ROSiE General Guidelines on Responsible Open Science

Rosemarie Bernabe, PhD,
University of Oslo
Project Coordinator, ROSiE Project

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

https://zenodo.org/records/10046520

Why the Need for RE/RI Guidelines on Open Science?

RoSiE
Responsible Open Science in Europe

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

Universal Declaration of Human Rights

Article 27

1. Everyone has the right freely to participate in the cultural life of the community, to enjoy the arts and to share in scientific advancement and its benefits.

2. Everyone has the right to the protection of the moral and material interests resulting from any scientific, literary or artistic production of which he is the author.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

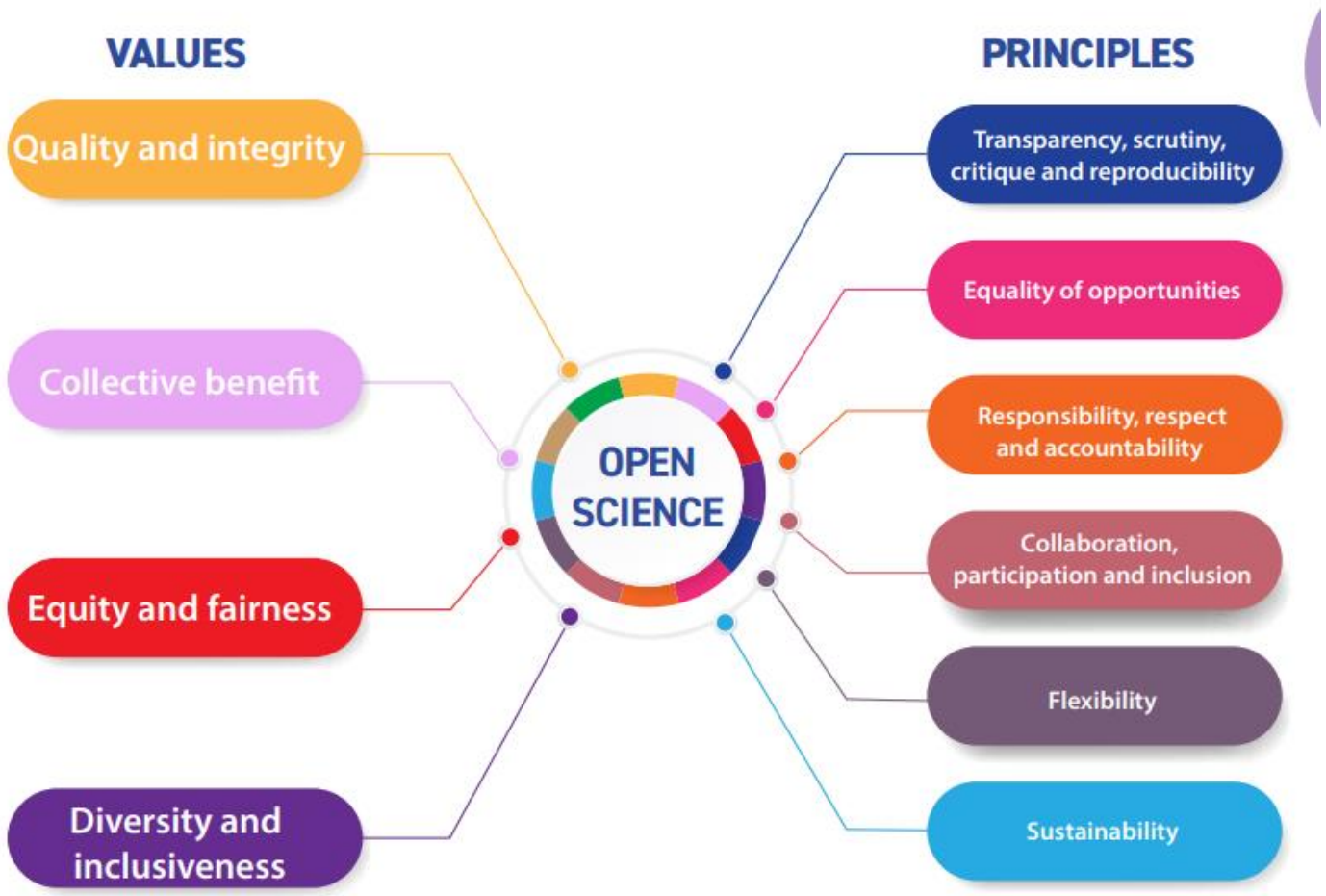
- ***Reliability*** in ensuring the quality of research, reflected in the design, methodology, analysis, and use of resources.

- ***Honesty*** in developing, undertaking, reviewing, reporting, and communicating research in a transparent, fair, full, and unbiased way.

- ***Respect*** for colleagues, research participants, research subjects, society, ecosystems, cultural heritage, and the environment.

- ***Accountability*** for the research from idea to publication, for its management and organisation, for training, supervision, and mentoring, and for its wider societal impacts.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430



UNESCO Recommendation on Open Science

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

OS has a number of sources of normativity – what counts as good or responsible OS – and this is one source of challenges in itself, i.e., in satisfying some objectives, it is seen as failing others.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

Sections of the Guidelines (and of this presentation)

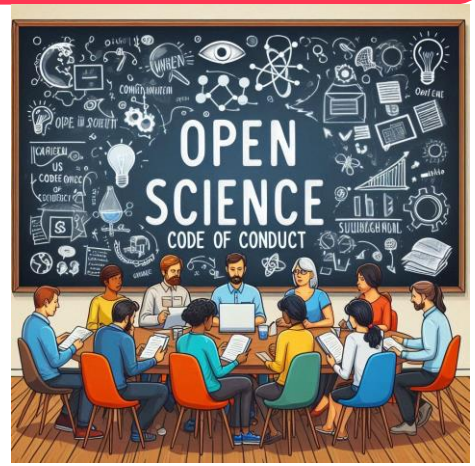
- Research Environment and Infrastructures
- Protection of Research Participants, the Environment, Ecosystems, and Cultural Heritage
- Open and Reproducible Research Practices
  - Open research practices
  - Open data
  - Open methods and tools
  - Open access publication
- Researcher evaluation
- Citizen science
- Training and Education
- Inclusivity

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

Research Environment and Infrastructures

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

Codes of Conduct not ready for Open Science



Open science and codes of conduct on research integrity

The purpose of this article is to examine the conceptual alignment between the ethical principles of research integrity and open science. Research integrity is represented in this study by four general codes of conduct on responsible conduct of research (RCR), three of them international in scope, and one national. A representative list of ethical principles associated with open science is compiled in order to create categories for assessing the content of the codes. According to the analysis, the current understanding of RCR is too focused on traditional publications and the so called FFP definition of research misconduct to fully support open science. The main gaps include recognising citizen science and societal outreach and supporting open collaboration both among the research community and beyond its traditional borders. Updates for both the content of CoCs as well as the processes of creating such guidelines are suggested.

Laine, H. (2018). Open science and codes of conduct on research integrity. *Informaatiotutkimus*, *37*(4). https://doi.org/10.23978/inf.77414

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

D1.1: Report on the relationship (tensions, challenges, overlaps) between RI, the wider RE perspective and OS

Authors: Ana Sofia Carvalho & Maria Strecht Almeida

3. None of the evaluated Codes are in contradiction with the values of open science, but only the Austrian and French Codes of conduct can be said to actively support the values and norms of open science. This may be justified since both Codes have been issued or updated after EU Open Science Policy (2019). Also open science discussion was already in the science policy mainstream in Europe during its drafting, and was a high priority for the European Commission, which has since adopted the Code for projects funded through its Horizon 2020 and Horizon Europe instruments.

4. However, even if some other Codes have been published or updated after 2019 (GERMANY, SPAIN, SWITZERLAND) the issues of OS are not tacked explicitly, and offer very little in terms of defining what it means, or guidance on how to practice it.

5. It should also be noted that not all the categories were included in the Codes from FRANCE and AUSTRIA; citizen science is only addressed in the Austrian Code and open education in the French Code.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

Guideline 2. Research Environment and Infrastructures

2.1. **Policy reforms and OS advocacy are crucial** for creating a culture that promotes, supports, and rewards OS. A policy environment conducive to responsible OS requires **aligned action on the European, national, and institutional levels.**

project has received funding from the European Union's Horizon
20 research and innovation programme under GA No 101006430

Protection of Research Participants, the Environment, Ecosystems, and Cultural Heritage

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

Article | Open Access | Published: 23 July 2019

Estimating the success of re-identifications in incomplete datasets using generative models

Luc Rocher, Julien M. Hendrickx & Yves-Alexandre de Montjoye

Nature Communications **10**, Article number: 3069 (2019) | Cite this article

While rich medical, behavioral, and socio-demographic data are key to modern data-driven research, their collection and use raise legitimate privacy concerns. Anonymizing datasets through de-identification and sampling before sharing them has been the main tool used to address those concerns. We here propose a generative copula-based method that can accurately estimate the likelihood of a specific person to be correctly re-identified, even in a heavily incomplete dataset. On 210 populations, our method obtains AUC scores for predicting individual uniqueness ranging from 0.84 to 0.97, with low false-discovery rate. Using our model, we find that 99.98% of Americans would be correctly re-identified in any dataset using 15 demographic attributes. Our results suggest that even heavily sampled anonymized datasets are unlikely to satisfy the modern standards for anonymization set forth by GDPR and seriously challenge the technical and legal adequacy of the de-identification release-and-forget model.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

Patterns

CellPress
OPEN ACCESS

Article

The risk of re-identification remains high even in country-scale location datasets

Ali Farzanehfar,[1] Florimond Houssiau,[1] and Yves-Alexandre de Montjoye[1,2,*]
[1]Department of Computing, Imperial College London, London SW7 2AZ, UK
[2]Lead contact
*Correspondence: demontjoye@imperial.ac.uk
https://doi.org/10.1016/j.patter.2021.100204

SUMMARY

Although anonymous data are not considered personal data, recent research has shown how individuals can often be re-identified. Scholars have argued that previous findings apply only to small-scale datasets and that privacy is preserved in large-scale datasets. Using 3 months of location data, we (1) show the risk of re-identification to decrease slowly with dataset size, (2) approximate this decrease with a simple model taking into account three population-wide marginal distributions, and (3) prove that unicity is convex and obtain a linear lower bound. Our estimates show that 93% of people would be uniquely identified in a dataset of 60M people using four points of auxiliary information, with a lower bound at 22%. This lower bound increases to 87% when five points are available. Taken together, our results show how the privacy of individuals is very unlikely to be preserved even in country-scale location datasets.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

Guideline 3. Protection of Research Participants, the Environment, Ecosystems, and Cultural Heritage

3.3. **Protecting the privacy and control interests** of research participants and their communities is essential in an open data environment. **Researchers, research ethics committees, RPOs, and policymakers** should **analyse the risks of reidentification and dual use** in different fields and **develop governance mechanisms and technical solutions** to address these risks. Exploring other approaches to protect privacy, other than anonymization, is increasingly becoming important and is thus recommended.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

Open and Reproducible Research Practices

- *Open research practices*
- *Open data*
- *Open methods and tools*
- *Open access publication*

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

D1.1: Report on the relationship (tensions, challenges, overlaps) between RI, the wider RE perspective and OS

Authors: Ana Sofia Carvalho & Maria Strecht Almeida

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

"the process of making data openly available (eg., by anonymization) may significantly reduce the value and usefulness of the data."

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

Guideline 4.2. Open data

4.2.5. **RFOs** should:

4.2.5.1. **incentivise data sharing** through, for example, the inclusion of open data requirement in data management plans.

4.2.6. **RPOs** should:

4.2.6.1**. institutionalise and incentivise data sharing** and compliance with FAIR and CARE principles.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

4.2.7. **Researchers** should:

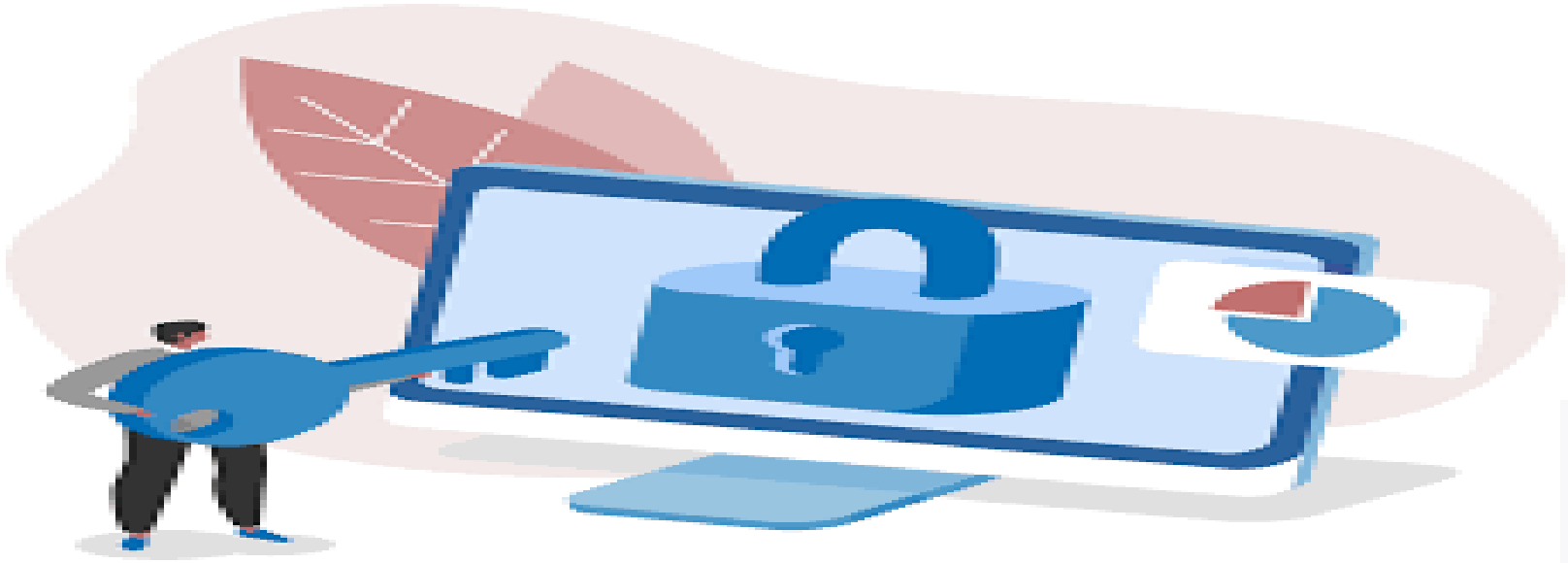
4.2.7.1. collect, analyse, and present research data with the **prospect of making it open and reusable**;
4.2.7.2. ensure that research data is **properly documented, formatted, and stored to facilitate findability and reusability**.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

Researcher Evaluation

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

Open Data



This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

Guideline 5. Researcher Evaluation

5.2. In accordance with DORA and CoARA recommendations, research performance assessment systems should **prioritise quality and openness of the research results** over the quantity of published papers. In the evaluation of a researcher's performance, RPOs and RFOs should use a **multidimensional approach**; bibliometric criteria should not be the only or main criteria. Evaluation systems should also consider the **potential positive societal impact of research** in accordance with OS and research integrity principles.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

Training and Education

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

OS Training 8.a. **Thorough training on responsible OS is not available/offered to researchers and citizen scientists.** In many countries, skills, and tools necessary for practicing responsible OS are often available only for people within institutions. In order to cover all training needs, national authorities and institutions should, when necessary, provide training for people outside institutions as well.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

Guideline 7. Training and Education

7.1. **Policymakers and RPOs should ensure that training and education in responsible OS focuses on the entire research lifecycle and start early-on**, integrating the development of relevant attitudes and skills into higher education and perhaps even high school curricula.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

Inclusivity

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

"OS may not be able to remedy the **historical injustices in knowledge production.** Previously excluded groups may not be able to participate in the co-production of knowledge…"

"OS reinforces existing **epistemic organization and governance**, e.g., by being shaped, governed, and reproduced by the institutions, countries, and agencies that govern existing knowledge production."

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

UNESCO Steering Committee Meeting: Need for Funding, Infrastructure, Institutions, and Capacities

"**Low investment in R&D**, **funding of Open Science infrastructures and capacity building** were raised as major challenges. "

"The respondents from Asian and Pacific States argued that in developing countries, **access to stable power supply and internet connectivity** still needed to be considered. Access to computer code for simulations, calculations, analysis, visualization, and general data processing would also be needed."

"The possible deepening of the **North-South digital and science divide** if the differences in **connectivity, capacities, and resources were not adequately addressed."**

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

Guideline 8. Inclusivity

8.1. **Researchers should be aware of potential biases** in research related to gender, ethnicity, age, disability, epistemological frameworks, and other factors and **act to ensure that responsible OS practices promote equity, inclusiveness, and diversity.**

8.2. **RPOs** should recognise potential global inequities in access to OS infrastructure and act to **promote global justice and support the needs of researchers in low- and middle-income countries** (LMICs).

There is a great need for **policymakers, RFOs, RPOs, and researchers from high-income countries** to provide **support to institutions from LMICs** in building their capacities, exchanging good practices, and establishing infrastructure conducive to OS.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

Citizen Science

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

D1.1: Report on the relationship (tensions, challenges, overlaps) between RI, the wider RE perspective and OS

Authors: Ana Sofia Carvalho & Maria Strecht Almeida

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

Guideline 6. Citizen Science

6.3. Citizen science offers a potential for socially relevant research and innovation, however, the involvement of citizen scientists without proper support can potentially be an ethics and integrity challenge. **Policymakers, RFOs, RPOs, and researchers** are **responsible for promoting and supporting citizen science**. This is done specifically by ensuring **support throughout the research lifecycle, through the provision of adequate funding, training, flexible grant structures that accommodate extended timeline research, and encouraging collaborations and building synergies between researchers and other stakeholders.**

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430

Thank you!

ROSIE-project.eu

This project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No 101006430