Enabling the re-use of research data:

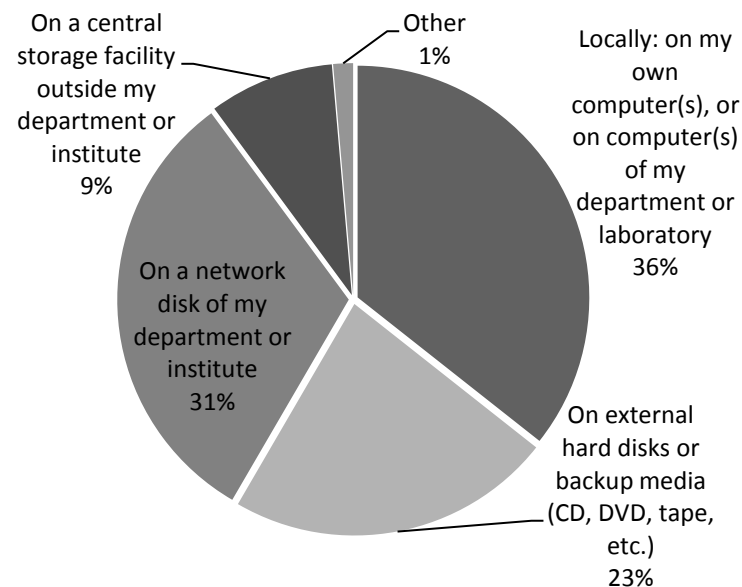# organising stakeholders and infrastructure in the Netherlands

Ingrid Dillo, deputy director DANS

RECODE conference, Athens 15-01-2015

Challenges in sharing data

- **Technical challenges:** standardisation, accessibility, discoverability, preservation and curation, data security, etc.

- **Legal and ethical challenges**

- **Cultural challenges**

"The coolest thing to do with your data will be thought of by someone else"

On a central storage facility outside my department or institute 9%

Other 1%

Locally: on my own computer(s), or on computer(s) of my department or laboratory 36%

On a network disk of my department or institute 31%

On external hard disks or backup media (CD, DVD, tape, etc.) 23%

Why not share?

- Those data are mine!

- Discredit my findings

- Still analyzing the data

- I cannot **trust** the data produced somewhere else

"Perhaps the biggest challenge in sharing data is trust: how do you create a system robust enough for scientists to trust that, if they share, their data won't be lost, garbled, stolen or misused?"

The Data Harvest:
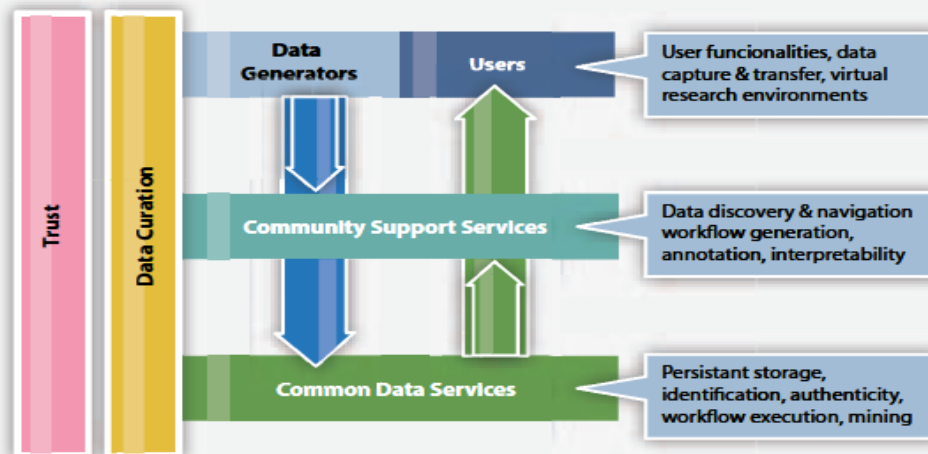How sharing research data can yield knowledge, jobs and growth

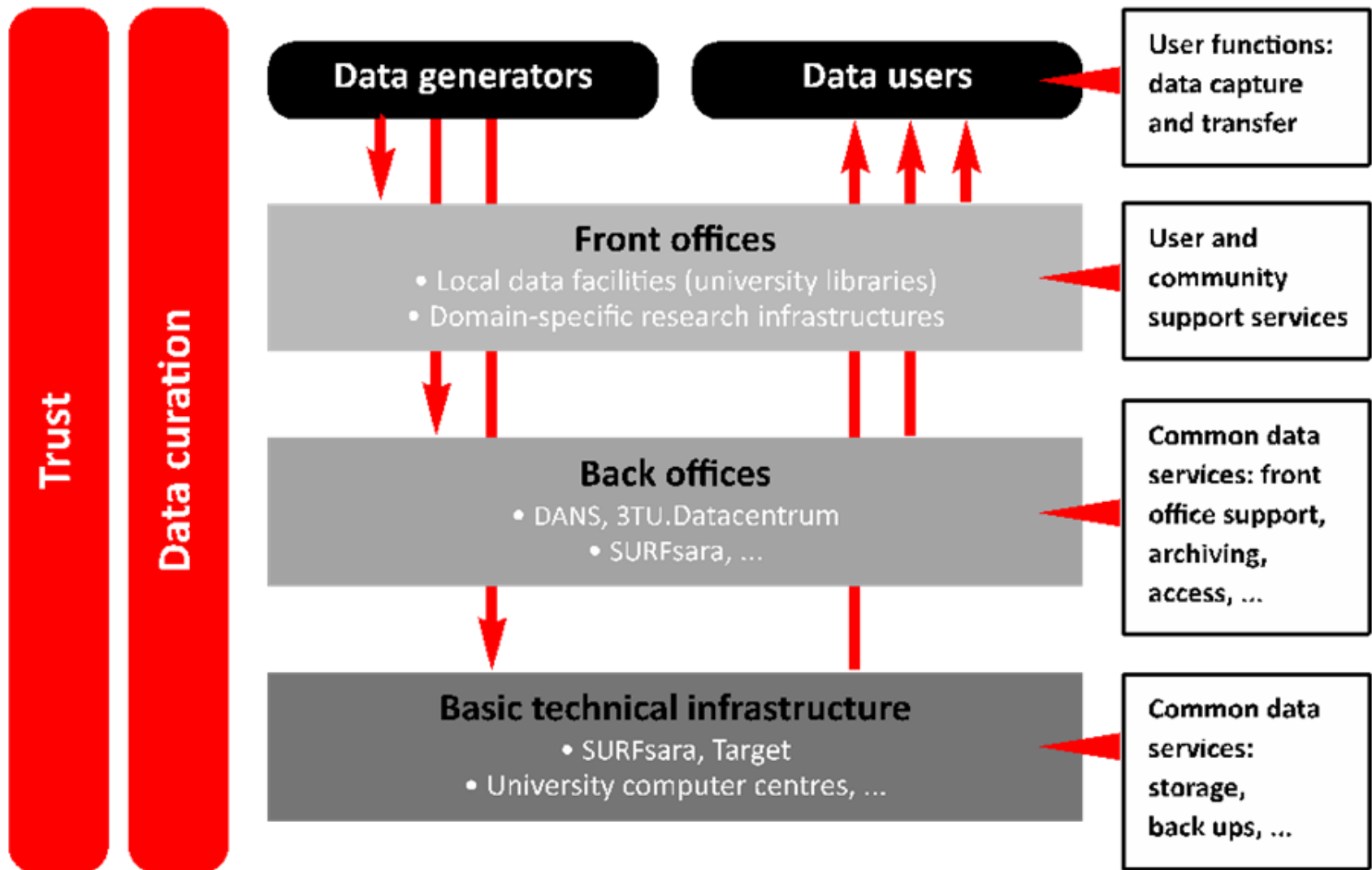An RDA Europe Report

December 2014

Riding the Wave (Oct 2010)

Riding the wave

How Europe can gain from the rising tide of scientific data

Final report of the High Level Expert Group on Scientific Data
A submission to the European Commission

October 2010

The Collaborative Data Infrastructure - a framework for the future

Trust

Data Curation

Data Generators — **Users**
User funcionalities, data capture & transfer, virtual research environments

Community Support Services
Data discovery & navigation workflow generation, annotation, interpretability

Common Data Services
Persistant storage, identification, authenticity, workflow execution, mining

1st recommendation: Develop an international framework for a Collaborative Data Infrastructure

Federated data-infrastructure in the Netherlands

Services in the FOBO model

- Information and awareness raising on data management and curation

- Training (data experts at FOs and researchers)

- Storage (during and after the research)

Back office: roles and responsibilities

Focus on expertise and long term storage:

- Expertise and innovation in the area of data curation, data management and re-use of data

  – Providing expertise to the research community: contributions to FO training courses for researchers
  – Providing expertise to the front office: training courses for data experts, consultancy, contact persons
  – Long term preservation of data in a **trustworthy** digital repository (certification; at least Data Seal of Approval)

Towards sustainable and trusted data repositories

There are 16 guidelines that together determine whether your data repository qualifies for the Data Seal of Approval.

Data Seal of Approval

- DANS initiative (2005/6)
- International Board
- 16 guidelines
- Self assessment
- Transparency
- 37 seals awarded since 2010; 35 on the way

Data Seal of Approval

The research data:
- can be **f**ound on the Internet
- are **a**ccessible (clear rights and licenses)
- are in a usable format (**i**nteroperable)
- are **r**eliable
- can be **r**eferred to (persistent identifier)

BO Organisations

- 3 organisations, serving different disciplines

rdnl — research data netherlands

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

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

Mission: promote and provide permanent access to digital research information
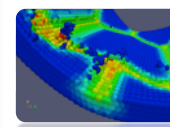
Data Archiving and Networked Services

DANS

EXPERIENCE 50 YEARS EXPERIENCE
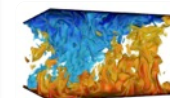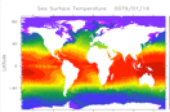
3TU.Datacentrum...

3TU.Datacentrum offers the knowledge, experience and the tools to archive research data in a standardized, secure and well-documented manner. It provides the research community with:

- A long-term archive for storing scientific research data
- Permanent access to, and tools for reuse of research data
- Advice and support on data management

National center for high performance computing & data services
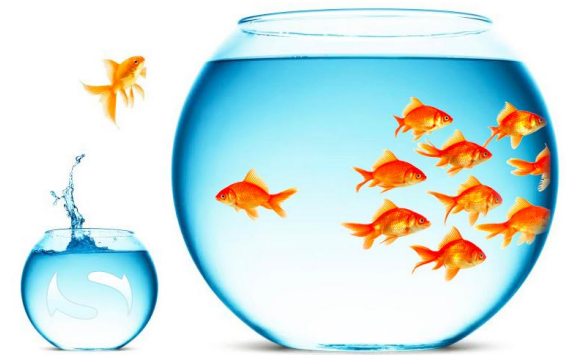
SURF SARA

- **Mission**: SURFsara supports research and education in the Netherlands by development and deployment of HPC-infrastructure, services and expertise

- **Services**: SURFsara offers integrated Dutch national ICT research infrastructure and provides **services** in the areas of high performance computing, cloud, **long-term data preservation**, big data analysis & visualization

- **SURF**: SURFsara is an operating company of SURF, the collaborative organization for higher education institutions and research institutes in the Netherlands, aimed at breakthrough innovations in ICT

Research Data Netherlands (RDNL)

- Mission: the promotion of sustained access and responsible re-use of digital research data

- BO collaboration in order to serve the FO better and more efficient

- training data experts, Dutch Dataprize, data acquisition, exchange of technical expertise

- Expanding areas of collaboration

- Open to other trustworthy digital repositories

Challenges

- Expanding the model over all universities (institutional agreements)

- Creating one single back office "desk" (RDNL)

- Creating a technical infrastructure for automatic data ingest

- Developing a business model to cover the costs

Concluding remarks

- National and international funders are formulating open data policies

- Need to enable researchers to comply with these policies through awareness raising, DM training, high quality and reliable storage facilities

- Meet these researcher needs as effectively and efficiently as possible

- Federated data-infrastructure to enable stakeholders to collaborate, coordinate activities and divide roles and tasks based on their specific expertise (economies of scale, selection of data).

- Small country

- The total picture is for more complex: international and discipline specific infrastructures

The work DANS does
summarized in a 4 minute
video, introduced by
Neelie Kroes

http://youtu.be/HJbo-OAaJ1I

Thank you for your attention

ingrid.dillo@dans.knaw.nl

www.dans.knaw.nl

http://www.researchdata.nl/en/

3TU.Datacentrum Data Archiving and Networked Service **DANS** **SURF SARA**

Research Data Netherlands is een samenwerking
van 3TU.Datacentrum, DANS en SURFsara.
www.researchdata.nl