

Why Open Science in the SKA era?



Credits: UNESCO

Lourdes Verdes-Montenegro
Susana Sánchez, Julián Garrido
IAA-CSIC

Swiss SKA Days – 6-8th Sept 2023

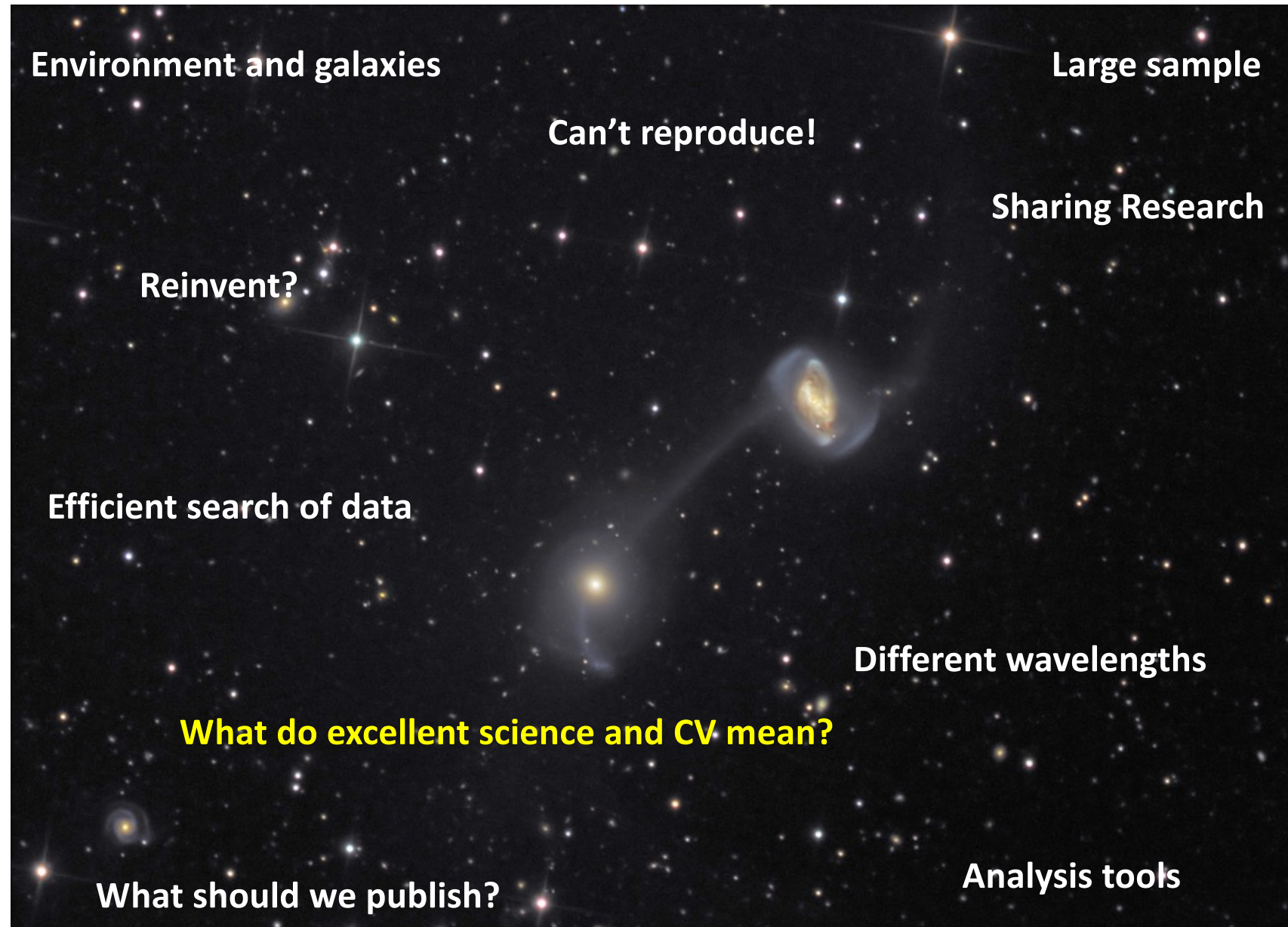


Zürich (1969)

Credits: Carlos Verdes-Montenegro



Motivation



NGC 5216: Keenan's System by Winder/Hager

Motivation

ERC Scientific Seminar Series

Prof. Lourdes Verdes-Montenegro

Instituto Astrofísica Andalucía, Granada, Spain

ERC Panel Chair



**Love for science
or 'academic prostitution'?**

12 April 2013
11.00 to 12.30

European Research Council
Established by the European Commission

Motivation

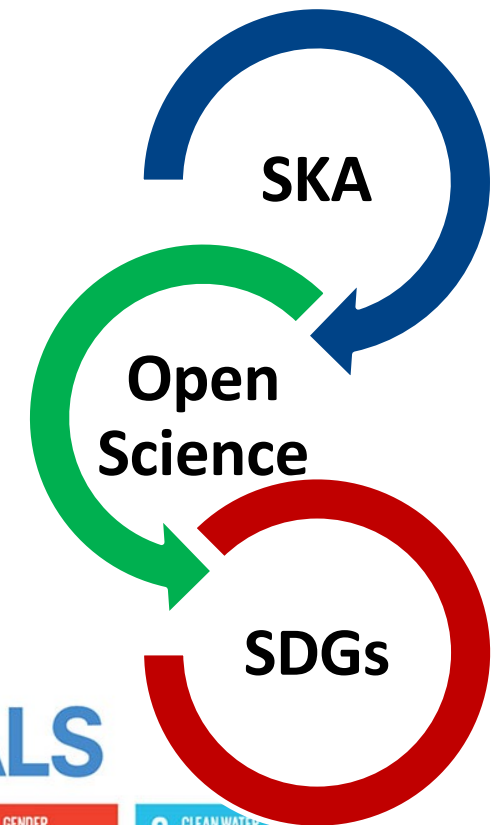
SKA **Science Digital @ UNGA 75**
SQUARE KILOMETRE ARRAY

The SKAO: A global Research Infrastructure for the 21st Century and beyond

Open Science for sustainability and inclusiveness:
the SKA role model

Lourdes Verdes-Montenegro, Susana Sánchez
IAA Severo Ochoa Centre of Excellence (CSIC)

Tuesday 29th September 2020



Credits: UNESCO

Advertisement



2 OPEN POSITIONS

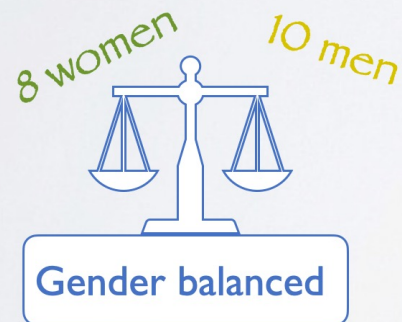
IAA-CSIC, GRANADA, SPAIN

Center of Excellence Award

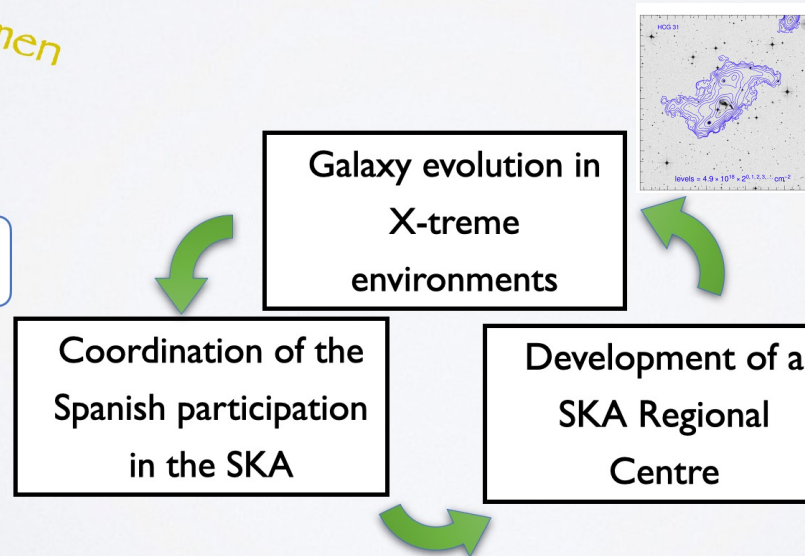


Candidates: (a) early/mid career postdoc & (b) advanced postdoc

- Expertise in detailed studies of galaxies with **HI interferometry (precursors/pathfinders)**, environment, kinematics



AMIGA team: 18 members



Contact: lourdes@iaa.es

Outline

- Open Science, a new concept?
- Is “Big Data science” possible without Open Science?
- Revised research assessments
- Impact
- Conclusions

Open Science, a new concept?

Open Science: a **new** concept?

- Too many **adjectives** for science:
excellent, high quality, trustable, ... Open

Open Science: a **new** concept?

- Too many **adjectives** for science:
excellent, high quality, trustable, ... Open
 - **Let's go back 1000 years in time...**

Scientific Reproducibility is a fundamental principle of the Scientific Method, a process pioneered by Ibn al-Haytham. In the XIth century, he proposed that a hypothesis must be supported by experiments based on **confirmable procedures** or mathematical evidence. Made special emphasis on reproducibility of results.

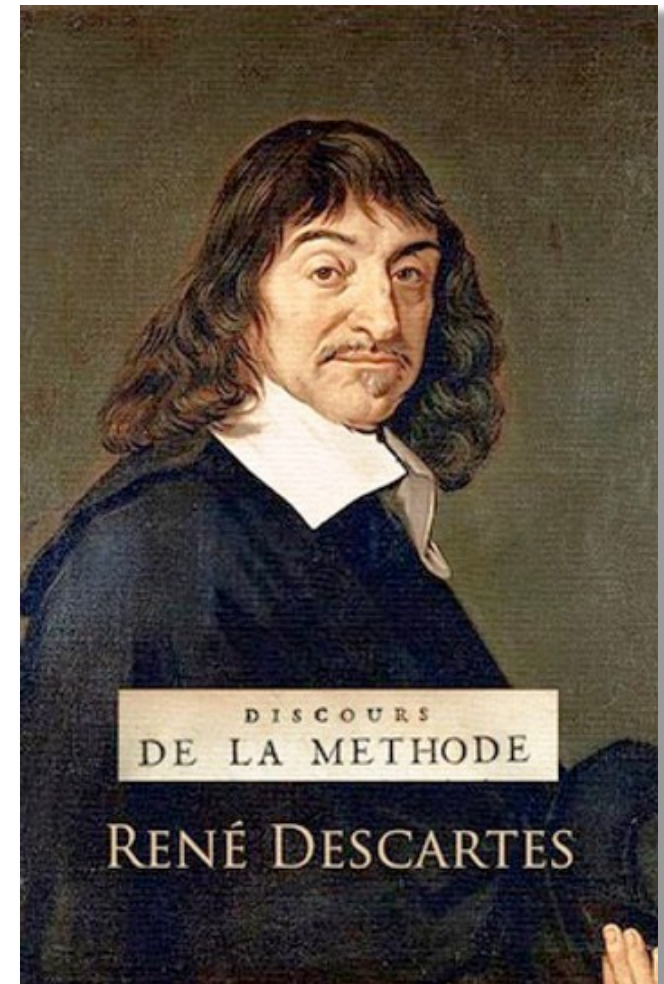


Ibn al-Haytham (965 – 1040)

Open Science: a **new** concept?

- Too many **adjectives** for science:
excellent, high quality, trustable, ... Open
 - Or let's go back 383 years in time...

Descartes reminded us in the 17th century that **Scientific Reproducibility** is a fundamental principle of the **Scientific Method**, and laid the foundations for the Philosophy of Science



- Science = Scientific Method = Reproducible = Open!

Open Science: a **new** concept?

- Too many **adjectives** for science:
excellent, high quality, trustable, ... Open
- Or let's go back 383 years in time...

Descartes reminded us in the 17th century that **Scientific Reproducibility** is a fundamental principle of the **Scientific Method**, and laid the foundations of the **Philosophy of Science**.

The concept is not new. The tools to implement Open Science practices are quickly moving forward



- Science = Scientific Method = Reproducible = Open!

Open Science: but then we already follow it, right?

- **We are scientists! We (want to) follow the Scientific Method!**

Open Science: but then we already follow it, right?

- **We are scientists! We (want to) follow the Scientific Method!**



Questionnaire on reproducibility (1500 scientists)

- 70% of researchers have tried and failed to reproduce another scientist's experiments
- **> 50% have failed to reproduce their own ones!**
 - Chemistry: 90% (**60%**)
 - Biology: 80% (**60%**)
 - Physics and engineering: 70% (**50%**)
 - Medicine: 70% (60%)
 - Earth and environmental science: 60% (**40%**)

Open Science: but then we already follow it, right?

- **We are scientists! We (want to) follow the Scientific Method!**



Questionnaire on reproducibility (1500 scientists)

- 70% of researchers have tried and failed to reproduce another scientist's experiments
- **> 50% have failed to reproduce their own ones!**
 - Chemistry: 90% (**60%**)
 - Biology: 80% (**60%**)
 - Physics and engineering: 70% (**50%**)
 - Medicine: 70% (60%)
 - Earth and environmental science: 60% (**40%**)

Open Science: but then we already follow it, right?

- **We are scientists! We (want to) follow the Scientific Method!**



Questionnaire on reproducibility (1500 scientists)

- 70% of researchers have tried and failed to reproduce another scientist's experiments
- **> 50% have failed to reproduce their own ones!**
 - Chemistry: 90% (60%)
 - Biology: 80% (60%)
 - Physics and engineering: 70% (50%)
 - Medicine: 70% (60%)
 - Earth and environmental science: 60% (40%)



Overly Honest Method
@OverlyHonestly

Maybe with this?



You can download our code from the URL supplied. Good luck downloading the only postdoc that can get it to run, though [#OverlyHonestMethods](#)

Open Science: then what happened since 1637?

- **Moving beyond the PDF**

40% Knowledge Burying in paper publication =

Rest In Paper

(S. Bechhofer 2011, Research Objects: Towards Exchange and Reuse of Digital Knowledge)



<http://www.clipartkid.com/rip-cliparts/>



Open Science: then what happened since 1637?

- **Moving beyond the PDF**

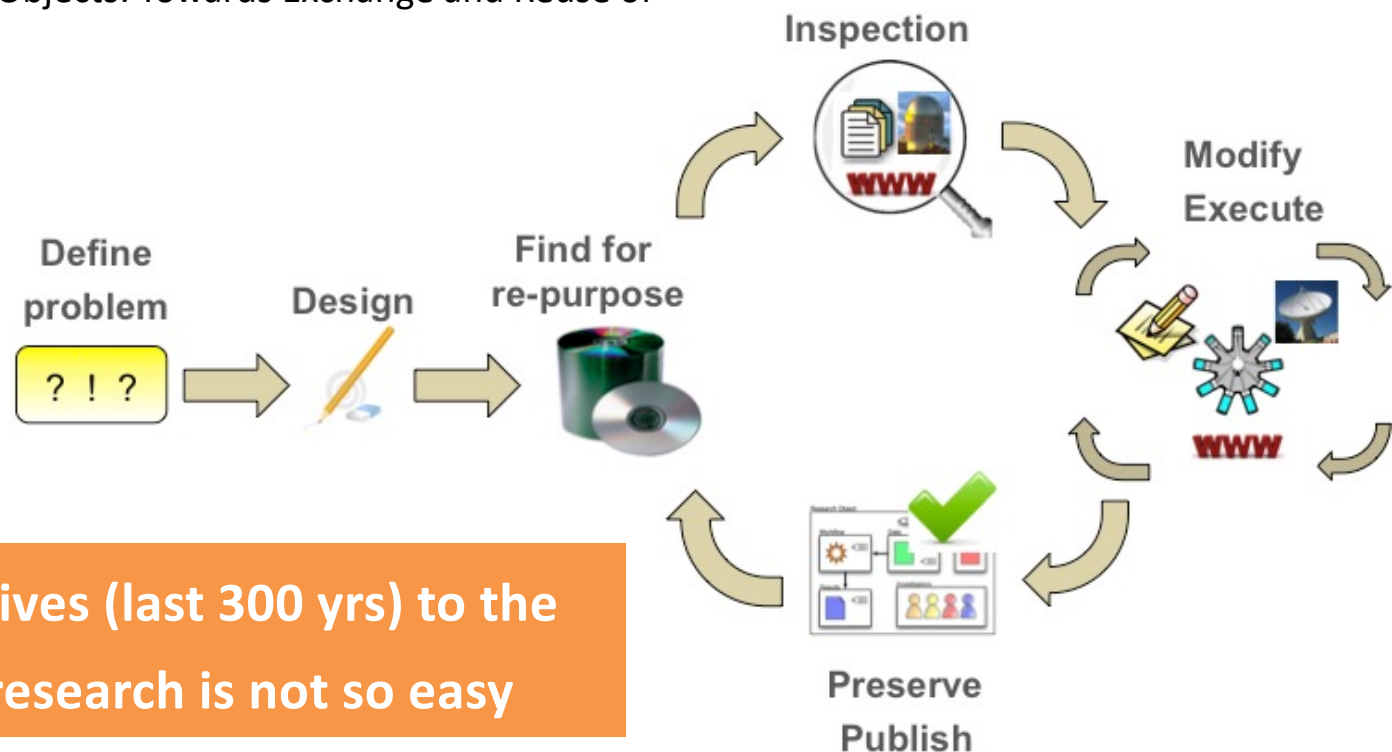
40% Knowledge Burying in paper publication =

Rest In Paper



(S. Bechhofer 2011, Research Objects: Towards Exchange and Reuse of Digital Knowledge)

In practice



Moving from narratives (last 300 yrs) to the actual output of research is not so easy

...indeed is not so easy

Big Data
preservation
& transfer

Primary (raw) data can not be
accessed in an automatic way

Standardized
catalogues

Processed data and images are only
publicly available in the paper PDF

Findable
code
Repositories

There are some scripts for processing
the data on a server somewhere, but
no one remembers where

Software
environment
preservation

The code is in a public repository, but
good luck trying to install/execute it.

...indeed is not so easy

Big Data
preservation
& transfer

Primary (raw) data can not be
accessed in an automatic way

Standardized
catalogues

Processed data and images are only
publicly available in the paper PDF

Findable
code
Repositories

There are some scripts for processing
the data on a server somewhere, but
no one remembers where

Software
environment
preservation

The code is in a public repository, but
good luck trying to install/execute it.

FAIR:

Findable

Accesible

Interoperable

Reusable

- FAIR (www.go-fair.org) is a multi-disciplinary bottom-up initiative to make scientific data reusable.

...indeed is not so easy

Big Data
preservation
& transfer

Primary (raw) data can not be
accessed in an automatic way

Standardized
catalogues

Processed data and images are only
publicly available in the paper PDF

Findable
code
Repositories

There are some scripts for processing
the data on a server somewhere, but
no one remembers where

Software
environment
preservation

The code is in a public repository, but
good luck trying to install/execute it.

FAIR:

Findable

Accesible

Interoperable

Reusable

- Effort is not always **rewarded**
- **Requires new advanced tools** to support scientists to fulfill FAIR

Is “Big Data science” possible without Open Science?

The Challenge: extraction of Scientific Knowledge

Huge and complex data volumes
Large teams distributed globally

A shared challenge for data-intensive research

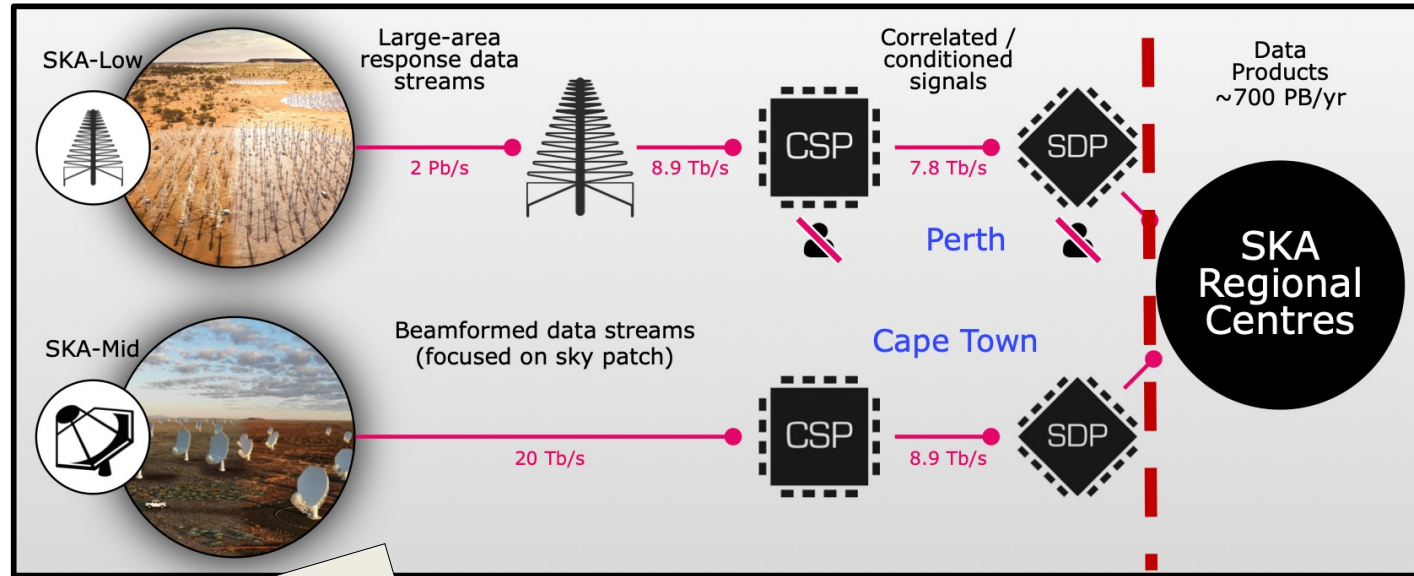
Computing / storage / network / human resources will be needed:

**Open Science
& e-Science**

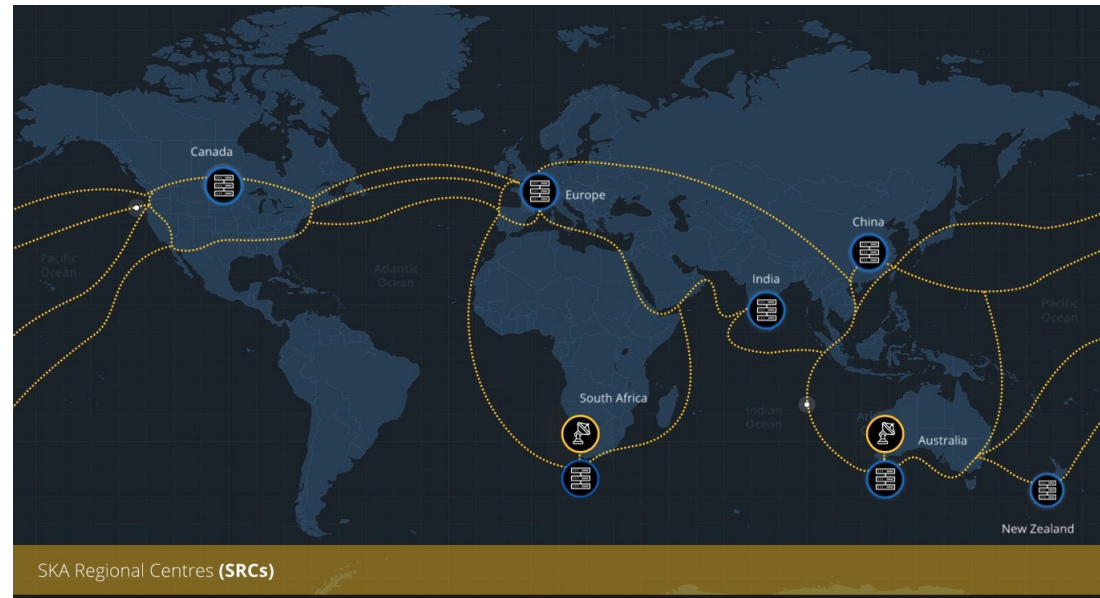
- Efficient exploitation of Distributed Computing Infrastructures
- Large international alliances of scientists
 - Tools to enhance scientific collaboration
 - Platforms to share data, methods and knowledge

Open Science is the Aim and also the Mean

The Square Kilometre Array “case”



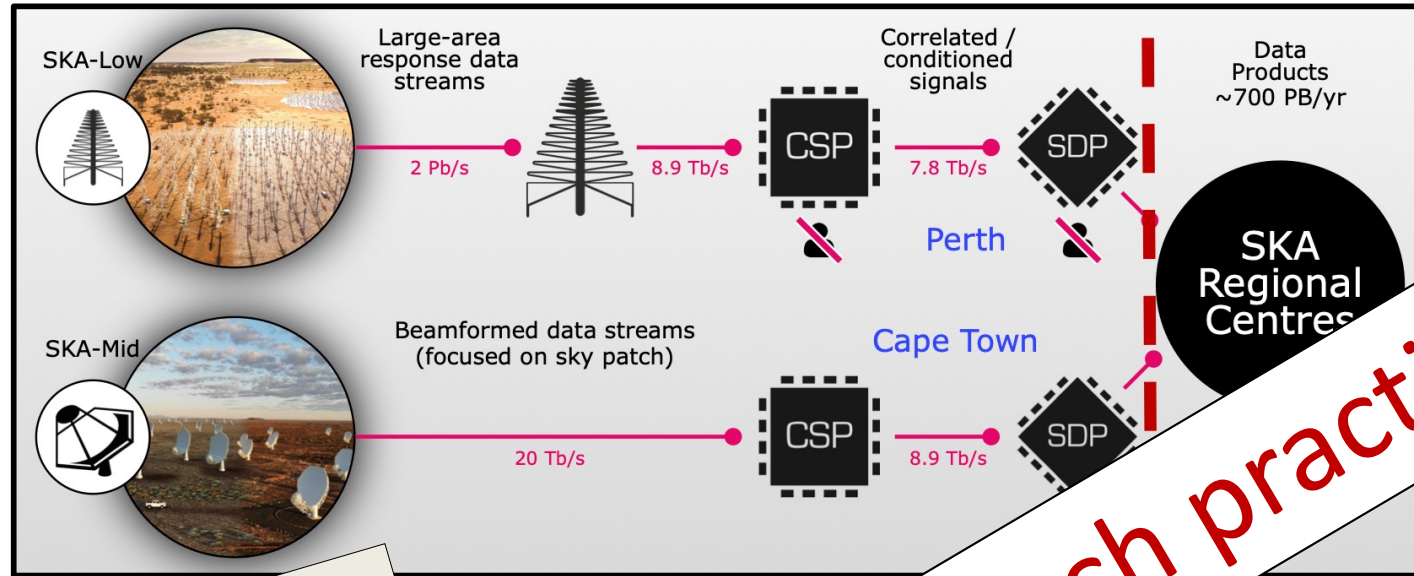
Mathieu's talk yesterday



Credits: AENEAS project

The SKA Regional Centres, the core of the SKA Science

The Square Kilometre Array “case”



Mathieu's talk yesterday

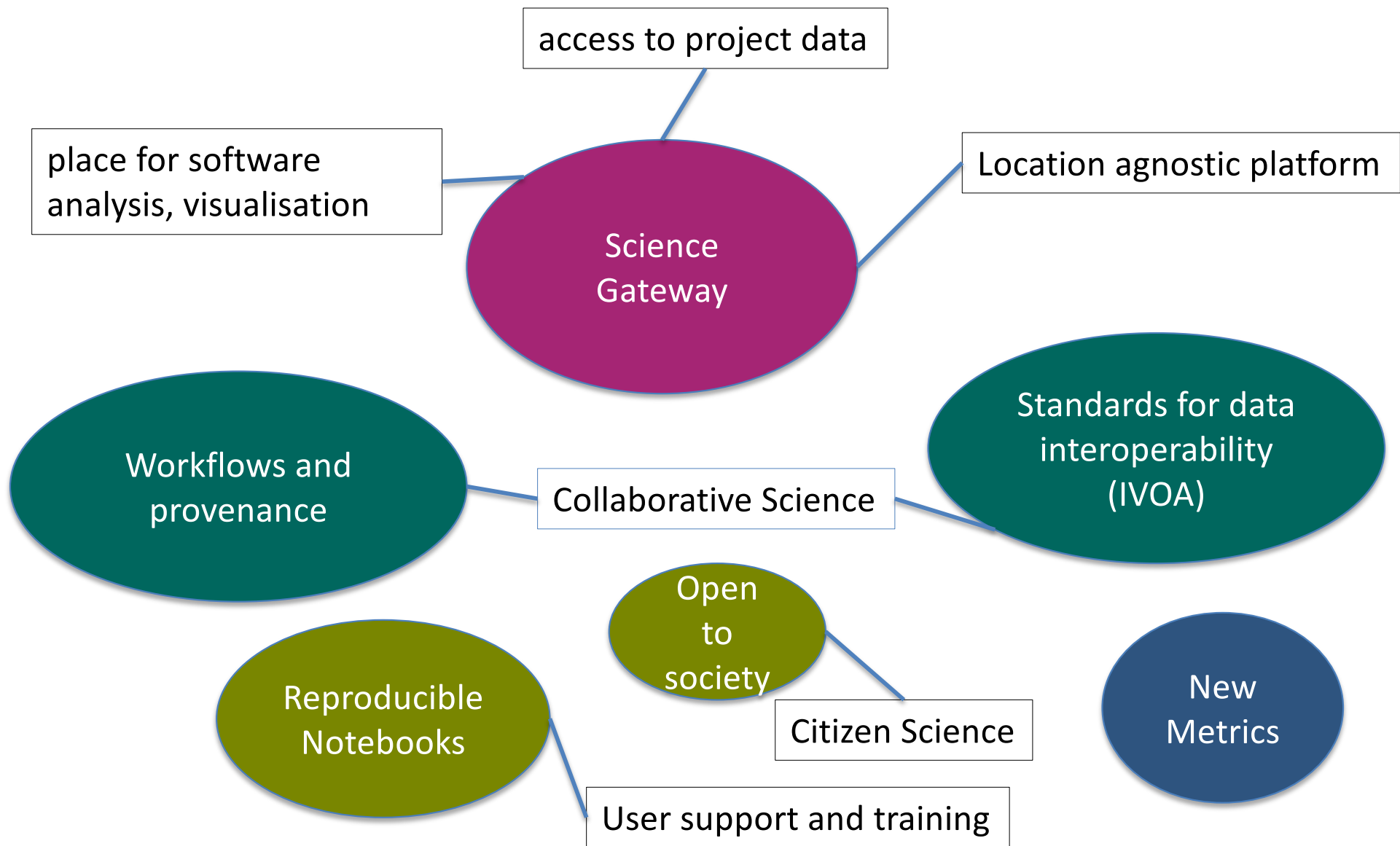
Global shift in research practices



Credits: AENEAS project

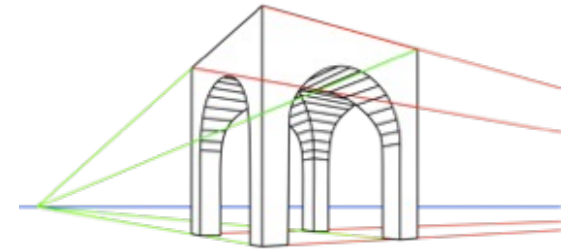
The SKA Regional Centres, the core of the SKA Science

Key ingredients of the SRCs to support Open Science

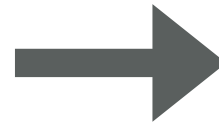


The challenge from different perspectives

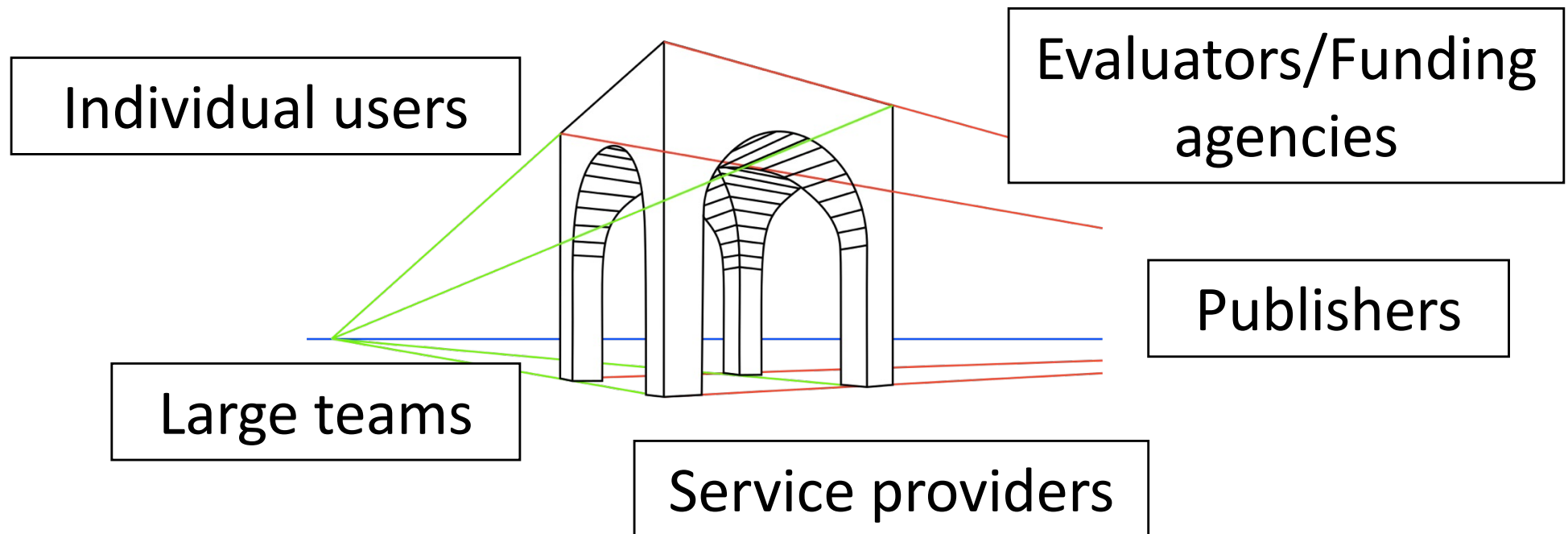
Implementation of Open, reproducible science is challenging, even more in this new framework:



new roles

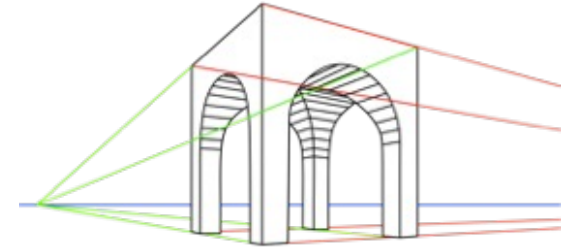


new perspectives



Data to the desktop: “individual scientist”

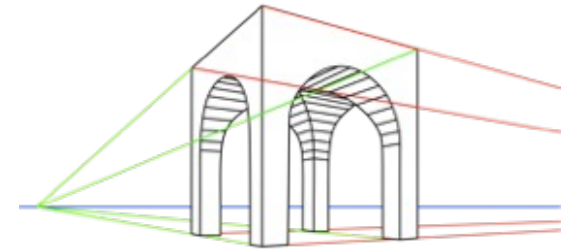
About trust



- I have the best code, which I know how to use and can do special things
- I do not trust any pipeline that you made
 - partly because I know better how to do it
 - partly because I read the news and there is a reproducibility crisis
 - well, and I can hardly reproduce the results of my own papers some years later...
- In general I want full control of the software and of the computational environment

Computation to data, providers perspective: Data Centres

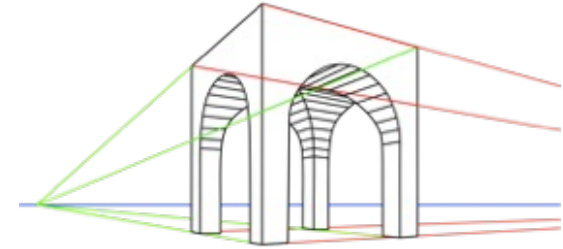
About technology



- We need to install your software in our platform. Can we trust it?
Can we run it? Environment, dependencies, etc
- Hey, we are offering services to the community, computation + tools. We would be grateful if you allow us to share it with other users (with proper credit)
- Mmmm, sharing is great, but, **putting the software in the platform is not enough**: you need to provide the context for people to be able to rerun the software on the same or other data

Large alliances of scientists

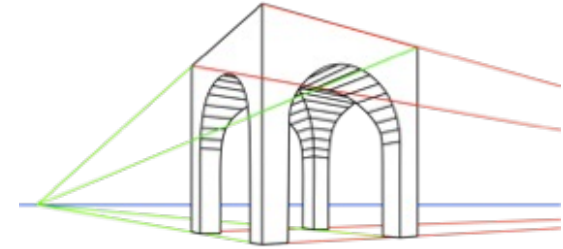
About metrics of research careers



- We have tools to generate Advanced Data Products, and we will put them there where the storage and computation is (Data Centres)
- But... we put effort on it, what would we gain if we make the ***additional effort* to make it reusable?** If we make it, then we will pave the way to competitors
- Well, maybe we will share in 4 yrs time (PhD typical time)

Publishers

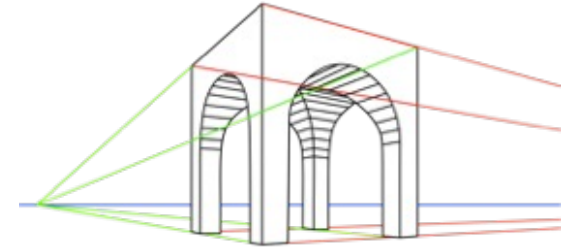
Publishing models



- Will we need **different profiles of referees** to evaluate the scientific discussion together with the data quality and the methods (aka. Reproducibility)?
- If the data and the methods (tools) will be in Data Centres, **will our referees need to become a “user” of the Data Centres** to be able to validate a paper?
- Will we be able to engage **so many referees** as may be needed?
- Will we need to validate the data, the tools, and the scientific analysis **separately**?

Policy makers / funding agencies

Evaluation

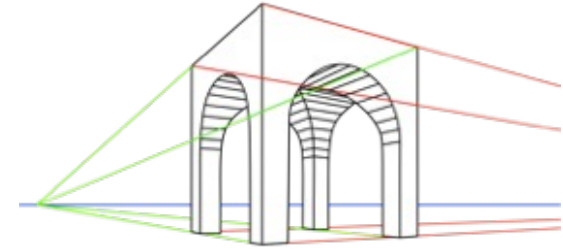


- How to measure reproducibility?
- How to weight it and/or aggregate with other indicators?

**See later on “Revised
research assessments”**

Infrastructures/facilities

About being an
example



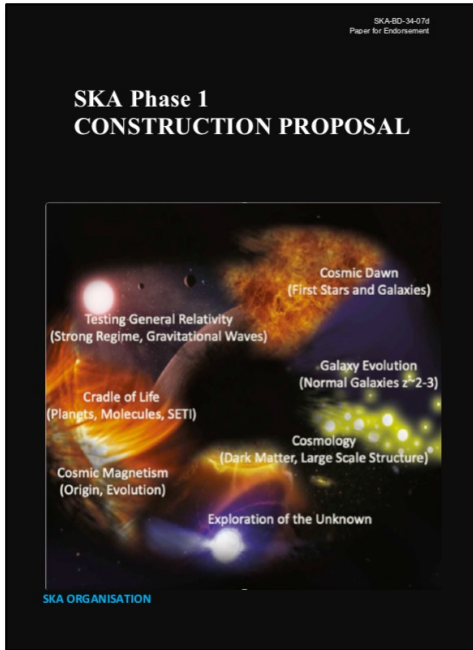
- For scientific facilities, adoption of Open Science is both a need and a duty.

The SKA and Open Science

3. Impact of the SKA 3.3.2 Open Science

Adoption of Open Science values

“Open Science, based on the precept of making scientific research collaborative, transparent and accessible to all, is rooted in SKA’s foundational principles. So is the related concept of scientific reproducibility, a fundamental aspect of the modern Scientific Method since the 17th century allowing independent teams to have access to methodology and tools to be able to confirm experiments and validate results.”



ENDORSED by the Council: Construction Proposal (CP) and Observatory Establishment and Delivery Plan (OEDP)

6. Observatory operations 6.1.2 Scientific success metrics

Reproducibility as a metric of success

*“Reproducibility of SKA science data products. This metric will measure how complete **the workflow description** is that is linked to each SKA data product. [...] must reflect completeness of the **provenance information** for each data product and accessibility of the software used. This is related to how well SKA science data products **adhere to the FAIR principles** .”*



Revised research assessments

Remember: Open Science started bottom-up

with manifestos authored by large sections of the scientific community
(Altmetrics-2010, DORA-2012, Metric Tide - 2015, Leiden Manifesto-2015,
Hong-Kong Principles – 2020)
(Astronomy = IVOA – 2002)



Wilsdon, J., et al. (2015). *The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management*. DOI: 10.13140/RG.2.1.4929.1363



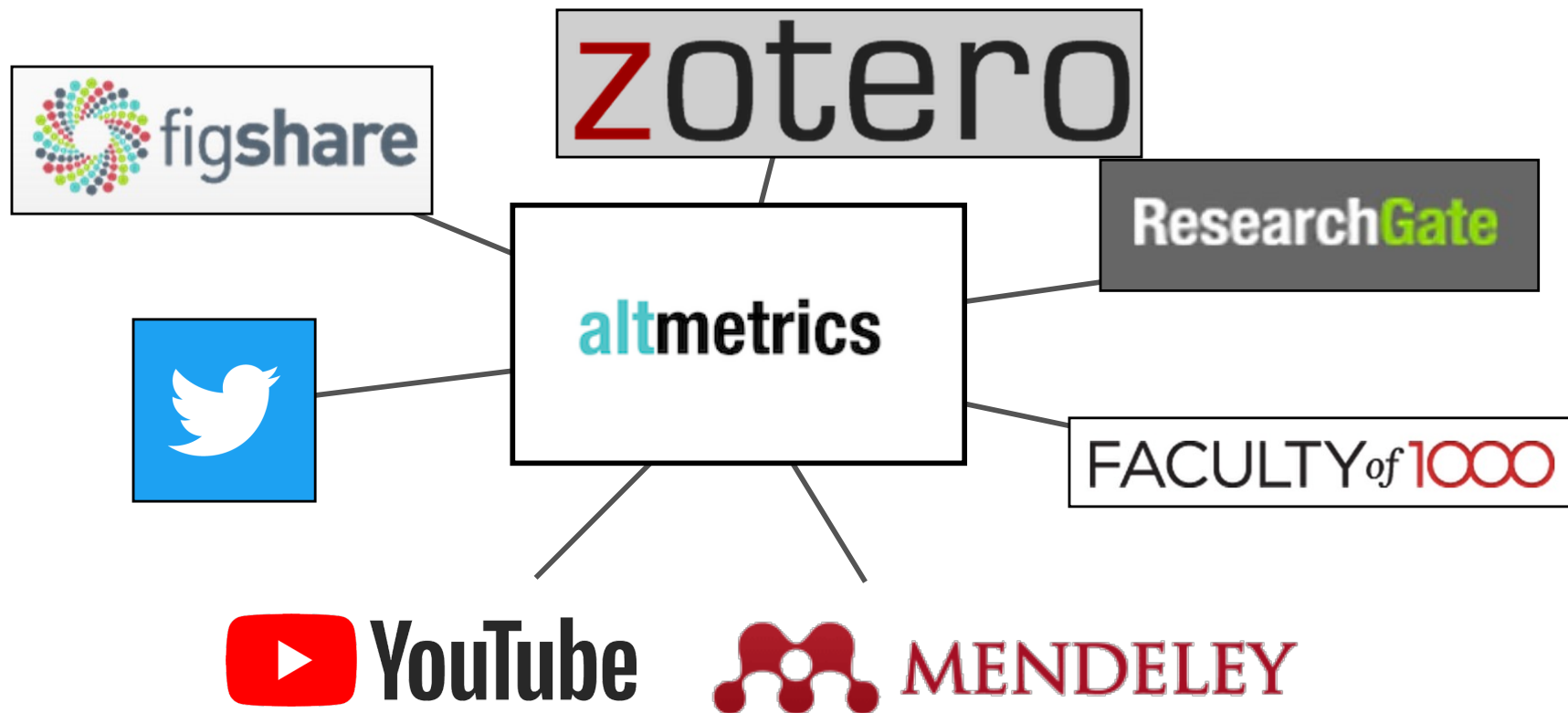
23 APRIL 2015 | VOL 520 | NATURE | 429

Almetrics

altmetrics

Altmetrics is the creation and study of new metrics based on the Social Web for analyzing, and informing scholarship.

<http://altmetrics.org/about/>



San Francisco Declaration on Research Assessment

There is a pressing need to improve the ways in which the output of scientific research is evaluated by funding agencies, academic institutions, and other parties. To address this issue, a group of editors and publishers of scholarly journals met during the Annual Meeting of The American Society for Cell Biology (ASCB) in San Francisco, CA, on December 16, 2012. The group developed a set of recommendations, referred to as the San Francisco Declaration on Research Assessment. We invite interested parties across all scientific disciplines to indicate their support by adding their names to this Declaration.

The outputs from scientific research are many and varied, including: research articles reporting new knowledge

[العربية](#)[Bahasa Indo](#)[中文](#)[Català](#)[Čeština](#)[Српски](#)

Evaluations Revisited



Next-generation metrics: Responsible metrics and evaluation for open science

Report of the European Commission Expert Group on Altmetrics

Not just citation of articles, various forms of social media shares, web-downloads, any other measure of the Q and impact of research outcomes



Mutual Learning Exercise

Open Science: Altmetrics and Rewards

Horizon 2020 Policy Support Facility

Thematic Reports: Types, use in the context of Open Science, Incentives and Rewards, Strategies, Experiences and Models, Final Report - Altmetrics and Rewards



Progress on Open Science: Towards a Shared Research Knowledge System

Final Report of the Open Science Policy Platform

April 2020

Evaluations Revisited

November 2021

Principles for assessment criteria

- Focusing research assessment criteria on quality
 - Openness of research, and results that are verifiable and reproducible where applicable, strongly contribute to quality
- Recognise the diversity of research and reward early sharing and open collaboration



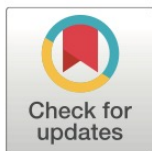
ESSAY

The Hong Kong Principles for assessing researchers: Fostering research integrity

David Moher^{1,2*}, Lex Bouter^{3,4}, Sabine Kleinert⁵, Paul Glasziou⁶, Mai Har Sham⁷, Virginia Barbour⁸, Anne-Marie Coriat⁹, Nicole Foeger¹⁰, Ulrich Dirnagl¹¹

1 Centre for Journalology, Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, Canada, **2** School of Epidemiology and Public Health, University of Ottawa, Ottawa, Canada, **3** Department of Epidemiology and Biostatistics, Amsterdam University Medical Centers, location VUmc, Amsterdam, the Netherlands, **4** Department of Philosophy, Faculty of Humanities, Vrije Universiteit, Amsterdam, the Netherlands, **5** The Lancet, London Wall Office, London, United Kingdom, **6** Institute for Evidence-Based Healthcare, Bond University, Gold Coast, Queensland, Australia, **7** School of Biomedical Sciences, LKS Faculty of Medicine, The University of Hong Kong, Pokfulam, Hong Kong SAR, China, **8** Queensland University of Technology (QUT), Brisbane, Australia, **9** Wellcome Trust, London, United Kingdom, **10** Austrian Agency for Research Integrity, Vienna, Austria, **11** Berlin Institute of Health, QUEST Center for Transforming Biomedical Research, Berlin, Germany

* dmoher@ohri.ca



OPEN ACCESS

Citation: Moher D, Bouter L, Kleinert S, Glasziou P, Sham MH, Barbour V, et al. (2020) The Hong Kong Principles for assessing researchers: Fostering research integrity. *PLoS Biol* 18(7): e3000737. <https://doi.org/10.1371/journal.pbio.3000737>

Published: July 16, 2020

Copyright: © 2020 Moher et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: PG is funded by an Australian National Health and Medical Research Council NHMRC

Abstract

For knowledge to benefit research and society, it must be trustworthy. Trustworthy research is robust, rigorous, and transparent at all stages of design, execution, and reporting. Assessment of researchers still rarely includes considerations related to trustworthiness, rigor, and transparency. We have developed the Hong Kong Principles (HKPs) as part of the 6th World Conference on Research Integrity with a specific focus on the need to drive research improvement through ensuring that researchers are explicitly recognized and rewarded for behaviors that strengthen research integrity. We present five principles: responsible research practices; transparent reporting; open science (open research); valuing a diversity of types of research; and recognizing all contributions to research and scholarly activity. For each principle, we provide a rationale for its inclusion and provide examples where these principles are already being adopted.

Impact



Vecteezy.com

Open Science for sustainability and inclusiveness

Open Science represents an approach to research that is collaborative, transparent and accessible

Open Science definition, European Commission, 2017, doi: 10.2777/75255

“Open Science embodies the need to transform, open and democratize the entire knowledge generation to ensure that every scientific challenge is faced and really drives and allows the achievement of the United Nations Sustainable Development Goals”

UNESCO and Open Science (2020) [1]

Open Science for sustainability and inclusiveness




Acceleration of knowledge transfer to Society, pandemics, sanitary crisis



- Speed up building of skills
- Teaching, e.g. how to access public archives, fostering collaborative practices
- Citizen science

Science hidden behind paywall barriers



-  Free access to research sources to the whole scientific community = limitations to science progress
- OS = Data and results more accessible and reliable
- OS = Promotion of scholarly exchange of ideas
- OS = Avoid duplication

Open Science for sustainability and inclusiveness

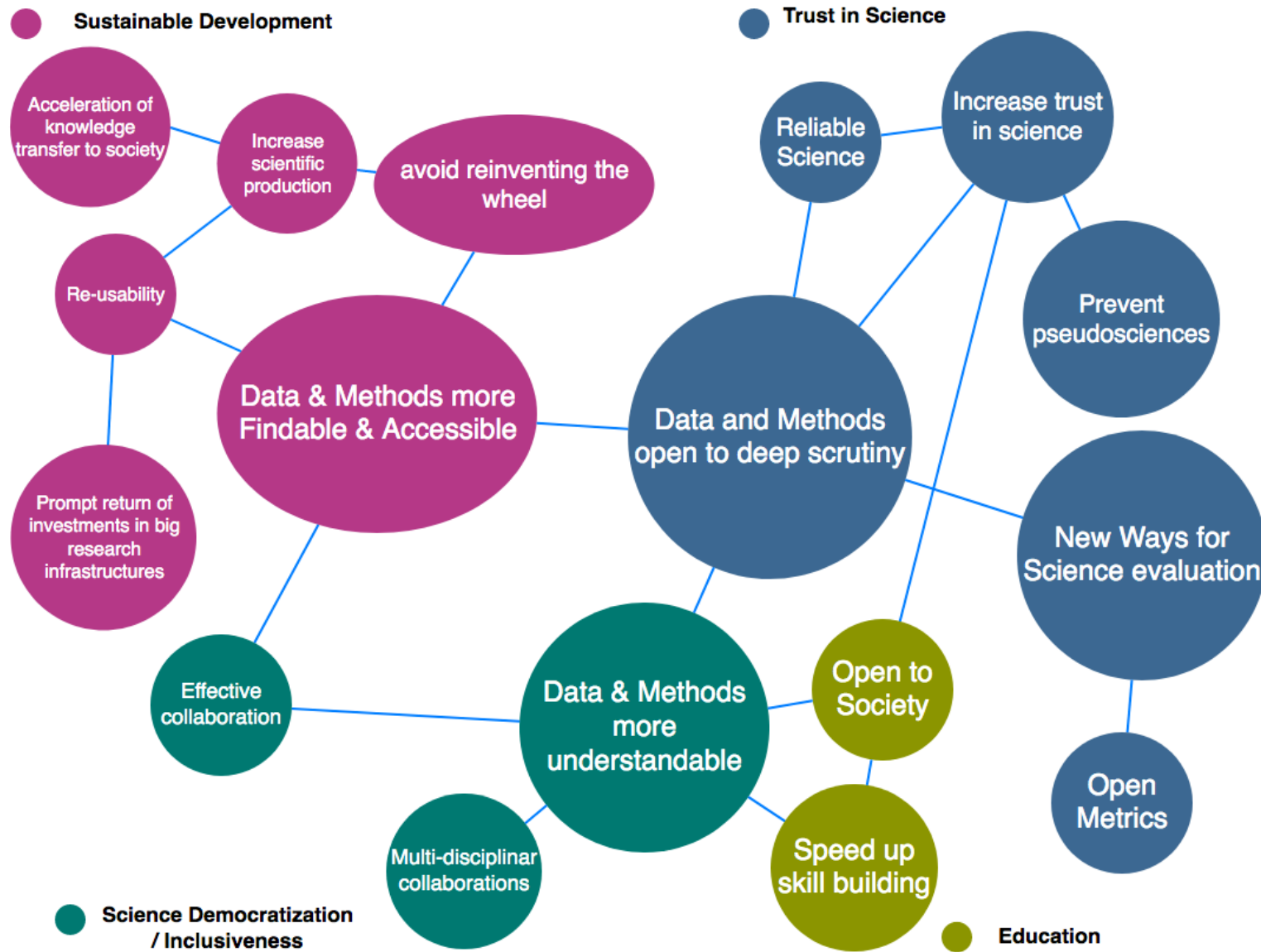


Promote equity, diversity and inclusion

- All previous items +
- A tool enabling an objective evaluation of work
- Barriers are even more emphasized to scientist women in places where their contribution tend to be ignored or anonymized.



Open Science



Conclusions

- **“Instead of playing the game it is time to change the rules”**

Chambers et al 2014, AIMS Neuroscience 1,4, 2014

- Astronomy is in a privileged situation as pioneer
- Open reproducible science is: a duty and a need
- We made a lot of progress in the last few years in all areas!

Conclusions

- **“Instead of playing the game it is time to change the rules”**

Chambers et al 2014, AIMS Neuroscience 1,4, 2014

- Astronomy is in a privileged situation as pioneer
- Open reproducible science is: a duty and a need
- We made a lot of progress in the last few years in all areas!

Infrastructures are key for implementing O.S.
(aka SKAO & SRCs)

Conclusions

- **“Instead of playing the game it is time to change the rules”**

Chambers et al 2014, AIMS Neuroscience 1,4, 2014

- Astronomy is in a privileged situation as pioneer
- Open reproducible science is: a duty and a need
- We made a lot of progress in the last few years in all areas!

Infrastructures are key for implementing O.S.
(aka SKAO & SRCs)

**In the end there should not be "good" science,
but only Science**

With financial support from



Junta de Andalucía
Consejería de Transformación Económica,
Industria, Conocimiento y Universidades



UNIÓN EUROPEA
Fondo Europeo de Desarrollo Regional