The data explosion

huge challenge and a gigantic opportunity

Margaret Louise Fotland, Adviser, UiO
Live Håndlykken Kvale, Senior Librarian, UiO

UiO : University of Oslo

Overview

Common ground	• Why share data? • Mandate
Report	• Policy guidelines • Data flow model
The road ahead	• Pilot • Collaboration

Foto: Wikimedia.org

The report proposes:

- Clear guidelines for data management at the UiO.
- A pilot that will provide for the establishment of a program to upgrade skills and good research support services.
- A clear division of roles, responsibilities and authority, in relation to the needs for technical infrastructure and implicitly the development of an offer for temporary storage and sharing of research data with metadata descriptions at UiO.
- UiO assisting on a number of key issues that require national coordination


Foto: Colourbox

Why share data?

- encourages scientific enquiry and debate
- promotes innovation and potential new data uses
- leads to new collaborations between data users and data creators
- maximises transparency and accountability
- enables scrutiny of research findings
- encourages the improvement and validation of research methods
- reduces the cost of duplicating data collection
- increases the impact and visibility of research
- provides credit to the researcher as a research output in its own right
- provides great resources for education and training

http://www.data-archive.ac.uk/create-manage/planning-for-sharing/why-share-data

Foto: F.H.Juell UiO

UiO

8 faculties
32 departments
2 museums
Library
Academic staff 3426
Adminstrative staff 1622
Academic support staff 934
Other 212
Total 6194

Foto:Stine Marie Barsjø, UiO

Collaborations numbers for UiO

(scival 2010-2014)

- In Norway totally 40 institutions, 24 academics institutions
- UiO and nordic countries except Norway: 90 total, 51 academic inst.
- UiO and Europe: 1335 inst. total, 779 inst. categorized as academic
- UiO global 2708 inst. (all sectors), 1789 inst. categorized as academic

Credit to Senior Adviser Herman Strøm, UiO for data on co-publications

Foto: Anders Lien UiO

Mandate

University Director, with the approval of the deans, gave 14 May 2014 mandate to:

- Map existing services and practices, and scientists' needs for storing and sharing research data at the UiO
- Develop draft principles and guidelines for storing and sharing / disclosure of research data which safeguard scientific employee rights
- Recommend solutions that corresponding funders increasing demands for storage and sharing, and the researchers' own needs

Which practises are out there?

- Disks
- Servers
- Dropbox
- USBs
- Mail

Foto: Colourbox

Other feedback

- Routines and training for learning good data practice
- PhDs do data collection, data is lost when they finish
- Positive towards sharing but requesting the tools
- Cross faculty similarities
- DOIs for data citation

Foto: Anders Lien UiO

Policy and guidelines

Foto: UiO

Foto: Hans Dalane-Hval

Ethics and Law

SHARING AND STORING RESEARCH DATA

A 4-level model of the data flow

The model presented here is developed from findings from a small exploratory project conducted in 2014/2015 mapping current practices and opportunities in research data management at the University of Oslo (UiO). The model aims to fill the gaps in the research data flow that were discovered in the project; from collection, through analysis, to publication and retrieval. It aims to optimize the needs at different levels in the research process.

4 LEVEL

DATA DISCOVERY

Level four is concerned with enabling data discovery. The quality of discovery services hinge on the quality of curated data, especially metadata. The importance of establishing all four levels should be emphasized: individual researchers and research communities will not be properly incentivized to produce and share quality research data if one of the four levels are missing. New research projects are enabled trough data discovery.

SOLUTION: INTERNATIONAL, NATIONAL OR DOMAIN SPECIFIC

3 LEVEL

ARCHIVED DATA

Level three is primarily associated with research that has already been published. For some disciplines, data might go directly from active data (level 1) to archived data; however a selection process should take place. Persistent identifiers, metadata descriptions and the data are no longer changeable. The degree of openness depends on whatever license or policy applies to it.

SOLUTION: NATIONAL OR DOMAIN SPECIFIC

2 LEVEL

PROJECT DATA

Level two represents a central solution for storage and sharing of research data with a high level of security and availability. At this level, the research data are typically undergoing changes and are in a pre-publication phase. Sharing "among friends" is essential to facilitate collaboration. Some descriptive metadata and version control helps researchers keeping track of their research data at this stage.

SOLUTION: INSTITUTIONAL OR INTER-INSTITUTIONAL

1 LEVEL

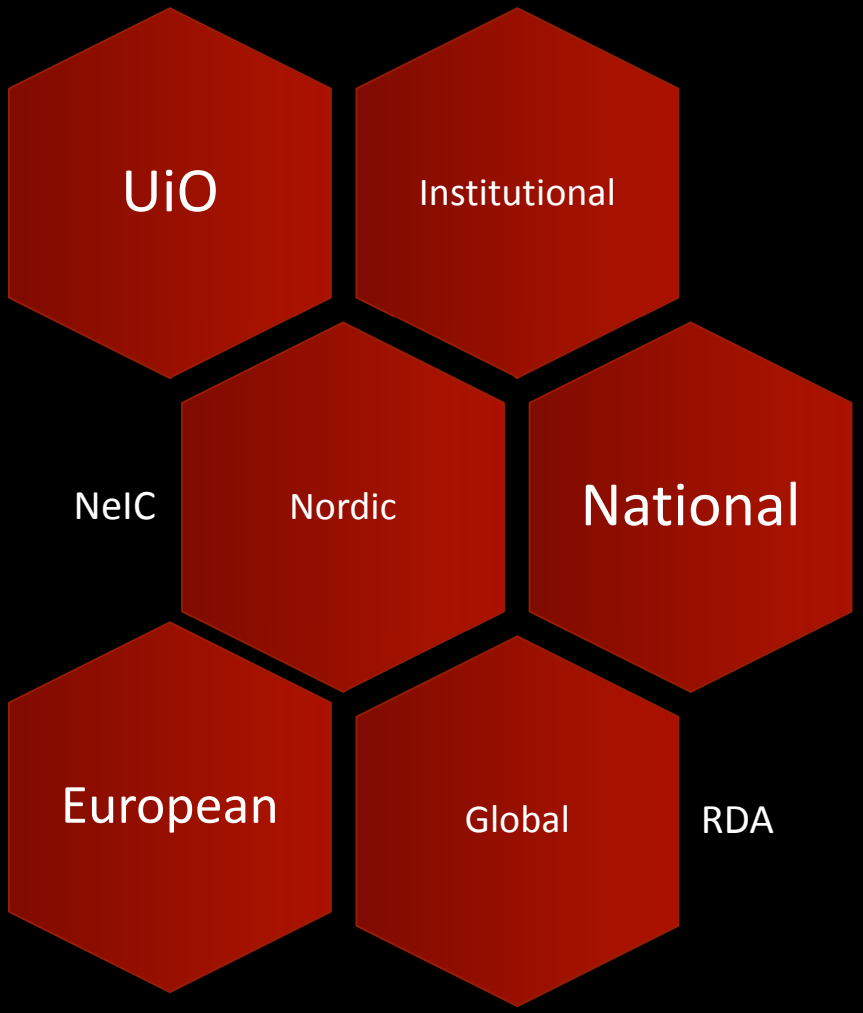
ACTIVE DATA

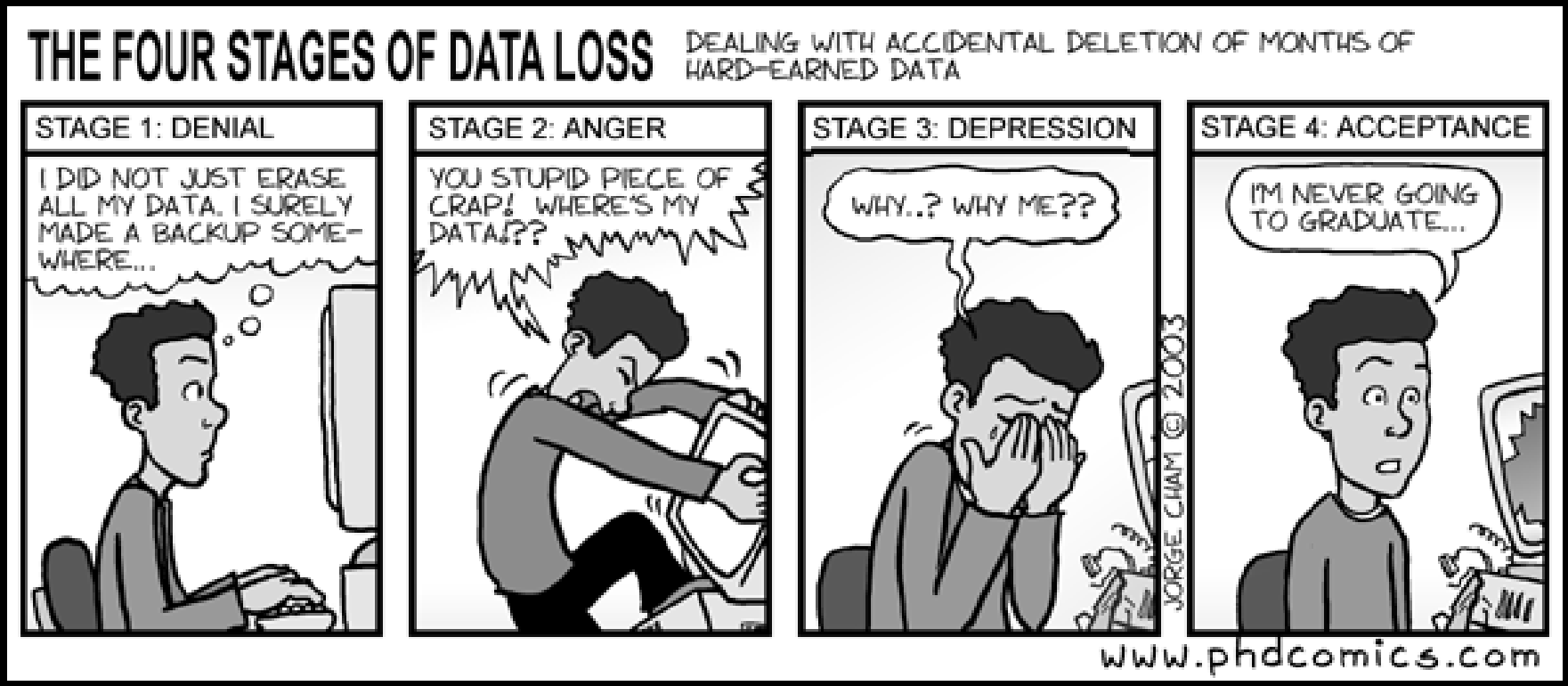
Active data research starts with a question or hypothesis based on existing understanding and observations with the aim of generating new insights. Researchers then collect data for interpretation and analysis. Researchers will need new domain-specific tools and infrastructure to manage their research and outputs with increased access to technology and digital data.

SOLUTION: INDIVIDUAL


Pilot
Technical infrastructure
Training
Web-pages
Foto: Colourbox

No reason to stand alone





Ill. Phd comics

Thanks to:

The working group


Svein, Head


Espen, KHM


Torben, MN


Hans, USIT


Einar, AF


Torgeir, UV


Katrine, MED


Tore, OD


Fridtjof, NHM


Alexander, HF


Tor, SV

And to Elin Stangeland, Gard Thomassen and Leon du Toit for continuing the work on research data as a collaboration among research support services at UiO

Thank you for listening

PLEASE VISIT OUR POSTER

m.l.fotland@admin.uio.no

l.h.kvale@ub.uio.no