KNOWLEDGE 2 CONNECT

UNIVERSITEIT GENT

UZ

Knowledge 2 Connect, KCGG - May 27, 2016

Research Data Management: an Introduction

—

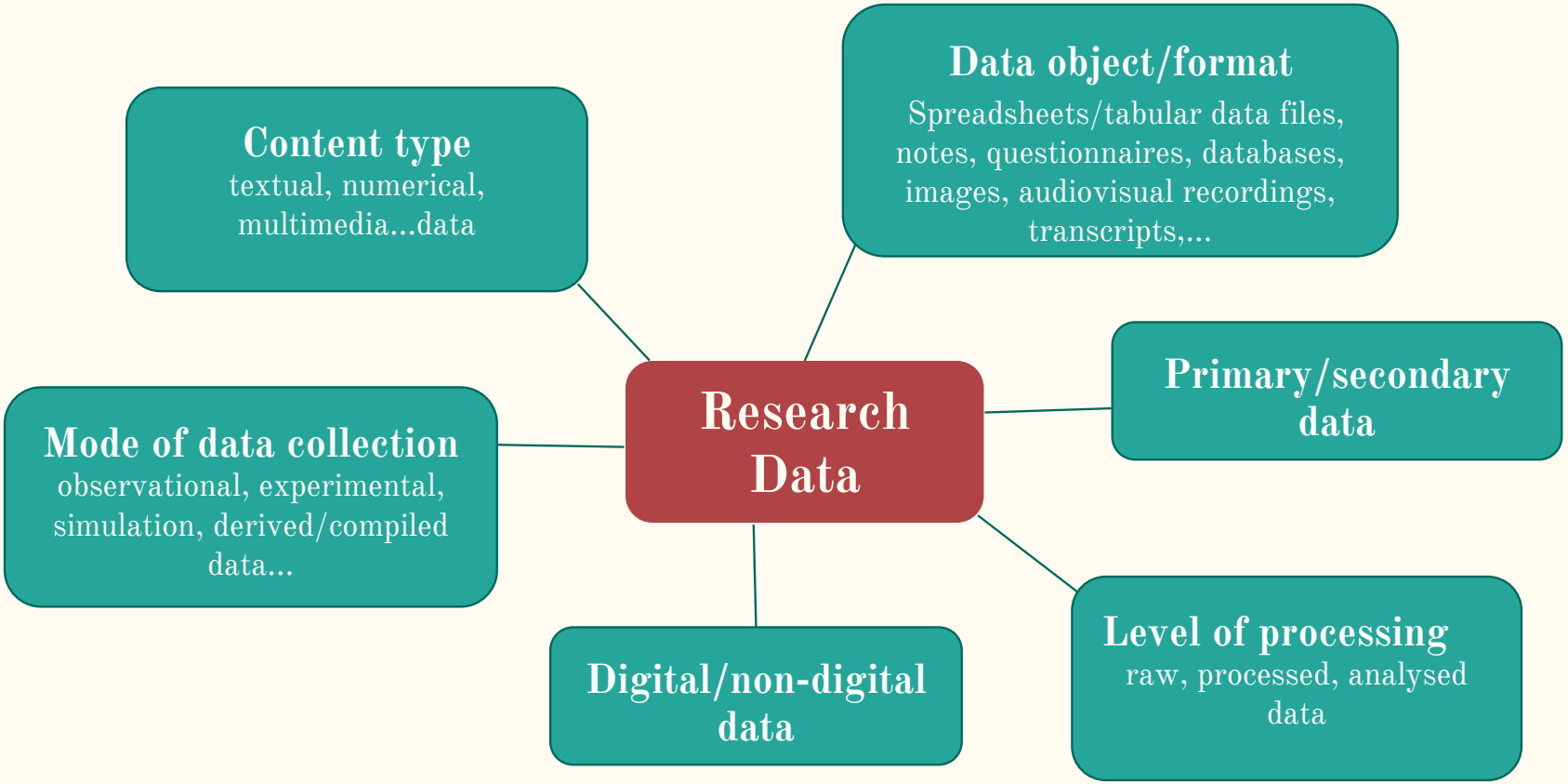
Dr. Myriam Mertens
Research Data Officer - Universiteitsbibliotheek Gent
myriam.mertens@ugent.be

What is RDM?

What are research data?

- Heterogeneous, discipline- & context-specific
- Smallest building blocks of research
- Any information collected/created for the purposes of analysis to generate & validate original scientific claims

Types of research data



PUBLICATIONS AND DATA

Shift in scholarly communication model

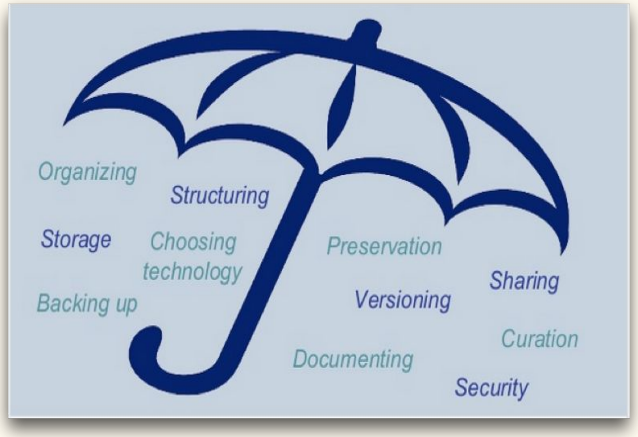
"An article (...) is not the scholarship itself, it is merely advertising of the scholarship."

(Buckheit & Donoho 1995)

Research data are valuable scientific output!

- Costly to create
- Needed to validate published scientific claims
- Often have value beyond original study

RDM helps you take care of your data



"Research data management is the compilation of many small practices that make your data **easier to understand, less likely to be lost,** and **more likely to be usable** during a project or ten years later."

(Briney 2015)

RDM is not an obligation to open data

- “Open means anyone can freely access, use, modify, and share for any purpose.” (opendefinition.org)
- “Closed” if necessary
- Share data under more restricted conditions
  - regulate access by user (group) and/or type of use
  - release data after embargo period

Why should you care about RDM?

External requirements

Good RDM helps you comply with...

- Privacy & confidentiality requirements
  - in legislation, ethical codes, contractual agreements, university regulations...
- Data management & sharing requirements
  - from research funders, publishers, institutions
  - increasing demands for RI
  - move towards OA to scientific research

H2020 Guidelines on Open Access

Availability of data, material and methods

An inherent principle of publication is that others should be able to replicate and build upon the authors' published claims. A condition of publication in a Nature journal is that **authors are required to make materials, data, code, and associated protocols promptly available to readers without undue qualifications**. Any restrictions on the availability of materials or information must be disclosed to the editors at the time of submission. Any restrictions must **also** be disclosed in the submitted manuscript.

After publication, readers who encounter refusal by the authors to comply with these policies should contact the chief editor of the journal. In cases where editors are unable to resolve a complaint, the journal may refer the matter to the authors' funding institution and/or publish a formal statement of correction, attached online to the publication, stating that readers have been unable to obtain necessary materials to replicate the findings.

Availability of data

Supporting data must be made available to editors and peer-reviewers at the time of submission for the purposes of evaluating the manuscript. The preferred way to share large data sets is via public repositories. (Details about how to share some specific data sets can be found in the sections below.)

Some of these repositories offer authors the option to host data associated with a manuscript confidentially, and provide anonymous access to peer-reviewers before public release. These repositories then coordinate public release of the data with the journal's publication date. This option should be used when possible but it remains the author's responsibility to communicate with the repository to ensure that public release is made on time for online publication of the paper. **For information about suitable public repositories, see sections that follow.**

http://www.nature.com/authors/policies/availability.html

UGent Research Integrity Policy Plan

mmertens | CAS Log out

Ghent University Contact | Campuses | Phone Book | Search

Home › News › News bulletin › Research Integrity Policy Plan to strenghten research quality and integrity

Research Integrity Policy Plan to strenghten research quality and integrity

(18-05-2015) As the first Belgian university Ghent University has drawn up a policy plan for research integrity. In the plan the university supplies a comprehensive number of tools to keep and strengthen the quality of scientific research.

The plan gives a clear overview of the general approach of Ghent University towards the topic of research integrity. In addition it also consists of a comprehensive number of tools and tips. All actions are aimed at one single goal; keeping and strengthening the quality and integrity of the daily research practice.

The main focus is on proactive prevention. On the one hand by stimulating 'good research practices'; on the other hand by improving the general quality culture, within but also outside our own organization. The Commission Research Integrity deals with those situations where things eventually do go wrong.

Amongst other things, the plan gives special attention to:

- **sensitization and communication**
  Ghent university reaches out, through its research integrity officer, to the entire research community to bring the issue of research integrity to the attention of all staff and students, in a positive way.
- **training**
  The Doctoral Schools Co-ordination Unit created an interactive transferable skills training to bring the topic of research integrity to the attention of young researchers.
- **data management**
  The quality of the research data is often intrinsic to the quality of the research. As a first step, a comprehensive web section tries to give guidance on the matter.

More information

- Stefanie.Vanderburght@UGent.be
- Research Integrity website

Delen 3 Share Tweeten

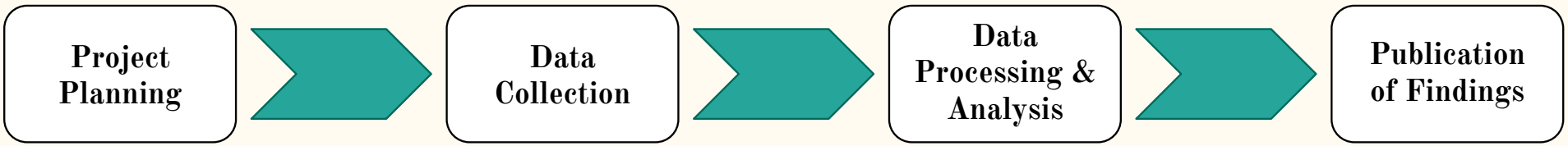
- Approved in March 2015
- "University-wide RDM system" among operational targets
  - support researchers via training, tools, infrastructure
  - faculties, researchers can develop own initiatives
  - list of concrete actions (including WG, policy, RDM webpages, planning tool...)
  - work in progress!

Benefits to researchers

- Minimise risk of data loss
- Increase research efficiency
- Improve transparency & demonstrate RI & quality
- Facilitate collaboration & new (forms of) knowledge creation
- Enhance research visibility (e.g. more citations)

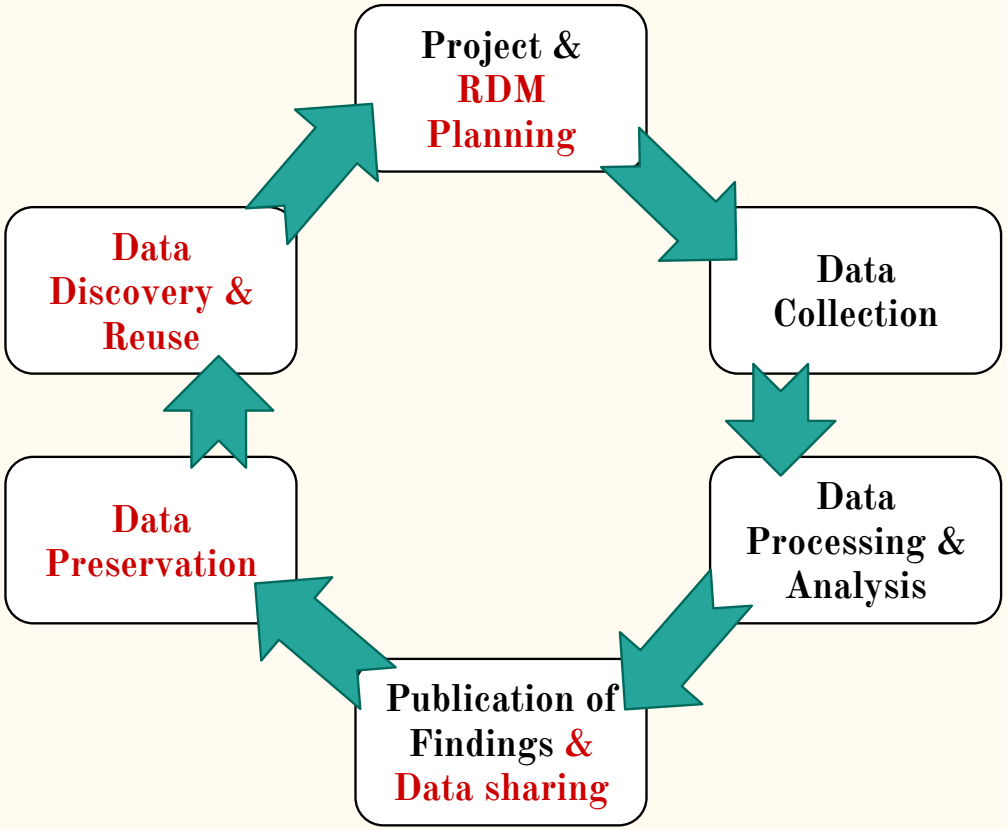
The research data lifecycle

Traditional view of the research process



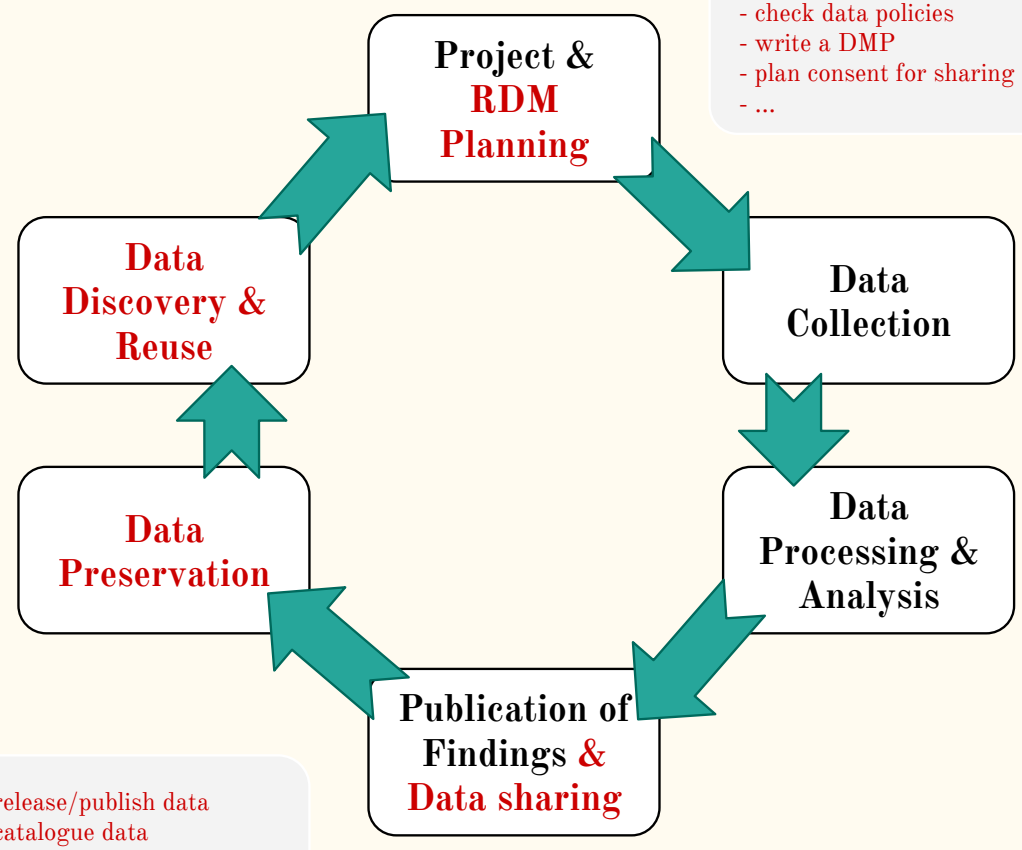
(Adapted from Briney 2015)

Research data lifecycle



(Adapted from Briney 2015 & Corti et al. 2014)

Research data lifecycle



- check data policies
- write a DMP
- plan consent for sharing
- ...

- organise data
- quality control
- document data
- store & backup data
- ...

- organise data
- quality control
- document data
- store & backup data
- ...

- release/publish data
- catalogue data
- license data
- control access
- ...

- select data for retention
- prepare data for preservation
- archive data
- ...

- search for data
- use data for research, teaching etc.
- cite data
- ...

(Adapted from Briney 2015 & Corti et al. 2014)

Data management planning

Why plan?

Planning takes time upfront, but helps you...

- Save time and problems later on
- Consider wide range of issues involved in good RDM
- Make better, more informed decisions about data
- Discuss data issues with supervisor, lab members, collaborators...

How to plan?

Write a **Data Management Plan** (DMP)

- Formal document specifying how research data will be managed during and after a project
- A “living” document
- Increasingly required by (international) research funders/institutions
- Good research practice even if not required

"(Data management) plans typically state what data will be created and how, and outline the plans for sharing and preservation, noting what is appropriate given the nature of the data and any restrictions that may need to be applied."

(Digital Curation Centre website)

What to cover in a DMP?

Common topics

1. Data collection & organisation
2. Data documentation
3. Ethical & legal issues
4. Data storage & backup
5. Data preservation (after research)
6. Data sharing
7. Responsibilities & resources

1. Data collection & organisation

What sort of data will be used in the project?

- data types
- data volume
- digital file formats (e.g. proprietary/open formats?)

How will data be captured & organised?

- data created during project, or reuse of existing data?
- data capture methods & standards
- file organisation system (hierarchical system or database, naming conventions, versioning…)
- quality control procedures

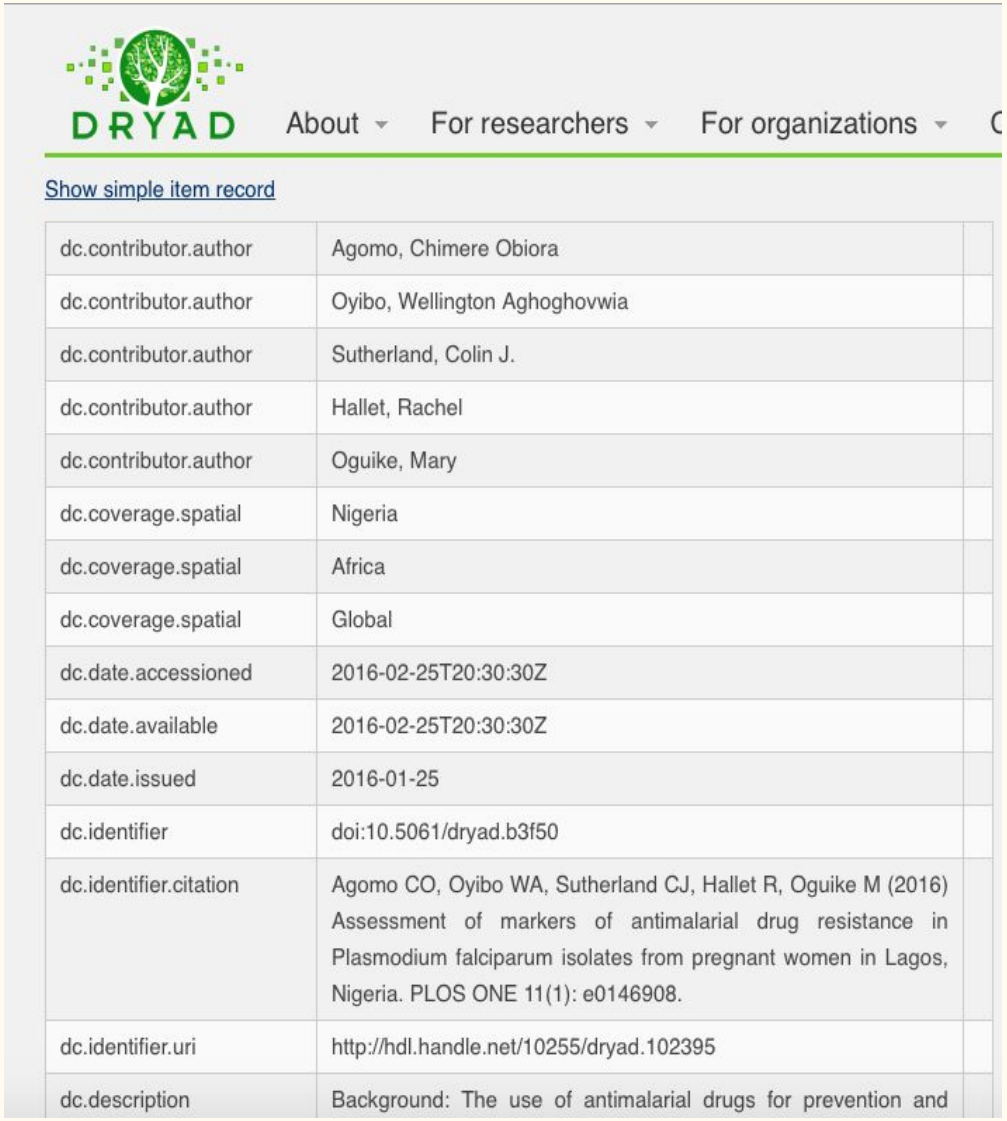
2. Data documentation (any information necessary to find, understand & (re)use data)

How will data be documented?

- what information should be recorded?
  - at data collection & lower levels
  - e.g. project background, methods of data collection/processing/analysis, relationships between data files, access & use conditions, data sources, data item descriptions
- how & where?
  - embedded within data files/databases
  - separate files (e.g. lab notebooks, readme.txt files)
- use of structured metadata?

Metadata ("data about data")

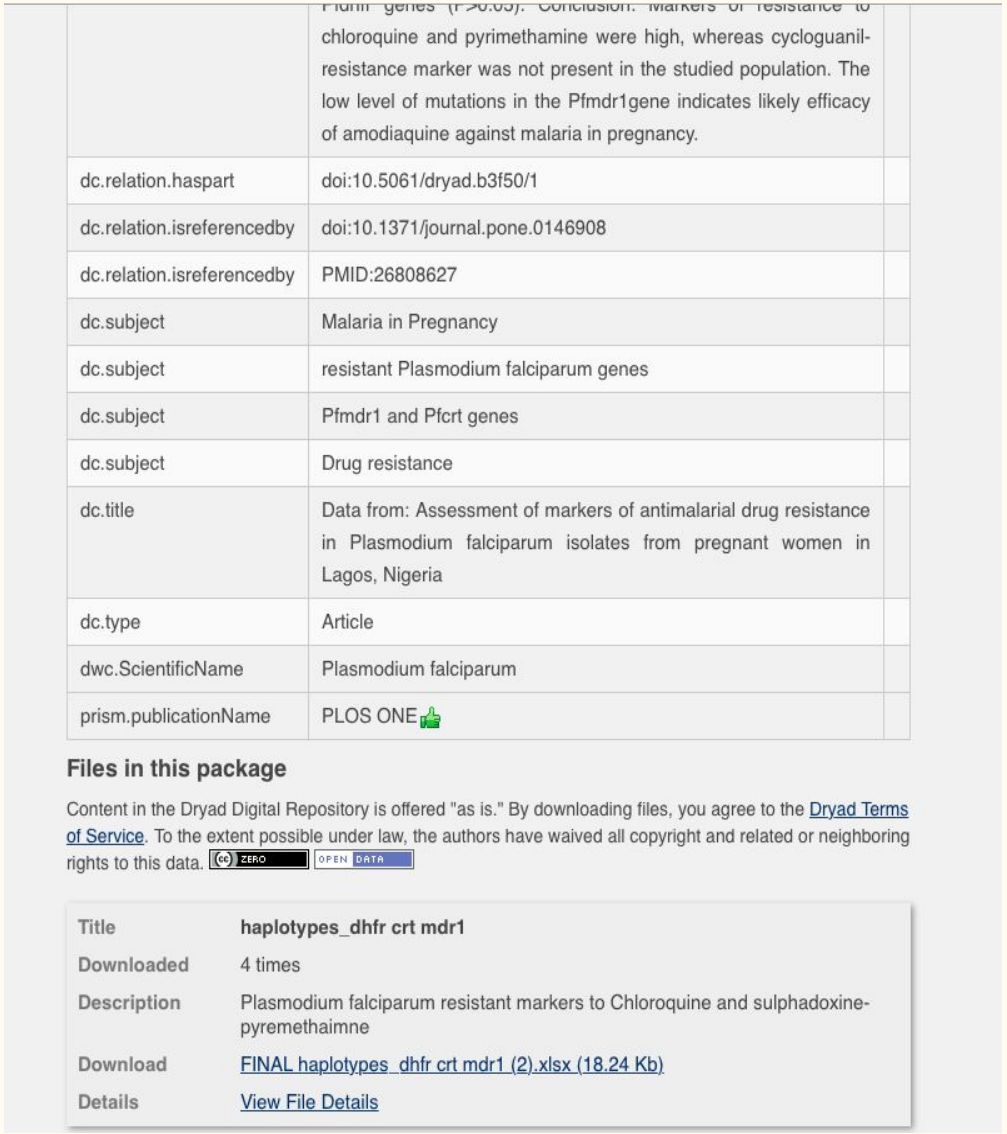
- highly structured & digital form
- human & machine readable
- useful for searching through large amounts of data
- used by data repositories for cataloguing
- elements from a controlled list, as defined by a metadata schema
  - e.g. date, title, subject, geographic location, version, instrument & protocol information, provenance, ...
  - various international schemas, some domain-specific



DRYAD About For researchers For organizations

Show simple item record

dc.contributor.author	Agomo, Chimere Obiora
dc.contributor.author	Oyibo, Wellington Aghoghovwia
dc.contributor.author	Sutherland, Colin J.
dc.contributor.author	Hallet, Rachel
dc.contributor.author	Oguike, Mary
dc.coverage.spatial	Nigeria
dc.coverage.spatial	Africa
dc.coverage.spatial	Global
dc.date.accessioned	2016-02-25T20:30:30Z
dc.date.available	2016-02-25T20:30:30Z
dc.date.issued	2016-01-25
dc.identifier	doi:10.5061/dryad.b3f50
dc.identifier.citation	Agomo CO, Oyibo WA, Sutherland CJ, Hallet R, Oguike M (2016) Assessment of markers of antimalarial drug resistance in Plasmodium falciparum isolates from pregnant women in Lagos, Nigeria. PLOS ONE 11(1): e0146908.
dc.identifier.uri	http://hdl.handle.net/10255/dryad.102395
dc.description	Background: The use of antimalarial drugs for prevention and



	Pfdhfr genes (P>0.05). Conclusion: Markers of resistance to chloroquine and pyrimethamine were high, whereas cycloguanil-resistance marker was not present in the studied population. The low level of mutations in the Pfmdr1gene indicates likely efficacy of amodiaquine against malaria in pregnancy.
dc.relation.haspart	doi:10.5061/dryad.b3f50/1
dc.relation.isreferencedby	doi:10.1371/journal.pone.0146908
dc.relation.isreferencedby	PMID:26808627
dc.subject	Malaria in Pregnancy
dc.subject	resistant Plasmodium falciparum genes
dc.subject	Pfmdr1 and Pfcrt genes
dc.subject	Drug resistance
dc.title	Data from: Assessment of markers of antimalarial drug resistance in Plasmodium falciparum isolates from pregnant women in Lagos, Nigeria
dc.type	Article
dwc.ScientificName	Plasmodium falciparum
prism.publicationName	PLOS ONE

Files in this package

Content in the Dryad Digital Repository is offered "as is." By downloading files, you agree to the Dryad Terms of Service. To the extent possible under law, the authors have waived all copyright and related or neighboring rights to this data. CC ZERO OPEN DATA

Title	**haplotypes_dhfr crt mdr1**
Downloaded	4 times
Description	Plasmodium falciparum resistant markers to Chloroquine and sulphadoxine-pyremethaimne
Download	FINAL haplotypes_dhfr crt mdr1 (2).xlsx (18.24 Kb)
Details	View File Details

http://datadryad.org/resource/doi:10.5061/dryad.b3f50

3. Ethical & legal issues

Will you collect personal or otherwise confidential data?

- informed consent/permissions for collecting, preserving, sharing.... data
- identity protection (e.g. anonymisation)
- data security measures required?

What intellectual property rights apply to the data?

- rights ownership (e.g. copyright, database right, …)?
- will you seek patents?
- permissions needed to use, share data?
- data licenses for reuse (e.g. standard licenses such as CC, ODC)?

4. Data storage & backup (during research)

How will data be stored & backed up during research?

- how many copies?
- storage media (e.g. hard drive, university network drive, cloud…) & locations (local/offsite)
- backup strategy (e.g. what, who, how often, full/incremental, automatic?)

How will data be kept secure?

- security risks (e.g. in terms of unauthorised access, editing, destruction…)?
- measures to manage access & security (physical/network/computer system & file security)?

5. Data preservation (after research)

What will happen to data from completed research in the long term?

- appraisal & selection of data for retention/destruction
- period of preservation
- preservation plan (e.g. outsourcing to a data archive/repository?)

Storing data during research is not the same as **preserving** them (i.e. keeping them findable, accessible, understandable and (re)usable for the future)!

6. Data sharing

Will you share any data beyond your research team?

- any restrictions on sharing required (e.g. to protect personal or otherwise confidential data, to seek patents first…)?

How will you share data?

- when, with whom, under what conditions?
- method of sharing (e.g. upon request, via a project website/journal/data archive or repository…)
- discoverability
- persistent identifiers (e.g. DOI)

7. Responsibilities & resources

Who is responsible for implementing (each element of) the DMP?

- responsibilities within research teams?
- within projects with external partners?
- what happens with data when you leave your research group/the university?

Will you need additional resources (beyond institutional provision)?

- e.g. specialist training, soft- or hardware, funds for storage/preservation...

Further tips for writing a DMP

Be aware of relevant data policies

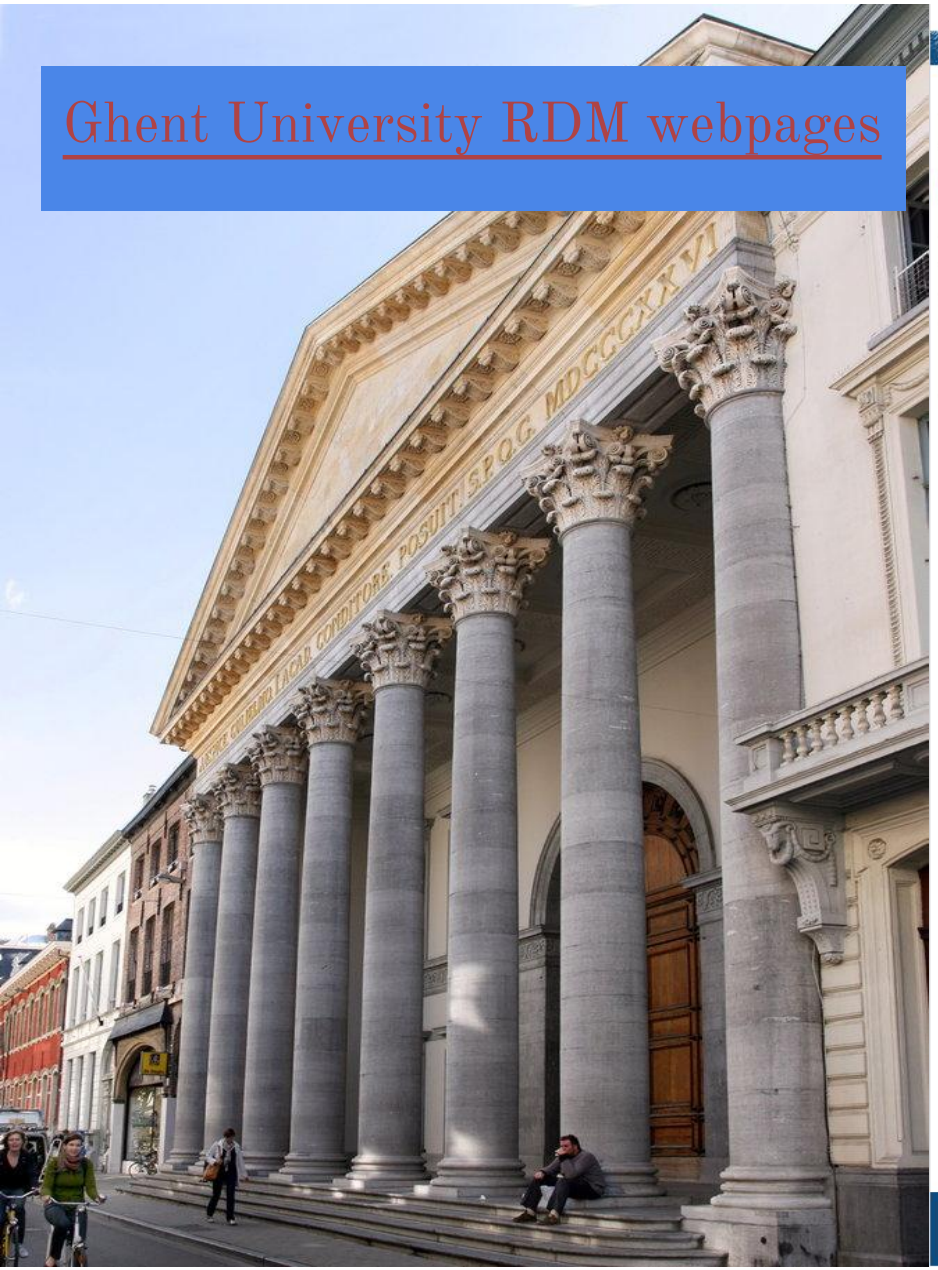
- Larger policies applying to your data
  - regarding privacy, IPR, data preservation & sharing…
  - legislation
  - ethical codes
  - disciplinary norms
  - funder/publisher/institutional requirements
- Local RDM procedures in your faculty, department and/or research group/centre

Familiarise yourself with RDM

Terminology, issues, (disciplinary) standards & best practices...

- E.g. check RDM webpages on Ghent University portal
- Many other online resources & training materials available
  - some domain-specific

Ghent University RDM webpages



NL | EN

mmertens | CAS Log out

UNIVERSITEIT GENT

Ghent University | Contact | Campuses | Phone Book | Search

Home › Research › For staff › Research policy and organisation › Data management

- About Ghent University
- Education
- Research
- Facilities
- Working at UGent
- Living in Ghent

Information for

- >> Exchange student
- >> Degree student
- >> Staff and PhD students

Info Je bent aangemeld. Info op jouw maat vind je op de studentensite of op het intranet voor personeel.

Data management

Good research data management has many advantages: for yourself, for your research institute, for your discipline and for the world around us:

- Ensuring the integrity of your research
- Increasing the impact of your research
- Supporting future use
- Satisfying in- and external requirements

What is datamanagement?

Data management includes all actions needed to make research data discoverable, accessible and understandable in the long term: organization, documentation, storage, sharing and archiving.

Questions? Look for an answer via one of the dedicated points of contact.

Data Management Plan (DMP)

A data management plan makes research data management concrete: what are you going to do and how?

- Creating a DMP
- Requirements of external funders

Privacy and Data safety

- Information security
- Confidential data
- Privacy protection for academic research and academic services

Sharing of digital research data

Data can be shared very openly or sometimes not at all:

- Tips on sharing research data
- Restrictions on data sharing
- Open data
- Sharing code (GitHub)

Related info:

- Citing data

Storage of digital data

There are countless ways to lose data. How do you keep your data findable and useable?

- Store your research data on the central infrastructure
- Store a completed dataset linked to a publication in the academic bibliography
- Store your datasets in a suitable repository

Related info:

- Metadata
- Requirements of publishers

Related content

- OECD Principles & Guidelines for Access to Research Data from Public Funding
- Glossary
- FLAMES: Flanders Training Network for Methodology and Statistics

Last update by Department of Research Affairs on 20-01-2016 13:45

Feedback | Accessibility | Cookies | Disclaimer

Online resources, for example:

Foster Training Portal

Digital Curation Centre ("How-to Guides & Checklists")

MANTRA - Research Data Management Training

UK Data Archive ("Create & Manage Data")

DATUM for Health - RDM training for health studies

Biosharing.org (portal of data standards, databases & policies in life, environmental & biomedical sciences)

Seek guidance

- Ask colleagues, supervisors, university support staff for advice
- Have a look at examples from completed DMPs
  - e.g. via DCC website
  - hand-out

Keep it simple

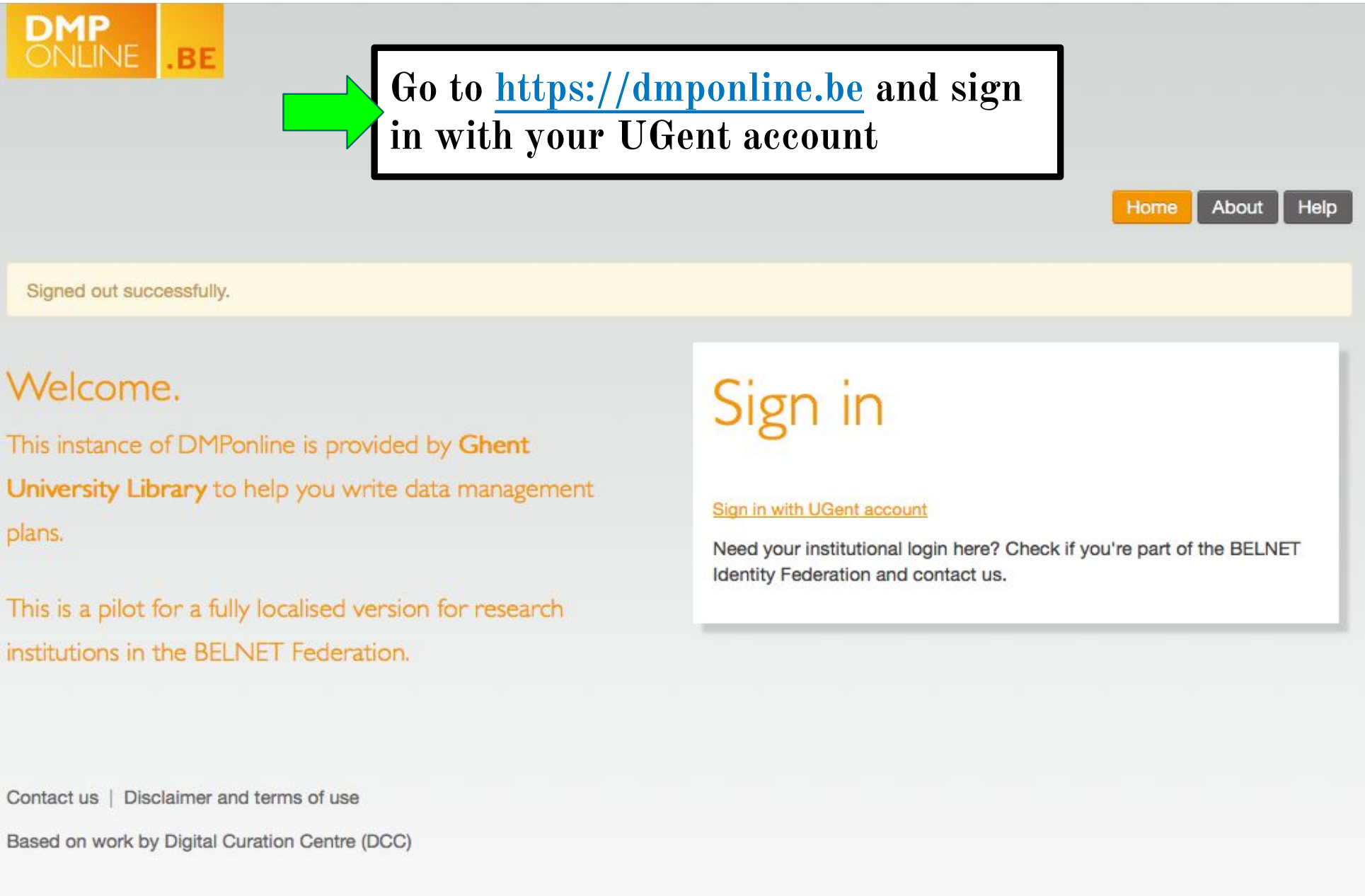
But...

- Be as specific as possible
- Justify your decisions
- Note down what you need to follow up

Use an online planning tool

DMPonline.be

- Local version of open source software by DCC
- Consists of DMP templates & guidance
  - H2020
  - Psychology & Educational Sciences
  - Generic


DMP ONLINE .BE
Go to https://dmponline.be and sign in with your UGent account
Home
About
Help
Signed out successfully.
Welcome.
This instance of DMPonline is provided by Ghent University Library to help you write data management plans.
This is a pilot for a fully localised version for research institutions in the BELNET Federation.
Sign in
Sign in with UGent account
Need your institutional login here? Check if you're part of the BELNET Identity Federation and contact us.
Contact us | Disclaimer and terms of use
Based on work by Digital Curation Centre (DCC)

Writing a DMP is just the first step!

- Be prepared to implement it
- What counts is what you do afterwards

Questions?

- myriam.mertens@ugent.be

KNOWLEDGE 2 CONNECT