



Facilitate Open Science Training for European Research

Data and Openness in the Arts and Humanities

Martin Donnelly
Digital Curation Centre
University of Edinburgh

Liber.ac Fair, University of Ljubljana
22 May 2018





OVERVIEW

1. Background: Open Access, Open Data, Open Science, Open... Art?
2. Differences between Artistic and Scientific methodologies
3. Differing definitions of 'data'
4. Particular strengths and weaknesses in the Arts and Humanities
5. Policy developments: towards a more holistic/inclusive approach?
6. What can we do?



1. Background: Open Access, Open Data, Open Science, Open... Art?

- Open Access came first, intended to speed up research and remove barriers to engagement with it
- As OA became established, RDM became the next hot topic in scientific communication
- In time RDM morphed (in many places) into Open Data
- Together, Open Access + Open Data have become known as Open Science (alongside things like Open Source Software, Open Workflows, Open Notebooks, etc)
- EU Open Data policy is now arranged according to the FAIR principles, expressed in science-centric language
- Is it any wonder Arts and Humanities people tend to feel marginalised/disengaged?!



An old slide of mine...

Timeline: Open Access and Data Sharing

- 1987: New Horizons in Adult Education launched by the Syracuse University Kellogg Project. (An early free online peer-reviewed journal.)
- 1991: The "Bromley Principles" Regarding Full and Open Access to "Global Change" Data, in Policy Statements on Data Management for Global Change Research, U.S. Office of Science and Technology Policy
- 2001: The term "Open Access" (OA), the free online availability of research literature, is first coined at an Open Society sponsored meeting in Budapest, Hungary.
- 2004: Ministerial representatives from 34 nations to the Organisation for Economic Co-operation and Development (OECD) issue the Declaration on Access to Research Data From Public Funding.
- 2006: The Scientific Council of the European Research Council (ERC) pledges to adopt an OA mandate for ERC-funded research "as soon as pertinent repositories become operational".
- 2012: European Commission recognises research data is as important as publications. Announces in July 2012 that it would experiment with open access to research data (see IP/12/790) http://europa.eu/rapid/press-release_IP-12-790_en.htm

(Derived from, *inter alia*, Peter Suber (2009) "Timeline of the open access movement", <http://www.earlham.edu/~peters/fos/timeline.htm>)

 FOSTER

 FOSTER

2. Differences between Art and Science methodologies

- The **scientific method** is a body of techniques for investigating phenomena, acquiring new knowledge, or correcting and integrating previous knowledge.
- To be termed scientific, a method of inquiry must be based on empirical and measurable evidence subject to specific principles of reasoning.
- The *Oxford English Dictionary* defines the scientific method as: “a method or procedure that has characterized natural science since the 17th century, consisting in systematic observation, measurement, and experiment, and the formulation, testing, and modification of hypotheses.”
- Source:
http://en.wikipedia.org/wiki/Scientific_method

An art methodology differs from a science methodology, perhaps mainly insofar as the artist is not always after the same goal as the scientist. In art it is not necessarily all about establishing the exact truth so much as making the most effective form (painting, drawing, poem, novel, performance, sculpture, video, etc.) through which ideas, feelings, perceptions can be communicated to a public. With this purpose in mind, some artists will exhibit preliminary sketches and notes which were part of the process leading to the creation of a work. Sometimes, in Conceptual art, the preliminary process is the only part of the work which is exhibited, with no visible end result displayed. In such a case the “journey” is being presented as more important than the destination.

Source:
http://en.wikipedia.org/wiki/Art_methodology

3. Definitions (i) - 'data' (& subsidiaries)

- Data
 - Facts
 - Truth
 - Reality



3. Definitions (i) - 'data' (& subsidiaries)

• Data

• Facts

• Truth

• Reality

data
/ˈdeɪtə/ ⓘ
noun

facts and statistics collected together for reference or analysis.
"there is very little data available"
synonyms: facts, figures, **statistics**, details, particulars, specifics, features; **More**

- the quantities, characters, or symbols on which operations are performed by a computer, which may be stored and transmitted in the form of electrical signals and recorded on magnetic, optical, or mechanical recording media.
- **PHILOSOPHY**
things known or assumed as facts, making the basis of reasoning or calculation.



3. Definitions (i) - 'data' (& subsidiaries)

- Data

- Facts

- Truth

- Reality

data

/ˈdeɪtə/

fact

/fakt/

noun

a thing that is known or proved to be true.

"the most commonly known fact about hedgehogs is that they have fleas"

synonyms: reality, actuality, certainty, factuality, certitude; [More](#)

• information used as evidence or as part of a report or news article.

"even the most inventive journalism peters out without facts, and in this case there were no facts"

synonyms: detail, piece of information, particular, item, specific, element, point, factor, feature, characteristic, respect, ingredient, attribute, circumstance, consideration, aspect, facet; [More](#)

• used to refer to a particular situation under discussion.

noun: the fact that

"despite the fact that I'm so tired, sleep is elusive"



3. Definitions (i) - 'data' (& subsidiaries)

- Data

data

/ˈdeɪtə/

- Facts

fact

/fakt/

- Truth

truth

/tru:θ/

- Reality

noun

the quality or state of being true.

"he had to accept the truth of her accusation"

synonyms: veracity, truthfulness, verity, sincerity, candour, honesty, genuineness; More

- that which is true or in accordance with fact or reality.

noun: the truth

"tell me the truth"

synonyms: the fact of the matter, what actually/really happened, the case, so; More

- a fact or belief that is accepted as true.

plural noun: truths

"the emergence of scientific truths"

synonyms: fact, verity, certainty, certitude; More



3. Definitions (i) - 'data' (& subsidiaries)

- Data

data

/ˈdeɪtə/

- Facts

fact

/fakt/

- Truth

truth

/tru:θ/

- Reality

reality

/rɪˈælɪti/

noun

1. the state of things as they actually exist, as opposed to an idealistic or notional idea of them.
"he refuses to face reality"
synonyms: the real world, **real life**, **actuality**; [More](#)
2. the state or quality of having existence or substance.
"youth, when death has no reality"



...and in the Arts and Humanities?

• Data “I’ve been asked to produce a data management plan, but I don’t *have* any data. [thinks] Unless you count my databases, I suppose...” – A historian I tried to help once

• Facts “You could not step twice into the same river” – Heraclitus, as reported by Plato (via Socrates)

• Truth Jesus answered, “You say that I am a king. In fact, the reason I was born and came into the world is to testify to the truth. Everyone on the side of truth listens to me.”

‘Truth?’ said Pilate. ‘What is that?’ – John 18:37-38

• Reality

“In science, one man’s noise is another man’s signal” – Edward Ng

“Methinks we have hugely mistaken this matter of Life and Death. Methinks that what they call my shadow here on earth is my true substance. Methinks that in looking at things spiritual, we are too much like oysters observing the sun through the water, and thinking that thick water the thinnest of air.” – Herman Melville



Definitions (ii) - 'research data'

- Definitions of research data vary from discipline to discipline, and from funder to funder...
- A science-centric definition:
 - “The recorded **factual** material commonly accepted in the scientific community as necessary to **validate** research findings.” (US Office of Management and Budget, Circular 110)
 - [Addendum: This policy applies to scientific collections, known in some disciplines as institutional collections, permanent collections, archival collections, museum collections, or voucher collections, which are assets with long-term scientific value. (US Office of Science and Technology Policy, Memorandum, 20 March 2014)]
- And another from the visual arts:
 - “**Evidence** which is used or created to generate new knowledge and interpretations. ‘Evidence’ may be intersubjective or subjective; physical or emotional; persistent or ephemeral; personal or public; explicit or tacit; and is consciously or unconsciously referenced by the researcher at some point during the course of their research.” (Leigh Garrett, KAPTUR project: see <http://kaptur.wordpress.com/2013/01/23/what-is-visual-arts-research-data-revisited>)
- Broadly speaking, it can be thought of as “anything which can be used to validate or reproduce/replicate a research conclusion, or enrich understanding of the research process”



RDM is “the active management and appraisal of data over the lifecycle of scholarly and scientific interest”

“The European Commission is now moving beyond open access towards the more inclusive area of open science. Elements of open science will gradually feed into the shaping of a policy for Responsible Research and Innovation and will contribute to the realisation of the European Research Area and the Innovation Union, the two main flagship initiatives for research and innovation”

<http://ec.europa.eu/research/swafs/index.cfm?pg=policy&lib=science>

The screenshot shows the top of a blog post from LSE (The London School of Economics and Political Science). The title is "Is withholding your data simply bad science, or should it fall under scientific misconduct?". Below the title are social media sharing icons for Twitter, Facebook (93), Email, LinkedIn (42), and YouTube. A small profile picture of Nicole Janz is shown next to the text: "A recent study sent data requests to 200 authors of economics articles where it was stated 'data available upon request'. Most of the authors refused. What does the scientific community think about those withholding their data? Are they guilty of scientific misconduct? Nicole Janz argues that if you don't share your data, you are breaking professional standards in research, and are thus committing scientific misconduct. Classifying data secrecy as misconduct may be a harsh, but it is a necessary step."

News

SHARE ↩

The 'FAIR' principles for scientific data management

Article dated: 8 June 2016

The 'FAIR' Guiding Principles for scientific data management and stewardship form the focus of an article in the Nature journal [Scientific Data](#) an open-access, peer-reviewed journal for descriptions of scientifically valuable datasets. The four principles — Findability, Accessibility, Interoperability and Reusability — provide a guideline for data producers and publishers to enhance the reusability of scientific data. The principles were developed by members of the FORCE11 community, who previously published the Joint Declaration of Data Citation Principles.

4. Strengths and weaknesses re. data in the Arts and Humanities (i)

- There's nothing new about 'data' re-use in the Arts and Humanities; it's an integral part of the culture/method, and always has been...
 - Shakespeare's plots, Kristeva's intertextuality, Barthes' "galaxy of signifiers", found sounds/objects/poems (e.g. Duchamp), variations on a theme, collage and intermedia art, T.S. Eliot, DJ culture (sampling/breakbeat), etc etc. ("Only connect" says Forster)
- However, it's often more fraught than data re-use in other areas (more so in the Physical Sciences than Social Sciences).
- Some characteristics of Arts and Humanities data are likely to require a different kind of handling from that afforded to other disciplines
- For starters, Artists/Humanists do not always naturally think of their sources/ influences/ outputs as 'data', and the value and referencing systems (and norms) may be quite different...

Strengths and weaknesses (ii)

- Digital ‘data’ emerging in the Arts is as likely to be an *output* of the creative research process as an *input* to a workflow (c.f. the UK AHRC policy)
- Furthermore, practice/praxis based research is more or less the sole preserve of the Humanities, and research/production methods are not always rigorously methodical or linear. This is at odds with the scientific approach, and the way in which most RDM resources are described/defined/oriented
- Arts ‘data’ is often personal, and creative data in particular may not be factual in nature. What matters most may not be the content itself, but rather the presentation, the arrangement, the *quality of expression*...
 - This variance in emphasis tends to be why the reason why Open Access embargoes are often longer in the Arts and Humanities than in other areas
 - Creative researchers also care a great deal about the way in which their work is presented, or ‘showcased’: standard repository installations don’t cut it!
- But what do Arts and Science data have in common? Both may be financially valuable and/or precious to their creators (owner-parent)
- People in the Arts are also used to sharing... indeed sometimes *over-sharing*



5. Recent and future (policy) directions

- Wellcome Trust holistic policy towards research outputs (“Policy on data, software and materials management and sharing”)
 - Replaces the former individual policies re. OA and data
 - Wellcome is a medical research funder, but it’s interesting that they also cover non-digital data, such as samples - relevant to the Arts in particular
- New EC recommendation (“COMMISSION RECOMMENDATION of 25.4.2018 on access to and preservation of scientific information”) also takes a more holistic approach, covering all components of Open Science (OA, data, and software/hardware) in a holistic approach
 - It’s only regrettable that they continue with the science-centric language!



6. What can we do?

- Be careful with our terminology
 - “Data” - be clear that this is not the dictionary definition, but rather shorthand for a variety of scholarly products/biproducts (see www.researchobject.org for examples)
 - Don’t use “science” and “research” interchangeably. Challenge those who do... (e.g. the FAIR principles!)
 - Some languages have words like ‘Wissenschaft’, which remove this issue
 - Perhaps if we focus more on talking about ‘research outputs’ as opposed to publications, data, etc - and organise our provision in the same way - the lack of engagement we have previously encountered
 - Perhaps we need to be more vague, less definitive, deliberately...
- Be mindful of the sometimes blurred lines between professional investigation and personal expression
 - Talk to researchers: understand their working methods, discover their needs, assuage their fears
 - Be gentle when introducing policies, and be mindful that the drivers behind RDM are not equally applicable when it comes to creative/humanities work - focus more on issues of preservation, re-use, scholarly quality benefits rather than speed, efficiency/VfM/Rol etc. (It’s not usually life and death in the Arts.)
 - Build bridges *before* they’re needed



Some recent events and initiatives

- Cultural Heritage Data Reuse Charter: Mission Statement
 - “Several European organisations (APEF, CLARIN, DARIAH, Europeana, E-RIHS) and European projects (Iperion-CH, PARTHENOS) involved in the cultural heritage domain are working together to set up principles and mechanisms for improving the use and re-use of cultural heritage data issued by cultural heritage institutions and studied and enriched by researchers.”
 - <https://docs.google.com/document/d/1yLDOyyTGbAn-zy9UAF49oXXb3GMAHMzvQtgEcSacssk/edit#>
- Survey of GLAM Open Access policies (covers publications AND data/related resources)
 - <https://is.gd/glamOApolicies>
- Conference on Research Data Management in Digital Humanities (17-18 April 2018, UCL Qatar)
 - <http://www.ucl.ac.uk/qatar/events/research-data-management-conference>



More resources

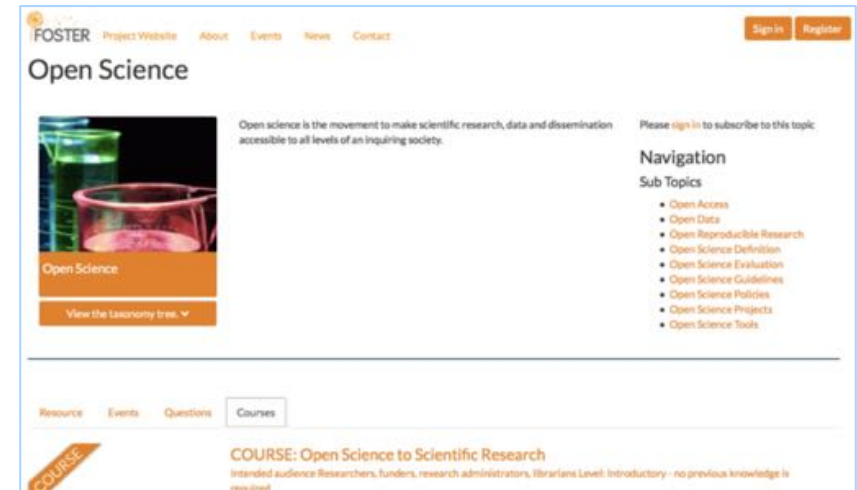
- Paper: Marieke Guy, Martin Donnelly, Laura Molloy (2013) “Pinning It Down: Towards a Practical Definition of ‘Research Data’ for Creative Arts Institutions”, International Journal of Digital Curation, Vol. 8, No. 2, pp. 99-110. URL: [doi:10.2218/ijdc.v8i2.275](https://doi.org/10.2218/ijdc.v8i2.275)
- Projects:
 - KAPTUR (2011-13) URL: <http://www.vads.ac.uk/kaptur/>
 - A consortial approach to building an integrated RDM system (2014-16) (Partners: CREST, University for the Creative Arts, ULCC, Leeds Trinity University, Arkivum). URL: <http://www.crest.ac.uk/welcome-to-the-crest-rdms-project-blog/>
- Event: “Research Data Management Forum #10: RDM in the Arts and Humanities”, September 2013, St Anne's College, University of Oxford. URL: <http://www.dcc.ac.uk/events/research-data-management-forum-rdmf/rdmf10-research-data-management-arts-and-humanities>
- Case study: Jonathan Rans (2013) “Planning for the future: developing and preserving information resources in the Arts and Humanities” URL: <http://www.dcc.ac.uk/resources/developing-rdm-services/dmps-arts-and-humanities>
- Blog posts:
 - Marieke Guy (2013) “RDM in the Performing Arts” URL: <http://www.dcc.ac.uk/blog/rdm-performing-arts>
 - Laura Molloy (2015) “Digital Preservation for the Arts, Social Sciences and Humanities - benefits for everyone” URL: <http://www.dcc.ac.uk/blog/digital-preservation-arts-social-sciences-and-humanities-benefits-everyone>
- Slides: Martin Donnelly (2013) “‘Found’ and ‘after’ - a short history of data reuse in the Arts” URL: <http://www.slideshare.net/martindonnelly/data-reuse-in-the-arts>



The FOSTER project

Facilitate Open Science Training for European Research

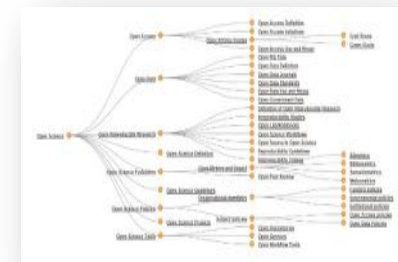
- Phase 1 (2014-2016): Spread the Seeds of Open Science and Open Access
- Creation of **Open Science Taxonomy**
- **2000+** training materials, categorized in the **FOSTER Portal**
- More than **100 f2f training events** in **28 countries** and **25 online courses**, totalling more than **6300 participants**



The FOSTER project

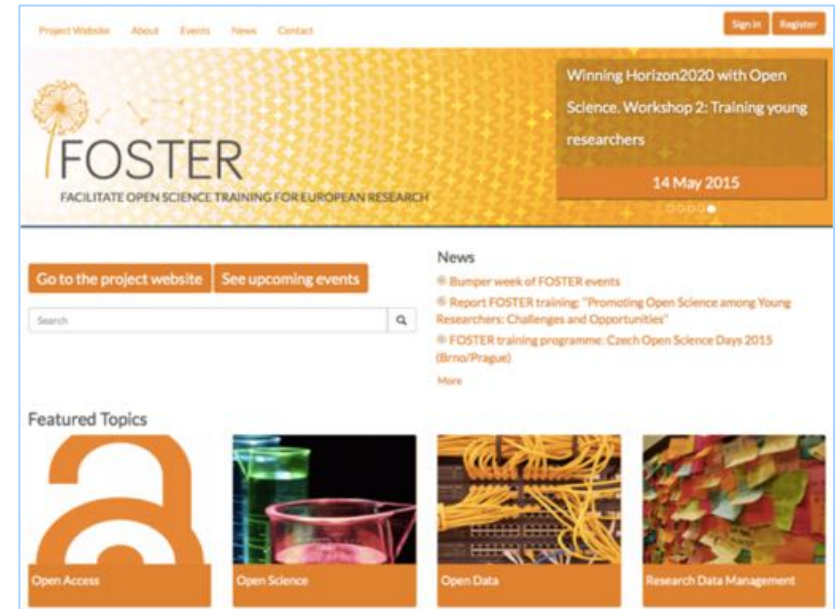
Facilitate Open Science Training for European Research

- Phase 2 (2017-2019): Let the Flowers of Open Science Bloom
- Focus on:
 - Training for the **practical implementation** of Open Science (face to face and online) including RDM and Open Data
 - Developing intermediate/advanced level/discipline-specific **training resources** in collaboration with three disciplinary communities (and related RIs): Life Sciences (ELIXIR), Social Sciences (CESSDA) and Humanities (DARIAH)
 - Update the FOSTER Portal to support moderated learning, badges and gamification
- In concrete terms:
 - **150** new training resources
 - **Over 50 training events** (outcome-oriented, providing participants with tangible skills) and **20 e-learning courses**
 - Multi-module **Open Science Toolkit**
 - Trainers Network, Open Science Bootcamp, Open Science Training Handbook, and more...



Contact details

- For more information about the FOSTER project:
 - Website: www.fosteropenscience.eu
 - Principal investigator: Eloy Rodrigues (eloy@sdum.uminho.pt)
 - General enquiries: Gwen Franck (gwen.franck@eifl.net)
 - Twitter: @fosterscience
- My contact details:
 - Email: martin.donnelly@ed.ac.uk
 - Twitter: @mkdDCC
 - Slideshare: <http://www.slideshare.net/martindonnelly>



This work is licensed under the Creative Commons Attribution 2.5 UK: Scotland License.