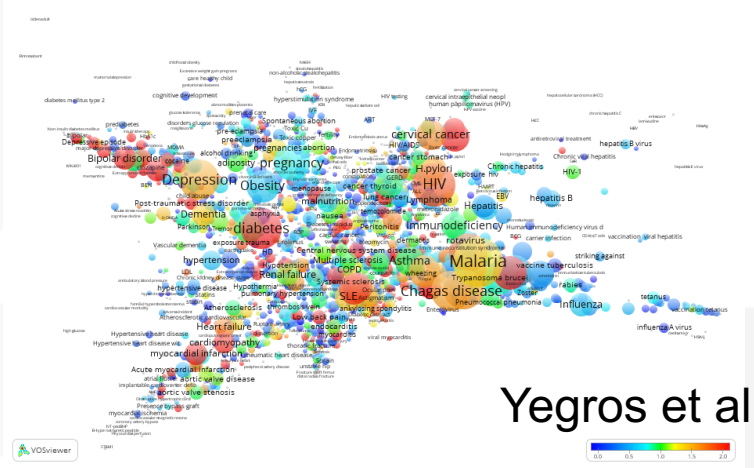




← Publication profile

Citation impact ↓

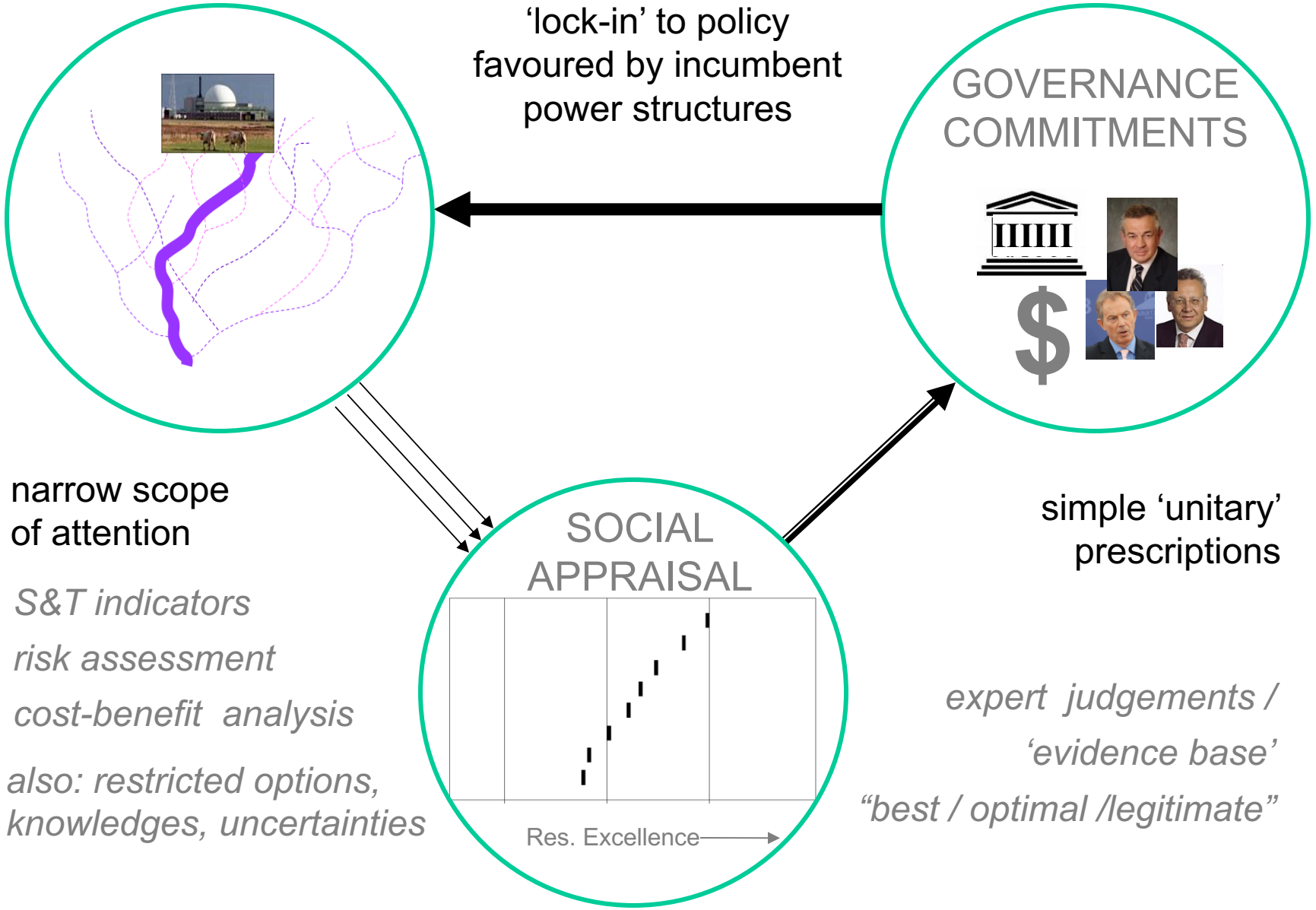




Yegros et al. (2016)

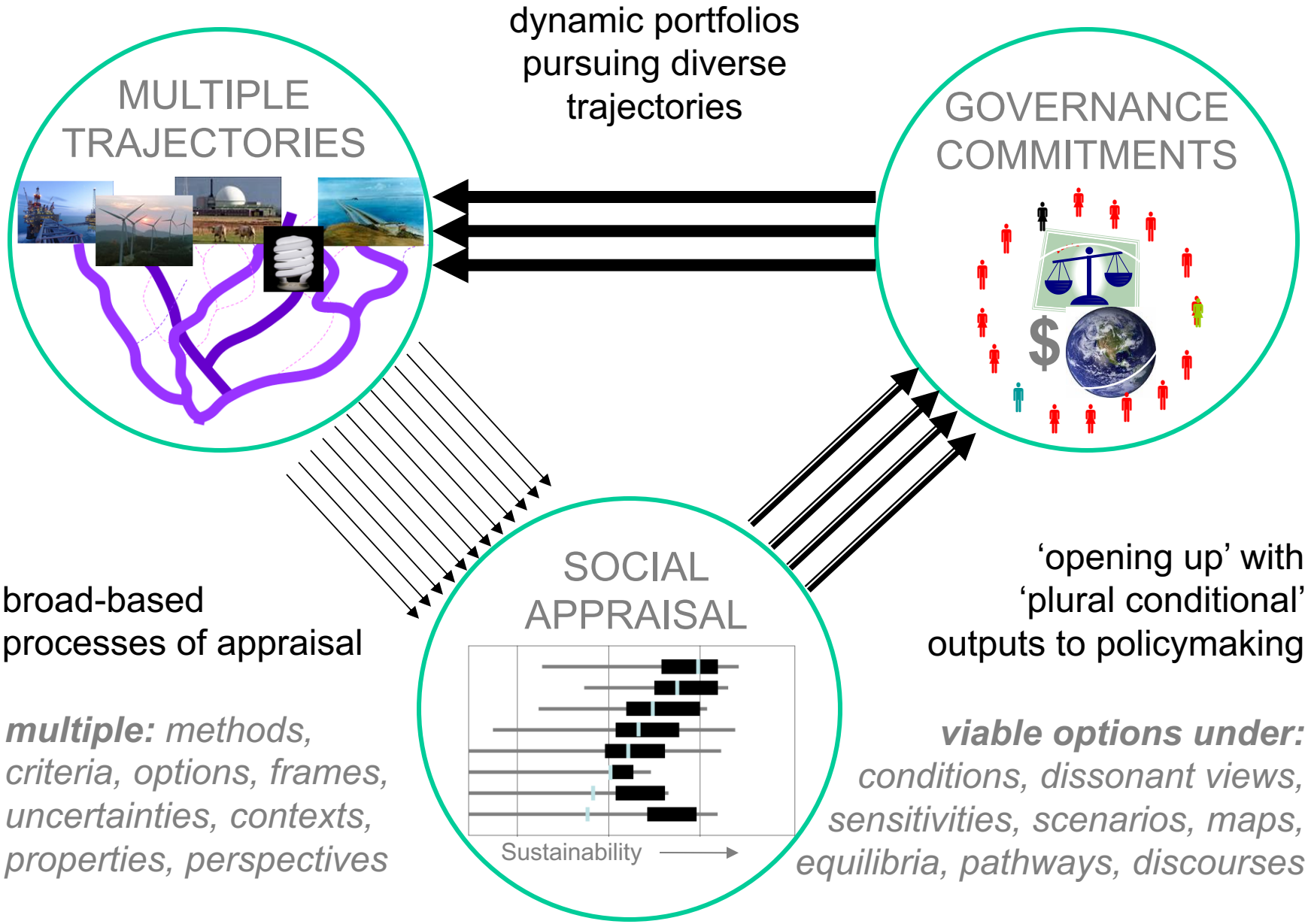
Closing down in policy dynamics

Stirling (2010)

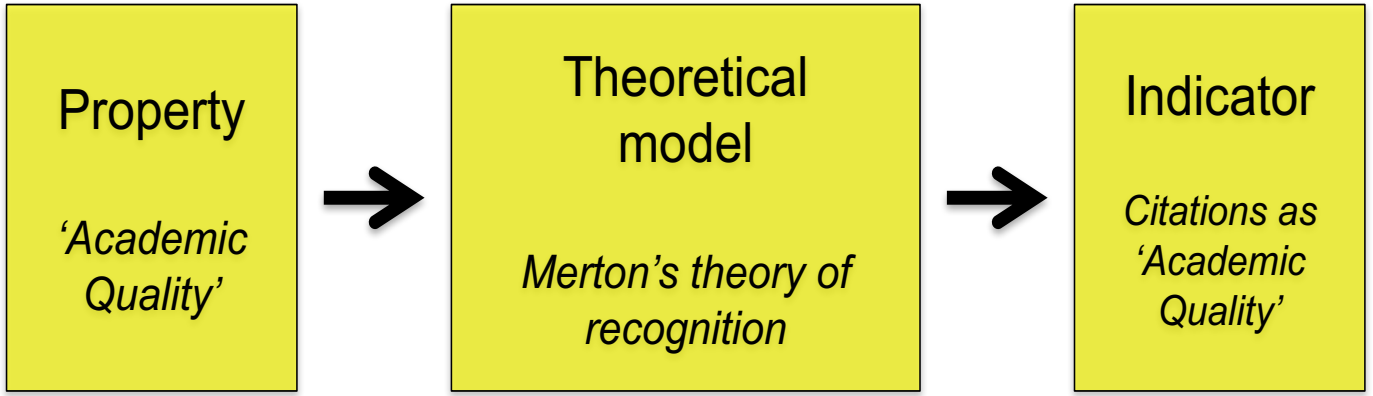


Breadth, Plurality and Diversity

Stirling (2010)



Properties, models and indicators



Sources of misrepresentation	Indicators	Model	Property
Unstable Parameters	UNSTABLE	Correct	Relevant
Incorrect Model	Stable	INCORRECT	Relevant
Irrelevant Property	Stable	Correct	IRRELEVANT

Why should WE engage with indicators in the wild?

Rationales for pluralisation and participation (Stirling, 2004)

1. **Substantive:** IndWild produce more socially robust knowledge
   More thorough scanning of knowledge. Inclusion of plural perspectives.

2. **Normative:** Under a democratic view, pluralisation is good on its own
   From a tool to project 'the perspectives' of incumbent institutions, towards becoming an 'honest broker', facilitating deliberation.

3. **Instrumental/Strategic:** IndWild provide credibility and legitimacy.
   Indicators for research impact assessment as a window of opportunity to reposition quantitative studies of science.

Big companies' services) are taking over consultancy services on indicators. In the face of a simplistic delivery of indicators of RIA… (e.g. Altmetrics) **academia can offer socially responsible research assessment.**

Re-shaping **design and use** of indicators in evaluation

- **Indicators may be harming research**
  - Current indicators are only (partially) appropriate for some types of science.
  - Biases against and potential suppression of creative and valuable types of research (agro-, health,…). Threat to diversity.
- **Not only more, but other types of indicators needed**
  - Making visible other contributions (e.g. IDR) and other types of research (e.g. action research, co-creation)
  - Enhancing visualisation of metrics for “opening up” perspectives rather than facilitating “closing down”
- **Towards different uses of indicators**
  - New embedding in assessment or policy context
  - Indicators used to pluralise (opening up) perspectives, as tools for interpretation and deliberation, not a substitute for judgement

Caveats

1. Use of the terms ‘indicators’ in a very broad sense
    - **Quantitative studies used in policy appraisal.**

2. Indicators as **tools for informing decision-making**.
    - Statements may not apply to indicators for pure research.
    - The point is precisely that we should not use lab experiments for policy
        - Clinical testing vs. experimental cell biology

3. The forms engagement proposed to pluralise have been present, **but relatively marginal** and **ad hoc**.
    - Discussion with experts (e.g. field delineation) and policy makers (on meaning of indicators) – ‘validation’ (Noyons, Laurens)
    - Use of other data sources (Leydesdorff comparing news, policy and pubs)

A research and policy agenda

- **Pluralising indicators (supply)**
  - *Broadening out:* Create more diverse indicators
    - Indicators of open science, RRI, hidden, social innovation
    - Improve representation of SSH scholarship, languages other than English, the “South”,…
  - *Opening up:* develop more pluralistic toolkits that present contrasting perspectives
    - Multi-ranking tools
    - Interactive visualisations

- **New embedding of indicators in assessment (demand)**
  - Develop new social processes on use of indicators
    - Indicators to inform decisions, not a substitute for judgement
    - STI indicators as tools for interpretation and deliberation

… towards S&T indicators as tools for **deliberation**

New forms of quantitative evidence that fosters plural reflection rather than justifying or reinforcing dominant perspectives

Indicators as tin-openers (Carter, 1989; De Rijcke et al., 2016)

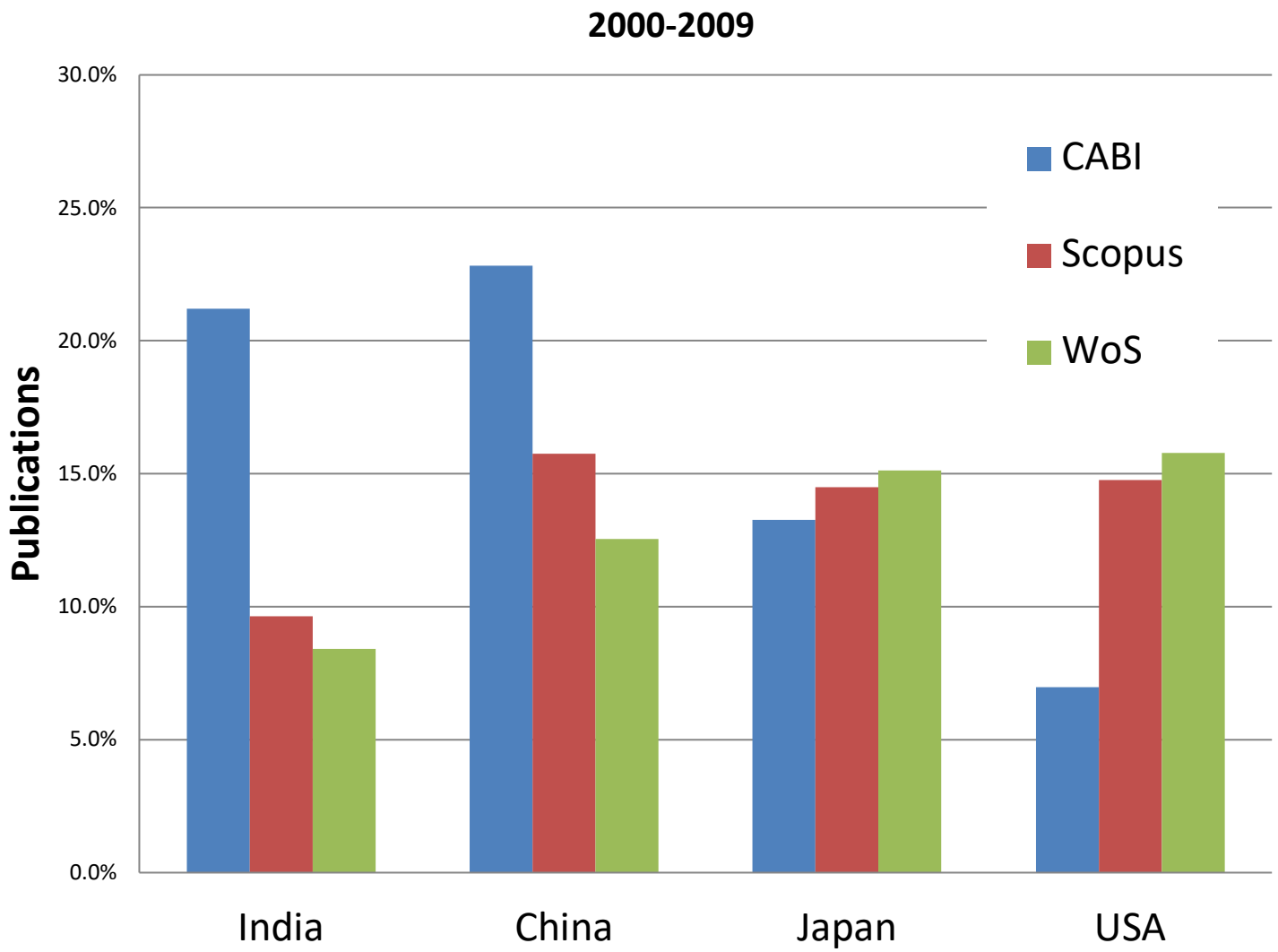
This shift is facilitated by trends pushed by ICT and visualisation tools

- Multidimensional outputs (interactive maps)
- Institutional repositories
- Multiple solutions -- highlighting variation, confidence intervals
- More inclusive and contrasting classifications (by-passing private data ownership? Pubmed, Arxiv)
- More possibilities for open scrutiny (new research groups)

2. Conceptual framework:

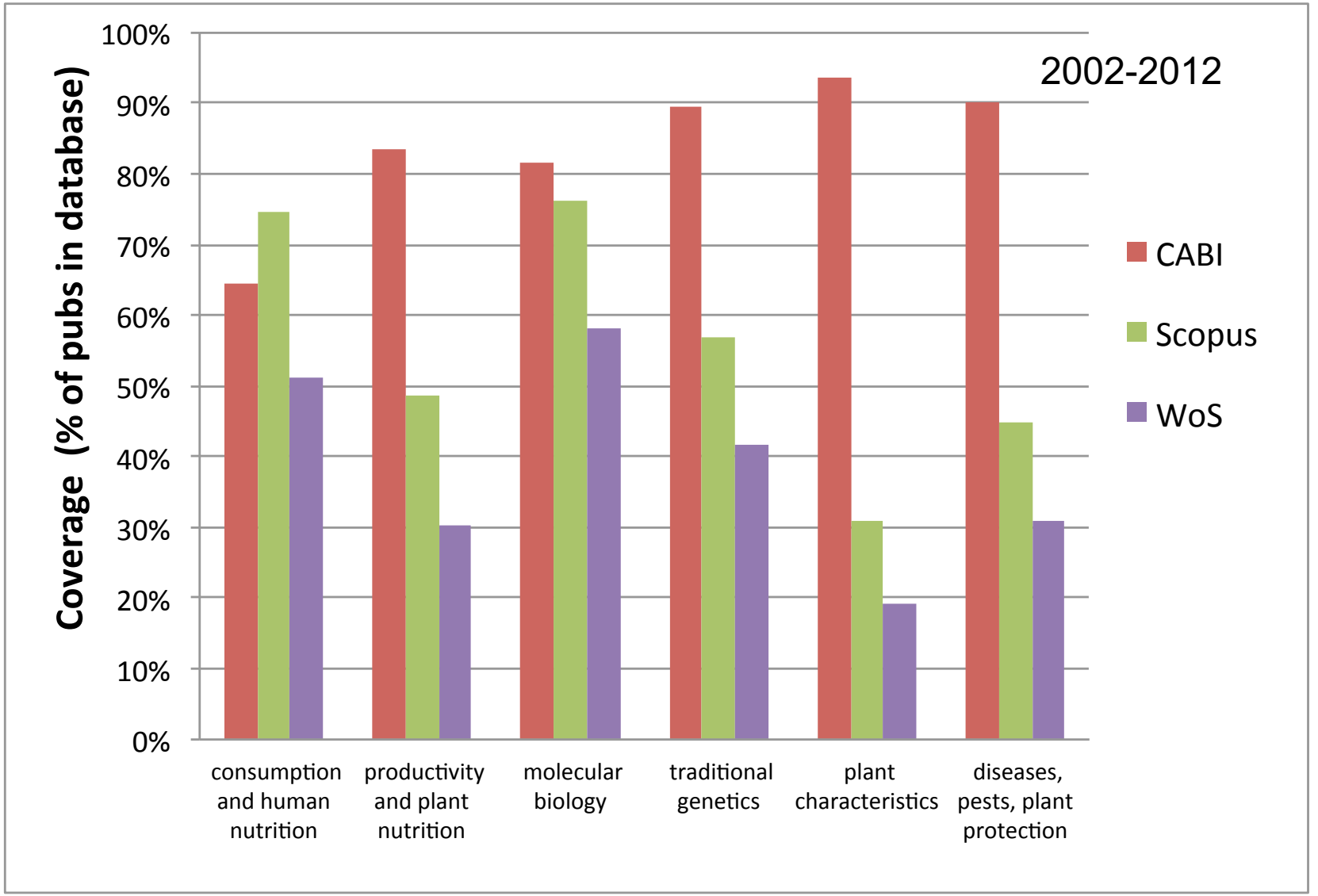
“Broadening out” vs. “Opening up” policy appraisal

Coverage: Knowledge production on rice by country



Rafols, Ciarli and Chavarro (in preparation)

Coverage: Knowledge production on rice by topic



Rafols, Ciarli and Chavarro (in preparation)

Examples of Opening Up with same data

Exploring options

- focus on directions / vectors instead of scalars
  - Maps or networks
- examine over time

Strategies for opening up indicators

- From prescriptive indicators to quantitative evidence
  - Deliberation on indicators and “indicators” for **informing** deliberation processes (Barré)
- Incorporating relevant invisible dimensions
  - Activities and outcomes so far marginalised
- Presenting contrasting perspectives
  - At least TWO, in order to allow choices
- Simultaneous visualisation of multiple dimensions / options
  - Maps, networks Allowing the user take its own perspective
- Exploration of multiple realisations of same concepts
  - Avoiding misplaced concreteness
- Interactivity for checking conditions
  - Allowing the user give its own weigh to criteria / factors
  - Allowing the user manipulate visuals

S&T indicator as a tools to open up the debate

- ‘Conventional’ use of indicators (‘Pure scientist ‘--Pielke)
  - Purely analytical character (i.e. free of normative assumptions)
  - Instruments of objectification of dominant perspectives
  - Aimed at legitimising /justifying decisions (e.g. excellence)

  → Unitary and prescriptive advice

- Opening up scientometrics (‘Honest broker’ --Pielke)
  - Aimed at locating the actors in their context and dynamics

    → Not predictive, or explanatory, but exploratory
  - Construction of indicators is based on choice of perspectives

    → Make explicit the possible choices on what matters
  - Supporting debate

    → Making science policy more ‘socially robust’

  → Plural and conditional advice

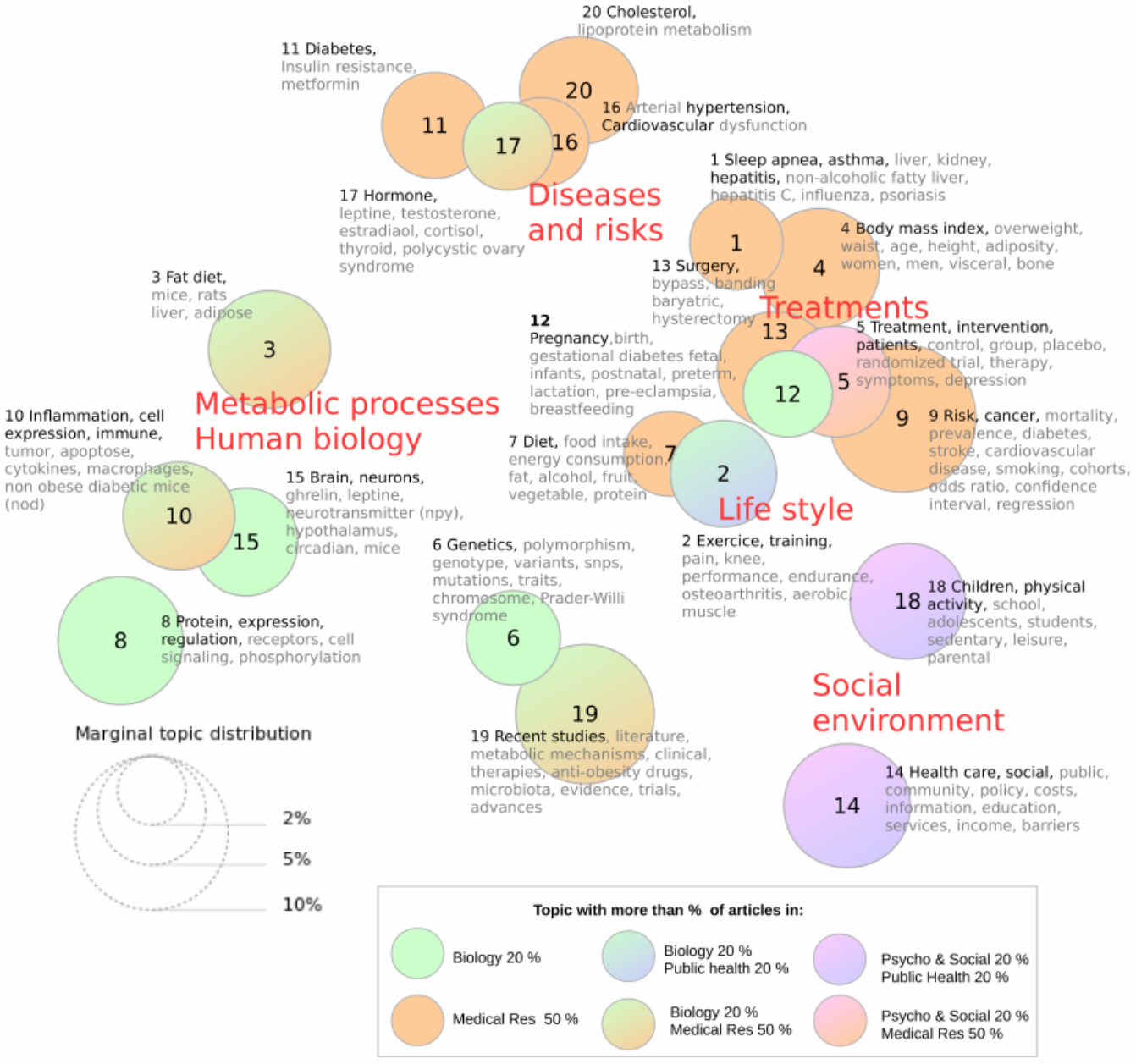
Barré (2001, 2004, 2010), Stirling (2008)

EXPLORING DEMAND:

THE CASE OF OBESITY

MAPPING BOTH SUPPLY AND DEMAND

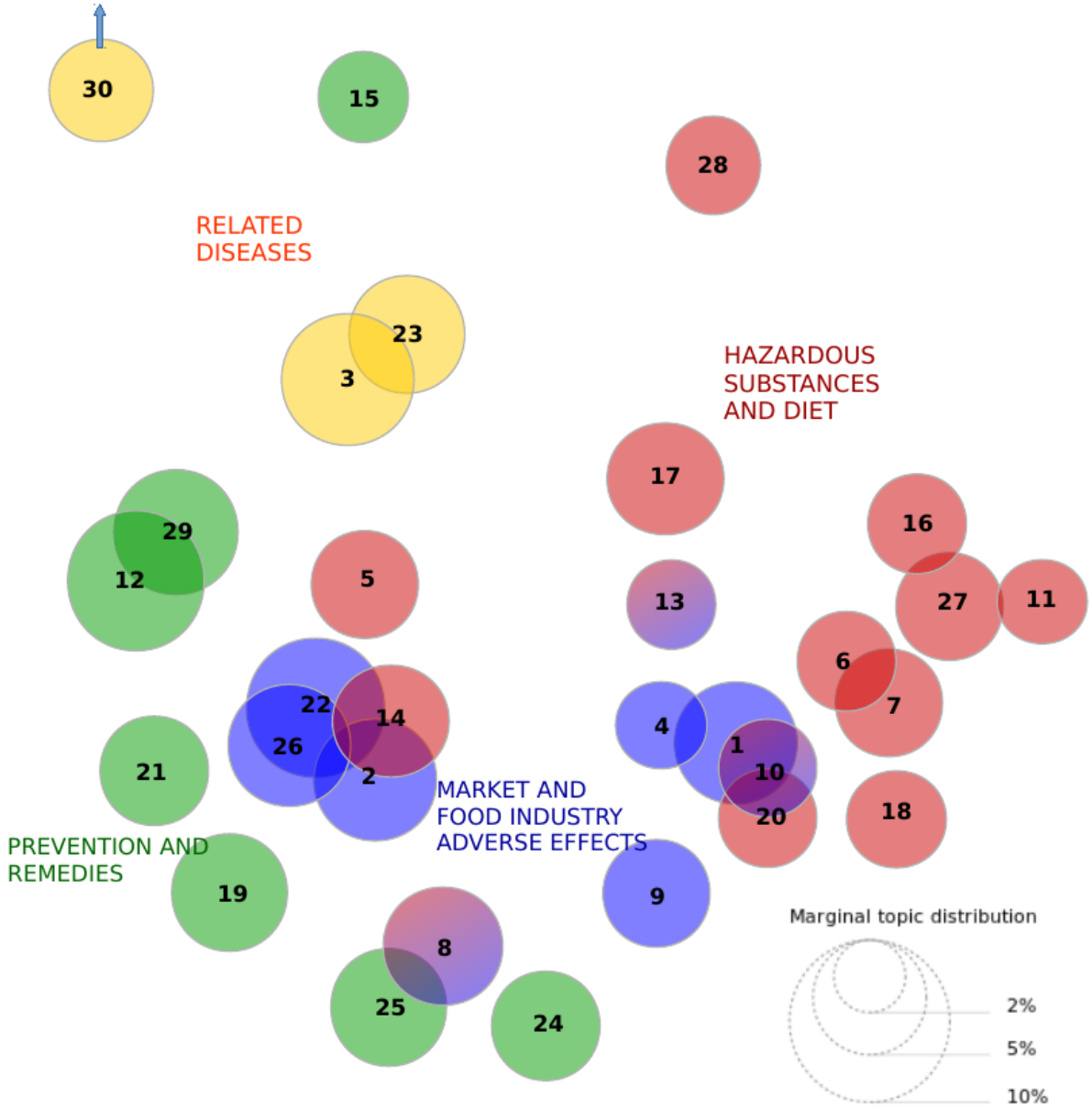
Science supply map using topic modelling



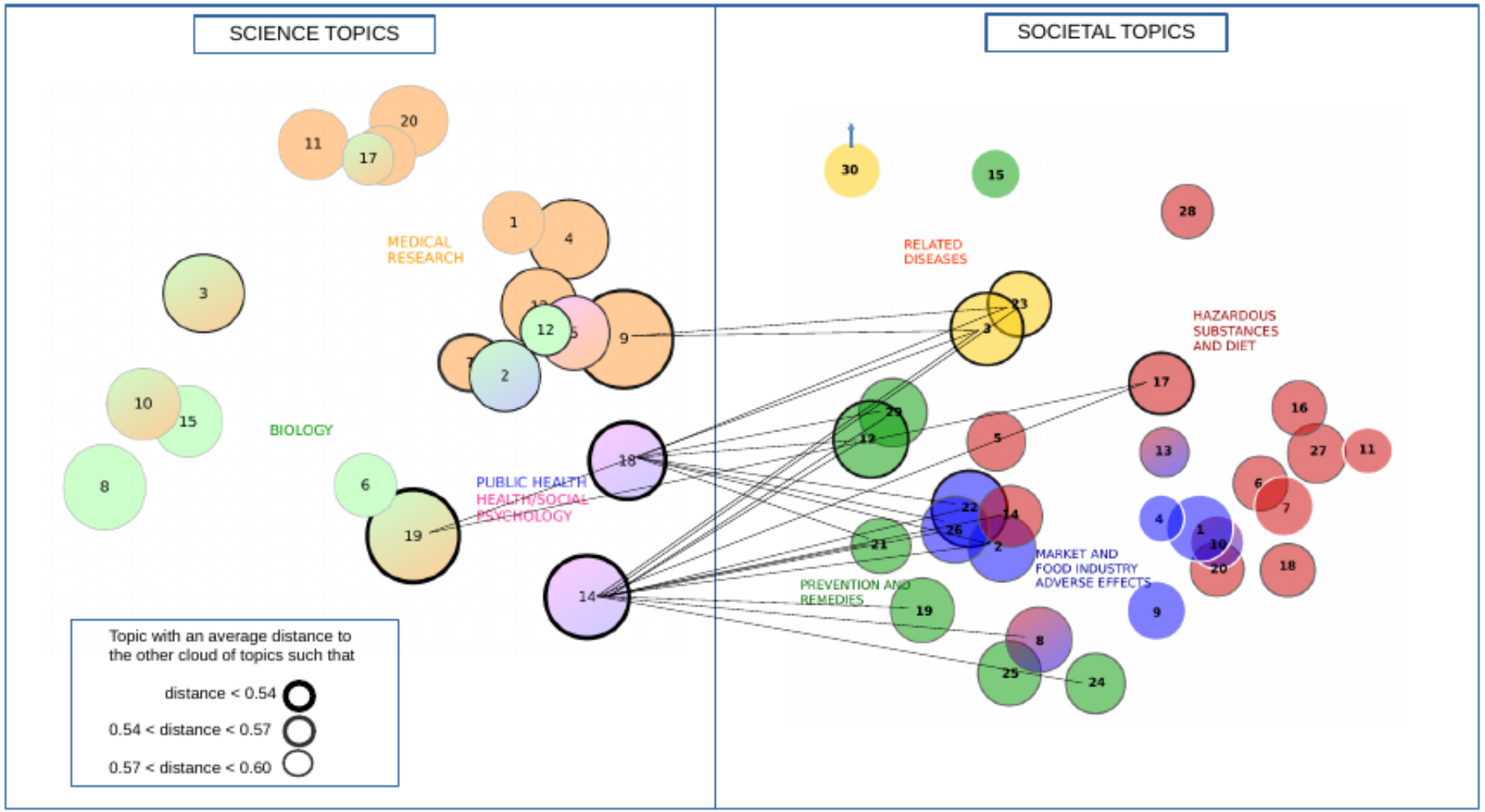
EU parliament themes (experiment)

Topic modelling of parliamentary questions (about 200 in about 2003-2013)

Not obvious how to relate policy themes with research themes



Semantic relation between supply (left) and demand (right)



Pressing demands of research management and evaluation

- Increasing size of HE and research
  - 1.5 M papers per year only in Web of Science (3x in 20 years)
  - Globalisation. China, Brazil…
- Increasing competition– globally and locally
  - Competition to attract int'l students.
  - Success rates of research calls are very low in the US, EU (10%-20%)
- Increasing societal demands
  - Interactions with industry and social actors (NGOs)
  - Grand challenges (climate change, epidemics, water & food security)
- New public management (efficiency & acceleration)

Traditional qualitative techniques of management cannot cope.

Hope that use of indicators can help...

Monitoring the Evolution and Benefits of Responsible Research and Innovation - MoRRI

Limits in measurements

Ingeborg Meijer, CWTS

Final Event – Discussion on policy aspects

Date: 7 March 2018

Location: Fondation Universitaire – rue d'Egmont 11, Brussels

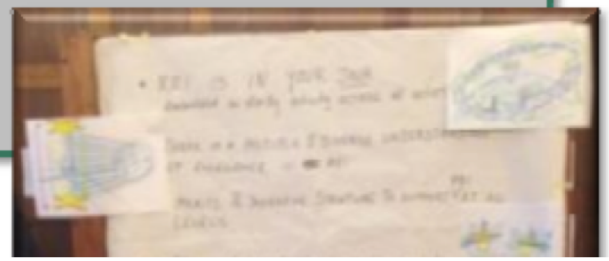
Agenda

- Vision in 2015
- Current status Monitoring
- Limitations of measurements
- The way forward/open questions...

- Interactive approach

Visioning workshop

- RRI is in your DNA embedded in daily activity across all actors
- There is a multiple and diverse understanding of excellence in research and innovation
- Merits and incentive structure to support RRI at all levels
- RRI is a creative activity or opportunity instead of a burden
- In all steps of the research process – agenda setting, evaluation, implementation – society is actively involved



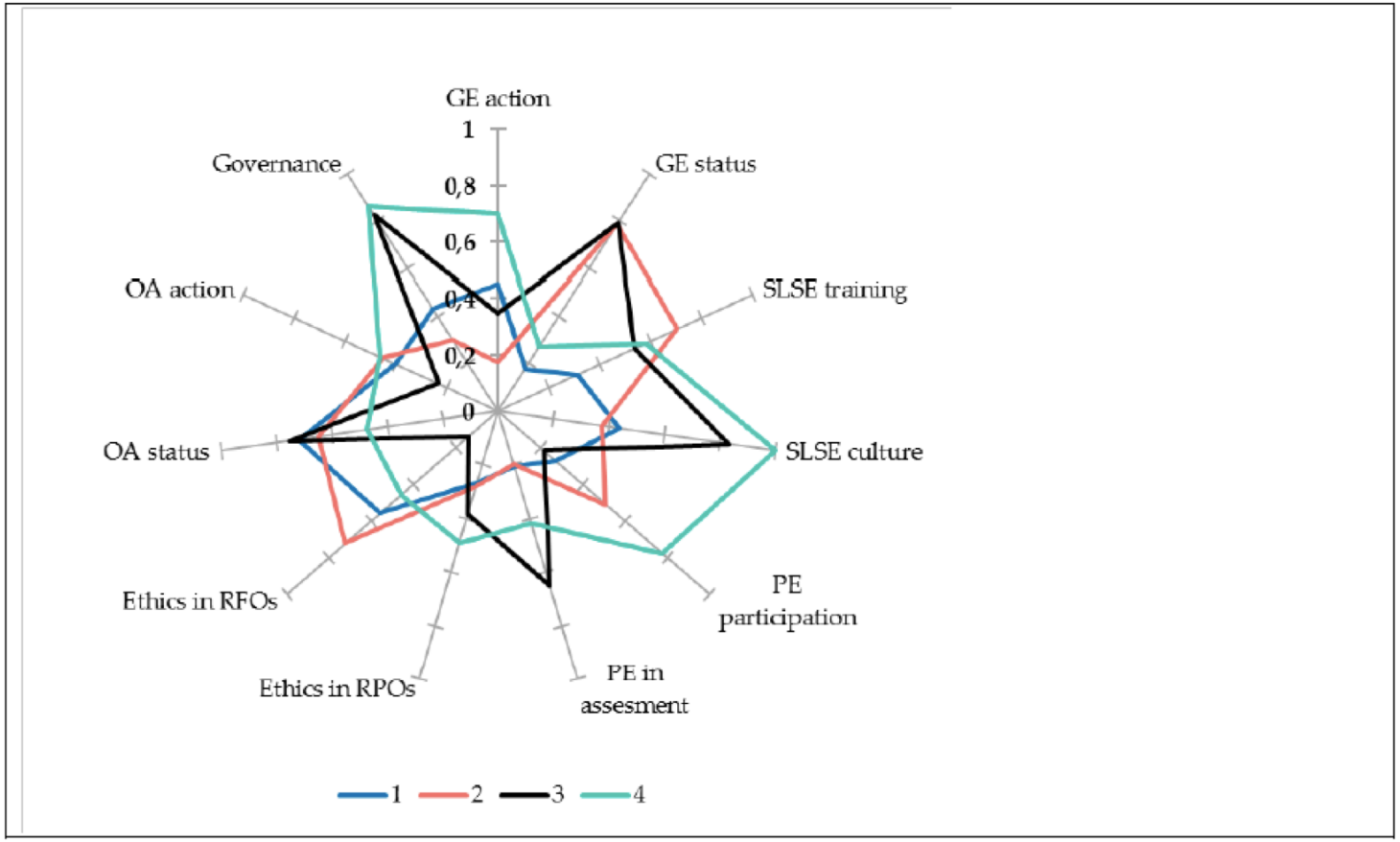
QUESTION

- When will this vision be a reality??

Monitoring current status

- 36+ Indicators, representing…
- 11 dimensions and
- 25 indicators
- 4 clusters

Figure 2 RRI characteristics of four country clusters



Calculation: Aarhus University

Mainstreaming potential?

Score board driven policy??

Indicator	BE	BG	CZ	DK	DE	EE	IE	EL	ES	FR	HR	IT	CY	LV	LT	LU	HU	MT	NL	AT	PL	PT	RO	SI	SK	FI	SE	UK
OA1.1: Share of Open Access publications	🟢	🟢	🟢	🔴	🔴	🔴	🟢	🟡	🟡	🔴	🟡	🟡	🟢	🟢	🟡	🟡	🟡	🟡	🟡	🟢	🟡	🟢	🟢	🔴	🔴	🔴	🟢	🔴
OA1.2: Citation scores for OA publications	🟡	🔴	🟡	🟢	🟢	🟡	🟡	🔴	🟡	🟢	🔴	🟡	🟢	🔴	🔴	🟡	🔴	🔴	🟢	🟢	🔴	🟡	🔴	🟡	🔴	🟢	🟢	🟢
OA4: Public perception of Open Access	🟢	🔴	🟢	🔴	🔴	🟡	🟢	🟢	🟡	🟡	🟢	🟡	🟢	🟢	🔴	🔴	🟡	🔴	🟡	🔴	🔴	🔴	🔴	🟡	🟡	🟢	🟡	🟢
OA5: Funder mandates	🟢	🟡	🟡	🟢	🟢	🟡	🟢	🟡	🟢	🟢		🟢	🟡	🟢	🟢	🟡	🟢	🟡	🟡	🟢	🟡	🟡	🟡	🟢	🟡	🟡	🟢	🟢
OA6: Research performing organisations' support structures for researchers as regards incentives and barriers for data sharing	🟢		🟢	🔴	🟡	🟢	🟡	🟡	🟡	🟡		🔴		🔴	🟢		🔴		🟢	🔴		🟡	🟡			🔴	🔴	🟢

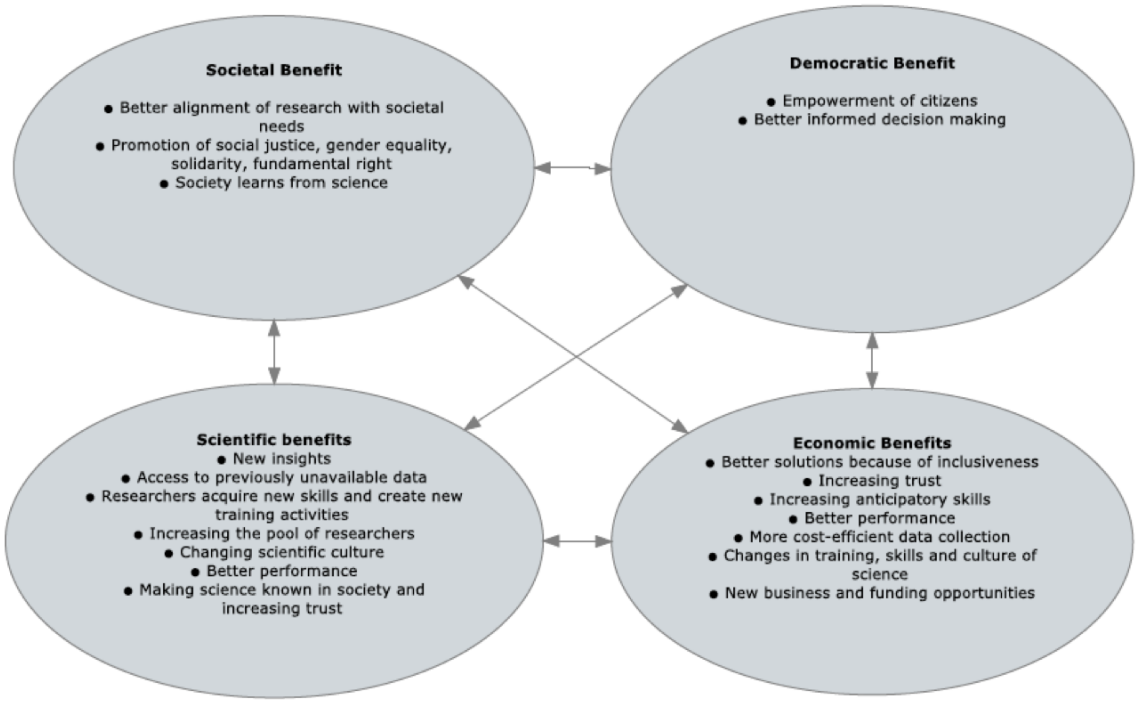
International learning while avoiding a horserace

technopolis [group] AARHUS UNIVERSITY Universiteit Leiden INSTITUT FÜR HÖHERE STUDIEN INSTITUTE FOR ADVANCED STUDIES Vienr ingenio CSIC-UPV Instituto de gestión de la innovación y del conocimiento Fraunhofer ISI US University of Sussex UCL HELLENIC REPUBLIC National and Kapodistrian University of Athens

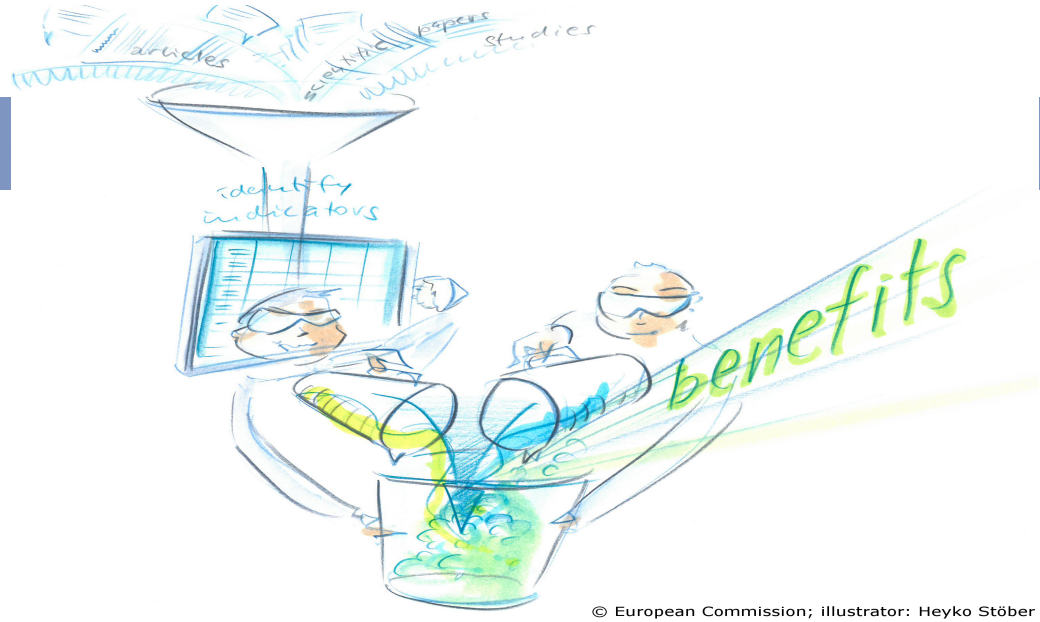
Challenges for RRI benefits

- Impact pathways to benefits
- Link between indicators & benefits

Figure 5 Societal, democratic, economic and scientific benefits of RRI



Source: MoRRI 2016, D.5.2



QUESTION

- What is the message that you get out of this monitoring activity? (also taking into account the responsible metrics and contextual effects)

Limitations of measurements

3 types of data sources

- Secondary data
  - Old data
  - No control
- Primary data - Survey based
  - Response/non-reponse
  - No of observations per category
  - Who is answering? On behalf of?
- Register data (database)
  - You measure what is there, it is quantitative
  - But may lack context

Data sources	Indicators
Science in Society actor survey	PE9, PE10
RPO-survey	GE1, GE5, GE8, GE9, SLSE2, SLSE 4, PE5, PE6, OA6, E1, GOV2, GOV3
RFO-survey	GE3, PE7, PE8, E3, GOV2, GOV3
Register data (database)	GE10, OA1, OA2, OA3
Qualitative, desk-research	SLSE1
Eurostat; She Figures; MORE2; MASIS; Eurobarometers; EPOCH; SATORI; Openaire.eu	GE2, GE4, GE6, GE7, SLSE3, PE1, PE2, PE3, PE4, E2, OA4, OA5, GOV1

QUESTION

- What kind of alternative or innovative data collection procedures would you suggest? (that is sustainable preferably)

The way forward/open questions...

- Blind spots in the keys
- Blind spots between the keys
- Scientific benefits?
- Sustainability
- Policy recommendations
- Alignment with other policies
- Reward and incentives
- FP9 reporting requirements
- Mixed methods