Institutional tasks implied by the UK Open Research Data Concordat

Executive Summary

This UKRN working paper outlines the tasks that are either explicitly or implicitly proposed for research institutions by the UK Open Research Data Concordat (2016). It also reflects on how those tasks are relevant to the recommendations from the 2018 report issued by the UK Government from the Open Research Data Task Force. The aim of this paper is to describe more clearly what UK research institutions might be expected to do to support open research data, which may be a helpful basis for sector-wide discussion on both those expectations and how they are being met. The tasks are grouped under the headings of organisational support, infrastructure, recognition, restrictions on open data, costs, and strategy.

The tasks were derived by a group of experts from international centres of expertise based in the UK and from UK institutions. Because of both the ambiguity in some parts of the Concordat, and the passage of time since it was published, there is inevitably room for debate about whether the tasks described in this report represent a true and complete picture of the expectations on research institutions based on the Concordat. Furthermore, in several areas good practice in supporting open research data now goes beyond that described in the Concordat; no attempt is made in this paper to reflect that.

2

Contents

Background and aims

This is a short account of the derivation of a set of tasks that are implied for institutions within the text of the UK Concordat on Open Research Data (Concordat Working Group, 2016). It also references the recommendations from the 2018 Open Research Data Task Force. The work was undertaken over 2023 and early 2024 by a group of research data management experts from the Digital Curation Centre (DCC), institutions, FAIRSharing.org, and UKRN.

The aim was to set out a list of concrete tasks that could be the basis for:

1. Institutions to review their implementation of the principles agreed in the Concordat
2. Regular reviews of progress to take place under Principle #10 of the Concordat, such as the UKRN STAR Project

While the output presented here is an expert consensus, the authors are conscious that this is the work of a small group and may (and probably should) evolve as it is tested and used.

In several areas (for example, with respect to the 'FAIRness' of data), good practice in supporting open research data now goes beyond that described in the Concordat; no attempt is made in this paper to reflect those changes.

Methods

The Concordat itself contains ten principles "with which all those engaged with research should be able to work" (Concordat Working Group, 2016, p. 3).

The derivation of tasks happened in three phases:

1. UKRN commissioned the DCC to undertake work to both derive a first draft of the tasks, and to use this to analyse the REF5a statements and other evidence from a sample of institutions, to inform planning for the UKRN STAR project.
2. A wider group of experts reviewed this first draft, drawing in some cases from their own previous analyses of the Concordat.
3. UKRN moderated these reviews and generated a consensus version of the tasks, which is what is presented below.
4. UKRN reviewed the tasks against the recommendations of the Open Research Data Task Force.

In more detail:

In phase 1, as a guide in identifying tasks, DCC and UKRN looked for relevant activities that, as far as possible, might aspire to S.M.A.R.T. criteria (Doran, 1981, pp. 35-36):

- Specific: targeted at a certain area
- Measurable: not so much quantifiable, but able to suggest an indicator of progress.
- Assignable: where possible, with a specified "owner" responsible.
- Realistic: Can credibly be achieved
- Time-related: In some cases, where a specified target period or time when results are achieved.

Discussion between UKRN and the DCC identified 16 tasks from the Concordat. This number was obviously greater than the ten principles, but that was because some of the principles have more than one task identified in them.

In phase 2, the wider group of experts drew from analysis by, for example, the University of Bath, and from insights arising from a knowledge of the FAIRSharing metadata schema and its mapping to other structures.

In phase 3, UKRN checked that each entry in the resulting list of tasks could be justified as reasonably following from the full text of the Concordat principles, rather than – for example – being an expression of good practice that has been recognised since the Concordat was adopted. UKRN then grouped the tasks thematically, based on an impressionistic and, inevitably, contestable set of themes.

In phase 4, UKRN mapped the tasks derived using the method outlined above to the recommendations in the Open Research Data Task Force 2018 report.

Findings

The tasks are presented here in three ways: mapped to the Concordat principles; grouped thematically; and mapped to the recommendations of the Open Research Data Task Force. The thematic grouping is because, inevitably perhaps, a task implied by one principle was often rather closely related to – or even almost identical to – a task implied by another principle. The mapping to the Concordat principles provides an audit trail demonstrating the derivation of the tasks, while the thematic grouping is likely to be more useful for those undertaking reviews of policy or practice. The mapping to the recommendations of the Open Research Data Task Force relate these tasks to an official release by UK Government.

Tasks mapped to the Concordat principles

Principle (summary)	Tasks for institutions implied by the full text of that principle
1. Open access to research data is an enabler of high-quality research, a facilitator of innovation and safeguards good research practice.	a. Institutions should have policies, practices and culture that recognise the value of open data and of the work done to enable it throughout the research process. b. Institutions should provide appropriate access to infrastructure systems and services to enable their researchers to make research data open and usable.
2. There are sound reasons why the openness of research data may need to be restricted but any restrictions must be justified and justifiable.	c. Institutions should have verifiable and transparent processes of oversight to guide decisions on whether, when and how data may be made available in some form, to reflect situations where (e.g.) data relates to or is derived from individuals, or has commercial sensitivity, or third-party intellectual property rights. d. Institutional policies and guidance should acknowledge that on a case-by-case basis sharing be proportionate to the level of risk associated with the data.
3. Open access to research data carries a significant cost, which should be respected by all parties.	e. Institutions should acknowledge that infrastructure, training, staff support, and progress reviews for open and FAIR research have significant costs and they should have procedures in place to ensure funding of costs within the dual support system.
4. The right of the creators of research data to reasonable first use is recognised.	f. Institutional research data policies must allow the possibility of an initial period of exclusive use or an embargo period before data availability and reuse. g. Data management plans should address requests for embargoes or periods of exclusive use.

Principle (summary)	Tasks for institutions implied by the full text of that principle
5. Use of others' data should always conform to legal, ethical, and regulatory frameworks including appropriate acknowledgement.	h. Institutions should provide support (such as training, guidance, advice and policies) on legal, ethical, and professional frameworks of research integrity as related to issues of open data. This support should cover both appropriate citation and acknowledgement of others' work including datasets, and provision of a preferred citation for underlying data and other relevant outputs including research software. i. Institutional policies should recognise contributions to shared/open datasets in academic appointment, promotion, research assessment and research funding decisions. j. Formal recognition should make use of responsible metrics for tracking research data use and impact, such as data citations.
6. Good data management is fundamental to all stages of the research process and should be established at the outset.	k. Institutions should provide access to the necessary infrastructure to enable researchers to manage their data effectively throughout the research process. l. Institutions should provide guidance to researchers throughout the research process on the correct and relevant data management and storage methodologies for that research field. m. Institutions should support and make use of shared research infrastructure. n. Institutions should provide researchers with training and support in research data management.
7. Data curation is vital to make data useful for others and for long-term preservation of data	o. Where institutions provide access to data repositories, these repositories should be able to guarantee the persistence of the datasets for a reasonable time period. p. Institutions should support provision of sufficient metadata for data evaluation and reuse. q. Institutions should encourage the use of non-proprietary formats.
8. Data supporting publications should be accessible by the publication date and should be in a citeable form.	r. Where institutions provide access to data (and, where possible, software) repositories, these repositories should enable datasets and related software to be citable, for example by providing them with a persistent identifier. s. Institutional policies should require that data underpinning publications should as default be retained for at least ten years in a citable form. t. Institutions should provide guidance on relevant preservation policies and/or retention schedules.

Principle (summary)	Tasks for institutions implied by the full text of that principle
9. Support for the development of appropriate data skills is recognised as a responsibility for all stakeholders.	u. Institutions should provide researcher training opportunities in an organised and professional manner. v. Institutions should support data science through skills development and career pathways.
10. Regular reviews of progress towards open research data should be undertaken.	w. Institutions should undertake regular reviews to monitor their progress in implementing their responsibilities under the Concordat, register issues to be addressed, and identify and share best practice. x. Institutions should demonstrate a long-term commitment to open data.

Tasks grouped thematically

The numbering in this section corresponds to the principles / tasks from the table above.

Organisational support

1a. Institutions should have policies, practices and culture that recognise the value of open data and of the work done to enable it throughout the research process.
5h. Institutions should provide support (such as training, guidance, advice and policies) on legal, ethical, and professional frameworks of research integrity as related to issues of open data. This support should cover both appropriate citation and acknowledgement of others' work, and provision of a preferred citation for underlying data and other relevant outputs including research software.
6n. Institutions should provide researchers with training and support in research data management.
9u. Institutions should provide researcher training opportunities in an organised and professional manner.
9v. Institutions should support data science through skills development and career pathways.
8s. Institutional policies should require that data underpinning publications should as default be retained for at least ten years in a citable form.
8t. Institutions should provide guidance on relevant preservation policies and/or retention schedules.
6l. Institutions should provide guidance to researchers throughout the research process on the correct and relevant data management and storage methodologies for that research field.
7q. Institutions should encourage the use of non-proprietary formats.

Infrastructure

1b. Institutions should provide appropriate access to infrastructure systems and services to enable their researchers to make research data open and usable.

7p. Institutions should support provision of sufficient metadata for data evaluation and reuse.

6k. Institutions should provide access to the necessary infrastructure to enable researchers to manage their data effectively throughout the research process.

6m. Institutions should support and make use of shared research infrastructure.

7o. Where institutions provide access to data repositories, these repositories should be able to guarantee the persistence of the datasets for a reasonable time period.

8r. Where institutions provide access to data (and, where possible, software) repositories, these repositories should enable datasets and related software to be citable, for example by providing them with a persistent identifier.

Recognition

5i. Institutional policies should recognise contributions to shared/open datasets in academic appointment, promotion, research assessment and research funding decisions.

5j. Formal recognition should make use of responsible metrics for tracking research data use and impact, such as data citations.

Restrictions on open data

2c. Institutions should have verifiable and transparent processes of oversight to guide decisions on whether, when and how data may be made available in some form, to reflect situations where (e.g.) data relates to or is derived from individuals, or has commercial sensitivity, or third-party intellectual property rights.

2d. Institutional policies and guidance should acknowledge that on a case-by-case basis sharing be proportionate to the level of risk associated with the data.

4g. Data management plans should address embargo requests.

4f. Institutional research data policies must allow the possibility of an initial embargo period before data availability and reuse.

Costs

3e. Institutions should acknowledge that infrastructure, training, staff support, and progress reviews for open and FAIR research have significant costs and they should have procedures in place to ensure funding of costs within the dual support system.

Strategy

10x. Institutions should demonstrate a long-term commitment to open data.

10w. Institutions should undertake regular reviews to monitor their progress in implementing their responsibilities under the Concordat, register issues to be addressed, and identify and share best practice.

Tasks grouped according to the recommendations from the UK Open Research Data Task Force report 2018

Recommendation from the ORDTF	Relevant institutional tasks listed above
Research organisations strengthen the provision of specialist support services within research organisations, and increase capacity in data stewardship, research software and data science.	5h, 6n, 9u, 9v, 8t, 6l
Funders, research organisations and publishers establish clear expectations on preservation of data and software, including the repositories to be used	8s, 8t
Research organisations and funders take steps to ensure that all researchers have access to user-friendly services, both generic and domain-specific	1b, 6k, 6m
Research organisations and funders develop, with support from Jisc, a set of principles for negotiation with commercial providers of ORD infrastructure to maximise interoperability, retain data ownership and reduce the risk of 'lock-in'	7q
Research organisations and funders take active steps to sustain and strengthen UK participation in international ORD services and initiatives.	(10x)
UKRI and other stakeholders work together to review the costs, business and funding models of current data services	3e
UKRI, funders and research organisations review levels of funding for ORD to ensure these remain appropriate to an increasingly data-rich research landscape	3e

Discussion

The Concordat is written as "principles." Principles are the source from which something originates or from which it is derived; they are not the results to be measured. Therefore, attempting to identify tasks to measure from the principles is a contestable challenge. The authors acknowledge that other interpretations of the principles are possible, but the set of tasks outlined above is likely to be an adequate basis for further work focused on institutions' implementation of the Concordat.

Grouping the tasks derived from the Concordat principles thematically reveals them to be largely focused on practical support measures that institutions should take. Comparing them with the recommendations from the Open Research Data Task Force, the Concordat tasks perhaps lack a certain breadth, for example being relatively light on funding and cost issues and on international engagement.

Conclusions

The Concordat dates from 2016 and the Open Research Data Task Force from 2018. Furthermore, all UK sector concordats and similar agreements are the subject of increasing discussion as attention is drawn to the broad, complicated and dynamic challenges faced by those trying to improve the conditions in which research is undertaken in the UK, and thereby to enable research and researchers to be the best that they can be. National and international funder policies and research assessment regimes are evolving, and it seems likely that the Open Research Data Concordat will be affected by these changes over the medium term. We hope that the work presented here will support the reviews flagged by the Concordat itself as being good practice for research organisations, and that those reviews will inform changes in national policy, assessment and funding.

References

Concordat Working Group (2016) *Concordat on Open Research Data*
https://www.ukri.org/wp-content/uploads/2020/10/UKRI-020920-ConcordatonOpenResearchData.pdf

Doran, G. T. (1981). There's a S.M.A.R.T. way to write management's goals and objectives. *Management Review*. 70 (11): 35-36. https://community.mis.temple.edu/mis0855002fall2015/files/2015/10/S.M.A.R.T-Way-Management-Review.pdf

Open Research Data Task Force (2018) Realising the Potential: Final Report of the Open Research Data Task Force. https://www.gov.uk/government/publications/open-research-data-task-force-final-report