Enabling data sharing in the Netherlands

Ingrid Dillo
Deputy Director DANS

Open Science/Open Data: challenges and best practices in Scientific Data Management
FAPESP, Sao Paulo, 20 September 2018
Outline

• DANS

• Data sharing in the Netherlands

• The European context
Mission: promote and provide permanent access to digital research resources

Institute of Dutch Academy and Research Funding Organisation (KNAW & NWO) since 2005

First predecessor dates back to 1964 (Steinmetz Foundation), Historical Data Archive 1989

DANS: the organisation
DANS: 60 enthusiastic colleagues
DANS is about keeping data FAIR

EASY
Certified Long-term Archive

DataverseNL to support data storage during research until 10 years after

NARCIS
Portal aggregating research information and institutional repositories

https://dans.knaw.nl
DANS centre of excellence: training and consultancy

Benefit from our knowledge on research data management by our training sessions, consultancy and information material.

TRAINING
DANS supports researchers (indirectly) in data management by providing training sessions. »»

CONSULTANCY
DANS assists in developing data management policy and obtaining certification. »»

INFORMATION MATERIAL
Watch the video 'Why share data' or download other information material. »»
Projects

FAIRSFair proposal
International collaborations
Outline

• DANS

• Data sharing in the Netherlands

• The European context
Why data sharing is important

Transparency and replication of research (scientific integrity)
Fraud Case Seen as a Red Flag for Psychology Research

A well-known psychologist in the Netherlands whose work has been published widely in professional journals falsified data and made up entire experiments, an investigating committee has found. Experts say the case exposes deep flaws in the way science is done in a field, psychology, that has only recently earned a fragile respectability.
Why data sharing is important

Re-use of data (efficiency, return on investment, standing on the shoulders of others)
The 19th-century whaling logbooks that could help scientists understand climate change

The public are being enlisted to read through detailed logs of whaling ships which include records of ice flows and weather conditions.

Hunters with their catch, circa 1900. Whalers' logbooks are providing detailed information about the climate and ice patterns many years ago. Photograph: Hulton Archive/Getty Images

Maritime historians, climate scientists and ordinary citizens are coming together on a project to study the logbooks of 19th-century whaling ships to better understand modern-day climate change and Arctic weather patterns.
16 year study suggests air temperature is external trigger for heart attack

28 Aug 2017

Topic(s): Environmental and Cardiovascular Disease;
Barcelona, Spain - 28 Aug 2017: A 16 year study in more than 280 000 patients has suggested that air temperature is an external trigger for heart attack. The findings are presented today at ESC Congress. (1)

“There is seasonal variation in the occurrence of heart attack, with incidence declining in summer and peaking in winter,” said first author Dr Moman A. Mohammad, from the Department of Cardiology at Lund University, Skane University Hospital, Lund, Sweden. “It is unclear whether this is due to colder temperatures or behavioural changes.”

This nationwide, 16 year, observational study led by Prof David Erlinge from Lund University, is the largest to investigate the association between heart attack incidence and weather conditions such as air temperature, sunshine duration, precipitation, and air pressure.

Using the Swedish myocardial infarction registry (SWEDHEART), all consecutive heart attacks treated at a coronary care unit between 1 January 1998 and 31 December 2013 were included in the study. The investigators studied the specific weather conditions during which heart attacks occurred using local meteorological data from hundreds of weather stations in the Swedish Meteorological and Hydrological Institute (SMHI).

https://www.escardio.org/The-ESC/Press-Office/Press-releases/16-year-study-suggests-air-temperature-is-external-trigger-for-heart-attack
THE DATA ARE MINE!

IVORY TOWER OF SCIENCE

SCENE FROM THE PAST?
To share or not to share?

Motivations not to share:

To share or not to share?

Motivations to share:

- the norms that researchers are exposed to within their research circle or discipline

http://repository.jisc.ac.uk/5662/1/KE_report-incentives-for-sharing-researchdata.pdf
To share or not to share

External drivers:
• Publisher requirements (DAPs)
• Funder policies/mandates

Enabling the researcher to comply with open data requirements:
• awareness raising, training and support for data management
• infrastructure for preservation of and long-term access to the data
The Netherlands Open Science ambition

Key ambitions for 2020:
• 100% open access publishing
• optimal reuse of research data
• corresponding evaluation systems

Make research data optimally suited for reuse:
*The aim of open science is that researchers reuse other parties' research data and services where possible and make their own data available as far as possible.*

https://www.openscience.nl
Code of Conduct Research Integrity

• Published on 14 September 2018
• To keep up with international developments
• More prominent place for research data
• KNAW, NWO, universities, university medical centres, and institutes of applied research

http://www.vsnu.nl/files/documents/Netherlands%20Code%20of%20Conduct%20for%20Research%20Integrity%202018.pdf
Standards for good research practice:

• **describe** how the collected research data are organized and classified so that they can be verified and reused.

• **make** research findings and research data **public** subsequent to completion of the research. If this is not possible, establish valid reasons for their non-disclosure.

• contribute, where appropriate, towards making data findable, accessible, interoperable and reusable in accordance with the **FAIR principles**.

• the research must be described in sufficient detail for it to be possible to replicate the data collection and its analysis.

• **do justice to everyone who contributed** to the research and to obtaining and/or processing the data.
Code of Conduct Research Integrity

Institutions’ duties of care:

• As a supervisor, principal investigator, research director or manager, provide for an open and inclusive culture in all phases of research.
• Provide a research infrastructure in which good data management is the rule and is facilitated.
• Ensure that, as far as possible, data, software codes, protocols, research material and corresponding metadata can be stored permanently.
• Ensure that all data, software codes and research materials, published or unpublished, are managed and securely stored for the period appropriate to the discipline(s) and methodology concerned.
• Ensure that, in accordance with the FAIR principles, data is open and accessible to the extent possible and remains confidential to the extent necessary.
• Ensure that it is clear how data, software codes and research material can be accessed.
Funder requirements

All calls for proposals include the data management protocol.

The data management protocol consists of two steps:

• A data management paragraph in the research proposal in which the researcher should answer a number of short questions;
• A data management plan that must be submitted after the proposal has been awarded funding. The approval of this plan is a prerequisite for NWO disbursing the grant.

The costs of data management are eligible for funding and should be included in the project budget.

https://www.nwo.nl/en/policies/open+science/data+management
Dutch (infra)structure in cheese & art

Emmenthaler

Miro

Gouda

Mondriaan
Federated infrastructure for research data

Data Archiving and Networked Services

doi:10.2218/ijdc.v9i2.333
Front office-Back office model

FO-BO model
- Division of labour
- Economies of scale

Back office
- Curation and preservation expertise
- Training of local data experts
- Long-term preservation infrastructure

Front office
- RDM guidance and training for local research community
- Data storage during research phase
Back office: Research Data Netherlands

https://researchdata.nl/en/

- DANS
- 4TU.Centre for Research Data https://researchdata.4tu.nl/home/

mission: to promote sustainable access to and responsible reuse of scientific research data in all phases of the research.
What do we have to offer?

- Certified trustworthy long-term data repositories
- Advice on policies and implementation of data stewardship
- Training, e.g. course “Essentials 4 Data Support”
- Dutch Data Prize
Developing RDM Policies

- Make data FAIR
- Write a DMP
- Guidelines for storing the data during and after the research

https://www.edugroepen.nl/sites/RDM_platform/Universiteit/Sitepages/Home.aspx (in Dutch)
Problems solved?

- Fragmentation and overlap in activities and approaches
- Re-inventing the wheel; not taking advantage of expertise and experiences
**Objective**

to support the most effective use of data in research and education by strengthening the knowledge pooling and knowledge sharing on all aspects of research data management

**Approach**

- obtaining a good overview of existing initiatives
- sharing successful approaches and results and promoting their re-use
- identifying gaps and putting them on the agenda
<table>
<thead>
<tr>
<th>1. Facilities and data infrastructure</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2015</strong></td>
</tr>
<tr>
<td>Q4</td>
</tr>
<tr>
<td><strong>Preparations:</strong> Put together team, current situation, planning</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2. Legal aspects and ownership</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Preparations:</strong> Put together team, current situation, planning</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3. Financial aspects</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Preparations:</strong> Put together team, current situation, planning</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4. Support of the researcher</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Preparations:</strong> Put together team, current situation, planning</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>5. Awareness / Engagement</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Preparations:</strong> Put together team, current situation, planning</td>
</tr>
</tbody>
</table>

*Data Archiving and Networked Services*
Community Building

72 working group members, dealing with 17 topics in the 5 focus areas

Visible, knowledgeable and efficient RDM community has been built, under the LCRDM umbrella
Concrete products

over 40 products: inventories, advices, tools, models, reference cards etc.

<table>
<thead>
<tr>
<th>TOPICS</th>
<th>PRODUCTS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Financial Aspects of RDM:</strong></td>
<td>• Position Paper Financial Aspects RDM</td>
</tr>
<tr>
<td>• Putting Financial Aspects of RDM at governance level</td>
<td>• Round Table Financers &amp; Governance</td>
</tr>
<tr>
<td><strong>Research Support &amp; Advise:</strong></td>
<td>• Inventory DMP templates</td>
</tr>
<tr>
<td>• DMP</td>
<td>• Document DMP Basic Criteria</td>
</tr>
<tr>
<td>• FAIR data</td>
<td>• Inventory DMP support (tips &amp; tricks)</td>
</tr>
<tr>
<td>• Training</td>
<td>• Inventory DMP reviewers (via questionnaire)</td>
</tr>
<tr>
<td>• Safe data</td>
<td>• Inventory &amp; Evaluation DMP tools</td>
</tr>
<tr>
<td>• Safe data</td>
<td>• Advise DMP Tooling</td>
</tr>
<tr>
<td><strong>Engagement:</strong></td>
<td>• Inventory FAIRness Data-repositories in NL (TU Delft)</td>
</tr>
<tr>
<td>• RDM-eisen funders</td>
<td>• Position Paper FAIR Data</td>
</tr>
<tr>
<td>• Incentives</td>
<td>• Inventory Training – wiki</td>
</tr>
<tr>
<td>• Use Cases</td>
<td>• Overview common elements curricula –wiki</td>
</tr>
<tr>
<td>• Shades of Open</td>
<td>• Advise on Profiles related Training portfolio’s</td>
</tr>
<tr>
<td>• Engagement Sessions</td>
<td>• RDM-training MOOC</td>
</tr>
<tr>
<td></td>
<td>• Software sustainability – advise</td>
</tr>
<tr>
<td></td>
<td>• Informed Consent Examples</td>
</tr>
<tr>
<td></td>
<td>• Decision Tree Privacy related research &amp; Data sharing</td>
</tr>
<tr>
<td></td>
<td>• Overview RDM-requirements funders</td>
</tr>
<tr>
<td></td>
<td>• Position paper Incentives</td>
</tr>
<tr>
<td></td>
<td>• Overview Use-Cases</td>
</tr>
<tr>
<td></td>
<td>• Teaching modules, information material</td>
</tr>
<tr>
<td></td>
<td>• Inventory &amp; interviews</td>
</tr>
<tr>
<td></td>
<td>• Mind Map</td>
</tr>
<tr>
<td></td>
<td>• As appears</td>
</tr>
</tbody>
</table>
Problems solved?

- Adoption of outcomes by stakeholders
- Creation of a structural RDM collaboration
- Funding
- Human resources in the data domain
- Connection with the researchers
Structure of my talk

- DANS
- Data sharing in the Netherlands
- The European context
4 September 2018: cOAlition S

- 11 national research funding organisations, with the support of the European Commission and the European Research Council (ERC)

Target:
- In 2020 funding agencies will require grant holders to publish only in journals that offer immediate open access, and under a licence that enables anyone to freely reuse and distribute the material.
Financing European research and innovation projects

Period 2014 - 2020; the budget is € 80 billion

Goal: Increasing competitive position of Europe, and find solutions for societal challenges

http://ec.europa.eu/programmes/horizon2020

Netherlands: over 3,300 H2020 projects

Promotes open science by mandates/pilots
Horizon 2020 open science policy

CHALLENGE
Wider access to scientific facts and knowledge helps researchers, innovators and the public find and re-use data, and check research results:

- offers better value for EU research funds
- encourages research across scientific fields
- a public benefit
- essential for solving today's complex societal challenges

SOLUTION
Horizon 2020 already mandates open access to all scientific publications

From 2017, research data is open by default, with possibilities to opt out
Requirements Open Research Data

1. **Data Management Plan** required within six months after project grant. DMP regarded as living document.

2. **Deposit** the data (alongside metadata and possible tools) in a trustworthy research data repository.

3. The data should be deposited as “open data”, in other words, data that can be accessed and reused free of charge to the user. Open if possible, closed if necessary.

4. The data need to be **FAIR**: Findable, Accessible, Interoperable and Reusable.
European Open Science Cloud

2015: vision of the EC of a large infrastructure to support and develop open science and open innovation in Europe and beyond.

to become a reality by 2020 and will be Europe’s virtual environment for all researchers to store, manage, analyse and re-use data for research, innovation and educational purposes.

FAIR Data Action Plan
Interim recommendations and actions from the European Commission Expert Group on FAIR data

June 2018

https://doi.org/10.5281/zenodo.1285290
Research Data Alliance

- global membership organization
- 7,000+ members in 137 countries
- 95 groups working on data related challenges, producing relevant outputs:

**Practical Policies Recommendations** designed to support data sharing and interchange between communities.

**RDA/WDS Workflows for Research Data Publishing Model:** assisting research communities in understanding options for data publishing workflows and increases awareness of emerging standards and best practices.

**23 Things: Libraries For Research Data:** An overview of practical, free, online resources and tools that users can immediately take advantage of to incorporate research data management into the practice of librarianship.

**Eleven Quick Tips for Finding Research Data:** to educate and train research students and early career researchers, and to help researchers more effectively and precisely discover data that meets their specific needs.
Main takeaways

- Netherlands in the European setting as an inspiring use case for your discussions on data sharing

- Open science and data sharing as part of that will only grow in importance

- Data sharing will make science more cost effective, will advance science and will help us deal with the big societal challenges worldwide

- In order to achieve this we need to:
  - change academic cultures and create incentives for researchers to start sharing their data
  - enable researchers to share their data through data management policies, training and guidance, and LTP infrastructure
Thank you for listening

ingrid.dillo@dans.knaw.nl
www.dans.knaw.nl