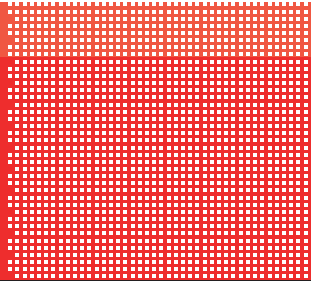




research
data
netherlands



A federated data infrastructure for the Netherlands: the front-office – back-office model

A federated data infrastructure for the Netherlands: the front-office – back-office model

Sustainable storage and sharing of research data is becoming increasingly important and it involves an increasing number of parties. In response to these developments Research Data Netherlands, an alliance of DANS, 3TU.Datacentrum and SURFsara, is introducing a new way of collaborating and dividing roles: the front-office – back-office (FO-BO) model.

The importance of storing and sharing research data

Sharing research data is important. Data availability enables replication of research and thus enhances scientific transparency. Reusability of data also contributes to scientific progress and the advancement of knowledge.

Research funders acknowledge the importance of data sharing and require a high degree of open access. NWO and KNAW argue that research data obtained with public funding should be publicly accessible as much as possible. The Dutch universities attach great importance to responsible research data management. The European Commission wants to increase the impact of publicly funded research by making its results available in open access. In the context of the Horizon 2020 program the EC is examining ways in which open access to research data can best be realised.

Towards a federated data infrastructure

Good, timely data management and secure data storage, both during and after completion of a research project, are essential prerequisites for sharing those data. Universities and colleges of higher education have a clear say in the definition of policies regarding data management within their own organizations. In order to implement these data policies good support and an

adequate technical infrastructure are indispensable. In the influential report *Riding the wave*, which has been embraced enthusiastically by European Commission Vice-President and European Commissioner for the Digital Agenda, Neelie Kroes, such a 'Collaborative Data Infrastructure' is touted as a framework for the future. At the initiative of Research Data Netherlands a federated data infrastructure is emerging in the Netherlands, including services, roles and responsibilities for various parties concerned. It was decided to develop the reference model outlined above into a four-layer structure (see page 4).

In this model, the Common Data Services layer from the EU framework has been split. The foundation is a basic technical infrastructure, which facilitates data storage and back-up. Above that is a layer of back-office data services, providing facilities and support for long-term archiving and accessibility. The highest level includes the front-office services. They provide for the first-line contacts, supporting, advising and training researchers and students in responsible data management. The front office can rely on the expertise of the back office.

Benefits of the FO-BO model

The FO-BO model offers benefits for all stakeholders because it provides for an optimum division of labour based on the respective expert competencies of those stakeholder and their various roles in the data infrastructure.

The figure on page 6 summarises these benefits for both back and front office and the research community.

Those involved in the data infrastructure

In this federated and layered data infrastructure the various stakeholders have specific responsibilities stemming from their respective positions and competencies.

- The **basic technical infrastructure** is provided by computer centres (Shared Service Centres, SSC-ICT), an area where parties like SURFsara, Target or university ICT services have a coordinating role on a national, regional or institutional level.

- The **back-office** functions are carried out by organizations with a national role to play in the field of long-term accessibility of data in trusted digital repositories, such as DANS and 3TU.Datacentrum (and SURFsara from September 2014 onwards), collaborating in Research Data Netherlands. Together they have expertise on data from the humanities, sciences and social studies.

- The **front offices** are located at universities (mostly at libraries), research/knowledge institutes, colleges of applied science, national and international research infrastructures (ESFRI/National Roadmap); for some features they can also be found with the funders (NWO, ZonMw, Ministries or independent governing bodies). All these organizations have in common that they are primarily responsible for the quality assurance of the data produced and processed by them or for them.

Services in the FO-BO model

In the federated data infrastructure roles can be divided according to the front-office - back-office model (FO-BO model). The services provided in this model are all related to data management and storage.

They fall roughly into three groups:

1. Awareness raising and information provision.
2. Training (focusing on data librarians/experts and researchers).
3. Data curation, management and storage (during and after research projects).

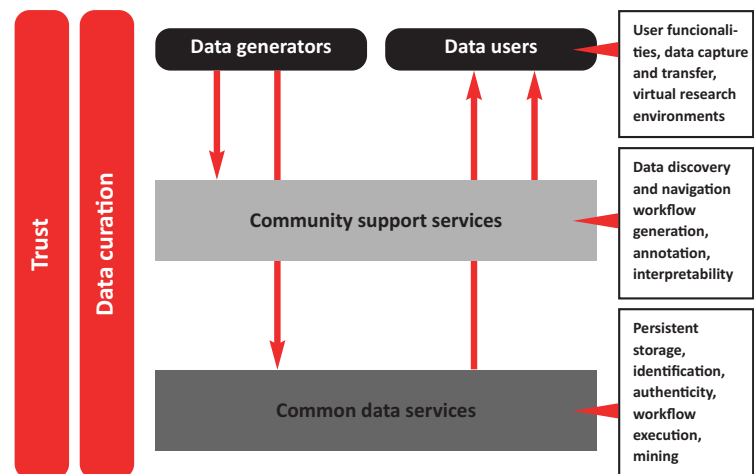
Roles and responsibilities in the FO-BO model

The focus of the **front office** is on supporting its own research organization. In the area of data management the front office takes care of awareness raising, providing information and training its researchers. In addition, the front office features so-called Virtual Research Environments or Data Labs, which offer research tools and secure temporary storage facilities (Sharepoint, Dataverse, etc.) for the organization's researchers. In consultation with the back office, the front office also facilitates the transfer of data to a trusted back-office digital repository after the research has been completed. Facilities that are shared by several universities, including Dataverse, can be hosted and supported by the back office.

The federated data infrastructure

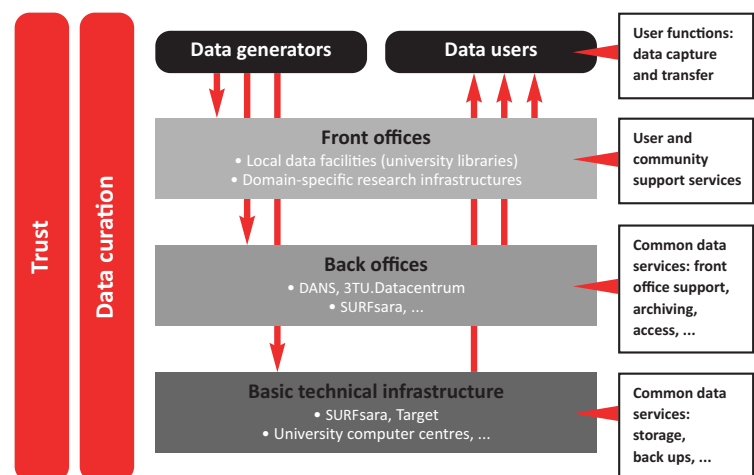
The top diagram, taken from the report "Riding the wave: How Europe can gain from the rising tide of scientific data" (October 2010), gives a general impression of how the various actors, data types and services should be interconnected in a global e-infrastructure for science. In this system each layer includes the right facilities for managing data in a sustainable manner and guaranteeing their reliability.

The bottom diagram shows how Research Data Netherlands has adapted the system recommended in the report Riding the wave.



Model 1: Federated data infrastructure according to Riding the Wave

Source: 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



Model 2: Federated data infrastructure in the Netherlands according to Research Data Netherlands

Data acquisition within its research community is another front-office duty. In all its tasks the front office will, if necessary, maintain contact with the back office. The focus of the **back office** is on the expertise surrounding data governance and data stewardship, including long-term storage and accessibility of the research data.

The back office is responsible for training the data librarians / experts employed by the front office and providing the front office with substantive support via dedicated contacts. Back-office employees may act as experts and contribute to the front-office training activities for researchers.

Where needed, the back office also provides consultancy services to the front office. In other words, the back office acts as a centre of expertise and innovation. Furthermore, the back office ensures the sustainable and secure storage and retrieval of data upon completion of the research project. For this purpose the data are transferred through the front office to the back office.

In the acquisition, support, consultancy and training services, the duties of front and back offices may overlap. Coordination and definition of responsibilities will therefore be necessary. Specific responsibilities will vary from organization to organization, but it is important to have clear agreements on e.g. data acquisition, NWO data contracts and the use of data management plans.

Organizations differ in subject, size and type of research data, staffing capacity, ambition, etc. These differences may lead to different designs of the FO-BO model. There will be institutes performing front-office tasks only, while outsourcing their back-office tasks to a data archive, but there will also be organizations with both front-office and back-office tasks (e.g. 3TU.Datacentrum).

DANS's cost model

DANS will charge the basic data storage costs including back-up. Storage is currently supplied by Vancis, a subsidiary of SURFsara. There is a Service Level Agreement with Vancis in place which guarantees a very high level of security and data availability.

In return for a one-off payment of these costs for five years in advance, DANS ensures conservation of the data "forever". This will enable the long-term safekeeping of data from projects with temporary funding.

This arrangement assumes that data and metadata are supplied in the agreed format. Where this is not the case, DANS will charge €75 per hour for processing data and organizing documentation. For larger consultancy projects DANS will charge an hourly fee of €100.

3TU.Datacentrum's cost model

In its pricing policy, 3TU.Datacentrum distinguishes between its partners, contracted customers and customers without a contract (individual depositions). Partners of 3TU.Datacentrum are members of the 3TU.Datacentrum consortium. Both partners and contracted customers who make their data accessible via 3TU.Datacentrum pay a one-off fixed rate per TB with various options for retention periods. Partners of 3TU.Datacentrum pay a lower rate than customers do. For self uploads (up to 4GB) by individual researchers not employed with a customer or partner, no fee is charged.

3TU.Datacentrum purchases the research data storage, including two back-ups, from TU Delft's Shared Service Center ICT. The services have been defined in a Service Level Agreement which guarantees secure and reliable storage.

Benefits for back office, front office and research community



Benefits for researchers

- + Researchers benefit from increased **data curation knowledge** at FO
- + Researchers have **better access** to sustainable storage of their data
- + Researchers **gain time** by engaging in data management in an early stage



Benefits for front office (FO)

- + FO benefits from **data storage facilities** at BO
- + FO benefits from BO's **knowledge** on data curation
- + FO supports library's mission to be a **partner in research**



Benefits for back office (BO)

- + BO benefits from **well trained FO contact persons** for researchers
- + BO benefits from **direct contact** with their target group through FO
- + BO **acquires more data** that are better prepared for archiving and reuse

The collaboration between back office, front office and research community benefits all parties.

The business model of Research Data Netherlands

Research Data Netherlands (RDNL) is a partnership of trusted digital repositories with a back-office function. RDNL offers a one-stop shop for back-office services, which may be carried out by several service providers. Current participants in RDNL are DANS and 3TU.Datacentrum. In September 2014 they were joined by SURFsara.

Based on its business model as well as the FO-BO model, RDNL seeks to conclude institution-wide framework agreements with university front offices. Such a framework will include more detailed agreements concerning the data management and storage services provided. They may involve the division of the respective responsibilities, but also the establishment and maintenance of the required technical infrastructure. Separate agreements are made with the RDNL partners. Both DANS and 3TU.Datacentrum use their own cost model.*

** Since RDNL is not a legal entity, no pricing agreements shall be made between the RDNL partners, who are self-financing. SURFsara's cost model will be included in a subsequent edition and also be published on our website.*

About Research Data Netherlands

Research Data Netherlands (RDNL) is an alliance between 3TU.Datacentrum and Data Archiving and Networked Services (DANS). In September 2014 SURFsara has become the third alliance member. With this coalition, the three data archives join forces in the area of long-term data archiving. Thus they create one point of contact for the purchasers of their services, usually front offices of universities and other institutions. In this way RDNL promotes sustainable archiving and reuse of research data in the Netherlands. RDNL has been operational since May 2013 and is open to other partners with a back-office function in data archiving.

More information

www.researchdata.nl/en
info@researchdata.nl

Partners of Research Data Netherlands

3TU.Datacentrum

Data Archiving and Networked Services

DANS

SURF SARA