What does a modern laboratory need (e.g., tools/infrastructure) in order to be prepared to implement the policy?

Routine Data RDM & Sharing at the Investigator level

Professor Carole Goble
The University of Manchester, UK
ELIXIR European Research Infrastructure
carole.goble@manchester.ac.uk

SESSION III. ENCOURAGING UPTAKE OF DATA SHARING IN THE SCIENTIFIC COMMUNITY

Data and Metadata Journey

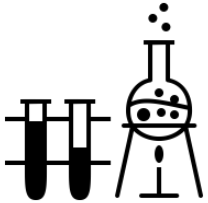
(Access to) support for (most of) the steps of data's life cycle and flow.

Multiple tools/platforms

(Ideally) managed platforms and resources that are smoothly joined up.

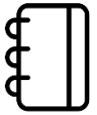
Mix of local & cloud based services to make an interlinked ecosystem of tools (a Research Commons).

Sociology of staged sharing



Lab-level RDM

Files from facilities, instruments, trials
File stores, LIMS, Secure stores
Repositories,
Long term Preservation, Data movement

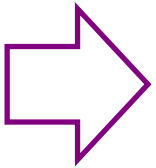


ELNs

Team-level Cataloguing & Sharing

Catalogue & organise across repositories & studies
Metadata tracking, consent tracking
Deposition to shared (public) repositories
Data Brokering, Secure Gate-keeping

Data and metadata flows



Metadata Templates & Metadata Transactions

Spreadsheets, scripts and protocols
Data staging points, metadata standards
Reporting for you, friends & strangers

Sharing flows

Labs -> Team Hubs -> Public
Data Hubs

Project Hub

Health Atlas | Browse | Help | Search here... | Search | Register | Log in

Home / Projects Index / GLA - German Lymphoma Alliance

GLA - German Lymphoma Alliance

Description:

B-cell lymphoma is the most frequent lymphoid neoplasia accounting for one third of all lymphomas. R-CHOP (rituximab, cyclophosphamide, doxorubicin, vincristine and prednisolone) is the current treatment standard, leading to high cure rates especially in younger patients. Although the 3-year event-free survival is about 60% even for elderly patients between 60-80 years old, the remainder of patients eventually relapse and the majority die of their disease. T
German Non-Hodgkin's Lymphoma Study Group (DSHNHL) was founded in 1993 and has since conducted more than 25 national and international trials including more than 8000 patients in first line and salvage therapies. The DSHNHL is a network of haematooncologists, pathologists, radiologists, radiotherapists, nuclear physicians, biologists, geneticists and biometricians. All results of clinical trials performed by the DSHNHL were published in international journals and could lead a contribution to improve the treatment results.

Programme: NHL

Health Atlas ID: 7RX4165T77-8

Funding codes:
- Deutsche Krebshilfe
- BMBF

Public web page: https://www.dshnhl.org

Organisms: Homo sapiens

Human Diseases: Non-hodgkin lymphoma, B-cell lymphoma

Health Atlas PALs: Markus Löffler

Project Coordinators: *No Project coordinators for this Project*

Project created: 16th Apr 2019

Project Hierarchy

- GLA - German Lymphoma Alliance
  - MMML - Molecular mechanisms in malignant lymphoma
  - MMML-MYC-SYS
  - MMML Demonstrators - Molecular Mechanisms in Malignant Lymphomas - Demonstrators of Personalized Medicine

Related items

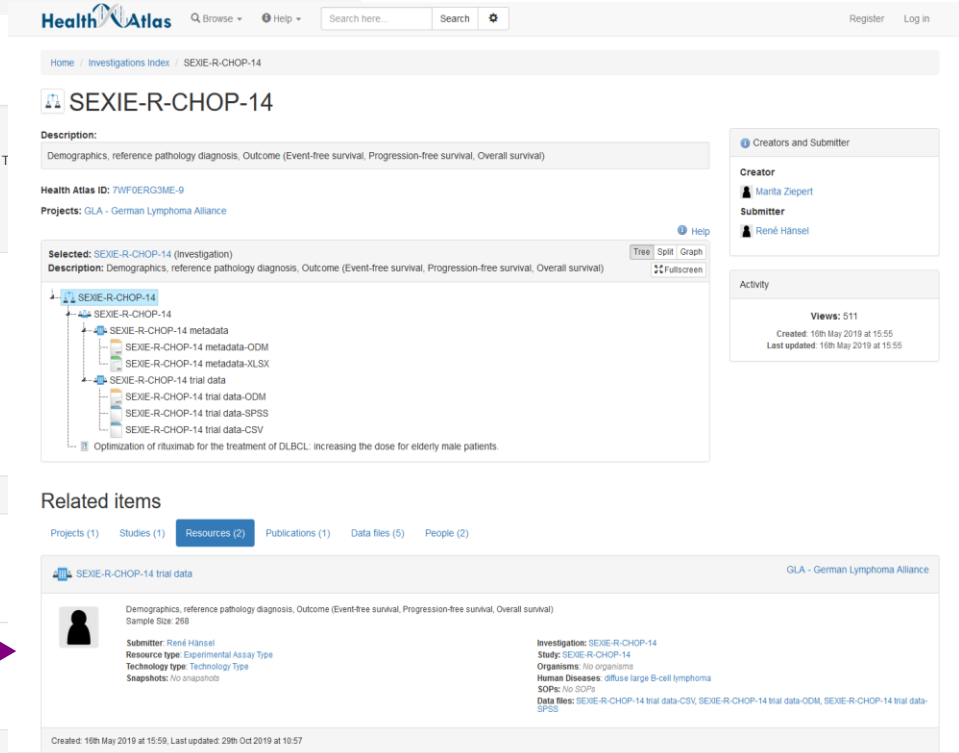
Institutions (3) | Investigations (17) | Studies (15) | Resources (16) | Publications (74) | Data files (91+1) | Models (2) | People (10)

Institute for Medical Informatics, Statistics and Epidemiology

imise.

Country: German
City: Leipzig
Web page: http://www

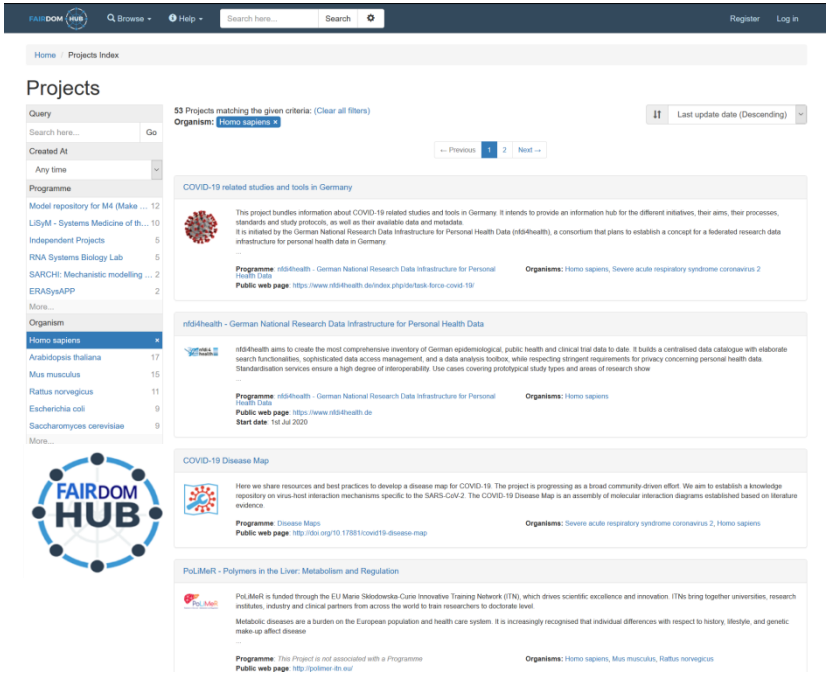
Interdisciplinary Center for Bioinformatics



Investigations (17) | Studies (15) | Resources (16) | Publications (74) | Data files (91+1) | Models (2) | People (10) | Human Diseases (3)

https://www.health-atlas.de/ **(FAIRDOM-SEEK platform https://fair-dom.org)**

FAIRDOM



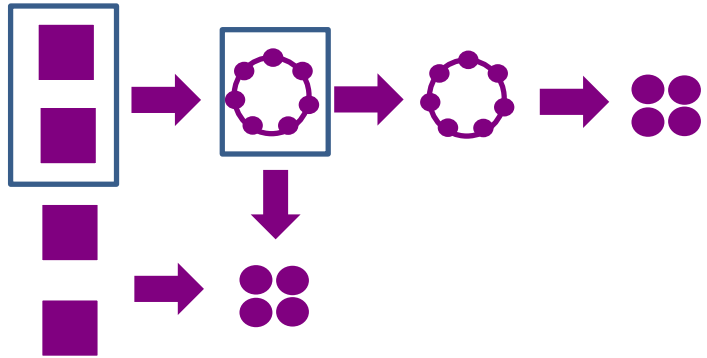
Investigation-Study-Assay/Analysis

FAIRDOM

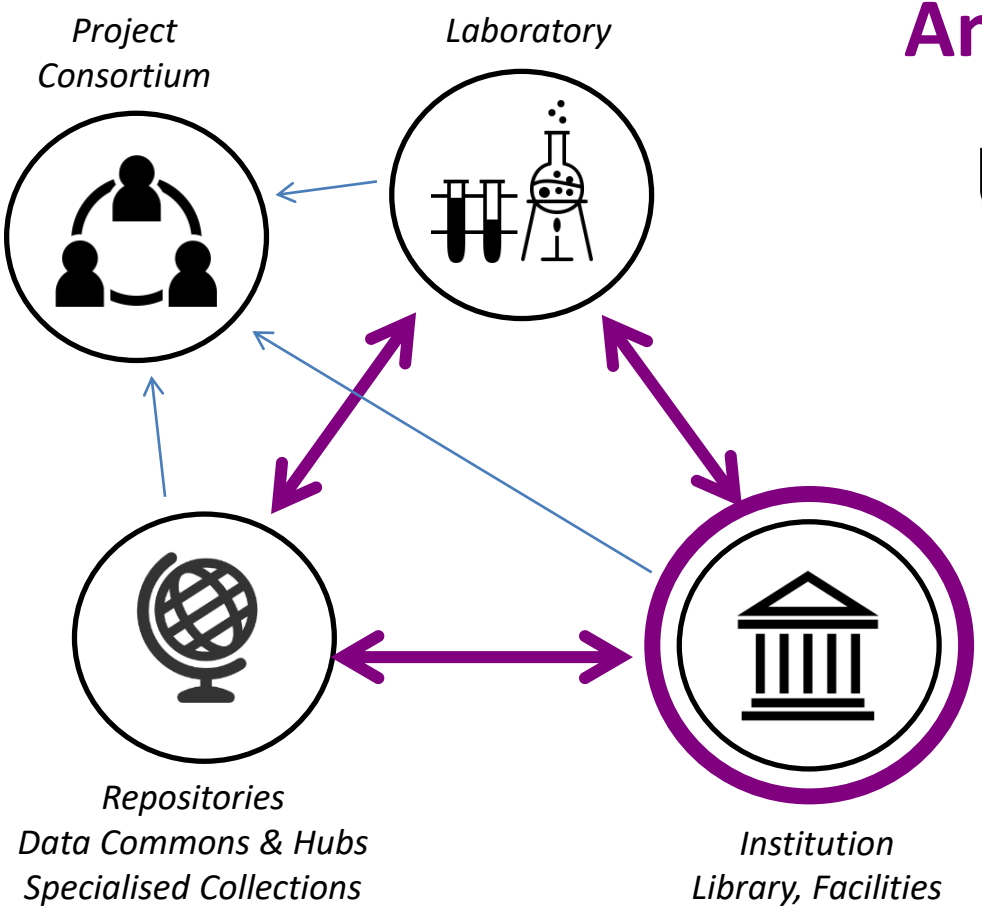
Support and encourage **incremental staged sharing** & permission control

- To get trust in data sharing and support team sharing

Support Data Staging & pipelines



Private → Public



An interlinked triangle

Trusted Research Environments

Access to analysis platforms

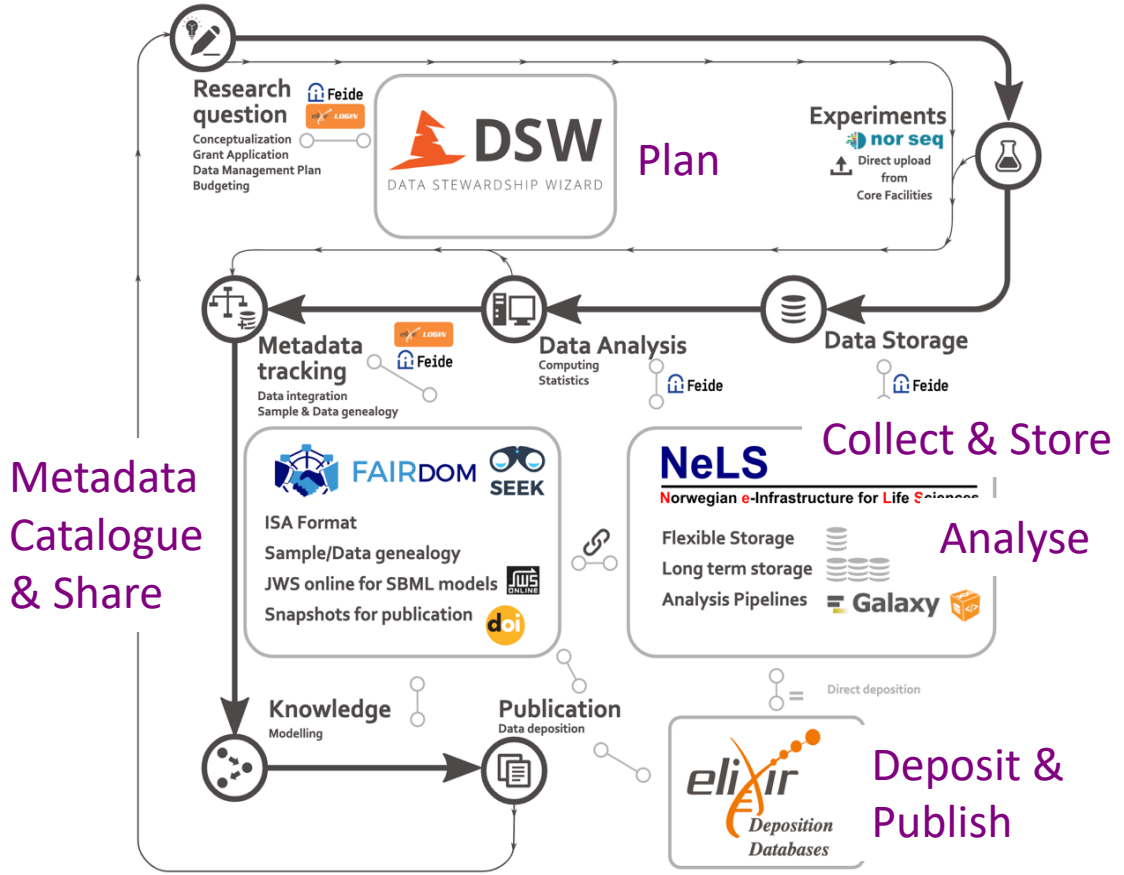
Cloud Store, Synch and Share services, AAI

Frictionless connections

Reduce the friction of data brokering, data visiting

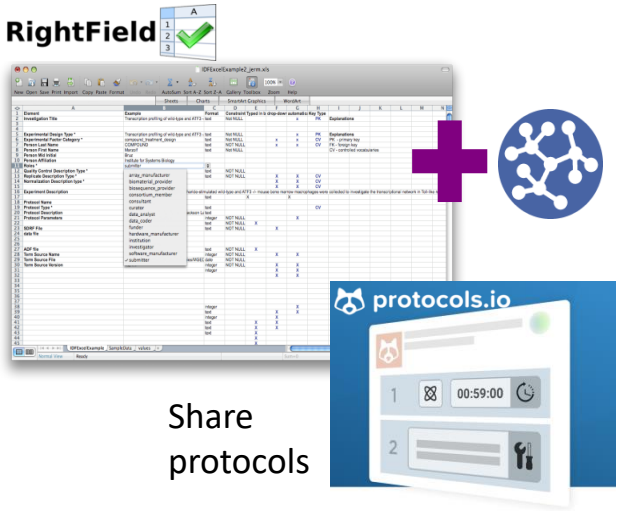
Leverage the resources and people

Blueprints & Examples



Norwegian e-Infrastructure for Life Sciences Data Management Tool Assembly

Nudge Points

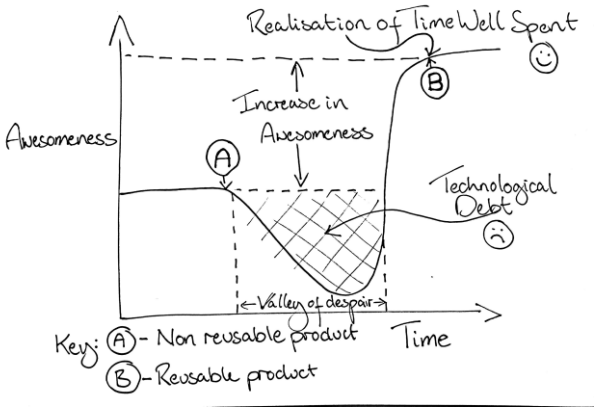


Share protocols

http://www.rightfield.org.uk
http://protocols.io

Data Stewardship in the Lab

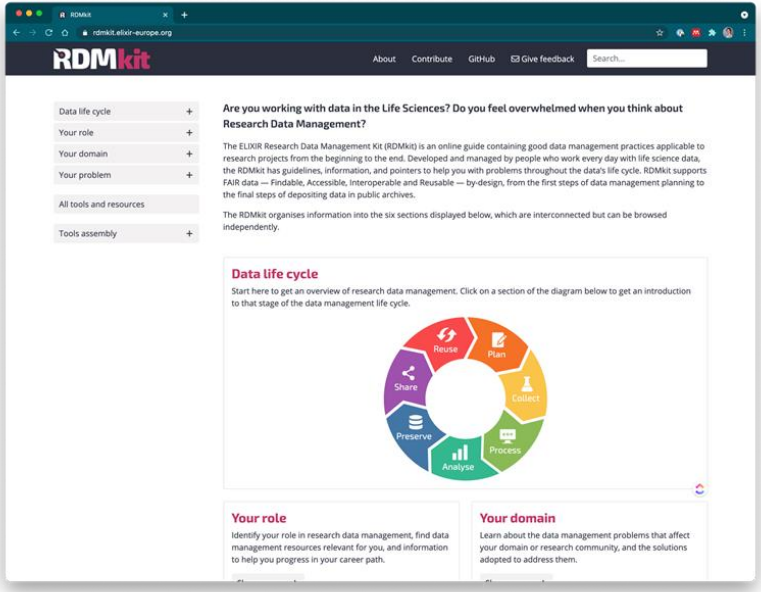
What can be automated?
Understand the (meta)data and its flows
Build in the processes and get them followed



Some help

RDMkit

https://rdmkit.elixir-europe.org



Guides, examples, tools, pointers, assembled and written by Life Science researchers and data stewards

elixir CONVERGE

Data life cycle -

Your role +

Your domain +

Your problem +

All tools and resources

Tools assembly +

Reusing

Edit me

- What is data reuse?
- Why is data reuse important?
- What should be considered for data reuse?
- Problems to be addressed at this stage
- Where can training materials and events about data reuse be found?

What is data reuse?

Data reuse means using data for other purposes than it was originally collected for. Reuse of data is particularly important in science, as it allows different researchers to analyse and publish findings based on the same data independently of one another. Reusability is one key component of the FAIR principles.

Data that is well-described, curated and shared under clear terms and conditions is more likely to be reused. Integration with other data sources is also important, since that can enable new, yet unanticipated, uses for the data.

Data life cycle +

Your role +

Your domain -

Plant sciences

Marine metagenomics

Human data

Biomolecular simulation data

Intrinsically disordered proteins

Microbial biotechnology

Your problem +

All tools and resources

Tools assembly +

Human data

Edit me

- Introduction
- Planning for, and collection of, human research data
- Processing and analysing human research data
- Preserving human research data
- Sharing & reusing of human research data
- Relevant tools and resources
- Training materials on the management of human-subject data

Introduction

When you do research on data derived from human individuals, there
the data life cycle. Note that much of the topics discussed on this pag

Data life cycle +

Your role +

Your domain +

Your problem +

All tools and resources

Tools assembly -

NeLS

Covid-19 Data Portal

COVID-19 Data Portal

Edit me

- What is the European COVID-19 Data Portal?
- How the portal is useful for researchers and how it is supposed to fit into their processes?
- What are the components for the COVID-19 Data Portal?
- What country specific instances are there and how new instances are deployed?

What is the European COVID-19 Data Portal?

The European COVID-19 Data Platform was launched to facilitate the urgent need to share and analyse COVID-19 data and thus accelerate research that will provide responses and build solutions, such as vaccines, treatments and public health interventions. The Platform comprises three core components, the SARS-CoV-2 Data Hubs, the Federated European Genome-phenome Archive and the COVID-19 Data Portal. The COVID-19 Data Portal brings together and continuously updates relevant COVID-19 datasets from a breadth of analytical platforms. Data are submitted using the SARS-CoV-2 Data Hubs functions or via other major centres

DSW
DATA STEWARDSHIP WIZARD

elixir CONVERGE

Take home points

- Know (meta)data flows and identify metadata nudge points
- Stores AND Catalogue
- Processes + Templates go with the Tools
- Lab + Institution + Public Commons triangle
- Support the staging of sharing
- RDM help by folks who in labs and facilities RDMkit