

FAIR Principles, Reproducibility, and Open Science

Robin Rice Data Librarian and Head of Research Data Support at the University of Edinburgh

What does research data management have to do with data quality? Research data management can be thought of as the sum of the actions taken over the lifecycle of a dataset to make it intelligible for the creator and other users. In addition to making decisions about storage and security, data transformations, documentation, access, and long-term preservation (preferably as part of a data management plan); a data creator's responsibility for research data, particularly that which underlies a research publication, is to ensure its 'FAIR-ness'.

The FAIR Principles were created by a group of Open Science adherents in 2016 (Wilkinson et al¹), and have since been adopted by most research funders and many research performing organisations. They have been highly influential and cited nearly 25,000 times. The FAIR Principles – comprising “findable” or discoverable, accessible, interoperable and reusable – set high standards for what data creators and future data custodians such as repository managers need to do to make their data – and metadata – ‘FAIR’.

Focusing on the R of Reusable, there are a range of reasons a reader, or data user, might wish to reuse a dataset and its code, from simply scrutinising the results in order to understand them better, to reproducing the analysis in order to validate the findings, to fully replicating the analysis in a different setting to verify the original conclusions. The “replication crisis” of the 2010s referred to common failures to reproduce published scientific results, which led to the shift toward open science practices.²

Open Science is often referred to as an umbrella of research practices that improve transparency and openness of science and scholarship. From publishing preprints to sharing data and code, to open access publications and open methods, the benefits are seen to accrue beyond the research practitioner/team who may gain from increased citations. Those who benefit range from other scientists and scholars and their students, research funders, institutions and libraries, to the general public who through taxes pay for much of the research and need to be able to trust it on some level, and citizen scientists unaffiliated with any institution. This movement is epitomised by the UNESCO Recommendation on Open Science³, adopted by 194 countries in 2021 as a global framework.

Data quality cannot be achieved without also pursuing research integrity, data stewardship incorporating the FAIR Principles, and aspects of Open Science including data sharing. Where open sharing is not possible due to sensitive content, other forms of transparency may be pursued, such as making metadata, and possibly code, available, and linking these to the original research using persistent identifiers.

¹ Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018 (2016). <https://doi.org/10.1038/sdata.2016.18>

² CASRAI. (2026). Guides Replication Crisis (Page). CASRAI Dictionary v2026.1. <https://casrai.org/guides/replication-crisis>. CC-BY 4.0.

³ UNESCO. (2021). UNESCO recommendation on open science. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000379949>