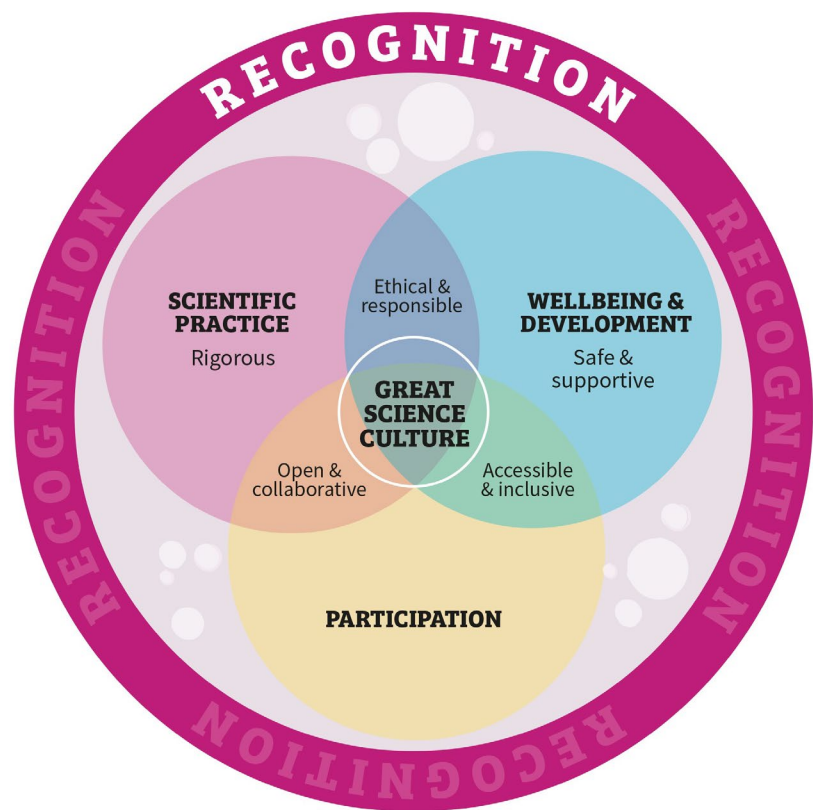




Data sharing at the RSC

- Open Science and positive science culture
- Open Data Standards
- Data Sharing
- Data Availability Statements
- Digital Discovery examples

Open science and a positive science culture



One of the three foundations of a great science culture is Scientific Practice.

Good scientific practice is achieved when science and innovation are conducted in a rigorous, ethical and responsible way. Effective science is also **open and collaborative**, involving the sharing of knowledge and expertise across disciplines, sectors and countries.

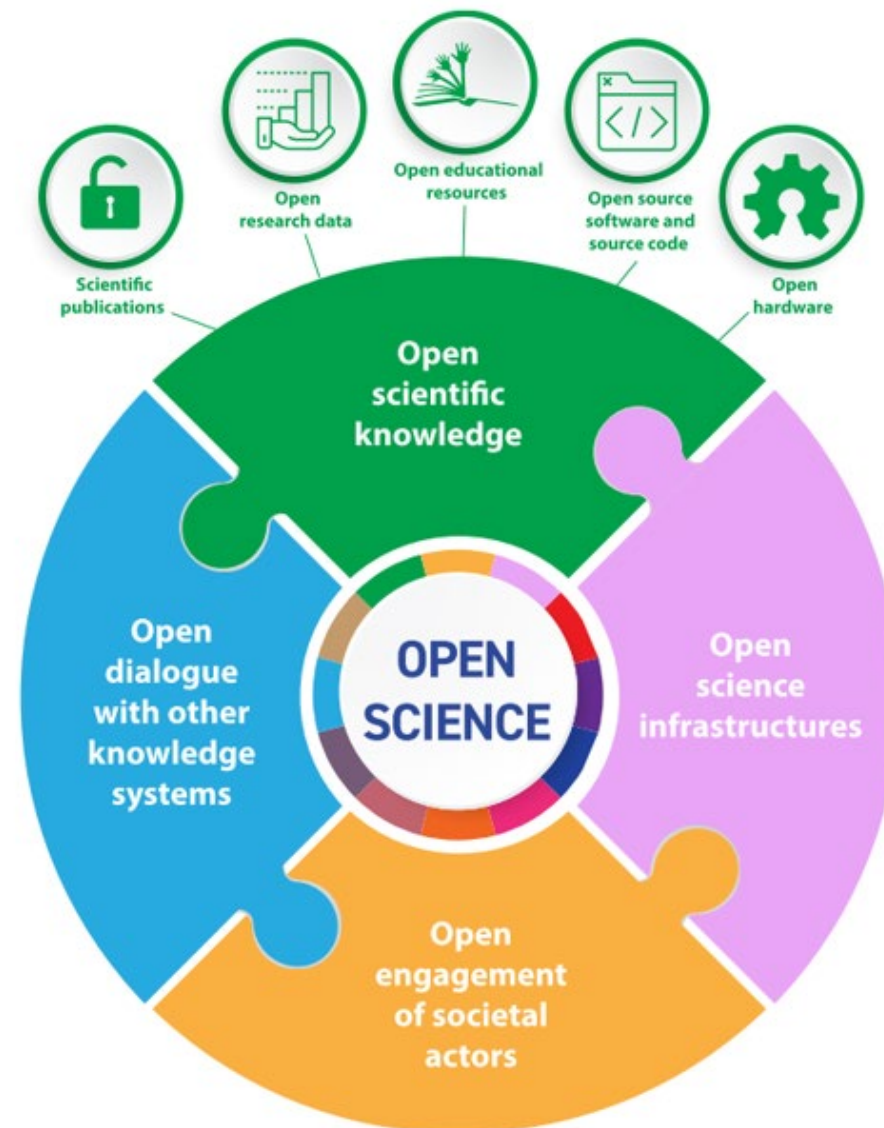
Scientific data is **deposited in an accessible repository** whenever it can be useful to others. Scientists ensure that practices such as peer review are fair and transparent, that tools and data are presented in a useful way, and that **data is FAIR (Findable, Accessible, Interoperable, Reusable)**.

What is Open Science?

"scientific knowledge **openly available, accessible and reusable** for everyone, to increase scientific collaborations and sharing of information for the benefits of science and society" – United Nations

"The principle and practice of making **research products and processes available to all**, while respecting diverse cultures, maintaining security and privacy, and **fostering collaborations, reproducibility, and equity.**" - U.S. Office of Science and Technology Policy and the National Science and Technology Council

"as open as possible, as closed as necessary."



Our Commitment to OS

Our purpose is to help the chemical science community make the world a better place; we envision a world in which the chemical sciences fulfil their potential as a force for good. We strive to **work with our community to break down barriers to advancing the chemical sciences and to ensure appropriate equitable, global access to knowledge and data.**

Read the full statement here:

<https://www.rsc.org/journals-books-databases/open-science/>

Our commitment to open science

Our purpose is to help the chemical science community make the world a better place; we envision a world in which the chemical sciences fulfil their potential as a force for good. We strive to work with our community to break down barriers to advancing the chemical sciences and to ensure appropriate equitable, global access to knowledge and data.

We believe that the best ways to enable chemical scientists to make the world a better place are through enabling greater access to knowledge, ensuring that tools and data are presented in a useful way, and that data is FAIR (Findable, Accessible, Interoperable, Reusable), that practices such as peer review are transparent, that research is carried out ethically and with integrity, that collaboration is encouraged and that ultimately more people are enabled and encouraged to participate in science.

As a group, these principles form the basis of the open science agenda – an agenda that is currently being defined by stakeholders from around the world. Our goal as a learned society, professional body, and publisher is to work with our community and stakeholders, including chemical science researchers across sectors, funders, and policymakers - to develop our vision for Open Chemistry. Together, we'll uncover barriers, assess motivations, pinpoint catalysts for change, and explore various pathways to an open future that best serves chemistry and society.

Our open science journey so far and our current initiatives

Open access →

Open access can lead us to a fairer society by making impactful research available to everyone. No matter who you are or where you live, you deserve to access and benefit from new discoveries.

Data sharing →

Discover why data sharing is important and why we believe that where possible, all data associated with the research in a manuscript should be Findable, Accessible, Interoperable and Reusable (FAIR).

Inclusion and diversity →

Our goal is to increase the diversity of people choosing and fulfilling their potential in the chemical sciences to create a truly inclusive community. Learn more about our strategy, activities and resources.

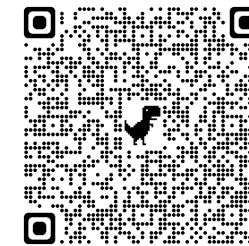
Supporting open data standards

We are actively **supporting the development of open data standards**
- we developed open chemistry ontologies, and are a founding member of InChI Trust, a structure-based chemical identifier



Data sharing in RSC journals

We provide recommendation and guidance on data sharing on our website:



[RSC data policies](#)

Our data sharing policy

The Royal Society of Chemistry believes that where possible, all data associated with the research in a manuscript should be Findable, Accessible, Interoperable and Reusable (FAIR), enabling other researchers to replicate and build on that research.

We strongly encourage authors to deposit the data underpinning their research in appropriate repositories.

For all submissions to Royal Society of Chemistry journals, any data required to understand and verify the research in an article must be made available on submission.

To comply, we suggest authors deposit their data in an appropriate [repository](#). Where this isn't possible, we ask authors to include the data as part of the article Electronic Supplementary Information (ESI).

Some journals may have additional subject requirements for both sharing and/or publishing supporting data, so please ensure you check the [journal specific guidelines](#).

Follow FAIR guidance

Deposit data where possible

Data supporting the reported findings must be shared at submission

Choose repositories over SI

Look out for journal-specific guidance

Choosing a repository



RSC recommended repositories



Selection criteria (basic)

- Register datasets with a persistent identifier (e.g. DOI)
- Provide support to include as much documentation as possible
- Facilitate linking the dataset to related items, such as a publication
- Support attribution and credit, e.g. by linking to ORCID, ROR, RAiD, etc.
- Have licensing options that are as open as possible



Tools and support

- Explore community repository registries such as

re3data.org
REGISTRY OF RESEARCH DATA REPOSITORIES

FAIRsharing.org
standards, databases, policies

- Check publisher recommendations and requirements
- Consult institutional guidelines, local librarians or research data support services
- Check funder policies and guidance

ROR: Research Organisation Registry; RAiD: Research Activity Identifier

Repository guidance at the RSC



RSC recommended repositories

Subject specific repository guide

Structural	Software, code & models	Computational simulations	Spectroscopy	Images	
Materials	Omics	Sequence data	Environmental	Life sciences	Supramolecular
Data type	Repository	URL	File / standard		
Crystal structure (organic / organometallic / metal organic) Required for all RSC journals	Cambridge Structural Database (CSD) – managed by the Cambridge Crystallographic Data Centre (CCDC)	Visit the CCDC website	Crystallographic information file (.cif)		

Generalist repository guide

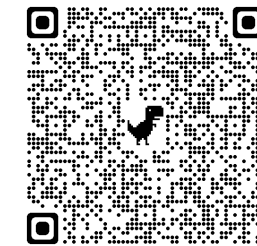
Repository Name	Information on costs	URL
Dryad Digital Repository	Fees apply	Visit the Dryad Digital Repository website

Choosing a repository

Where possible, choose **discipline-specific**, community-recognised repository

Alternatively, explore **institutional** or **generalist** repositories

Data Availability Statements at the RSC



RSC DAS guidance

“To maintain high standards of transparency, research reproducibility, and to promote the reuse of new findings, we **strongly encourage** authors to include a Data Availability Statement (DAS) as part of the final published article.”

Why DASs?

- To summarise which data is available to support your findings, where it is stored and how it can be accessed
- To demonstrate compliance with relevant funder, institution or journal data policies
- To support reproducible research and ensure availability of data for reuse

Digital Discovery



Publishes research on high-throughput methodologies for accelerated discovery of chemical and biological entities.

- AI and machine learning
- Quantum computing
- Automation

Data availability and reproducibility

Accessibility of data and code is absolutely essential for the reproducibility of computational research.

- authors strongly encouraged to deposit as much data and code
- data and code must be made available for Editors and Reviewers at the time of peer review

Data reviewer

A third researcher in addition to two regular reviewers who checks the data and code and their usability.

Data reviewer checks:

- accessibility of data
- data cleaning
- data representation
- model training and validation
- reproducibility of the code

Examples – data availability

Faster and more
loop reinforcement

The authors can



ion with double
S

Editor and reviewers

Examples – data availability

Accelerated n

by GPT

Work builds or



Examples – data availability

Issue 5, 2023

Previous Article | Next Article

Data available

All the training data is available conditionally generated. <https://github.com>

ditionally and
ub:

