CWTS
Meaningful metrics

New Research Indicators and their meaning for Open Science

Paul Wouters
Centre for Science and Technology Studies (CWTS),
Leiden University

Universiteit Leiden

STM Innovations Seminar
5 December 2018 London, UK

Centre for Science and Technology Studies (CWTS)

- New programme 2017-2022
  Valuing science & scholarship
- CWTS studies scientific research and its connections to technology, innovation, and society.
- Our research, bibliometric and scientometric tools, and evaluation expertise provide a solid basis for supporting research assessment and strategic decision making and for developing science policy.

CWTS

Definition of open science

In May 2016, the Competitiveness Council adopted conclusions on 'The transition towards an Open Science system' where it acknowledges that "Open Science has the potential to increase the quality, impact and benefits of science and to accelerate advancement of knowledge by making it more reliable, more efficient and accurate, better understandable by society and responsive to societal challenges, and has the potential to enable growth and innovation through reuse of scientific results by all stakeholders at all levels of society, and ultimately contribute to growth and competitiveness of Europe".

CWTS

1. Reward and incentive systems
2. Measuring quality and impact
3. Future of scholarly publishing
4. FAIR open data
5. Open Science Cloud
6. Research integrity
7. Citizen Science
8. Open education and skills

CWTS

OSPP recommendations – general

- Appoint national coordinators and task forces for the implementation of Open Science.
- Ensure the scholarly infrastructure is highly interoperable (identifiers).
- Give credit to all participants.
- Align HR strategies to open science codes.
- Foster open science literacy.
- Raise awareness through a campaign.

CWTS

OSPP recommendations – research indicators and next generation metrics

- Do not use journal brand or IF for individual researcher assessment as proxy for quality
- Develop indicators that capture full range of contributions
- Do pilots to check validity of these indicators as part of Horizon2020
- Apply ORCID and develop CV best practices
- All metadata should be open

CWTS

Open Science Monitor (in development)

- monitor for Europe and global observatory of open science trends
- reference point for the open science community
- determine impacts of OS
- structured analysis of policy relevant trends in OS
- comprehensive
- inclusive and open for comments

CWTS

Rewards working group

- Research Performing Organisations (RPOs) should be strongly encouraged to include OS practices in the evaluation of performance and of career development
- Research Funding Organisations (RFOs) should be strongly encouraged to include OS practices in the evaluation criteria for funding proposals and as part of the assessment of the researchers.
- The Career Evaluation Matrix as central tool

CWTS

Skills working group

- Policy: making Open Science skills an integral part of the next framework programme (FP9) with dedicated funding
- Guidelines to implement Open Science
- Raise awareness of Open Science policy initiatives
- Train Researchers for Open Science
- Provide Support for Open Science
- Career development for Open Science

CWTS

MLE 4th report: common ground

- Achieve Open Access to publications
- Go beyond Open Access
- Shift to multiple indicators for research assessment
- Clarify responsibilities: decisions need to be taken
- Evaluate investments in infrastructure and resourcing
- Clarify legal framework relating to Open Science
- Coordinate with European governance and other member states
- Involve researchers and research organisations in all aspects of Open Science
- Prioritise public engagement in Open Science activities
- Enhance research outputs and quality,
- Support early-career researchers and prevent the brain drain
- Monitor the transition to Open Science

CWTS

MLE 4th report: key principles

- Respect for diversity
- Collaboration
- Accountability
- Transparency
- Social responsibility and engagement
- Fairness
- Impact

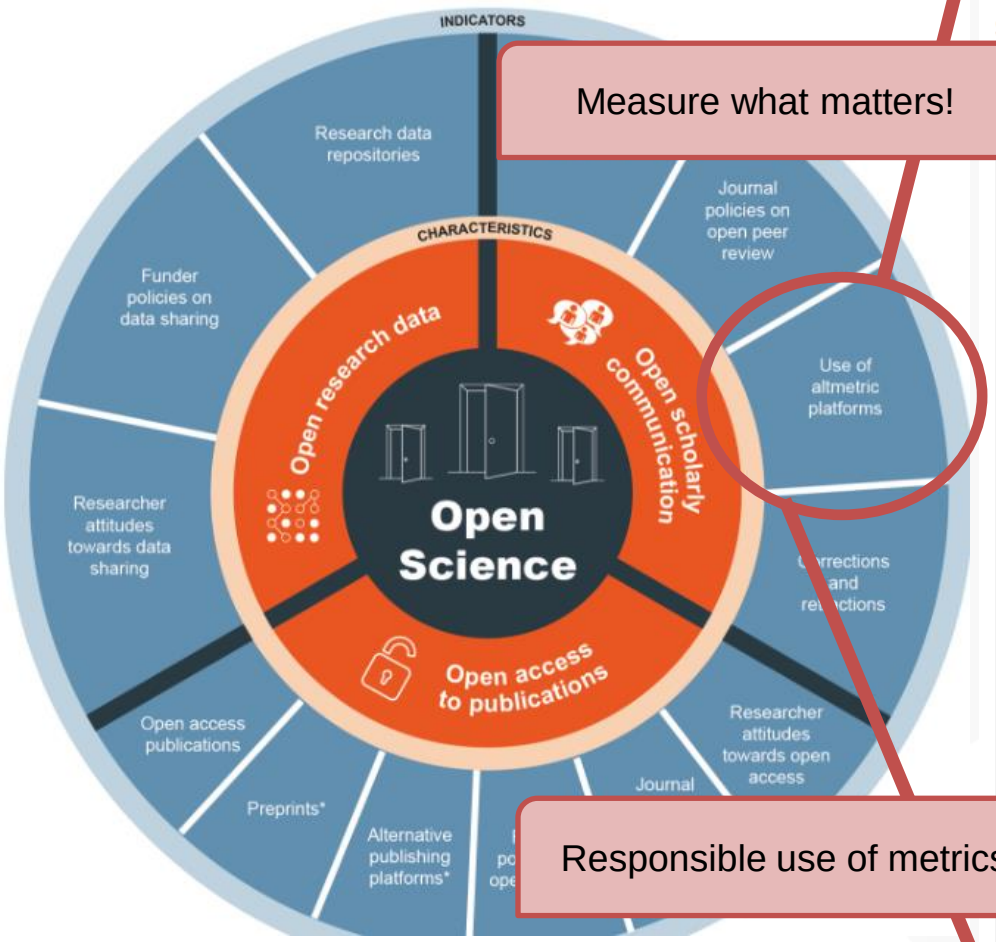
CWTS

MLE report 4: lessons learnt

- Comprehensive approach: use OS to reorganize the science system
- How to deal with the costs of OS
- Role of commercial publishers disputed
- Valorising non-English languages
- Monitoring the transition
- Consistent EU policies:
  - OS as key in FP9
  - IP and OS
  - Infrastructures (EOSC)

CWTS

Handshake needed!



Measure what matters!

Responsible use of metrics

Explore the indicators related to open scholarly communication

Select an indicator to see its description, visualise the data, understand its limitations, identify the data sources.

Open peer reviews

- Percentage of peer reviews that are published
- Percentage of publications in PeerJ that use open peer review

Journal policies on open peer review

- Journal policies on open peer review

Use of altmetric platforms

- Number of mentions of publications in media and social media

Corrections and retractions

- Corrections and retractions recorded in Web of Science

Preprints

- Number of preprints

Alternative publishing platforms

- Articles published before peer review

https://ec.europa.eu/research/openscience/index.cfm?pg=home&section=monitor

Next Generation Metrics

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.

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

23 APRIL 2015 | VOL 520 | NATURE | 429

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) (2015) *Nature* 520: 429–31. doi:10.1038/520429a

CWTS

Initiative for Open Citations (I4OC)

Key findings

- No perfect metrics: neither alternative, nor traditional
- Responsible use of metrics is key
- Open science requires open metrics

Selected recommendations

Recommendations	Short Term Goals	Long Term Goal
Ground an open science system in a mix of expert judgement, quantitative, and qualitative measures	Provide guidelines for responsible metrics in support of open science	**Fostering open science**
Make better use of existing metrics for open science	Assess suitability of indicators, encourage development of new indicators	

CWTS

Selected recommendations

Recommendations	Short Term Goals	Long Term Goal
Open, transparent and linked data infrastructure for metrics in open science	Use open metrics and reward adoption of open science principles and practices	**Removing barriers to open science**
Measure what matters	Highlight how inappropriate use of indicators can impede open science	

CWTS

Evaluation for fostering researchers' engagement with Open Science

EC Expert Group on Indicators for Researchers' Engagement with OS
(P. Wouters (chair), B. Holbrook, M. Jacob, Lynn Kamerlin, A. Oancea, I. Ràfols)

New indicators and evaluation frameworks so as to:

- overcome existing barriers posed by evaluation to OS;
- reform the academic reward system to incentivize OS;
- and thus enhance capabilities to facilitate OS;

First notice: the solution IS NOT replacing old indicators by new indicators

Attention:

- Indicators for fostering OS are NOT necessarily indicators of OS
- Evaluation and indicators that facilitate OS (but OS is diverse)

- Different from EC Open Science Monitor

  → looks at progress (not evaluation) in specific activities related to OS

Indicators for engagement with Open Science

What is Open Science?

- "a new approach to the scientific process based **on cooperative work and new ways of diffusing knowledge by using digital technologies and new collaborative tools**. The idea captures a systemic change to (...) towards **sharing and using all available knowledge at an earlier stage in the research process**." (p. 33)

A large diversity of practices such as

- Open access
- Open data
- Open source development
- Collaboration with stakeholders (co-creation)
- Citizen science (e.g. crowd-sourcing)
- Science communication
- (...)

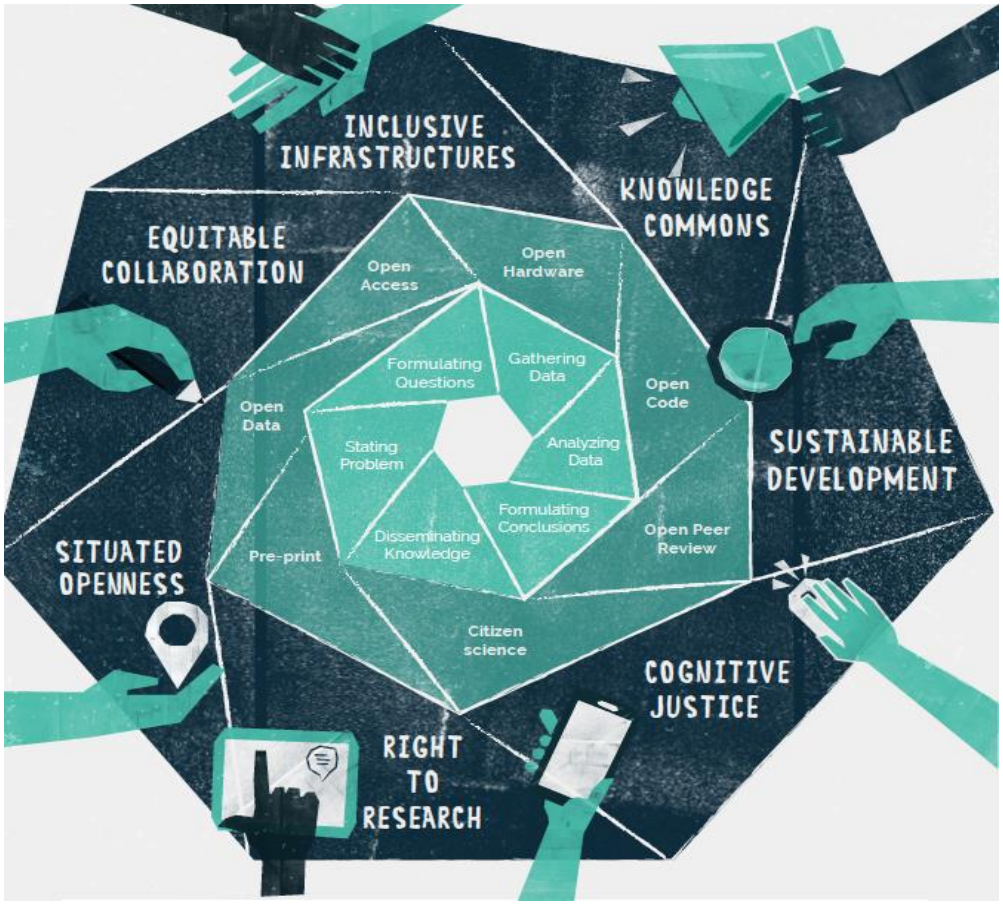
Manifesto of OCSDNet (Open and Collaborative Science in Development Network)

For all the claimed benefits of OS...

... current model is NOT making science a more **inclusive** practice.
... many scientists continue to be **underrepresented and excluded**
... new technologies exclude those with **limited digital rights** .
... citizens rarely get to **shape the research agenda.**

Principles

- knowledge commons
- cognitive justice
- situated openness
- right to research
- equitable collaboration
- inclusive infrastructures
- use knowledge as a pathway to **sustainable development**



What do we understand by Open Science?

OS as Open Knowledge Practices

The focus is not about the form of the **outputs** (e.g. related to accessibility)
The focus is on **processes of knowledge exchange** (the knowledge flows)

Processes of Open Science are based on:

- **CommunicationS and interactionS** within and between researchers and stakeholders. Often not linear but **interactive** and **iterative.**

- **The "qualitieS" of communicationS and interactionS**, not about the quantity.
  - e.g. FAIR -- Findable, Accessible, Interoperable, Re-usable.
  (meta-data is crucial in data sharing)

- **Contexts of use are key and diverse**

Assess the processes, rather than counting outputs.
Just 'counting' can be analytically wrong, unfair and harmful in evaluation

Evaluation informed by indicator frameworks

Dimensions to consider in evaluation

- *Goal of evaluation*
- *Research Mission*
- *Level of assessment*
- *Epistemic cultures and research approaches*
- *Potential stakeholders, audiences and beneficiaries*
- *Research environment:*

For a given configuration of relevant dimensions, certain indicators will be relevant while others will not.

Between FULLY TAYLORED - .....- UNIVERSAL

Prêt à porter indicators

What properties are relevant in assessment for Open Science?

Conventional evaluation:

Objectives → **Outputs** → Outcomes (too late) → Impacts (too late)

In practice: Assess **Outputs** and assume that outcomes/impacts will follow.

Assessement for engagement with OS:

-- More diverse types of Outputs (database, research materials, etc.)

-- **but valuable when** knowledge is shared

Need to **assess the processes of knowledge exchange**

Outputs → Outcomes

Process of sharing

What matters is
NOT # of outputs

But QualitieS of
interactions

Including capabilities in assessment

According to Sabina Leonelli (2017, MLE report, p. 14)
-- **OS tends to focus on high-resource**, well-recognised research environments.

Example of Open Data repositories

mainly display output from English-speaking labs in prominent research institutions, which have:

- funds to curate contents
- participate in the development of expensive equipment and software;
- visibility to determine dissemination formats/procedures;
- resources and confidence to build on data donated by others.

- Indicators of Open Science activity may **mainly** reflect resources availability rather than effort towards OS

Therefore:
- **Capabilities in research environment should be taken into account**

This applies :
- Global South, European ‘periphery’, relatively marginalised groups (gender?)

Two models of indicator frameworks

Accountability Model:
(For accountability, aggregate level, mainstream countries/sciences, high resource)

- Suite of many (20-50) relevant indicators:
  - Selection of indicators of counts
    - # citations from pats
    - # meetings with stakeholders
    - # news mentions
    - # ...

However: Data is unreliable. More counts does not mean more OS.
→ Need to contextualise and interpret the landscape of quant indicators

Learning Model (more appropriate given the goal of this report)

- Suite of relevant dimensions of **knowledge exchange and capabilities**
  - Self-reported or by expert judgement illustrated by relevant evidence (e.g. by metrics

    Examples:
    - Engagement with stakeholders (0-5 scale). Evidenced by # meetings per year
    - Public comms capability (0-5). Evidence. Trained or experienced employee
    - Dissemination to specialised audience (0-5). Evidence by # blogs in professional society website.

Possibly more reliable but less policy purchase!!

Summary of the argument

1. Why is EVALUATION such an issue in OPEN SCIENCE?
   - Incentives and rewards structures seen as barriers preventing researchers‘ engagement with OS.
     - Elitist peer reviews & indicators (jour. hierarchy & JIFs) closing OS activities

2. Demands and expectations for new indicators to improve evaluation
   - Altmetrics (e.g. twitter, news, policy mentions), Open Acces, Open Data stats
   - Promises of universal indicators break down:
     - **Research is diverse** -- it cannot be discribed with general indicators (this may lead to goal displacement, task reduction and gaming)
     - Counting outputs **does not** necessarily reflect **qualitieS**

3. **Indicator frameworks:** towards plural and conditional assessment
   - Evaluation processes should depend on and take into account
     - **Missions, eval. goals, assessment levels, espistemic cultures, stakeholders and environments of research**
     - **Focus on processes of knowledge exchange and capabilities with qualitative indicators**

•

Thank you for your attention

CWTS