

# RDMkit walkthrough

Flora D'Anna, PhD – Data steward at ELIXIR Belgium



ELIXIR-CONVERGE is funded by the European Union's Horizon 2020 research and innovation programme under grant agreement No 871075.



# Outline

- **What is RDMkit**
- How to use it: possible scenarios starting from
  - data life cycle stage
  - topics related to a role
  - a specific domain
  - the search bar
- How to contribute

# RDMkit: Research Data Management toolkit for life sciences

**RDMkit**

---



- Research data management (DM) best practices and guidelines.
- What you need to consider or take into account when performing data management tasks.
- Practical approaches to navigate the plethora of available RDM resources (tools, standards, databases, registries, etc).

# RDMkit content



Data management

About

Contribute

 GitHub



Search RDMkit



## Data life cycle

Data management aspects to be considered for each stage of the data life cycle.



## Your tasks

Considerations, solutions and tools for tackling common data management tasks.



## Your domain

Best practices and tools used in your domain or by your research community.



## Tool assembly

Examples of tools or software assembled into an ecosystem to cover research data management tasks.



## National resources

Pointers to Country-specific information resources and national best practices and guidelines.



## Your role

Useful guidelines and resources for your job and career path.

# The “what and why” of data management throughout the stages of the data life cycle



Data management

About

Contribute

GitHub



Search RDMkit



## Data life cycle

Data management aspects to be considered for each stage of the data life cycle.



# What to consider when performing common data management tasks and possible solutions



Data management

About

Contribute

 GitHub



Search RDMkit



## Data life cycle

Data management aspects to be considered for each stage of the data life cycle.



## Your tasks

Considerations, solutions and tools for tackling common data management tasks.

# Domain-specific best practices, tools and resources



Data management

About

Contribute

 GitHub



Search RDMkit



## Data life cycle

Data management aspects to be considered for each stage of the data life cycle.



## Your tasks

Considerations, solutions and tools for tackling common data management tasks.



## Your domain

Best practices and tools used in your domain or by your research community.

# Examples of tools assembled into an ecosystem for research data management



Data management

About

Contribute

 GitHub



Search RDMkit



## Data life cycle

Data management aspects to be considered for each stage of the data life cycle.



## Your tasks

Considerations, solutions and tools for tackling common data management tasks.



## Your domain

Best practices and tools used in your domain or by your research community.



## Tool assembly

Examples of tools or software assembled into an ecosystem to cover research data management tasks.

# Summary or overview of data management resources of interest for a specific Country



Data management

About

Contribute

 GitHub



Search RDMkit



## Data life cycle

Data management aspects to be considered for each stage of the data life cycle.



## Your tasks

Considerations, solutions and tools for tackling common data management tasks.



## Your domain

Best practices and tools used in your domain or by your research community.



## Tool assembly

Examples of tools or software assembled into an ecosystem to cover research data management tasks.



## National resources

Pointers to Country-specific information resources and national best practices and guidelines.



# Guidelines, tools and resources of interest for a specific role and career path



Data management

About

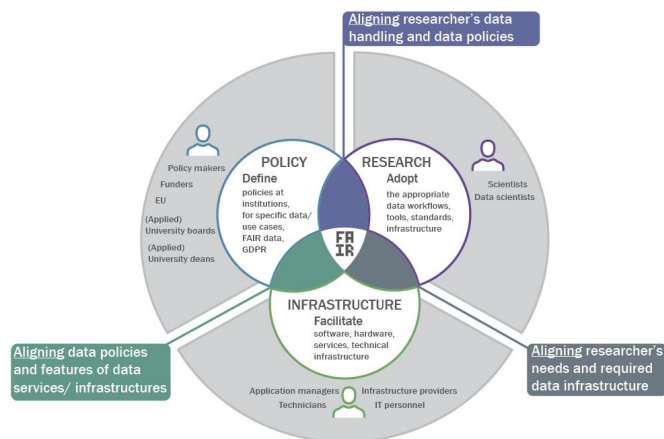
Contribute

GitHub



Search RDMkit

## Implementation areas for data stewardship



### Your role

Useful guidelines and resources for your job and career path.

Mijke Jetten, & Celia W.G. van Gelder. (2021, May 18). ELIXIR webinar: towards professionalising data stewardship. Zenodo. <https://doi.org/10.5281/zenodo.4769858>



# Use the search bar to find what you are looking for




# Outline

- What is RDMkit
- **How to use it: some possible scenarios starting from**
  - data life cycle stage
  - topics related to a role
  - a specific domain
  - the search bar
- How to contribute

# What you need to consider when transferring/sharing data

**Data management**

Data life cycle



Your role

Your domain

Your tasks

Tool assembly

National resources

Data life cycle

## Sharing

- What is data sharing?
- Why is data sharing important?
- What should be considered for data sharing?
- Related pages
- More information

### What is data sharing?


Sharing data means making your data known to other people.

You can share your data with collaboration partners in the context of a research project, or you can publish your data to share it with the community and society at large.

# What you need to consider when transferring/sharing data

**Data management**

Data life cycle



Your role

Your domain

Your tasks

Tool assembly

National resources

Data life cycle

## Sharing

- What is data sharing?
- Why is data sharing important?
- What should be considered for data sharing?
- Related pages
- More information

### What is data sharing?

Sharing data means making your data known to other people.

You can share your data with collaboration partners in the context of a research project, or you can publish your data to share it with the community and society at large.

## Related pages

**Your tasks**

**Data protection**  
How to make research data compliant to gdpr.

**Data publication**  
Prepare data and find repositories for publication.

**Data transfer**  
How to transfer data files.

**Identifiers**  
How to use identifiers for research data.

**Licensing**  
How to license research data.

**Documentation and metadata**  
How to document and describe your data.


**Sensitive data**  
How to identify different research data types.

**Data storage**

# What you need to consider when transferring/sharing data

**Data management**

Data life cycle



Your role

Your domain

Your tasks

Tool assembly

National resources

Data life cycle

## Sharing

- What is data sharing?
- Why is data sharing important?
- What should be considered for data sharing?
- Related pages
- More information

### What is data sharing?

Sharing data means making your data known to other people.

You can share your data with collaboration partners in the context of a research project, or you can publish your data to share it with the community and society at large.

## Related pages

**Your tasks**

Data protection

How to make research data compliant to gdpr.

Data publication

Prepare data and find repositories for publication.

Data transfer

How to transfer data files.

Identifiers

How to use identifiers for research data.

Licensing

How to license research data.

Documentation and metadata

How to document and describe your data.

Sensitive data

How to identify different research data types.

Data storage

Your tasks

## Data transfer

- How do you transfer large data files?
- More information
- Relevant tools and resources

### How do you transfer large data files?

#### Description

Often, research in Life Sciences generates massive amounts of digital data, such as output files of 'omics' techniques (genomics, transcriptomics, metabolomics, proteomics, etc). Large data files cannot be sent by email because they exceed the file size limit of most common email servers. So, how can large data files be transferred from a local computer to a distant one?

#### Considerations

There are many aspects to consider when dealing with data transfer.

- The **size or volume** of the data and the **capacity or bandwidth of the network** that links your local computer with the distant computer are crucial aspects. Data size and bandwidth are tightly linked since transferring large volumes of data on a low bandwidth network will be so time consuming that it could be simpler to send the data on a hard drive through carrier services.
- Ensure that you are aware of the **legal and ethical implications** of your data transfer
  - For data concerning persons, you have to safeguard compliance with various legal and ethical frameworks, including

# What you need to consider when transferring/sharing data

**Data management**

Data life cycle



Your role

Your domain

Your tasks

Tool assembly

National resources

Data life cycle

## Sharing

- What is data sharing?
- Why is data sharing important?
- What should be considered for data sharing?
- Related pages
- More information

### What is data sharing?

Sharing data means making your data known to other people.

You can share your data with collaboration partners in the context research project, or you can publish your data to share it with the community and society at large.

## Related pages

**Your tasks**

Data protection

How to make research data compliant to gdpr.

Data publication

Prepare data and find repositories for publication.

Data transfer

How to transfer data files.

Identifiers

How to use Identifiers for research data.

Licensing

How to license research data.

Documentation and metadata

How to document and describe your data.

Sensitive data

How to identify different research data types.

Data storage

**Your tasks**

## Data transfer

- How do you transfer large data files?
- More information
- Relevant tools and resources

### How do you transfer large data files?

**Description**

Often, research in Life Sciences generates massive amounts of digital data, such as output files of 'omics' techniques (genomics, transcriptomics, metabolomics, proteomics, etc). Large data files cannot be sent by email because they exceed the file size limit of most common email servers. So, how can large data files be transferred from a local computer to a distant one?

**Considerations**

There are many aspects to consider when dealing with data transfer.


- The **size or volume** of the data and the **capacity or bandwidth of the network** that links your local computer with the distant computer are crucial aspects. Data size and bandwidth are tightly linked since transferring large volumes of data on a low bandwidth network will be so time consuming that it could be simpler to send the data on a hard drive through carrier services.
- Ensure that you are aware of the **legal and ethical implications** of your data transfer
  - For data concerning persons, you have to safeguard compliance with various legal and ethical frameworks, including


## More information


**Training**


 Training in TeSS

**Links to other ELIXIR resources**

 Step-by-step process for: Transferring data with SFTP

 Step-by-step process for: Downloading data with Aspera

 Support for DMP on: How will the raw data be transported?

 Support for DMP on: How will your first data come in?

# What you need to consider when transferring/sharing data

**Data management**

**Data life cycle**

**Sharing**

- What is data sharing?
- Why is data sharing important?
- What should be considered for data sharing?
- Related pages
- More information

**What is data sharing?**

Sharing data means making your data known to other people.

You can share your data with collaboration partners in the context research project, or you can publish your data to share it with the community and society at large.

## Related pages

**Your tasks**

**Data protection**  
How to make research data compliant to gdpr.

**Data publication**  
Prepare data and find repositories for publication.

**Data transfer**  
How to transfer data files.

**Identifiers**  
How to use identifiers for research data.

**Licensing**  
How to license research data.

**Documentation and metadata**  
How to document and describe your data.

**Sensitive data**  
How to identify different research data types.

**Data storage**

**Your tasks**

**Data transfer**

- How do you transfer large data files?
- More information
- Relevant tools and resources

**How do you transfer large data files?**

**Description**

Often, research in Life Sciences generates massive amounts of digital data, such as output files of 'omics' techniques (genomics, transcriptomics, metabolomics, proteomics, etc). Large data files cannot be sent by email because they exceed the file size limit of most common email servers. So, how can large data files be transferred from a local computer to a distant one?

**Considerations**

There are many aspects to consider when dealing with data transfer.

- The **size or volume** of the data and the **capacity or bandwidth of the network** that links your local computer with the distant computer are crucial aspects. Data size and bandwidth are tightly linked since transferring large volumes of data on a low bandwidth network will be so time consuming that it could be simpler to send the data on a hard drive through carrier services.
- Ensure that you are aware of the **legal and ethical implications** of your data transfer
  - For data concerning persons, you have to safeguard compliance with various legal and ethical frameworks, including

## More information

**Training**

Training in TeSS

**Links to other ELIXIR resources**

Step-by-step process for: Transferring data with SFTP

Step-by-step process for: Downloading data with Aspera

Support for DMP on: How will the raw data be transported?

Support for DMP on: How will your first data come in?

## Relevant tools and resources

Tool or resource	Description	Related pages	Registry
<b>Amazon Web Services</b>	Amazon Web Services	Data storage   Data analysis	Training
<b>Aspera Fasp</b>	With fast file transfer and streaming solutions built on the award-winning IBM FASP protocol, IBM Aspera software moves data of any size across any distance	Data Steward: Infrastructure	
<b>Box</b>	Cloud storage and file sharing service	Data storage   Data Steward: Infrastructure	Training
<b>cURL</b>	command line tool and library for transferring data with URLs	Data Steward: Infrastructure	
<b>DropBox</b>	Cloud storage and file sharing service	Data storage   Data Steward: Infrastructure	
<b>FileZilla</b>	A free FTP solution	Data Steward: Infrastructure	
<b>Globus</b>	Globus lets you efficiently, securely, and reliably transfer data directly between systems separated by an office wall or an ocean. Focus on your research and offload your data transfer headaches to Globus	Data Steward: Infrastructure	
<b>Google Drive</b>	Cloud Storage for Work and Home	Data storage	
<b>ICloud</b>	Data sharing	Data storage   Data analysis	
<b>Image Data Resource (IDR)</b>	A repository of image datasets from scientific publications	Microbial biotechnology   Data publication	Tool Info   Standards/Databases

## National resources

**Belnet**

Belnet is the privileged partner of higher education, research and administration for connectivity. We provide high-bandwidth internet access and related services for our specific target groups.

Data Steward: research  
Researcher  
Data Steward: Infrastructure

**NIRD**

The National Infrastructure for Research Data (NIRD) infrastructure

Data storage | NelS

**View national resources** 2

# Policy-related topics: measuring compliance

## Data management

Data life cycle ▾

Your role ▲

Researcher

Data Steward:  
policy

Data Steward:  
research

Data Steward:  
infrastructure

Your role

## Data Steward: policy ✎ ! 🔗

- [Description](#)
- [Focus](#)
- [Learning path](#)
- [Related pages](#)
- [More information](#)
- [Relevant tools and resources](#)

## Related pages

### Your tasks

**Compliance monitoring & measurement**  
Measure compliance to data management regulations and standards.

**Licensing**  
How to license research data.

**Data management plan**  
How to write a data management plan (dmp).

**Data protection**  
How to make research data compliant to gdpr.

**Sensitive data**  
How to identify different research data types.

Your tasks

## Compliance monitoring & measurement ✎ ! 🔗

- [How can you measure and document data management capabilities?](#)
- [How can you ethically access genetic resources of another country?](#)
- [Related pages](#)
- [Relevant tools and resources](#)

## Related pages

### Tool assembly

#### TransMed

Transmed from elixir luxembourg supports projects in translational biomedicine for clinical and translational projects.

## Relevant tools and resources

Tool or resource ⓘ	Description	Related pages
<a href="#">FAIR Cookbook</a>	FAIR Cookbook is an online resource for the Life Sciences with recipes that help you to make and keep data Findable, Accessible, Interoperable and Reusable (FAIR)	<a href="#">Data Steward: research</a> <a href="#">TransMed</a>
<a href="#">FAIR Evaluation Services</a>	Resources and guidelines to assess the FAIRness of digital resources.	<a href="#">Data Steward: research</a> <a href="#">Data Steward: policy</a>
<a href="#">FAIRassist.org</a>	Help you discover resources to measure and improve FAIRness.	<a href="#">Data Steward: research</a> <a href="#">Data Steward: policy</a>

# What standards are used for bioimaging?

## Data management

Data life cycle



Your role



Your domain



Plant sciences

Marine metagenomics

Human data

Biomolecular simulation data

Intrinsically disordered proteins

Microbial biotechnology

Epitranscriptome data

Proteomics

Toxicology data

Bioimaging data

Structural bioinformatics

Your domain

## Bioimaging data

- [Introduction](#)
- [What constitutes bioimage data](#)
- [Standard \(meta\)data formats](#)
- [\(Meta\)Data collection](#)
- [Data publication and archiving](#)
- [Related pages](#)
- [Relevant tools and resources](#)

## Introduction

Bioimaging specialists are acquiring an ever growing amount of data: However, image data management often does not receive the attention it deserves since it is considered a burdensome task. At the same time, storing images is no longer an option, assuming it ever was! Data volume is exploding and acquired images need storing but potentially processed images will be alongside the original images. It is critical to proactively identify where the data will be stored, who will cover the cost of the hardware, and who will cover the cost of the software. It is critical to proactively identify where the data will be stored, who will cover the cost of the hardware, and who will cover the cost of the software. It is critical to proactively identify where the data will be stored, who will cover the cost of the hardware, and who will cover the cost of the software.

## Related pages

### Your tasks

#### Data management plan

How to write a data management plan (dmp).

#### Data organisation

Best practices to name and organise research data.

#### Data publication

Prepare data and find repositories for publication.

#### Existing data

How to find and reuse existing data.

### Tool assembly

#### OMERO

OmERO is a software platform for managing, sharing and analysing images data.

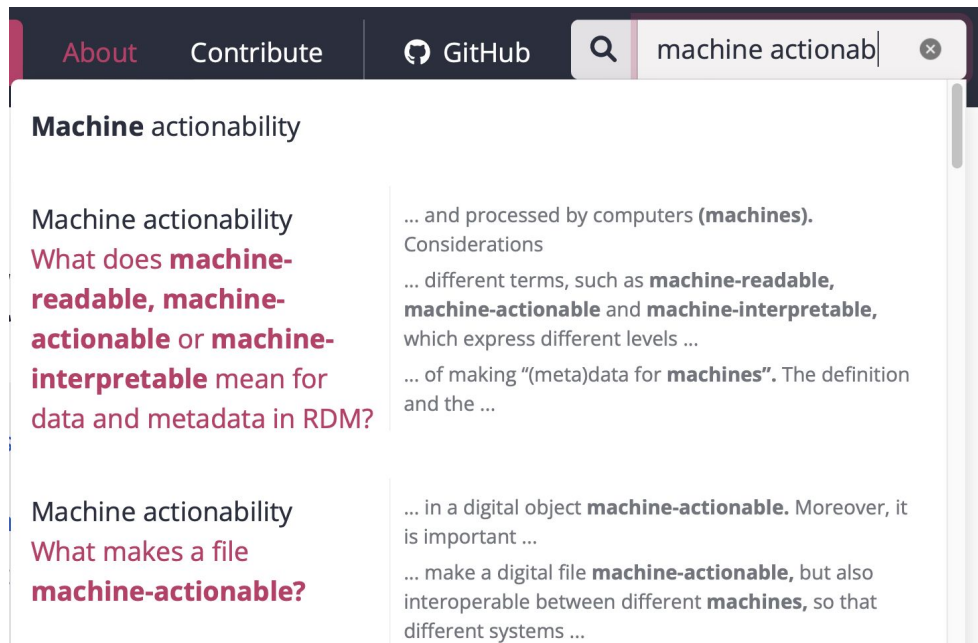
#### XNAT-PIC

Xnat for preclinical imaging centers (xnat-pic) is a set of tools to store, process and share preclinical imaging studies built on top of the xnat imaging informatics platform.

## Relevant tools and resources

Tool or resource	Description	Related pages	Registry
<a href="#">b2share</a>	Store and publish your research data. Can be used to bridge between domains	<a href="#">Data storage</a> <a href="#">Data publication</a>	<a href="#">Standards/Databases</a>
<a href="#">Bio-Formats</a>	Bio-Formats is a software tool for reading and writing image data using standardized, open formats	<a href="#">OMERO</a>	<a href="#">Tool info</a> <a href="#">Training</a>
<a href="#">BioImageArchive</a>	The BioImage Archive stores and distributes biological images that are useful to life-science researchers.	<a href="#">Data publication</a>	<a href="#">Standards/Databases</a>
<a href="#">BisQue</a>	Resource for management and analysis of 5D biological images	<a href="#">Data organisation</a> <a href="#">Data Steward: research</a> <a href="#">Data analysis</a>	<a href="#">Tool info</a>
<a href="#">Cytomine-IMS</a>	Image Data management	<a href="#">Data Steward: research</a>	

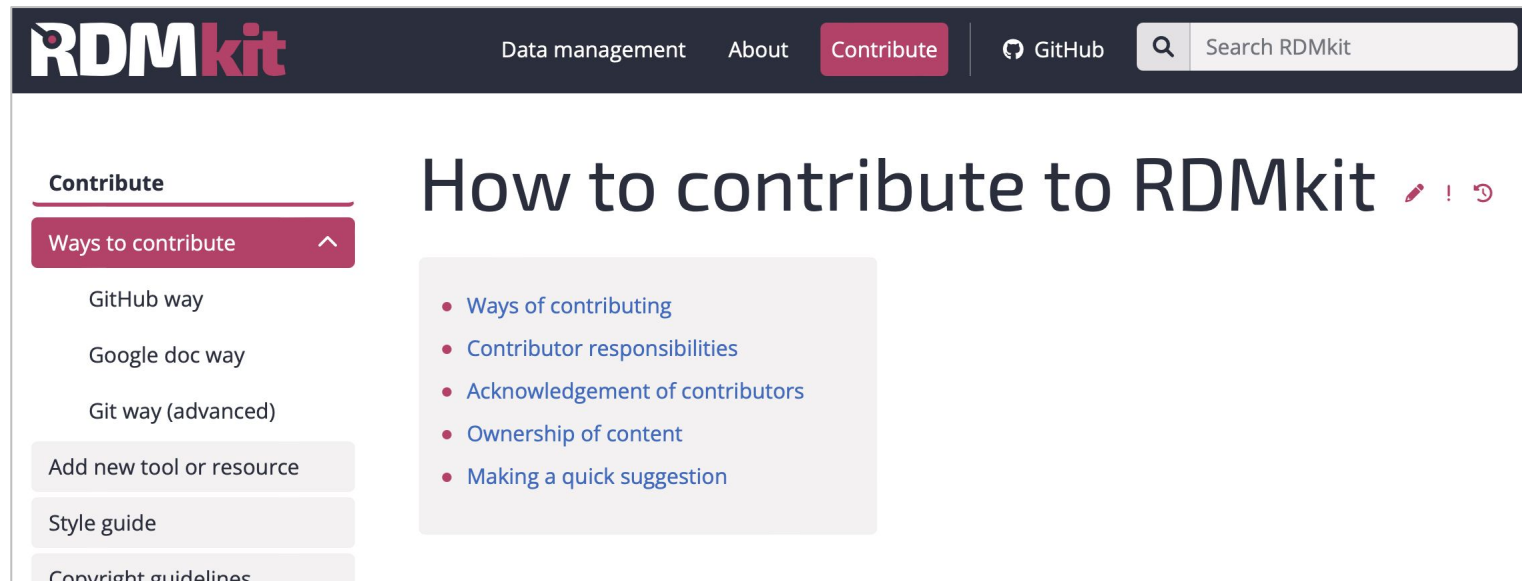
# What does “machine-actionable” mean in life sciences research data management?



# Outline

- What is RDMkit
- How to use it: some possible scenarios starting from
  - data life cycle stage
  - topics related to a role
  - a specific domain
  - the search bar
- **How to contribute**

# How to contribute to RDMkit



The screenshot shows the RDMkit website with a dark blue header. The header contains the RDMkit logo, navigation links for 'Data management', 'About', and 'Contribute' (highlighted in a pink box), a 'GitHub' link, and a search bar labeled 'Search RDMkit'. The main content area has a title 'How to contribute to RDMkit' with edit, warning, and refresh icons. A left sidebar under the 'Contribute' heading lists 'Ways to contribute' (expanded), 'Add new tool or resource', 'Style guide', and 'Copyright guidelines'. The 'Ways to contribute' section lists: 'GitHub way', 'Google doc way', 'Git way (advanced)', 'Add new tool or resource', 'Style guide', and 'Copyright guidelines'. A central box lists five bullet points: 'Ways of contributing', 'Contributor responsibilities', 'Acknowledgement of contributors', 'Ownership of content', and 'Making a quick suggestion'. To the left of the sidebar are two pink icons: a GitHub cat and a document.

**RDMkit** Data management About **Contribute** GitHub Search RDMkit

## How to contribute to RDMkit

**Contribute**

Ways to contribute ^

- GitHub way
- Google doc way
- Git way (advanced)
- Add new tool or resource
- Style guide
- Copyright guidelines

- Ways of contributing
- Contributor responsibilities
- Acknowledgement of contributors
- Ownership of content
- Making a quick suggestion

RDMkit editors: [rdm-editors@elixir-europe.org](mailto:rdm-editors@elixir-europe.org)

# Open contribution and editorial process:

 Join RDMkit!

## RDMkit in numbers

**132**

### Contributors

The force behind RDMkit

**390**

### Tools & resources

Explained in the context of  
real world problems

**89**

### Pages

Helping you with data  
management

**Recommended in the Horizon Europe Program Guide** as the resource for  
“Data management guidelines and good practices for the Life Sciences”.

# Acknowledgements

- RDMkit editorial board
- RDMkit contributors
- Member of ELIXIR-CONVERGE project
- Funders



This project has received funding from the European Union's Horizon 2020 programme under grant agreement No 871075 (ELIXIR-CONVERGE).



This work is licensed under the Creative Commons Attribution 4.0 International License, except where otherwise noted. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.



# Thank you

Any questions?



*ELIXIR-CONVERGE is funded by the European Union's  
Horizon 2020 research and innovation programme under  
grant agreement No 871075.*

