Focus on FAIR

DANS: Dutch national centre of expertise and repository for research data

Author:
Henk Wals, Director

December 2020

>>>

Focus on FAIR

This year, DANS is celebrating its 15th anniversary. We are proud of our achievements. In 2020, with more than 150,000 datasets, DANS is in the top five of similar repositories in the world. We helped shape a European data infrastructure and developed an internationally recognized quality seal for repositories. In the process, DANS blossomed into a worldwide centre of expertise on FAIR data. We became good at executing and managing European projects. That is one of reasons why we have grown. In 2005, we started out as a team of 15; now there are 60 of us.

It is our mission to enhance the reusability of research data and thus the quality of scientific research. The use of FAIR data improves the verifiability and reproducibility of research. It also enhances efficiency because building datasets is expensive. What is more, linking data can lead to new discoveries and insights. Reusing and connecting data accelerates knowledge circulation and increases our ability to solve complex issues. In short, making better use of data leads to better science, which matters to society.

In other words, the work of DANS is socially relevant. And there is still a lot to do. Most datasets of published research in the Netherlands still reside on local servers or disks, have a limited shelf life and are inaccessible to the outside world. FAIRness of research data is in need of a solid boost, both financially and in terms of collaboration and coordination. To this end, *the National Plan for Open Science* has been set up. Being one of the major players in the Dutch data landscape, DANS can make a powerful contribution to this initiative.

Drawing up our fourth 5-year programme we have not allowed ourselves to be guided by the question of how DANS alone can continue its success. Instead, we started from the question of what the Dutch data landscape needs to be successful and how DANS can make its most effective contribution. Our answer is that we can do so in three ways: as a versatile data repository, as a centre of expertise for FAIR data, and as a partner. This has resulted in an innovative programme that we will briefly explain below. The keepers of DANS, i.e. KNAW and NWO, have expressed their support.

Henk Wals
Director of DANS
December 2020

Three programme pillars

Centre of expertise for FAIR research data

DANS will expand and renew its range of expert services. There clearly is a growing need for FAIR data expertise to which DANS can make a substantial contribution.

DANS has participated in dozens of international projects to develop standards, techniques, methods and practices for improved sharing, reuse and archiving of research data. In the meantime, as a data repository, we have learned from the complicated issues we have been confronted with in practice. It is no false claim if we state that DANS has grown into a straightforward centre of expertise through this combination of theory and practice. In the coming period we are going to make this function explicit. In this way, the demand for expertise on FAIR data and Research Data Management will be given a clear focus. A team of approximately 20 Research Data Management Consultants and R&D engineers will be engaged in the following activities.

- Providing expertise as a participant in research and data infrastructure projects, especially at the European level. Demand for this kind of participation is expected to increase. We see a new challenge in the transfer of newly developed knowledge from international projects to the national level.

- Taking administrative responsibility within international organisations, as we are currently doing in *the Research Data Alliance*, *CoreTrustSeal*, *the World Data System*, *DataCite*, *DARIAH*, amongst others.

- Providing consultancy on demand and as a member of national and international advisory bodies.

- Training data professionals and researchers. Data-handling knowledge is not yet an established part of university curricula. In general, it can be said that in research circles there is a lack of data literacy. There is also a growing need for advice, training, and refresher courses among data stewards.

- To strengthen the foundation of its own strategy formation, DANS will set up a new team to acquire business intelligence (BI). The BI team will focus on collecting and analysing information about users, competition, the market, and trends in data processing within scientific research. Most of these data and analyses will be shared in an open-format monitor. Such a monitor seems particularly useful to us, because at the moment this information is largely missing while it is highly relevant for shaping national and European policies - and assessing their effects.

Versatile data repository

DANS is going to radically overhaul its data services offering. Starting point of this change is that requirements among the scientific disciplines vary considerably. The nature and dimensions of the datasets are also greatly different. These differences are reflected in the diversity of the national landscape. There are excellent data services that are close to the research, but they do not cover everything. That is the task of more generic repositories such as DANS.

The downside to generic catch-all repositories is their one-solution-fits-all approach. Indeed, researchers can do little more than deposit or download datasets. There is a need for data services that seamlessly integrate with *virtual research environments* (VREs) and provide editing, version management and analysis capabilities using external tools. Obviously, long-term preservation of the dataset and its usability must also be possible. However, as VREs and research practices in various disciplines vary widely, different solutions will have to be found for different research disciplines.

No generic repository has yet been successful in this respect, but DANS is taking up the challenge. Applying a new, modular services structure, we will provide tailor-made services to various user groups, while at the same time our infrastructure will grow with these new developments and continue to offer stability and reliability.
DANS will offer a combination of 'reuse' and 'long-term preservation' services. The technological foundation of the DANS infrastructure will be the *Dataverse* platform, developed by an international community led by Harvard University.

Multiple instances of *Dataverse* - referred to as Data Stations - will be customised to the needs of various disciplines and research communities. The Data Stations are intended to be used with datasets that are still being worked with and added to. Researchers can bring their data to one of the DANS Data Stations: *Archaeology*, *Social Sciences and Humanities*, *Life Sciences*, *Nature- and Technical Sciences*. Each
station is assigned a Data Station Manager, who will establish and maintain contacts with the relevant research community.

In addition to the Data Stations, DANS will continue to offer *DataverseNL*, the platform service which allows universities, universities of applied sciences, and research institutions to establish their own repositories.

Data contained in Data Stations and *DataverseNL* will automatically finds its way to the DANS Data Vault, a secure, reliable, and certified long-term preservation repository. The Vault also contains all datasets that have previously been entrusted to DANS. In addition, the Vault is a service provided to institutions and companies looking for long-term reliable storage of their data archives. DANS performs this service in a uniform manner based on international standards and best practices.

In order to realise the new set-up, DANS is making drastic changes to its organisation. A team of Data Station Managers and an R&D team will be established. All technological units at DANS are going to be managed by a Chief Technology Officer, who will drive constant innovation from the board of directors.

Active collaborator

DANS is going to invest more in collaboration and coordination. The task of raising the share of FAIR research data in the Netherlands to an acceptable level and facilitating its analysis requires a joint approach. Enabling awareness campaigns, coherent training programmes, the connection of datasets from various research disciplines: no organisation in the Dutch data landscape can achieve these goals on its own.

DANS wants its services to be complementary - and if possible connected - to the services of other providers. In the coming period DANS will work to promote collaboration and coordination, partly in the context of *the National Plan for Open Science* (NPOS). We advocate a pragmatic approach in which existing organisations take their responsibility in consultation with the NPOS board. At first, this should involve the larger discipline-generic organisations with sufficient resources and strength. DANS is such an organisation, as are the eScience Center and SURF. The main focus areas of these three organisations are more or less complementary: data, software, infrastructure. Connections between these domains are, however, also needed. If the three national organisations succeed in coordinating their activities and then start to facilitate coordination within the landscape, much will have been achieved.

At the same time, DANS is convinced that concrete bottom-up initiatives on a smaller scale can also be effective. We will be happy to collaborate in projects that lead to improvement of services and better coordination. We will initiate such projects ourselves. We intend to continue the Research Data Netherlands (RDNL) partnership with 4TU.ResearchData and SURF.

<<<

This document presents the 5-year programme (2021-2025) by DANS. The programme is based on the question of what the Dutch data landscape needs to be successful and how DANS can make its most effective contribution. We can do so in three ways: as a versatile data repository, as a centre of expertise for FAIR data, and as a partner. This resulted in the innovative programme that is explained in this document. The keepers of DANS, i.e. KNAW and NWO, have expressed their support.

Data Archiving and Networked Services (DANS)

DANS is the Dutch national centre of expertise and repository for research data. We help researchers make their data available for reuse. This allows researchers to use the data for new research and makes published research verifiable and reproducible. With more than 150,000 datasets and a staff of 60, DANS is one of the leading repositories in Europe.

DANS is an institute of KNAW (Royal Netherlands Academy of Arts and Sciences) and NWO (Dutch Research Council).