



Running order

1. Introduction to DARIAH

What is DARIAH? What is a research infrastructure, and why does it matter to you?

Working groups

Policy and foresight

Training, Education, and careers

The DARIAH Mission Statement

2. A (relatively) short introduction to identifying and managing humanities research data

Introductions

Exercise: what are (humanities) data? What are your data?

Open science and open research data

Exercise on open science

Open Data Management Best Practices

Some resources: print and digital

e.g. 'Future Proof and FAIR Research Data: Open Data Management Best Practices and First Steps', Ulrike Wuttke

Data management plans



What is DARIAH? — "The Digital Research Infrastructure for the Arts and Humanities"

OR: a network to enhance and support digitally enabled research and teaching across the Arts and Humanities

What is a Research Infrastructure?

"...can be defined as shared, unbounded, heterogeneous, open, and evolving socio-technical systems comprising an installed base of diverse information technology capabilities and their user, operations, and design communities."

The important thing about these definitions is the manner in which they encourage us to think of infrastructure as **knowledge**, **as networks**, **as people**, **as tools**, **as data**, **and indeed perhaps as spaces as well**, but always in the service of aggregating resources to make us better connected and more informed



Mission: to empower arts and humanities research communities with digital methods to create, connect and share knowledge about culture and society.

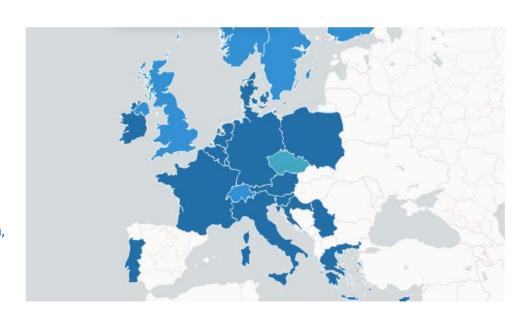
17 member countries

Austria, Belgium, Bulgaria, Croatia, Cyprus, Denmark, France, Germany, Greece, Ireland, Italy, Luxembourg, Malta, Netherlands, Poland, Portugal, Serbia, Slovenia

and 10 cooperating partners

Czech Republic, Finland, Hungary, Israel, Norway, Romania, Spain, Sweden, Switzerland, United Kingdom







Why does Research Infrastructure matter to us?

Building and using shared infrastructure does the following:

- Provides access to networks, data and knowledge for researchers
- Brings greater efficiency and insight in work
- Enhances pathways for research visibility, reuse and impact
- Facilitates better alignment with shared standards and policy frameworks, such as Open Science
- Increases opportunities for seeking collaborative funding
- Promotes long-term sustainability of research outputs



Put simply, DARIAH is a network of people, services, tools and working groups that allow us to share:

- Knowledge
- Tools
- Data
- Collaborations, connections, and networks
- And.... a voice to be heard in European and national policy debates, for instance about the future of scholarly communication



Want to know more?

Check out the the PARTHENOS Training Suite

"Introduction to Research Infrastructures"

https://training.parthenos-project.eu/sample-page/intro-to-ri/





DARIAH Working Groups:

Transnational Scientific Working Groups on Emergent Themes

The Working Groups:

- Ensure that we are meeting the emergent needs of research communities in and with the humanities
- Give us a platform for engagement with them
- Give researchers a non-competitive, non-time limited lightweight, transnational mechanism to organise themselves





More on the Working Groups...

https://www.dariah.eu/activities/working-groups-list/

Sustainable publishing of (meta-)data

https://www.dariah.eu/activities/working-groups/sustainable-publishing-of-metadata/

The goal of this working group is to bring together, and communicate about, knowledge, hands-on expertise and tools concerning sustainable publishing of (meta)data by integrating experiences from research infrastructures, such as EHRI, Europeana, APE and others, and in dialogue with the archival institutions.

Ethics and Legality in the Digital Arts and Humanities (ELDAH)

 $\frac{https://www.dariah.eu/activities/working-groups/ethics-and-legality-in-the-digital-arts-and-humanities-eldah/}{}$

In the age of digital information technology and the constant availability of information through the Internet, it is not only important to have democratic access to knowledge, but also essential to consider the potential that lies in the critical production and expansion of knowledge.

Community calls are a way to engage with stakeholders that work in dispersed or virtual teams (like the DARIAH WGs). They are usually open to anyone interested in a particular topic, but in the DARIAH context the calls are open to the Working Group coordinators



Policy and Foresight

- The average researcher might not care about policy (it is a specialist area)
- BUT their future work conditions will be determined by policy decisions
- The digital is a particularly pronounced place for research policy to be focussed





DARIAH and Open Science Policy

Do you publish openly?

Do you deposit and share your data?

How does your work look according to 'Next generation metrics'? (https://ec.europa.eu/research/openscience/pdf/report.pdf)

Can you access training?



The Open Science Policy Platform (OSPP)

https://ec.europa.eu/research/openscience/index.cfm?pg=openscience-policy-platform



DARIAH Activities to promote humanities in Open Science policy and Open Science in the Humanities



Erzsébet Tóth-Czifra DARIAH-EU Open Science Officer

Data Reuse Charter
Open Access Publication Policy that is meaningful and relevant for humanists
Liaison with projects/bodies like OPERAS, HIRMEOS, OPEN AIRE +, FOSTER, the OS MOOC, EOSC Hub, EGDF, OSPP
Citizen Science training materials





Training, education, skills and careers in DARIAH-EU



Photo credit: Francesca Morselli

- Infrastructures create knowledge, differently from universities
- They provide training and skills development, but again, differently (online training, internships etc)
- They are a place where careers grow (as in the concept of the #altac or #postac)



Training questions and values in DARIAH

Inclusiveness...

On many different levels, and with many layers

Balancing the need to remain at the cutting edge of discussions around digital research whilst prioritising the need to incorporate and encourage new learners (i.e. those who are completely new to digital research/DH)?

Asking: how can we 'preach beyond the choir'?



Oxford, Bodleian Library, MS 264



Effort

How can the different providers/venues of training and education in digital humanities coordinate their efforts most efficiently and effectively?

...avoid unproductive duplication

...form collaborations?

...bear in mind the pressures on researchers, especially the precariat



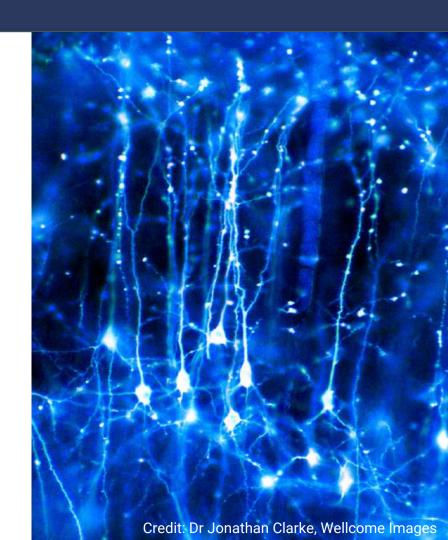
Synergies...

Addressing and spanning the different ways that people need/want to learn

Creating synergies between face-to-face training and online training

Between education and training

Between the training and education resources created by different individuals/groups, for different occasions and/or contexts





Infrastructure as a (new) place of knowledge creation



Login | Register

Statistic

City

Institution

Education

Objects

Home | Contact | About

Search Options: Country

- all -

- all -

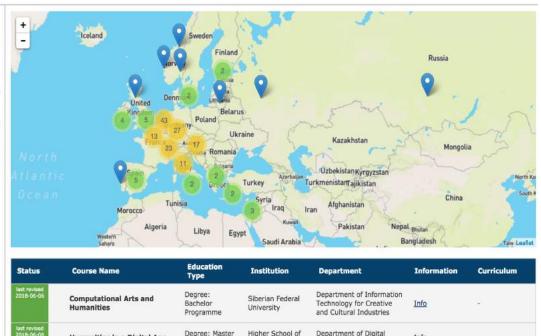
- all -

- ali -

-- none selected --

Disciplines - none selected --

Techniques -- none selected --



https://registries.clarin-dariah.eu/courses/





Consideration of learning and teaching needs and best practice **by humanists**, and **for the humanities**

"When it comes to senior scholars, results of our previous research show that humanists favor and best learn in practice, when instruction closely follows their area of study and when it unfolds organically through collaboration with colleagues and students" - S. Antonievic



Future directions in training and education and Research Infrastructures?

- We can expect to see the role of infrastructures continue to rise in importance as a locus for building skills
- Hierarchies for knowledge creation are shifting anyway (citizen science)
- "Problem-" or "Mission-based" research is on the rise
- New career paths for the research-trained continue to emerge





DARIAH-ERIC Mission Statement, 2019 (Preliminary)

Connecting the arts and humanities;
Complementing the arts and humanities;
Creating for the arts and humanities;

...and always for the arts and humanities.



Identifying and managing humanities research data



Introductions



- Exercise: what are (humanities) data? What are your data?
- Open science and open research data
- Exercise on open science
- Open Data Management Best Practices
 - some resources: print and digital
- 'Future Proof and FAIR Research Data: Open Data
 Management Best Practices and First Steps', Ulrike Wuttke
- Data management plans



Exercise:

Q1. what are (humanities) data?

Q2. what are your data?



DARIAH Annual Event 2019: Humanities Data

'What is humanities data?': https://www.youtube.com/watch?v=WiKwdienmjo

Primary sources (texts, pictures), secondary sources, theoretical texts, digital tools (software), annotations, etc;

Digitized sources and born digital sources;

Annotations, corpora, code, XMLs, collection descriptions, recordings, interviews etc (often based on primary and secondary sources);

Various formats and types (pictures, texts, multimedia, measurements, etc.)

From 'Future Proof and FAIR Research Data: Open Data Management Best Practices and First Steps', Ulrike Wuttke: https://www.fosteropenscience.eu/node/2603



- Humanities are a very broad research discipline, including many specific research contexts, but also increasingly interdisciplinary research
- Humanities research lives from the enrichment of data (layers of interpretation)
- It's problematic to distinguish between primary data (raw data) and secondary data
- There are barriers to openness: issues with ownership of the data (cultural heritage institutions, publishers)

From 'Future Proof and FAIR Research Data: Open Data Management Best Practices and First Steps', Ulrike Wuttke: https://www.fosteropenscience.eu/node/2603



Open Science and Open Research Data



Exercise:

- What is open science?
- What are the advantages of open science?
- What are the challenges?
- What tools, networks, people, and/or readings help you to do more open research?



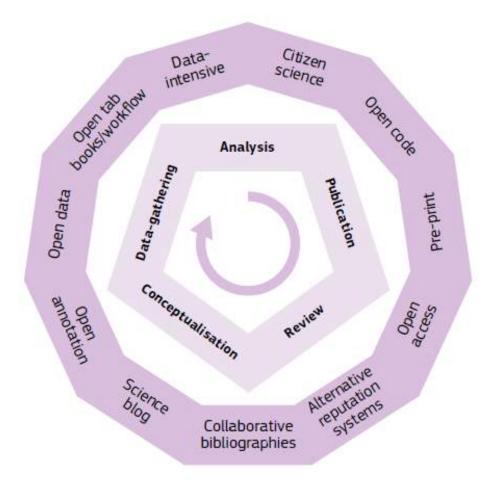


Figure 1: Open Science opens up the entire research enterprise (inner circle) by using a variety of means and digital tools (outer circle) (EC, 2016, p. 36).

'Next-generation metrics:
Responsible metrics and evaluation
for open science'
Report of the European Commission
Expert Group on Altmetrics
(https://ec.europa.eu/research/openscien
ce/pdf/report.pdf)



What is open data?

(Research) data that is freely available online for (re)use and republish for everyone, provided that the data source is attributed

"Open access contributions include original scientific research results, raw data and metadata, source materials, digital representations of pictorial and graphical materials and scholarly multimedia material." Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities (2003)

Ideal: data with no restrictions from copyright, patents, or other control mechanisms > transparent results However: "as open as possible, as closed as necessary"

- From 'Future Proof and FAIR Research Data: Open Data Management Best Practices and First Steps', Ulrike Wuttke: https://www.fosteropenscience.eu/node/2603
- See "Manage, Improve, and Open up your Research Data", PARTHENOS



What does Open Data involve?

- Sharing is not giving away.
- Working in an open environment benefits all, especially the data sharer:
 - reach as many people as possible
 - be cited more often
 - build cooperation and collaborations
 - provides the incentive to check your data
- Poses challenges, e.g. interoperability and documentation
- Some aspects are discipline specific > e.g. Humanities
- · Essential: Data Management Planning
 - From 'Future Proof and FAIR Research Data: Open Data Management Best Practices and First Steps', Ulrike Wuttke: https://www.fosteropenscience.eu/node/2603
 - See "Manage, Improve, and Open up your Research Data", PARTHENOS



FAIR Principles

TO BE FINDABLE:

- F1. (meta)data are assigned a globally unique and eternally persistent identifier.
- F2. data are described with rich metadata.
- F3. (meta)data are registered or indexed in a searchable resource.
- F4. metadata specify the data identifier.

TO BE ACCESSIBLE:

- A1. (meta)data are retrievable by their identifier using a standardised communications protocol.
- A1.1. the protocol is open, free, and universally implementable.
- A1.2. the protocol allows for an authentication and authorisation procedure, where necessary.
- A2. metadata are accessible, even when the data are no longer available.



See "Manage, Improve, and Open up your Research Data", PARTHENOS:

https://training.parthenos-project.eu/sample-page/manage-improve-and-open-up-your-research-and-data/introduction-to-research-data-management/the-fair-principles/



TO BE INTEROPERABLE:

- I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- 12. (meta)data use vocabularies that follow FAIR principles.
- 13. (meta)data include qualified references to other (meta)data.

TO BE RE-USABLE:

- R1. meta(data) have a plurality of accurate and relevant attributes.
- R1.1. (meta)data are released with a clear and accessible data usage license.
- R1.2. (meta)data are associated with their provenance.
- R1.3. (meta)data meet domain-relevant community standards.



"Top 10 FAIR Data & Software Things, Humanities: Historical research

A resource that DARIAH developed in collaboration with the Centre for Digital Scholarship at Leiden University

"A concise overview of the ten topics that are most essential for scholars in the field of historical research who aim to publish their data set in accordance with the FAIR principles. In historical research, research data mostly consists of databases (spreadsheets, relational databases), text corpora, images, interviews, sound recordings or video materials."

https://librarycarpentry.org/Top-10-FAIR/2018/12/01/historical-research/



Questions we need to ask from ourselves...

...in paving the way of an Open Humanities agenda

How to increase the profile of humanities within the Open Science paradigm?

How to close the implementation gap between high-level principles of Open Science and community practices in the humanities?

Does Digital Humanities equal Open Science? How do open innovations from the humanities impact Open Science?

What kind of standards are we referring to?

How do we identify community needs and gaps in open workflows?

How do the values of Open Science manifest themselves in present-day humanities research practices..? ...and how do
these values help
to reassess and
reshape our
fundamental
knowledge
creation
mechanisms?

What are the barriers for establishing a culture of open sharing in the humanities?

How do we make our research outputs more visible?

Credit: Erzsébet Tóth-Czifra



In many cases arts and humanities scholars already doing OS without explicitly branding their research practices as such.

We believe that Open Science works best when it is well-anchored in the different disciplinary settings and community practices.

In many cases there is a **huge implementation gap** between the high-level principles of Open Science and actual research practices.

In many cases, although humanities scholars see the value in transparency, accountability, and reusability we frequently **encounter difficulties or even limitations** when aiming to translate these principles into their everyday scholarly practices.





Why don't we share data more?

- We are "fearful, distrustful and a little lazy"?
- We don't want to?
- Research cycles are long
- We want to protect our ideas
- We don't believe we can share
- We don't recognise the term in our work
- Our data is hybrid, paratextual, epistemically marked, nonstandard...

https://www.dariah.eu/activities/open-science/data-re-use/



University of Leeds, "Find, reuse and cite data"

- Data sources
- Reuse data
- Cite data
- Social media as data

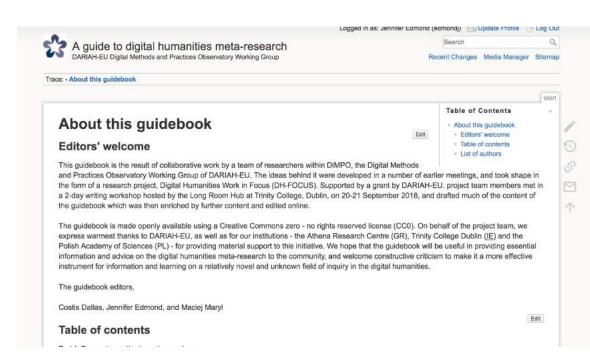




DARIAH-EU Digital Methods and Practice Meta-Research – 'DIMPO' Working Group

'A guide to digital humanities meta-research'

Coming soon!









'Open Data for Humanities, A Pragmatic Guide

Open Data for Humanists, A Pragmatic Guide

So, why do I care?

Most humanists would agree that sharing knowledge with other researchers is a cornerstone of academic life. Many will also fear that sharing too much, too early can be professionally damaging, however. And many also would not find much resonance between how they see their work and the discourses of Open Data, with its emphasis on particular approaches to Data Management Planning that have been adapted from other, more data intensive, disciplines.

The conflict between these positions is in part a semantic one. The many things that would be seen as data in another discipline are often called something else in the humanities. We resist the blanket term 'data' for the very good reason that we have more and precise terminology (e.g. primary sources, secondary sources, theoretical documents, bibliographies, critical editions, annotations, notes, etc.) available to us to describe and make transparent our research





YouTube Videos

'An Ontologist and a Data Scientist walk into a bar: Data in Research Projects'



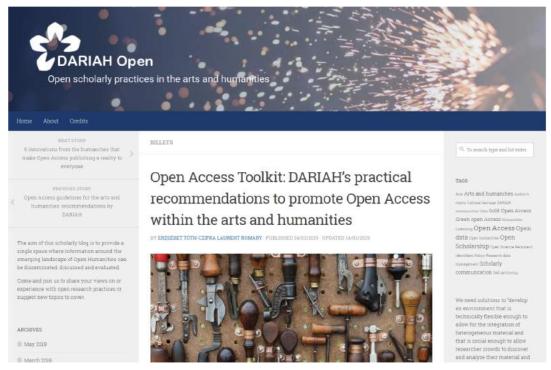
An Ontologist and a Data Scientist walk into a bar: Data in Research Projects

https://www.youtube.com/watch?v=WNG1iLB4KtA&inde x=1&list=PLKq1g7snsFGc7f1_Aidypmz62d7i6Uh4x



A place to explore pathways to the open research culture as they specifically pertain to research communities in the arts and humanities

Building digital sovereignty across our networks and connecting them with fair Open Access service providers



News, resources, expert interviews, success stories but also debates on discipline-specific challenges from around Open Humanities

The blog is also a platform to amplify the good work that done in our national communities or working groups

https://dariahopen.hypotheses.org



Get involved

Come and join us to share your views or your experience with open research practices or suggest new topics to cover

All the posts are licensed under CC-BY license which means that you are free to reuse and build on them

For instance, we were really glad to see one of our posts has been translated into Spanish



The Open Science Helpdesk

Through this service scholars can turn to us with questions and challenges regarding all things open.

Later we also aim to develop a knowledge base from these questions.

...Alternatively, send in your questions and challenges regarding all open issues via the DARIAH Helpdesk







A resource for bringing together, highlighting and community reviewing all formats of openly available digital publications (blogs, preprints, research articles, videos etc.)

Facilitating cultural changes: spreading Open Access information about how to use tools, how to work with data

OpenMethods: highlighting Open Access content

about Digital Humanities methods and tools



Bridging the flow of knowledge across multilingual Digital Humanities communities, newcomers and advanced scholars alike, to discover DH work

The real power of the platform lies in the community around it: our Editorial Team currently comprises 25 editors from 11 countries who speak a total of 16 languages. This team decides upon suitable online content to republish and enrich

https://openmethods.dariah.eu/

Follow/feed the blog: @openmethods_dh





DARIAH-EU Self-archiving and open data management flyers.

DOI: 10.5281/zenodo.2657248







Face-to-face training

Providing guidance to researchers at all career stages and levels of expertise

- Summer schools, winter schools, webinars, masterclasses
- Learn more at: <u>https://www.dariah.eu/tools-services/dariah-training/</u>
- https://www.dariah.eu/events/

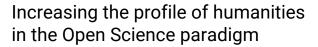


Bringing humanities perspectives to global Open Science events, training and organizations...

....and bringing Open Science to humanities events



Launch of the <u>Open Humanities and Social</u>
<u>Sciences advocacy group</u> at OpenCon
2017. Image source, Katie Steen, twitter.



Reaching out to other communities who are interested in further shaping an Open environment that works for humanities research.



How we can build a global community of humanities scholars who are committed to open scholarship?



Open Humanities session at the Open Science Barcamp 2019. You can find a coverage of session here:

https://genr.eu/wp/humanitiesprogressing-in-open-science/ Image source: Ralf Rebmann, CC BY 4.0 license.



A focus on...

'Future Proof and FAIR Research Data: Open Data Management Best Practices and First Steps'

by Ulrike Wuttke, University of Applied Sciences Potsdam

e.g. Response to "what are the advantages of Open Science for researchers?"

- Higher transparency of research methods and evaluation
- Higher reproducibility of research findings
- Researchers and research institutions save money and time
- Higher (societal) impact of research(ers)
- Open Science gets research(ers) out of the Ivory Towers!
- Open Science as part of Good Scientific Practice





https://www.fosteropenscience.eu/node/2603



Tips on sharing research data

Make use of discipline specific, institutional or European repositories to deposit data/publications (e.g. Zenodo: https://zenodo.org/)
Use tools to register research data (e.g. re3data: https://www.re3data.org/) and to find a repository (Directory of Open Access Repositories: https://v2.sherpa.ac.uk/opendoar/), for humanities e.g.:

- DARIAH (https://de.dariah.eu/en/repository)
- CLARIN (https://www.clarin.eu/content/repositories)
- GESIS (<u>www.gesis.org</u>)

Data Deposit Recommendation Service: https://ddrs-dev.dariah.eu/ddrs/

From: "Future Proof and FAIR Research Data: Open Data Management Best Practices and First Steps, Ulrike Wuttke" https://www.fosteropenscience.eu/node/2603



The Standardization Survival Kit

An overlay platform dedicated to promoting a wider use of standards within the Arts and Humanities:

- Documenting existing standards by providing reference materials.
- Fostering the adoption of standards.
- 18 scenarios: Heritage science scenarios + "traditional" DH ones → Living memory of best practices
- Developed within the framework of the EU project PARTHENOS:

https://www.parthenos-project.eu/portal/ssk-2

What ARE standards?

http://training.parthenos-project.eu/sample-page/intro-to-ri/interoperability/what-are-standards/



CESSDA Training



- Discovering and Using Data, https://www.cessda.eu/Training/Training-Resources/Data-Discovery-and-Reuse
- Managing Research Data, https://www.cessda.eu/Training/Training-Resources/Research-Data-Management
- Preserving Data: CESSDA Tools and Services
- https://www.cessda.eu/Training/Training-Resources/Digital-Preservation-CESSDA-Tools-Services



A Short message on PIDS:

- Persistent Identifiers (e.g. DOI and ORCID):
 Slayer of the Error 404 message & champion of linked open data
- long-lasting, unambiguous reference to a digital object (journal article, dataset, scientific sample, artwork, PhD thesis, publication, or person)
- PID takes you to a metadata record containing information about an digital object or person (its current location for access or download)
- PIDs are stable: metadata of PID record can be updated (e.g. given new location)
- PIDs organisations: Crossref, DataCite and ORCID
- example ORCID: https://orcid.org/0000-0002-8217-402551

More resources on PIDs

- Watch the video "<u>Persistent identifiers and data citation explained</u>" by Research Data Netherlands.
- Watch the video "What are persistent identifiers" for an example about how they are used in digital heritage.
- If you don't have one, <u>request an ORCID</u>. Add all your information as completely as possible.
- Read Alice Meadow's blog post <u>Six Things to do now you have an ORCID iD</u>.
- Go to a data record and click on the DOI to see how the DOI can be resolved to current URL of the data set: http://dx.doi.org/10.17026/dans-x4b-uy8q.
- Read "<u>Digital Object Identifier (DOI) System for Research Data</u>".



Erzsébet Tóth-Czifra, "Sustainable and FAIR Data Sharing in the Humanities"

https://docs.google.com/document/d/1_JnqZatR0HV7RkFd9jmZPWjgqOuZ71H0CJEv4McMbis/edit#heading=h.umsl4u9hziyj



Data Management Plans



Q. Who has already written a data management plan?

Research Data Management Planning

- Often you will need a written and agreed Data Management Plan(DMP), esp. in case of external funding
- To help DMP, many funding agencies provide a model or template for a DMP
- DMP may seem an intimidating (or even unwelcome task), but in the end, it is just a tool for thinking systematically through your research process from a "data perspective"
- DMP helps you to maximize research value (high quality research data and research excellence) and prevents unpleasant surprises at the close of your project (and data loss!)



From 'Future Proof and FAIR Research Data: Open Data Management Best Practices and First Steps', Ulrike Wuttke: https://www.fosteropenscience.eu/node/2603



See "Manage, Improve, and Open up your Research Data", PARTHENOS



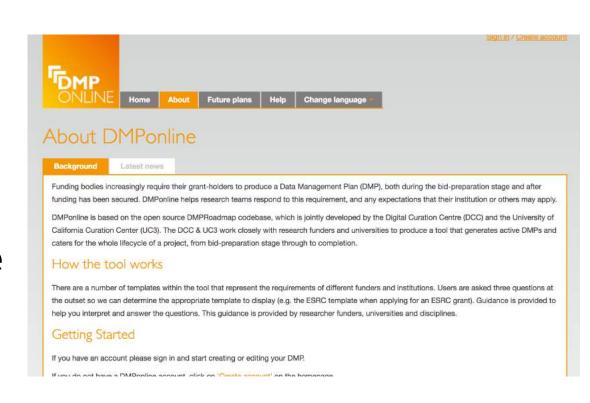
Sections of the DMP

- 1. Description of the Data
- 2. FAIR Data
- 3. Allocation of Resources
- 4. Data Security
- 5. Ethical Aspects
- 6. Other
- From 'Future Proof and FAIR Research Data: Open Data Management Best Practices and First Steps', Ulrike Wuttke: https://www.fosteropenscience.eu/node/2603
 - See "Manage, Improve, and Open up your Research Data", PARTHENOS



Use Tools for Data Management Planning

e.g. DCC DMPOnline https://dmponline.dcc.ac.uk/





Research Data Management Organiser

https://rdmorganiser.github .io/en/



The Research Data Management Organiser (RDMO) enables institutions as well as researchers to plan and carry out their management of research data. RDMO can assemble all relevant planning information and data management tasks across the whole life cycle of the research data.

RDMO is ready for application in smaller or bigger projects. In the next project/phase, which started November 2017, the RDMO tool will be extended and the project partners AIP, FHP, and KIT Library will collaborate with the RDMO users to improve its usage. The tool will be extended by enhancing its implementation of roles and interfaces to institutional infrastructure, e.g. repositories, ticketing systems, and the infrastructure for authentication and autorization. Tutorials, documentation and other material are planned for dissemination, and workshops for users and developers.



Discussion and Questions?





Contact me?

Dr Deborah Thorpe

Training and Education Officer

Deborah.thorpe@dariah.eu

- @DebsEThorpe
- @DARIAHeu

Acronyms in this presentation