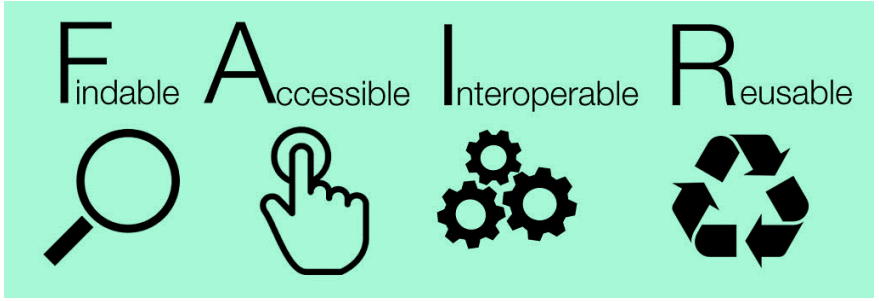
FAIR vs. GDPR: *which will win?*

Robin Rice
Data Librarian and Head, Research Data Support
University of Edinburgh
LIBER 2018: Lille

Two acronyms, two paradigms

- FINDABLE
- ACCESSIBLE
- INTEROPERABLE
- REUSABLE

- GENERAL
- DATA
- PROTECTION
- REGULATION



by SangyaPundir [CC BY-SA 4.0 (https://creativecommons.org/licenses/by-sa/4.0)], from Wikimedia Commons

FAIR paradigm: *Open by Default*

- **FINDABLE**: “Metadata and data should be easy to find for both humans and computers. Machine-readable metadata are essential for automatic discovery of datasets and services.”
- **ACCESSIBLE**: “Once the user finds the required data, she/he needs to know how can they be accessed, possibly including authentication and authorisation.”
- **INTEROPERABLE**: “The data usually need to be integrated with other data. In addition, the data need to interoperate with applications or workflows for analysis, storage, and processing.”
- **REUSABLE**: “The ultimate goal of FAIR is to optimise the reuse of data. To achieve this, metadata and data should be well-described so that they can be replicated and/or combined in different settings.”

GDPR paradigm: *Privacy by Default*

Six principles of the GDPR:

- a) Lawfulness, fairness and transparency
- b) Purpose limitation
- c) Data minimisation
- d) Accuracy
- e) Storage limitation
- f) Integrity and confidentiality (security)

Pluses for researchers:

Legal basis for processing not consent but either public task/public interest or legitimate interest.

Some limited exemptions apply for *"Archiving purposes in the public interest, scientific or historical research."*

DP challenges for human subject researchers

Concepts in the Law

- Privacy by Design and by Default
- Accountability 7th principle
- Personal data
- Special categories of personal data
- Legal basis for processing
- Privacy notices
- Data Protection Impact Assessment
- Data controllers, data processors
- Safeguards for data transfer outside the EEA
- Data subject rights
- Minimisation principle
- Anonymisation and Pseudonymisation
- Reporting of breaches, big fines

Support researchers require

- Handling personal data securely
- Selecting secure data systems designed for privacy
- Collecting sufficient personal data, special categories, but not more
- Transparently communicating data processing actions to human subjects (information sheets & consent forms)
- Understanding and documenting risks
- How to anonymise / pseudonymise data
- Knowing who is a data controller, data processor
- Creating legally binding data use agreements
- Dealing with breaches

What do librarian FAIR advocates have to say about DP? (Not much)

LERU Advice Paper (May 2018): Open Science and its role in universities: A roadmap for cultural change

"There are challenges to establishing responsible RDM practices. Some researchers feel challenged by the need for research data management plans and the **requirements of the General Data Protection Regulation** (GDPR) (p. 13 of 31)."

[Nothing in recommendations.]

LIBER Open Science Roadmap (July 2018)

"**ENGAGE** in the development of national and European legislation and policies which impact on Open Science.

When topics such as copyright, text and data mining, **data protection** and FAIR data are discussed, reinforce the importance of Open Science and the need to adopt frameworks which give maximum access to knowledge and resources" (p. 11 of 51).

[Also a brief mention in Uni of Southern Denmark case study.]

CONCERNS

- Will researchers get the support they need to share data based on human subjects, or will they be risk-averse and avoid sharing?
- Will the European Open Science Cloud and other FAIR-enabled infrastructure be built with data protection requirements in mind?
- Does open by default conflict with privacy by design?
- Will IT and Libraries help researchers who work with human subjects with their unique needs for data processing, archiving, and sharing?
- Will researchers in social and health sciences be able to take advantage of innovations in data science?
- If the open science agenda takes off, will human subject researchers be disadvantaged in terms of incentives and rewards?
- Can interdisciplinary, global grand challenges of the day such as climate change and inequality research be solved by the open science agenda and citizen science given the legal limitations on sharing of data about human subjects?

In short -

When it comes to human subject research, which will win out –

***FAIR* or *GDPR*?**

R.Rice@ed.ac.uk

@sparrowbarley