Self-assessment DANS
for periodic evaluation of the institute

Version 1.7
06-10-2017
Content

Preface: DANS and the Standard Evaluation Protocol .............................................. 2
Introduction .................................................................................................................. 3
1. Organisation ............................................................................................................. 4
  1.1. Mission ............................................................................................................. 4
  1.2. Governance ....................................................................................................... 4
  1.3. Composition and staff ..................................................................................... 5
  1.4. Finances .......................................................................................................... 5
  1.5. Organisation development .............................................................................. 5
2. Strategy .................................................................................................................... 7
3. Results achieved in the past six years (2011-2016) ............................................. 8
  3.1 Quality of DANS research data services .......................................................... 8
  3.2 Scientific quality of research at DANS ............................................................. 10
  3.3 PhD students .................................................................................................... 10
  3.4 Relevance to society ......................................................................................... 11
4 Relevant environmental developments, forecast of trends, and viability .......... 12
  4.1 Relevant environmental developments over the past six years ....................... 12
  4.2 Forecast of trends in the coming years .............................................................. 13
  4.3 Viability ............................................................................................................ 14
5 SWOT analysis and benchmarking of DANS ....................................................... 16
  5.1 SWOT Analysis ................................................................................................. 16
  5.2 Benchmarking .................................................................................................. 17
6 Integrity and standards ........................................................................................... 19
  6.1 Institutional and data integrity ........................................................................ 19
  6.2 International certification standards ................................................................. 20
7 Diversity .................................................................................................................. 21
Appendix 1: D1 – Output indicators ........................................................................ 22
Appendix 2: Datametric analysis of DANS Data Archive ........................................ 24
Appendix 3: D3a – Research staff ............................................................................. 32
Appendix 4: D3b – Research output ........................................................................ 32
Appendix 5: D3c – Total funding of the institute ..................................................... 32
Appendix 6: D3d – PhD candidates ........................................................................ 33
Appendix 7: List of abbreviations and definitions .................................................. 34
Appendix 8: A selection of national and international projects carried out by DANS in the past few years, and some recent data/research infrastructure proposals ............... 38
Appendix 9: Other relevant documents .................................................................... 39
Preface: DANS and the Standard Evaluation Protocol

The following quote from the previous evaluation of DANS still applies: "DANS’ position in the KNAW is unique, for two reasons. First, it is the only institute that is funded by the KNAW and NWO. Second, it is the only institute where the emphasis is on service, not research."

The Standard Evaluation Protocol (SEP) 2015-2021, which describes the methods used to assess research conducted at Dutch universities, NWO and Academy institutes every six years, as well as the aims of such assessments, does not apply very well to DANS. The SEP assessment criteria are grouped in three broad categories:
1. Research quality
2. Relevance to society
3. Viability

Doing research is a secondary activity at DANS. Our modest research is mostly applied and explicitly aims to support the development of our core services. We welcome an evaluation of the quality of our small research group, but assessing DANS on those criteria does not do justice to the main task of the institute, which is to encourage permanent access to digital research resources. To rank such services under “relevance to society” is also inadequate. We serve primarily a scientific audience: the majority of our users are scholars and scientists, and it makes sense to distinguish them from the wider public.

We have nonetheless attempted to apply the SEP criteria as much as feasible, but we suggest to interpret “research quality” as “quality of scientific services” when evaluating DANS.

---

2 The documentation on the previous evaluation of DANS in 2010-2011 is available here: https://www.knaw.nl/en/institutes/dans.
Introduction

Over the past ten years, about 2.5 million data files from 225 thousand datasets have been downloaded from the DANS data archives. The content of the archives grew from around 15,000 files (in 1,500 datasets) in 2007 to over 4 million files (in 35,000 datasets) today (August 2017). During this decade, half a million users visited the data archives, paying together over a million visits. The annual number of visits grew from 20,000 in 2007 to 160,000 in 2016. The annual number of users has been fluctuating around 60,000 a year since 2010. The total volume of data in the DANS EASY archive grew from 0.1 TB in 2007 to about 15 TB today.

Research data are now generally regarded as a valuable research output, much more so than at the time of the previous evaluation. The past few years have been quite turbulent where research data are concerned. Many new players in this area have appeared, although several newcomers have already given up and disappeared. An international registry counted just a few hundred research data repositories in 2012, about 1,500 in early 2016, and close to 2,000 now. DANS is regularly adapting both its organisation and its services to the rapidly changing demands and rising expectations.

In the Netherlands, virtually all self-respecting research organisations (funders, universities, institutes) have formulated data policies over the past five years. We see two interrelated explanations for the increased interest in research data, also at the political level:

- Big Data is offering great new opportunities for research.
- Open Science (including Open Data) has become a new paradigm: transparency of the full research cycle is both increasing scientific productivity and preventing malpractice.

Currently, institutional data policies are at different stages of implementation. DANS started providing back-office services to institutional repositories already a couple of years ago, offering expert advice and training, technical and organisational repository solutions, aggregating and disseminating the increasingly distributed data and research resources, and long-term archiving options to institutions, national and international. DANS is aware that it is a relatively small player from a country of modest size, and therefore puts a lot of energy in forging both national and international coalitions and collaborations. With our background in the Digital Humanities and Social Sciences, we have succeeded in becoming a respectable player internationally, even among Big Science facilities. We consider our international role vital for our future direction and indispensable for our national role.

DANS has recently started a complete overhaul of its service infrastructure, adjusting them to expected future requirements, focusing on our unique selling point of a combination of deep expertise and practical experience in providing data services, making sure the services are adaptive to the changing needs of both individual researchers and institutions. Research and Development projects underpin our innovative spirit. Like DANS was the cradle of the international Data Seal of Approval, internationally the most successful standard for trustworthy digital repositories, our ambition is to be the first to propose a practical implementation of the FAIR (Findable, Accessible, Interoperable, Reusable) data principles and a hands-on mechanism to establish the conditions for data sharing under the new European privacy regulation.

---

3 Users and visits are counted on a “daily unique” basis: two visits by the same user on the same day count as one, but on two different days they count as two.

4 Calculated here is: single storage of data for distribution; in reality, data is stored multiple times in at least two different geographic locations, and of certain file types archival versions (not for distribution), such as high-resolution images and videos, are stored separately.

http://www.re3data.org/
(GDPR), which will come into effect in May 2018. By collaborating with initiatives such as the Software Heritage Archive (at INRIA, France) and the Software Sustainability Institute (Edinburgh, UK), we will also be able to provide research software archiving services at minimal additional cost.

1. Organisation

As a joint institute of KNAW and NWO, DANS is governed by both scientific umbrella organisations. Organizationally DANS is part of the KNAW. The governance of the institute is described in a covenant between KNAW and NWO.

1.1. Mission

Data Archiving and Networked Services (DANS) is the Netherlands institute for permanent access to digital research resources. DANS encourages researchers to make their digital research data and related outputs Findable, Accessible, Interoperable and Reusable (FAIR). We provide expert advice and certified services. Our core services are: DataverseNL for short-term data management, EASY for long-term archiving, and NARCIS, the national portal for research information. By participating in (inter-)national projects, networks and research, DANS contributes to continued innovation of the global scientific data infrastructure. Open if possible, protected where necessary.

1.2. Governance

Twice a year DANS holds meetings with external committees, in different compositions, to monitor the general strategy and our key performance indicators:

- A Steering Committee consisting of one KNAW director and one NWO director meets with the DANS director to discuss the general strategy of the institute. The Chair of the Scientific Advisory Board attends these meetings as an advisor.
- The Scientific Advisory Board: offers solicited and unsolicited advice to the Steering Committee and DANS management.
- ‘Periodiek Bestuurlijk Overleg’ (PBO) with the KNAW director and a delegation of the board of the KNAW. Since 2016 the KNAW governs its institutes with the help of a convenant in which key performance indicators have been defined.
- For its key services: DataverseNL Advisory Board and NARCIS Advisory Board.

---

6 These principles also underly the Data Seal of Approval, which preceded FAIR about a decade.
7 See: https://dans.knaw.nl/nl/over/organisatie-beleid/informatiemateriaal/SamenwerkingsovereenkomstDANSKNAW2015.pdf (available in Dutch only)
8 See: https://dans.knaw.nl/nl/over/organisatie-beleid/informatiemateriaal (available in Dutch only)
9 More information on all boards and committees is available via https://dans.knaw.nl/en/about/organisation-and-policy/steering-committe-and-advisory-boards
1.3. Composition and staff

The organizational chart of DANS is as follows (Figure 1):

![DANS organizational chart]

The three departments of DANS are subdivided into six “competency groups”. The director and deputy-directors each head a department; coordinators lead the competency groups. The management team meets on a weekly basis, and a joint meeting of the directors with the coordinators takes place each month. The staff of DANS consists of about 50 people, including trainees and research students. There are also a number of volunteers who carry out specific tasks. DANS has carried out work satisfaction surveys in 2011 and 2014. The average “mark” (on scale 1-10) the DANS employees gave to the institute was a 7.8 in 2011 and 7.7 in 2014. Tables on the personnel composition are available in Appendix 3.

1.4. Finances

The funding of DANS has remained virtually stable over the past six years, with the following fluctuations:

- In 2011 DANS took over the NARCIS portal from the KNAW Bureau, together with the staff involved. This became visible in the budget received from KNAW in 2012, which rose from €2,222 to €2,577. From 2014 onward, the budget received from KNAW decreased slightly and has been fluctuating around €2.5.
- The budget received from NWO, which was €789 in 2011, declined with €100 in 2012 and remained virtually stable since then. This decline was related to a compensation in the housing costs in the years 2005-2011 (which stopped).
- The budget received from externally funded projects and paid data services grew from €899 in 2011 to €1,644 in 2017.\(^\text{10}\)
- Adding up everything, the total DANS budget grew from 3.8 M€ in 2011 to 4.9 M€ in this year (2017). In relative terms, the percentage of externally funded projects and services grew from about 23% to 35% (see Appendix 5).

By economizing IT costs and by the growth of the externally funded activities, DANS has managed to keep up financially with the steady increase of its data holdings. The joint contribution to the DANS budget by KNAW and NWO decreased from 77% to 65%. We have flagged the situation to our parent organisations and made a plea for a budget increase in our last two financial reports.

1.5. Organisation development

DANS has implemented several organisational adaptions, at approximately two-year intervals, in order to meet the requirements of a growing organisation, which is in the process of conquering and consolidating its place in the scientific information landscape. The basic principles underlying these adaptions included: restricting organisational overhead, closely connecting leadership to staff, linking the dynamic project organisation

---

\(^{10}\) Mid-year figure summer 2017; some €0,1 should be added from other sources of income.
to the more static line organisation, and creating a more efficient consultation structure without communication barriers amongst departments or between sections.

Next to the upkeep of our technical infrastructure (see also 3.3), participating in external national and European projects is an important part of the daily routine within DANS. Collaborating with colleague organisations around Europe and the rest of the world enables us to gain profound expertise and to improve the (inter-)national cohesion of research data services and infrastructure. In order to guarantee a sound financial and procedural progress, including timely completion of projects, DANS has established an internal project office several years ago. More recently a project acquisition team was organised, consisting of the general director, the policy (vice-)director and two project developers, to ensure a sustained flow of projects on a wide variety of policy and technical data topics, in line with the DANS mission. Project funding has appeared to be not only a source of increasing our budget, but also a means to engage in innovation and to be part of the strongest international infrastructure consortia.
2. Strategy

The time frame of this evaluation covers two strategic periods, the “DANS Vision and Strategy 2011-2015” and “Sharing data together: DANS strategy 2015-2020”. The strategy 2011-2015 was formulated after the previous external evaluation, taking the recommendations of the visitation committee and the boards of KNAW and NWO into account. Links to the documents related to the previous evaluations are provided in Appendix 9.

In spring 2014 DANS underwent a less formal midterm evaluation. At approximately the same time, Prof. H.J. van den Herik was asked to write an advice on “DANS and data”. This report and the midterm evaluation were taken into account when DANS formulated its strategy 2015-2020.

The changes in the formulation of the DANS mission from 2011 to 2015 were minor. The revisions explicitly addressed the take-over of the NARCIS portal by DANS from the KNAW central office (in 2011) and the support of data management during ongoing research, facilitated by the take-over of the DataverseNL from Utrecht University (in 2014).

Starting with the 2011 strategy, DANS formulated its priorities as much as possible in a quantitative way, and we set out our goals in time on a Roadmap, which were monitored and reported every three months. The 2011 roadmap counted four strategic priorities, which were operationalized into 13 concrete (“SMART”) targets. The four priorities were:

- DANS will strengthen its services by serving more users more efficiently;
- DANS will develop into a discipline-independent data organisation;
- DANS will conduct research to support and improve its services;
- DANS will be an important building block in data provision in Europe.

In our strategic update of 2015, we condensed the priorities to three, and the operational objectives to seven, although the differences were practical rather than fundamental:

- We reduced the level of detail in measuring user satisfaction and supply of services;
- We left out the explicit statement about the disciplinary coverage (after the creation of RDNL);
- We left out the explicit aim to do research on digital preservation;
- A new aim was formulated to promote software sustainability;
- A new aim was added to support RDM during research, replacing the aim of participating in projects with research communities.

Key performance indicators for measuring our success are represented in Appendix 1 and in a Datametric analysis in Appendix 2.

---

11 Both strategy reports can be found on our website: [https://dans.knaw.nl/en/about/organisation-and-policy/information-material](https://dans.knaw.nl/en/about/organisation-and-policy/information-material)
3. Results achieved in the past six years (2011-2016)

As mentioned in the Preface, assessing "scientific quality" in terms of the usual SEP-indicators for research institutes does not do justice to the DANS mission and the majority of our activities, which are to offer scientific data services. We therefore draw a distinction between the scientific quality of our limited research activities and the quality of our scientific data services.

3.1 Quality of DANS research data services

As described in section 2, DANS measures its success on the basis of strategic priorities and specific objectives, as defined in our strategic reports. We will present a summarizing narrative of our achievements of the past six years. The underpinning quantitative indicators are given in Appendix 1.

Objective 1: Prominent in federative data infrastructure

The core indicators we use for measuring our prominence in the federative data infrastructure are based on how we contribute to national and international data infrastructures and related projects to develop or strengthen them. The number of projects we participate in remains more or less stable. Financially, the proportion of externally funded activities has grown over the past six years from about 23 to 35% of the total DANS budget. DANS is among the most successful KNAW institutes in acquiring EU Horizon 2020 projects. In absolute terms, several of the bigger institutes in the life sciences acquired more project revenues, but taking the size of the institute into consideration, DANS scores better than any other KNAW institute. A selection of projects is presented in Appendix 8.

DANS staff also performs board, management and other vital governing functions in a plethora of infrastructure and data organisations, such as DARIAH (CIO), CESSDA (chair of General Assembly), NCDD (board member), RDNL (chair/board), LCRDM (chair/member of five working groups), RDA (secretary general, co-chair of three groups, member of technical and organisational advisory boards, host of RDA Plenary 2014), DRYAD (board), WDS (co-chair/board), DSA (chair/board/secretariat), EuroCRIS (secretariat, board), Science Europe Working Group on Research Data (chair).

Another important indicator of our prominence is that DANS regularly receives visitors from other data organisations (from countries including Sweden, Austria, China, Japan, South Africa), who want to learn from our experience.

Objective 2. Growth of services: supply of information resources

The growth of the supply of information resources is measured on the basis of the content in the three core service systems of DANS: EASY, NARCIS and DataverseNL. For every service, we measure two indicators.12 DANS has started to offer data vault services to other repositories which are less well-placed to guarantee long-term preservation and access. Prominent examples are Elsevier's Mendeley Data and Dryad.13 This means both an internationalisation of our services and a broadening in terms of disciplines covered.

Given our experiences as the host of DataverseNL, we are exploring the possibility to offer a similar service for emerging data repositories/archives in Europe, starting with aspiring CESSDA service providers. These services are offered on a paid basis.

Objective 3. Growth of use of services: demand

The growth of the use of the DANS services is calculated on the basis of two indicators per service as well. We have not carried out a user satisfaction survey recently, but in quantitative terms the use of our services is growing spectacularly. Especially the

---

12 Appendix 1 provides an overview of the figures of the six indicators. Moreover, we made a separate datametric analysis with graphs on the development of the core data archive EASY, provided in Appendix 2.

13 Mendeley data see: https://data.mendeley.com/; for DRYAD, see: http://datadryad.org
growing use of NARCIS as access portal for Dutch research resources, including the DANS data archives, is remarkable. In the datametric analysis a deeper analysis is made of the use of the data archive (see Appendix 2).

Originating in the humanities and social sciences, DANS takes a step-by-step route to expand its services to other disciplines, starting with the life sciences. We embark on this course in direct consultation with the relevant parties and institutes (such as the Dutch Techcentre for the Life Sciences, DTL), offering existing services, where there is a demand, acting as a niche supplier. Our focus is on the long-tail of research data rather than on “Big Data”.

**Objective 4. Support for “living” data and RDM**
DANS offers DataverseNL as back office service to universities, higher education institutions and research institutes. When we took over the Dutch Dataverse Network from Utrecht University in 2014, there were eight partners; by the end of 2016 there were 11 (4TU Federation counts as one), and more institutes have expressed their interest to become partner, of which several will do so in 2017.14
In addition, DANS provides courses, online webinars, written instructions, guidelines and consultancy on research data management.

**Objective 5. Software Sustainability**
Both in order to keep data accessible, and as valuable output of the scientific process in its own right, it is useful to keep software sustainable. DANS was never tasked to sustain software tools from scientific research. Yet, there is a growing demand for such services. Software archiving and maintenance are very costly, and DANS is not budgeted to do this. This is why in our strategy from 2015 onwards we formulated software sustainability as a goal that is conditional on funding. We carried out several exploratory projects and feasibility studies in this area. In 2016 DANS signed a MoU with INRIA in France to support the Software Heritage Archive (SHA) created there. Goal of this initiative is to collect, preserve, and share all software code that is publicly available. DANS will provide software archiving services to Dutch researchers in collaboration with SHA. We will continue our endeavour in the area of Software Sustainability. This is still largely a vacuum in the infrastructure landscape, and DANS aims to set up a joint European Infrastructure for Software Sustainability.

**Objective 6. Linking information types**
DANS aims to connect datasets, publications and research information and to offer them to users in mutual coherence. This is why the institute carries out research and development in the domain of Linked Data. Some of the results are:
- Dissertation on CEDAR project (linking historical census data) in spring 2016; several papers, articles, symposiums;
- Consultancy on European PRELIDA-project (Preserving Linked Data) 2013-2015;
- Project proposal *Digging into Data* (with University of Wisconsin, Milwaukee);
- CEDAR datasets archived in EASY;
- PARTHENOS project: contribution to formulation of standards for Linked Open Data in Humanities context;
- EHR2: “Innovative services for archives” - storing Holocaust information as Linked Data;
- CLARIAH: Core role in setting up a Digital Humanities Linked Data Cloud.15

---

14 Please see Appendix 1.
Objective 7. Innovation and integration of systems
The innovation of the DANS services is an ongoing process, including updating and upgrading the underlying software and adding new functionality. Some results of the past years:

- NARCIS: upgrade harvester and indexer techniques (Meresco); evaluate OpenAire CERIF format for exchanging project, organisation and person data;\(^{16}\)
- EASY: renewal of deposit and ingest functions. Mass ingest via SWORD2-service. Renewal of archival storage layer (from Fedora 3 to BagIt);\(^{17}\)
- DataverseNL: upgrade from version 3 to version 4, implementation of Handles and federated login using SURFconext.

In order to keep up with the ever-changing environment surrounding us, the technical infrastructure, on which the data services run, have to be substantially upgraded every few years. For the period 2017-2020, DANS has established the “DANS Data Services 2020” programme to bring the technological infrastructure up to a competitive level. We have drafted the outline of a plan to innovate and integrate our core services. The outline of this plan was approved by the KNAW in spring 2017. It entails a much-improved national portal for access to research resources, up-to-the-minute data ingest and access functions, provisions for archiving and exposing linked data, and a secure data vault for long-term preservation. For 2017-18 DANS has made available K€ 500 from its general reserve to begin this work.

3.2 Scientific quality of research at DANS
The small Research and Innovation group (R&I) of DANS focuses on the support and innovation of the DANS services. It studies how to make digital information available in a sustainable way during the different phases of the research cycle. R&I collaborates in different research projects with other parties.\(^{18}\)

Most members of the R&I team work only part-time on research tasks. Although 12 people are involved, the staff of DANS counted as scientific is five FTE. Visiting scholars and professors have contributed to the R&I work at DANS.\(^{19}\)

One of our researchers, Dirk Roorda, was winner of the Digital Humanities Award for the best tool in 2014, for his work on the SHEBANQ project (System for HEBrew Text: ANnotations for Queries and Markup).\(^{20}\) A quantitative overview of publications, conference papers, etc. is available in Appendix 4.

3.3 PhD students
DANS does not have a PhD programme, nor lump-sum funding to employ PhDs. Nevertheless, we aim at employing PhDs on external project funding whenever an opportunity arises. One PhD successfully finished his dissertation on Linked Data in the Dutch historical censuses in 2016 (Albert Meroño Peñuela, CEDAR project)\(^{21}\); another has started recently (Kathleen Gregory, Re-SEARCH project).\(^{22}\)

- **Context and quality assurance of PhD research**: in both cases the quality of the research proposals was guaranteed by the external refereeing of the projects (in open competition).
- **Participation in graduate/research schools**: Albert took part in the research group led by Prof. Frank van Harmelen; as long as it existed, he also participated in the eHumanities Group of KNAW.
- Kathleen, who has just started her PhD, participates in WTMC led by Prof. Sally Wyatt.\(^{23}\)

---

\(^{16}\) Common European Research Information Format; see [http://www.eurocris.org/cerif/main-features-cerif](http://www.eurocris.org/cerif/main-features-cerif)

\(^{17}\) Simple Web-service Offering Repository Deposit; see [http://swordapp.org/about/](http://swordapp.org/about/)

\(^{18}\) See: [https://dans.knaw.nl/en/about/research-and-innovation/research-projects](https://dans.knaw.nl/en/about/research-and-innovation/research-projects)

\(^{19}\) The team members are mentioned here: [https://dans.knaw.nl/en/about/research-and-innovation/team](https://dans.knaw.nl/en/about/research-and-innovation/team)

\(^{20}\) See: [https://shebanq.ancient-data.org/](https://shebanq.ancient-data.org/)

\(^{21}\) See: [https://www.cedar-project.nl/](https://www.cedar-project.nl/)

\(^{22}\) See: [https://dans.knaw.nl/en/projects/projects](https://dans.knaw.nl/en/projects/projects)

\(^{23}\) Netherlands Graduate Research School of Science, Technology and Modern Culture: [http://www.wtmc.eu/](http://www.wtmc.eu/)
• **Selection and admission procedures:** in both cases the selection was international, and an interdisciplinary hiring committee was set up with specialists from various institutions (universities and DANS).

• **Supervision of PhD candidates internally and guidance of PhDs to labour market:** the internal guidance was/is carried out by dr. Andrea Scharnhorst, coordinator of the DANS research group. Albert was also supervised by a DANS postdoc, a linked-data specialist. The promotor was Prof. Frank van Harmelen. After finishing his PhD Albert was employed by VU to work on the CLARIAH linked data cloud. Kathleen is guided externally by Prof. Sally Wyatt.

3.4 **Relevance to society**

• **Use of core services**

Although the services of DANS target a scientific audience, they have a relevance for the wider society as well. User surveys carried out in the past have shown that in particular NARCIS draws a broad audience. Especially policy makers and journalists/news media are important user groups. Moreover the “informed citizen” searching for specialists or further information on research information is a considerable user category.

The EASY long-term archive likewise is primarily intended for researchers. Yet, archaeological companies, market and public opinion research bureaus, and several public parties, among which the RCE, Statistics Netherlands (CBS), Planning Bureaus (such as SCP), the Research and Documentation Centre of the Ministry of Justice (WODC), the National Institute for Public Health and the Environment (RIVM) and more use the DANS data archive.

In 2015 DANS negotiated a contract with Elsevier’s Mendeley Data to offer a data vault solution for long-term preservation. In 2017, a similar contract was closed with the international not-for-profit research repository DRYAD.

• **Expertise centre:** Training, Advice/consultancy

The second main role DANS performs for society is as a centre of expertise in the area of long-term digital preservation. DANS provides training, advice and consultancy to a variety of partners in the public sector. DANS provides, with partners from RDNL, the course “Essentials 4 Data Support”. This is an introductory course for those who support researchers in storing, managing, archiving and sharing their research data. The training course was nominated for the DPC Digital Preservation Awards 2016. In addition to lectures on data management for PhD students, more than 200 data supporters have been trained so far.

DANS was the cradle of the Data Seal of Approval, a worldwide method of core certification of digital repositories, also applied outside of the research domain. In the Dutch Digital Preservation Coalition, DANS worked on certification with partners such as the Netherlands Institute of Sound and Vision, the cultural heritage sector, the National Archives and the National Library of the Netherlands.

• **Contribution to policy development on Open Science**

DANS has been a propagator of Open Access and Open Science before the terms became commonplace. Our motto is: “Open if Possible, Protected if Necessary”. NARCIS is an excellent laboratory to study the development of Open Access to publications over time, across disciplines, institutions and publication types.

DANS contributed and still contributes to initiatives, platforms, meetings and reports propagating a more transparent science system, such as the National Plan for Open Science. Currently we are the national Open Access Service Desk for the European Open Access portal OpenAIRE. We will continue to promote NARCIS as the best tool for monitoring the development of Open Access and Open Data.

---

24 See also: [https://www.openscience.nl/](https://www.openscience.nl/)
4 Relevant environmental developments, forecast of trends, and viability

4.1 Relevant environmental developments over the past six years

- Importance of data sharing has grown
  Research data sharing has become of much more importance to scholars, research funders, policy makers, publishers, and even to the general public. One of the goals of the European Horizon 2020 Programme is “open research data” for EU-funded research. This has a positive effect on stakeholders to formulate data management policies and on the willingness of researchers to share research data. This change towards open research data is also evident with the national research funders. The Dutch national research funder NWO wants that the research data from their funded research becomes open following the FAIR-principles “Findable, Accessible, Interoperable and Reusable”. Herewith researchers have to take into consideration aspects like privacy, public security, ethical limitations, property rights and commercial interests. From October 2016 NWO requires all researchers who apply for funding to write a data management plan and to inform how they will store and share their research data.

- Growing (inter)national collaboration, continued fragmentation
  We see that a culture change is taking place among researchers to share data in (inter-) national collaborations and research infrastructures. Examples are European research infrastructure projects like Holocaust studies (EHRI), ARIADNE (integration of archaeological datasets), PARTHENOS and CESSDA.\textsuperscript{26} Overarching infrastructures across domains like EUDAT and the European Open Science Cloud (EOSC) are emerging. The international Research Data Alliance (RDA), of which DANS deputy-director Ingrid Dillo is acting Secretary General, has in a short time developed into the largest global data organisation.\textsuperscript{28} Despite the multitude of coordinating bodies, including RDNL and LCRDM in The Netherlands, the national and international data landscapes are actually still rather fragmented.

- Growing number of data service providers and coordination efforts
  As more players (such as research libraries) became interested in offering research data services, DANS adapted its service policy and in 2013 started to provide back-office services to universities and research institutes, which had direct responsibilities for the researchers in their organisations. The front-office / back-office (FOBO) model is the DANS implementation of the Federated or Collaborative Data Infrastructure, promoted both in international and national policy reports outlining the future way of organizing and servicing the data landscape.\textsuperscript{29} In this model the local data management support offices (often connected to university libraries) acts as front offices, advising researchers on RDM, sustainable storage and data sharing. The front offices use the services of the back offices: certified data archives with at least the Data Seal of Approval. The advantage of this model is that all parties benefit optimally from one another’s knowledge and services. Both NARCIS and DataverseNL provide such collaborative or shared national services, and EASY was extended to receive data flows from institutional repositories.

In order to service a wider range of disciplines and to coordinate the back-office services with the most important providers, thus increasing the efficiency of the national data infrastructure, DANS teamed up with 3TU.Datacenter (now 4TU.ResearchData) and SURFsara to form Research Data Netherlands (RDNL) in 2013/4.\textsuperscript{30}

\textsuperscript{26} See also our list of abbreviations and definitions and our projects page: \url{https://dans.knaw.nl/en/projects/projects}
\textsuperscript{28} See for more information on RDA: \url{https://www.rd-alliance.org/group/rda-organisational-assembly/post/oa-gab-your-chance-lead}
\textsuperscript{29} The idea was described and explained extensively in a special report on the FOBO model: \url{http://www.researchdata.nl/fileadmin/content/RDNL_algemeen/Documenten/RDNL_FOBOmodel-UK-web.pdf}
\textsuperscript{30} See: \url{http://www.researchdata.nl/}
With the creation of the National Coordination Point for Research Data Management (LCRDM) in 2016 at the request of the Association of Universities in the Netherlands (VSNU), several RDNL tasks shifted to the new platform. However, the governance structure of LCRDM and the commitment of the participating institutions to implement the recommendations have yet to mature, and we think additional measures are needed in order to arrive at a nationally and internationally cohesive and collaborative research data infrastructure. DANS plays a vital role in LCRDM and is well-positioned to develop or adapt its future services in direct consultation with the biggest customers: the Dutch universities and their researchers.

4.2 Forecast of trends in the coming years

• **Open Science: EUDAT, OpenAIRE and the European Open Science Cloud**
  The EU has expressed its wish to develop EUDAT as a European-wide consortium of data service providers, OpenAIRE as the portal for access to European funded research resources, and the European Open Science Cloud to be the overarching infrastructure and consortium uniting the whole fabric needed for open science. As a partner in all three consortia, DANS is well-placed to perform an important role here, in particular serving the humanities and social sciences. Part of the EOSC vision is that it will be mandatory for researchers to make data from EU-funded research available under the FAIR principles. In January 2017, the Horizon 2020 European Open Science Cloud project started with a pilot (EOSCpilot). DANS is one of the 33 European partners in this pilot.

• **EU General Data Protection Regulation (GDPR)**
  The new EU General Data Protection Regulation that replaces the current Data Protection Directive may offer new impediments for sharing personal data, if no exemption is made for scientific research. The main problem foreseen is that the GDPR will be so strict that it will restrain research on individual data, and that it will restrict data sharing, especially for the biomedical and social sciences. It still is not sure what the consequences of the GDPR for data sharing will be. This will become clearer in the course of the next years. DANS has developed a prototype for tagging the sensitivity of data in accordance with the GDPR for EUDAT.

• **Governance / portfolio evaluation**
  In the context of the expected “portfolio evaluation” by the Ministry of Education, Culture and Science, concerning the functioning and governance of all NWO and KNAW institutes, DANS has started an exploration of the best possible options for the future. Pending the outcome of this exploration and the portfolio evaluation, DANS aims to strengthen its collaboration with the SURF organisation, in particular SURFsara.

• **Development of data journals**
  The expectation is that the number of data journals will grow, although this may be a temporary phenomenon. Data journals can be an important instrument to give researchers more credits for producing and sharing data. Making available datasets should also be rewarded as important scientific output. Together with Brill publishers, DANS publishes the “Research Data Journal for the Humanities and the Social Sciences” (RDJ). This is a peer-reviewed journal, which contains data papers that describe deposited datasets. The RDJ is e-only and open access.

• **More support for Research Data Management and Data Management Plans**
  Due to the requirements of EC and NWO funding there will be more expertise needed about RDM and DMPs. There will be more training necessary both for data supporters and for researchers.

DANS also took the initiative to make the system of Research Data Management plans more efficient, by propagating the alignment of core RDM policies and requirements.

32 See also: [https://www.surf.nl/en/lcrdm](https://www.surf.nl/en/lcrdm)
supplemented by domain protocols for data management. Science Europe has adopted this approach, and we are working with NWO, the ERC, and DG Research to bring this forward.

4.3 Viability

Looking back
Since around 2010 the attention for research data, its management, preservation and access by many actors in the science domain has increased. Recently Open Science has become a catchphrase capturing the imagination, which has also been taken up politically. The Dutch government has the ambition to play a leading role in promoting open science. NWO and KNAW also have active open science ambitions and policies. Many new service providers have entered the research data arena. The Re3Data registry currently reports around 1,900 data repositories worldwide: local, national, international, disciplinary and generic.

The demand for research data services has skyrocketed and with that the competition amongst service providers also increased. During the last few years DANS managed to strengthen its position as an internationally well-respected, knowledgeable, innovative and no-nonsense service provider. The size and stability of the archive make it attractive for new depositors (network effect), and once a user has found her way, it is attractive to keep using DANS as a repository (lock-in effect). Moreover, many data policies (e.g. by funders like NWO and by a growing number of scholarly journals) recommend the use of DANS.

Looking forward
We think DANS has a strong and healthy starting position for the future challenges. As a service organisation, new visions and strategic directions always start from the needs of the customers, building on the foundation of our past strengths, carefully selecting niches for further development and innovation. For the years to come, we will be forward-looking, anticipating trends in future user needs, in order to further consolidate our position as a valuable service and expertise centre catering core data requirements of science and scholarship.

The Unique Selling Point of DANS is to provide certified, reliable, long-term preservation and access to research resources, which are as open as possible. This will remain the basis of our operation in the next few years, as there is a clear and continued need for such a service, both nationally and internationally. We foster the combination of data expertise with practical services, where we aim to find the balance between flexibility (adapting to new future trends) and stability of our core services:

- Connect everything: research resources, organisations, services: DANS is well aware that we are a component in a much larger system, and that our services work best if we connect them to what other organisations (in other domains, in other countries) are providing. Aiming to reduce the fragmentation of the national and international data landscape, we will continue to forge collaborations, and if the opportunity arises, mergers are not excluded.

- Upgrade the technical infrastructure of our core services (see section 3.1, objective 7). Although this upgrade requires a considerable investment, DANS gains economies of scale because of its central functions. Adding users and datasets to the existing infrastructure is relatively cheap, even if they are from other disciplines.

- Provide access to all research resources: DANS already is a national aggregator and provider of research information and resources in the Netherlands, but this can be expanded to information provided by others and by providing access to worldwide research resources. By collaborating with, among others, DataCite (worldwide data search portal), OpenAire (European research portal), Software Heritage Archive (worldwide access to research software), we can simply use and pass on what others have developed.
Experiment with innovative additions to our services in niches where DANS can make a difference. Examples are: development of a FAIR data assessment tool (to review the quality and fitness for use of datasets in trustworthy repositories); exploring the potential of a DataTag tool to enable the secure sharing of data compliant with the GDPR; making an inventory and perhaps a registry of Knowledge Information Systems (research ontologies, domain vocabularies, thesauri, etc.).

- Be cost-effective and make sure growth of storage requirements is funded by financial contributions by institutions, much like the gold Open Access model for publications.
- Expand services to those disciplines that offer opportunities and that display a demand, filling the vacant niches.
- Leave data services that other parties can (and want to) deliver (better) to them and use the capacity that is freed by this for fulfilling new demands.
- Mine the rich information resources we offer to provide better metrics and insight for strategic and policy development.
5 SWOT analysis and benchmarking of DANS

5.1 SWOT Analysis

In the SWOT Diagram in this section we describe the strengths, weaknesses, opportunities and threats DANS is experiencing in five areas: services, expertise, network, finance & organisation, and vision & policies.

<table>
<thead>
<tr>
<th>Strengths</th>
<th>Opportunities</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Services: with certified long-term preservation as Unique Selling Point</td>
<td>• Services: good prospects for further development of</td>
</tr>
<tr>
<td>DANS services cover the whole research cycle including data management</td>
<td>long-term archiving services for national and</td>
</tr>
<tr>
<td>planning and data support during research.</td>
<td>international institutions and networks.</td>
</tr>
<tr>
<td>• Expertise: unique combination of</td>
<td>• Expertise: national science agenda and open</td>
</tr>
<tr>
<td>theoretical expertise on and practical</td>
<td>science policies (e.g. EOSC) present excellent</td>
</tr>
<tr>
<td>experience with long-term preservation and research data management.</td>
<td>climate to strengthen and exploit our expertise on</td>
</tr>
<tr>
<td>• Network: DANS is respected partner in a wide range of national</td>
<td>FAIR data assessment, software sustainability, data</td>
</tr>
<tr>
<td>and international data networks and infrastructures, and</td>
<td>protection, open data, and data management</td>
</tr>
<tr>
<td>occupies a number of leading positions in them.</td>
<td>support.</td>
</tr>
<tr>
<td>• Finance &amp; Organisation: financially healthy</td>
<td>• Network: rising demand from outside of the humanities</td>
</tr>
<tr>
<td>despite fixed budget, strong projects portfolio, flexible and</td>
<td>and social sciences, in particular from the life</td>
</tr>
<tr>
<td>responsive organisation, professional project bureau and</td>
<td>sciences.</td>
</tr>
<tr>
<td>project acquisition.</td>
<td>• Finances &amp; Organisation: DANS is regarded as a key</td>
</tr>
<tr>
<td>• Vision &amp; Policies: innovative, practical,</td>
<td>player (next to SURF and NLeSC) in the national</td>
</tr>
<tr>
<td>highly motivated staff, combination of theory and practice, keen</td>
<td>digital infrastructure, offering good prospects for</td>
</tr>
<tr>
<td>eye for opportunities and niches.</td>
<td>additional funding.</td>
</tr>
<tr>
<td></td>
<td>• Vision &amp; Policies: strategic choices give us a good</td>
</tr>
<tr>
<td></td>
<td>position for expanding activities in software</td>
</tr>
<tr>
<td></td>
<td>sustainability and as international data vault for</td>
</tr>
<tr>
<td></td>
<td>long-term archiving.</td>
</tr>
<tr>
<td></td>
<td><strong>Current Action:</strong> Leverage our &quot;head start&quot; position</td>
</tr>
<tr>
<td></td>
<td>internationally and in other disciplines.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Weaknesses</th>
<th>Threats</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Expertise: the background of the majority of our staff is in the</td>
<td>• Services: Increasing competition from a wide range of aspiring service</td>
</tr>
<tr>
<td>humanities and social sciences rather than in other domains, whereas</td>
<td>providers in a fragmented data landscape with no clear leadership or</td>
</tr>
<tr>
<td>there is an increasing demand from e.g. the life sciences.</td>
<td>coordination.</td>
</tr>
<tr>
<td>• Network: we experience a certain distance towards the universities,</td>
<td>• Expertise: Rapid technological developments turn existing knowledge of</td>
</tr>
<tr>
<td>that perhaps do not perceive of DANS as for and by them.</td>
<td>staff outdated; strict temporary: fixed personnel ratio makes it</td>
</tr>
<tr>
<td>• Finances &amp; Organisation: inelastic budget whereas data volumes,</td>
<td>challenging to keep experienced staff.</td>
</tr>
<tr>
<td>requirements and demands by users grow and change rapidly.</td>
<td>• Finances &amp; Organisation: funding of innovation is difficult when</td>
</tr>
<tr>
<td></td>
<td>budget remains stable while existing services must be maintained at</td>
</tr>
<tr>
<td></td>
<td>the same time.</td>
</tr>
<tr>
<td><strong>Current Action:</strong> We are strengthening our network and collaboration</td>
<td><strong>Current Action:</strong> We continue to invest in our employees to remain in</td>
</tr>
<tr>
<td>with other disciplines, universities and other service providers.</td>
<td>the forefront of data technology and are operating more and more as a</td>
</tr>
<tr>
<td></td>
<td>strong hub in an (inter)national ecology of service providers and</td>
</tr>
<tr>
<td></td>
<td>users.</td>
</tr>
</tbody>
</table>

**Current Action:** Leverage our competencies and research for continuous development, innovation and improvement of our services.
5.2 Benchmarking

It is one of DANS’ strategic aims to be a “leading building block” in the international collaborative data infrastructure (see objective 1 in section 3.1). Is DANS a leading data service provider among its peer organisations? How to benchmark a service organisation of which no second one exists nationally, and how to make international comparisons with data centres and service providers of very different organisational structure, target domains, and size? Which organisations can we consider as peers? What criteria to use?

Peer organisations:

- Other Dutch data service providers: SURFsara, 4TU.ResearchData, university repositories and RDM support offices, KB, National Archives of the Netherlands (NA), broadcasting archives like the Netherlands Institute for Sound and Vision (IBG). DANS is nationally well-respected; as a national service provider, it is treated on equal footing with e.g. KB, NA or IBG. For the humanities and social sciences, there is no other data archive with a comparable collection.
- Social science data archives: CESSDA service providers such as the data archives in UK, Germany, France, Scandinavia, Switzerland, and ICPSR in the US. In content and use, DANS is among the most sizeable CESSDA archives. In staff size, it is among the bigger ones, although ICPSR, UKDA, GESIS (Germany) and NSD (Norway) are bigger. DANS director Peter Doorn is chair of the CESSDA General Assembly. Statistics Netherlands provides secure remote access to microdata, but also uses DANS for disseminating anonymized data.
- Humanities data archives: e.g. CLARIN centres and Archaeology Data Service (ADS, UK) – DANS is acting as a CLARIN centre, although, as an archive serving many disciplines, it has hard for us to comply with certain domain-specific CLARIN criteria; The E-depot for Dutch Archaeology within DANS was modelled after the ADS, and this outstanding data service still is an example for us. We strive to be on equal footing with them. The CLARIAH project is providing a growing linked-data cloud, in which DANS plays an important role as service developer and provider. The KNAW Humanities Cluster provides specialised data services on research-related datasets. KNAW institutes with digital collections provide their own archiving and access services.
- Data expertise centres: e.g. Digital Curation Centre (DCC, UK) and Australian National Data Service (ANDS) – Unlike DCC and ANDS, DANS performs both advisory and expertise functions next to concrete data services. DANS is intensifying its collaboration with DCC in areas such as training and data management planning.
- Generic repositories such as Zenodo, Figshare, Mendeley Data and DRYAD – Zenodo is hosted by CERN, Figshare and Mendeley are private companies related to international publishers, DRYAD is not for profit with a focus on life sciences; it is hard to compare DANS with those, but we are proud that we were selected by Mendeley and DRYAD as background archives, and that we are in their (advisory) boards.

Criteria for comparison:

- Content and use of archive related to size of relevant academic user community – no systematic information on this is available for peer organisations, but the (user) statistics presented in this report show that DANS is serving its target audiences very well.
- Participation in international infrastructure projects – DANS has an excellent track-record of doing European projects, and is often invited by colleague organisations to contribute to collaborative projects
- Leading position in international data organisations, committees, etc. – DANS staff takes a number of chairing and other leading positions in international data organisations, as is reported elsewhere in this self-evaluation.
• Research output related to staff size (including papers at international data conferences) – information on our output is provided in Appendix 4, but we do not possess information on our peer organisations for comparison.

• Early adopter or developer of innovative services – DANS was perhaps the first archive to offer a self-deposit archiving system; we were the inventors of the Data Seal of Approval; probably the only data archive running a data journal; will be first to offer tools for FAIR data assessment and data tagging for secure sharing of privacy-sensitive data compliant with GDPR.

• Visits by peer organisations for advice on service practices: DANS has been visited by many sister archives from abroad that want to learn from our experience. In European infrastructures such as CESSDA and EHRI, DANS provides back-office services for organisations that are less equipped to provide their own data services.
6 Integrity and standards

6.1 Institutional and data integrity
KNAW and NWO both advocate Open Access to research outputs, including scientific data. Maximum access to data pre-eminently supports scientific methods in which researchers check findings made by others and build critically on one another’s work. In this context DANS contributed actively to the report Responsible research data management and the prevention of scientific misconduct by an advisory committee chaired by Prof. Kees Schuyt (2013).

DANS firmly supports The Netherlands Code of Conduct for Academic Practice of the Association of Universities in the Netherlands (VSNU). We promote the Code of Conduct among the depositors and users of research data by explicitly referring to it in our data licenses. DANS staff also stick to the code in our own research activities.

Internal peer review is executed by means of regular research group meetings, where research design and results are discussed and documents are shared. Peer-reviewed publishing in Open Access journals and monographs is particularly encouraged. Preprints of all publications are made available as open access by default. Datasets produced in DANS research are by default deposited in EASY. PhD Students participate in KNAW-wide training on research ethics and integrity.

Furthermore, all DANS staff, including guest researchers, trainees, et cetera, are obliged to sign the Declaration of Confidentiality for Employees, in which the employee states that he/she will observe and maintain the utmost secrecy with regard to all confidential information that is supplied or will be supplied to him/her by DANS or by persons designated by DANS. This also includes information about other persons as they might occur to be related to DANS services.

The institute monitors and documents conformity with relevant regulations concerning the ingest, archiving and use of digital objects. These include: data protection, protection of the rights of affected parties, confidentiality regulations, copy and usage rights, internal and external compliance.

When a researcher deposits data in EASY, a license agreement is agreed upon with DANS. This agreement is based on the most relevant laws and codes of conduct, such as the Copyright Act, Databases Act, Personal Data Protection Act and VSNU Codes of Conduct.

The rules for using data from EASY, in conjunction with the license conditions, are stated in the General DANS Conditions of Use. These conditions also take the above mentioned legal framework into account. The main points from the conditions include:

- The user should, particularly upon distribution or disclosure, respect any copyrights and/or database rights on the dataset;
- The user shall always include a bibliographical reference in the research results that he/she publishes, in whichever form, where use is made of digital data, originating from one of the datasets of DANS. This applies to datasets in all access categories;
- Datasets containing personal data, within the meaning of the Personal Data Protection Act, may only be used for scientific, statistical or historical research.

The institute has developed an information security policy based on the information security policy of KNAW. This policy is at the basis of the yearly work plans for the further improvement of security levels and is based on the standpoint that in principle everything should be protected, unless it needs to be open and employees should not have more authorizations than strictly needed for the performance of their duties.

35 See also https://dans.knaw.nl/en/about/organisation-and-policy/legal-information
6.2 International certification standards

Information managers and other stakeholders can rely on a framework of various international certification standards for digital repositories in order to assess and improve the quality of their work processes and management systems, which can be summarized by the term Trustworthy Digital Repository (TDR).

Three certification instruments, with increasing degrees of detail and depth, are available: Data Seal of Approval (DSA), nestorSeal (verification according to DIN 31644) and ISO 16363 certification.

The assessments vary in intensity from a peer review of completed documentation (self-assessment) in the case of DSA, to a prepared on-site visit by an external audit team in the case of ISO. These instruments are mainly used in Europe. Outside Europe, many archives receive International Council for Science - World Data System (ICSU-WDS) certification, which is similar to DSA certification. Data sponsors, producers and re-users may trust any managing body that has been certified according to one of the above standards.

As founding father of DSA DANS is very active in this area and it participates both in the DSA Board as well as the Scientific Committee of the World Data System. DANS has played an important role in unifying the two quality standards for digital repositories, which has recently resulted in the Core Trust Seal.

We acquired both the DSA and nestorSeal and are a certified regular member of WDS. These certification standards incorporate different aspects of institutional, repository and data integrity. The institute needs to provide evidence that e.g.:

- It ensures, to the extent possible, that data are created, curated, accessed, and used in compliance with disciplinary and ethical norms;
- It guarantees the integrity and authenticity of the data;
- It enables users to discover the data and refer to them in a persistent way through proper citation;
- The technical infrastructure of the repository provides for protection of the facility and its data, products, services, and users.
DANS recognises the value of a diverse and skilled workforce and is committed to creating and maintaining an inclusive and collaborative workplace culture that will provide sustainability for the organisation into the future. The DANS diversity policy is compliant with that of KNAW, and diversity is an element in the covenant the institute has with its mother organisation. We are committed to leveraging the diverse backgrounds, experiences and perspectives of our employees to provide excellent user services to an equally diverse community. DANS’ commitment to recognising the importance of diversity extends to all areas of our activities, including recruitment, talent development, skills enhancement, appointment to roles and functions, retention of employees, mentoring and coaching programs, flexible work arrangements, forms of leave available to employees, and training and development.

- **Workplace Conduct**
  DANS is committed to eliminating all forms of unlawful discrimination, unlawful harassment, bullying and victimisation of persons in the workplace. Moreover, we are committed to maintaining a workplace where all staff can work without fear of unacceptable conduct impacting them.

- **Appointment to Roles**
  To achieve our strategic aims through our commitment to a merit-based appointment process, it is essential that capable and diverse employees are attracted, retained and deployed in roles that maximise their contribution and potential. A fair and effective process for appointment to roles is essential to ensure we can access the broadest pool of diverse candidates.

- **Flexible Working Arrangements**
  We are committed to recruiting and retaining the best talent to help us achieve our vision. This means we need to be adaptable in the way we work to meet the needs of our people and our customers. Flexible work arrangements provide a way of recognising and accommodating individual circumstances whilst balancing the DANS operational requirements.

- **Leave Policies**
  DANS understands that our employees have changing needs during different life and career stages. A range of leave options are available to employees to help them manage those changes. These may include parental leave, leave without pay, career break, personal/career’s leave and community service leave.
# Appendix 1: D1 – Output indicators

## Research quality

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1c. Number of datasets available via NARCIS</td>
<td>21.025</td>
<td>26.360</td>
<td>29.430</td>
<td>141.212</td>
<td>149.249</td>
<td>161.245</td>
</tr>
<tr>
<td>1d. Number of publications available via NARCIS</td>
<td>666.742</td>
<td>738.753</td>
<td>851.342</td>
<td>965.785</td>
<td>1.204.449</td>
<td>1.296.670</td>
</tr>
<tr>
<td>1e. Number of studies stored in DataverseNL*</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>269</td>
<td>341</td>
<td>396</td>
</tr>
<tr>
<td>1f. Number of files stored in DataverseNL*</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>729</td>
<td>1.023</td>
<td>1.166</td>
</tr>
</tbody>
</table>

* DataverseNL became a DANS service in 2014

<table>
<thead>
<tr>
<th>D1.2 Use of research products by peers</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
<th>2014</th>
<th>2015</th>
<th>2016</th>
</tr>
</thead>
<tbody>
<tr>
<td>2a. Number of datasets downloaded from EASY</td>
<td>15.026</td>
<td>25.431</td>
<td>30.118</td>
<td>45.686</td>
<td>31.013</td>
<td>30.723</td>
</tr>
<tr>
<td>2b. Number of files downloaded from EASY*</td>
<td>373.553</td>
<td>152.255</td>
<td>180.666</td>
<td>248.145</td>
<td>425.539</td>
<td>627.669</td>
</tr>
<tr>
<td>2c. Number of NARCIS users</td>
<td>611.047</td>
<td>476.325</td>
<td>540.267</td>
<td>519.237</td>
<td>728.335</td>
<td>947.414</td>
</tr>
<tr>
<td>2d. Number of visits to NARCIS</td>
<td>957.261</td>
<td>784.633</td>
<td>888.718</td>
<td>901.407</td>
<td>1.516.890</td>
<td>2.196.143</td>
</tr>
<tr>
<td>2e. Number of downloads in DataverseNL**</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>2.123</td>
<td>2.990</td>
<td>4.014</td>
</tr>
<tr>
<td>2f. Number of DataverseNL partners**</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>8</td>
<td>9</td>
<td>11</td>
</tr>
</tbody>
</table>

* Download numbers fluctuate due to changing delivery methods for big file quantities

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>3a. Invited lectures</td>
<td>4</td>
<td>5</td>
<td>14</td>
<td>55</td>
<td>42</td>
<td>18</td>
</tr>
<tr>
<td>3b. Participating in externally funded national research infrastructures and related projects</td>
<td>12</td>
<td>13</td>
<td>15</td>
<td>12</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>3c. Participating in externally funded international research infrastructures and related projects</td>
<td>8</td>
<td>11</td>
<td>14</td>
<td>15</td>
<td>19</td>
<td>11</td>
</tr>
</tbody>
</table>
Relevance to society

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>4a. Reports (policy only)</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>4b. Articles in professional and popularizing journals for non-academic readers</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>4c. Number of datasets available via NARCIS - cumulative</td>
<td>21.025</td>
<td>26.360</td>
<td>29.430</td>
<td>141.212</td>
<td>149.249</td>
<td>161.245</td>
</tr>
<tr>
<td>4d. Number of publications available via NARCIS - cumulative</td>
<td>666.742</td>
<td>738.753</td>
<td>851.342</td>
<td>965.785</td>
<td>1.204.449</td>
<td>1.296.670</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>D1.5 Use of research products by societal groups</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
<th>2014</th>
<th>2015</th>
<th>2016</th>
</tr>
</thead>
<tbody>
<tr>
<td>5a. Number of datasets downloaded from EASY</td>
<td>15.026</td>
<td>25.431</td>
<td>30.118</td>
<td>45.686</td>
<td>76.624</td>
<td>107.385</td>
</tr>
<tr>
<td>5b. Number of files downloaded from EASY</td>
<td>373.553</td>
<td>152.255</td>
<td>180.666</td>
<td>248.145</td>
<td>425.539</td>
<td>627.669</td>
</tr>
<tr>
<td>5c. Number of NARCIS users</td>
<td>611.047</td>
<td>476.325</td>
<td>540.267</td>
<td>519.237</td>
<td>728.335</td>
<td>947.414</td>
</tr>
<tr>
<td>5d. Number of visits to NARCIS</td>
<td>957.261</td>
<td>784.633</td>
<td>888.718</td>
<td>901.407</td>
<td>1.516.890</td>
<td>2.196.143</td>
</tr>
</tbody>
</table>
Appendix 2: Datametric analysis of DANS Data Archive

1. Growth of data archive EASY

The number of datasets archived in the DANS Electronic Archiving System (EASY) increased from around 1,500 in 2007 to over 35,000 in 2017 (Fig. 1). A dataset usually consists of the data belonging to a particular project and can contain one or more data files.

In the early years, the annual growth rate was higher than in later years, due both to the relatively small numbers before 2010 and due to a retro-digitization project of archaeological data and reports in collaboration with the State Archaeological Service (Now: Cultural Heritage Agency of the Netherlands – RCE). Since 2012 the annual growth rate of the collection fluctuates around 10%. In the figure, distinction is made between data uploads by individual researchers and bulk upload of institutional collections or repositories. Bulk import was considerable in the time of the retro-project mentioned, and we expect it to become more important in the future, because of new institutional arrangements in which DANS serves as a background archive. In the first half of 2017 we are already seeing a growth of bulk import due to new institutional data arrangements.

Many datasets consist of multiple data files. The average number of files per dataset varies over time, but on average it is over a hundred. The total number of files in EASY increased almost linearly over time, from a little under 15,000 in 2007 to over 4 million today (August 2017; see Fig. 2).
The overwhelming majority of the datasets stored in the DANS archives comes from archaeology, the (other) humanities and the social sciences, where the average data sizes are modest (see also further down, Fig. 6). Before 2009, the mean size per dataset was less than 100 MB; from 2010-2013, the mean size was around 150-200 MB; and since then the mean size fluctuates around 400 MB per dataset. The total volume of data archived in EASY (displayed is single storage at one site, in reality there is multiple storage at two or more different locations) grew from 0,1 TB in 2007 to about 15 TB now.

Fig. 4. gives an impression of the size of the archive according to discipline over the past 10 years. The total number of datasets is inflated by about 3,000, because of datasets classified under more than one discipline. It is clear from the graph that the relatively young E-Depot for Dutch Archaeology (EDNA, started in 2004) quickly became the largest section of the archive, even though archaeology is a relatively small academic domain. The success of EDNA is largely explained by two factors: (1) the above-mentioned retro-digitization of archaeological data and reports; (2) the fact that RCE has made deposit at DANS obligatory for data belonging to every archaeological project carried out in The Netherlands. We do not know which proportion of data from other
areas is not archived at DANS, but it is likely that in most research fields only a small percentage of the data produced is archived for long-term access. The greater attention to data policies by research funders, universities and other institutions has not yet markedly altered this situation.

The large share of archaeological data sets in the archive is also apparent from Fig. 5. In the early years of DANS, the proportion of archaeological datasets grew from less than 20 to about 75%. Since 2009 the proportional distribution of archived datasets over the main disciplines has not changed much. Socio-cultural sciences and humanities both take about 10%, other social and behavioural sciences around 2% each and the life sciences about 1%.

The size distribution of the archived datasets is reflected in Fig. 6. It is clear from the graphs, that the overwhelming majority of the stored datasets at DANS are small, which is (still?) characteristic for the humanities and social sciences. DANS typically serves the "long tail of science". Although there is a tendency that datasets deposited in later years are somewhat bigger than those in earlier years, the effect on the size distribution is still small (see also the text accompanying Fig. 3). How time-related the term Big Data is, is illustrated by the example of the population census of 1960, the first one to be fully computerized. The 11 million punch cards of the original file take a little less than 0,5 GB of storage. In the 1970s, one statistical run with SPSS on this dataset cost the full annual computing budget of the faculty of social sciences of the University of Amsterdam.
Figure 7a and b. Archived datasets according to size - absolute figures and percentages

Figure 8 - Distribution of archived datasets according to type of access, 2012-2016
Looking at the datasets according to type of access, it is remarkable that the number of datasets that is publicly or openly accessible is rapidly increasing over time (Fig. 7). This is perhaps an indication of an increased awareness among researchers that sharing data is useful, and it can be interpreted as a sign of acceptance of open science principles, at least among researchers depositing their data in the DANS archive. Was in 2012 the percentage of openly accessible data in EASY less than 50%, in 2016 this had increased to 70%. The datasets requiring explicit permission for access by the depositor decreased both in absolute and in relative terms. The category "group access" pertains to the archaeological sector, where access to detailed data is limited to professional archaeologists in order to prevent the disturbance of heritage sites. Originally DANS required users to register and log in if they wanted to deposit and download data. Since 2016 DANS also allows full open access without registration (also known as CC0 or public domain dedication) for data without any copyright (although data citation is still required according to the academic code of conduct).

2. Growth of use of data archive
Between 2007 and August 2017, a total of 223,258 datasets, containing 2,361,588 files have been downloaded from EASY.

![Figure 9 - Datasets downloaded from EASY 2007-2016](image)

The reuse of datasets has increased in a fairly similar way as the growth of the archive until 2013. In 2007 only about 750 datasets were downloaded for reuse, a number which grew to around 30,000 in 2013. Since then, the number of downloads per year has stabilized. 2014 was an exceptional year, in which one user downloaded all openly accessible archaeological data, about 15,000 datasets in total. In further analyses and figures we have omitted this exceptional case.
Unsurprisingly, with a view on the size of the collection, archaeology is also the domain with the highest number of downloads in absolute terms: in recent years (since 2012), the number fluctuated from 20,600 to 24,000. Next are the socio-cultural sciences (between 5,700 and 7,800 downloads annually since 2012), the social sciences (between 2,000 and 3,000) and the behavioural sciences (1,600 – 2,700). The downloads in the humanities varied from 1,900 to 2,600 per year. Note that in Fig. 9 the total number of downloads is also slightly inflated because of datasets counted under more than one domain.

It is also instructive to look at the downloads relative to the size of the archive per domain (Fig. 10). If we divide the number of downloads by the number of archived datasets per discipline, we get an indicator of relative data reuse. The mean for all datasets is slightly above 1, meaning that every dataset in the archive is downloaded about once every year. The social sciences are clearly on top since 2010, with a reuse ratio of between 3 and 4.5. They are followed by a group of behavioural, socio-cultural and life sciences, with between 2 and 3 downloads per archived dataset per year.
Archaeology and the humanities score slightly below average, with a reuse ratio of resp. around 1 and 0.75.

Top 25 of downloaded datasets from EASY (since deposit)

<table>
<thead>
<tr>
<th>Rank</th>
<th>Title of Dataset</th>
<th>Dataset Downloads</th>
<th>Downloaded Files</th>
<th>Persistent Identifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Nationaal Kiezersonderzoek, NKO 2006</td>
<td>1125</td>
<td>5398</td>
<td>urn:nbn:nl:ui:13-4zd-x4e</td>
</tr>
<tr>
<td>2</td>
<td>De steentijd van Nederland</td>
<td>1108</td>
<td>1452</td>
<td>urn:nbn:nl:ui:13-tg4-mof</td>
</tr>
<tr>
<td>3</td>
<td>Brabant cohort - derived student file</td>
<td>884</td>
<td>2161</td>
<td>urn:nbn:nl:ui:13-zkg-gjv</td>
</tr>
<tr>
<td>5</td>
<td>Netherlands Longitudinal Lifecourse Study - NELLS First Wave - 2009 - versie 1.3</td>
<td>742</td>
<td>1920</td>
<td>urn:nbn:nl:ui:13-54c-uiie</td>
</tr>
<tr>
<td>9</td>
<td>WoON2012: release 1.0 - WoonOnderzoek Nederland 2012 (voor overheid, universiteiten en overige partijen)</td>
<td>603</td>
<td>5043</td>
<td>urn:nbn:nl:ui:13-60fd-6i</td>
</tr>
<tr>
<td>11</td>
<td>Nationaal Kiezersonderzoek, NKO 1971-2006 cumulatieve file</td>
<td>553</td>
<td>1306</td>
<td>urn:nbn:nl:ui:13-e9w-ig9</td>
</tr>
<tr>
<td>14</td>
<td>NLGis shapefiles</td>
<td>452</td>
<td>149040</td>
<td>urn:nbn:nl:ui:13-wsh-wv7</td>
</tr>
<tr>
<td>15</td>
<td>Arbeidsaanbodpanel 1985 t/m 2010</td>
<td>430</td>
<td>6577</td>
<td>urn:nbn:nl:ui:13-4js-jl3</td>
</tr>
<tr>
<td>16</td>
<td>Culturele Veranderingen in Nederland 2006 - CV'06</td>
<td>387</td>
<td>1488</td>
<td>urn:nbn:nl:ui:13-73o-mbh</td>
</tr>
<tr>
<td>17</td>
<td>Enquête Beroepsbevolking - EBB - jaargangen 1987 t/m 2012</td>
<td>381</td>
<td>3521</td>
<td>urn:nbn:nl:ui:13-sk6-fmg</td>
</tr>
<tr>
<td>18</td>
<td>Slachtoffers van oplichting en van poging tot oplichting (projectnummer 1742)</td>
<td>368</td>
<td>750</td>
<td>urn:nbn:nl:ui:13-h84-sjz</td>
</tr>
<tr>
<td>19</td>
<td>WoON2015: release 1.0 - WoonOnderzoek Nederland 2015</td>
<td>352</td>
<td>2599</td>
<td>urn:nbn:nl:ui:13-pv3u-84</td>
</tr>
<tr>
<td>20</td>
<td>VHMO-data Rotterdam, 1880/'81 en 1920</td>
<td>335</td>
<td>1900</td>
<td>urn:nbn:nl:ui:13-egj-jjh</td>
</tr>
<tr>
<td>23</td>
<td>Mobiliteitsonderzoek Nederland 2009</td>
<td>289</td>
<td>2358</td>
<td>urn:nbn:nl:ui:13-yii-0li</td>
</tr>
</tbody>
</table>

So far, we dealt with average downloads from the archive, in total and per discipline. But of course, the distribution of downloads is highly skewed: there are more and less popular datasets, and like there are books in a library that nobody ever borrows, there
are also datasets that are not (yet) downloaded. Nevertheless, of the 36,561 datasets archived in EASY at the time of writing, 31,924 datasets have been downloaded at least once: a gross reuse percentage of 87%!

The above table gives the top 25 datasets of all time, with the Dutch Parliamentary Election Survey of 2006 as the overall winner, with 1,125 dataset downloads (in total 5,400 files) until August 2017. However, if we took file downloads as the main success indicator, the list would look rather different, with the NLGis shapefiles as the winner: 150,000 files downloaded. This dataset contains the municipal boundaries of the Netherlands since the early 19th century till the present time.

![Figure 12 - Frequency of 0-30 dataset downloads](image)

When we look at the least frequently downloaded datasets, (only) 4,637 datasets have not (yet) been downloaded at all, 5,492 datasets have been downloaded once, 5,322 twice, and then the numbers begin to drop in an asymptotic way.
### Appendix 3: D3a – Research staff

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Scientific staff</td>
<td>2</td>
<td>4.25</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>5.25</td>
</tr>
<tr>
<td>Postdocs</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>PhD students</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total research staff</strong></td>
<td>2</td>
<td>6.25</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5.25</td>
</tr>
<tr>
<td>Support staff</td>
<td>35.6</td>
<td>40.6</td>
<td>43.4</td>
<td>38.85</td>
<td>30.25</td>
<td>36</td>
</tr>
<tr>
<td><strong>Total staff</strong></td>
<td>37.6</td>
<td>46.85</td>
<td>48.4</td>
<td>43.85</td>
<td>35.25</td>
<td>41.25</td>
</tr>
<tr>
<td>Visiting fellows</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

Note: staff numbers based on KNAW personnel registration; fluctuations exaggerated due to changing definitions and ways of registration of temporary project staff, trainees, etc.

### Appendix 4: D3b – Research output

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Refereed articles</td>
<td>2</td>
<td>7</td>
<td>3</td>
<td>5</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>Non-refereed articles</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Books</td>
<td>5</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Book chapters</td>
<td>2</td>
<td>7</td>
<td>9</td>
<td>13</td>
<td>11</td>
<td>6</td>
</tr>
<tr>
<td>Conference papers, refereed and non-refereed</td>
<td>5</td>
<td>6</td>
<td>14</td>
<td>27</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>Professional publications</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Publications aimed at the general public</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Reports non-refereed</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>6</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>Other professional output: software products, seminars, products over a (research) project</td>
<td>10</td>
<td>4</td>
<td>2</td>
<td>12</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>31</td>
<td>31</td>
<td>37</td>
<td>69</td>
<td>47</td>
<td>47</td>
</tr>
</tbody>
</table>

Note: Fluctuations exaggerated due to changes in registration system and interpretations of various categories of informal publications.

### Appendix 5: D3c – Total funding of the institute

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Funding in €</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Funding KNAW</td>
<td>2.142</td>
<td>2.577</td>
<td>2.587</td>
<td>2.503</td>
<td>2.535</td>
<td>2.546</td>
</tr>
<tr>
<td>Funding NWO</td>
<td>767</td>
<td>667</td>
<td>667</td>
<td>677</td>
<td>677</td>
<td>677</td>
</tr>
<tr>
<td>Direct funding</td>
<td>2.909</td>
<td>3.244</td>
<td>3.254</td>
<td>3.180</td>
<td>3.212</td>
<td>3.223</td>
</tr>
<tr>
<td>Project funding</td>
<td>858</td>
<td>1030</td>
<td>1178</td>
<td>1545</td>
<td>1024</td>
<td>1373</td>
</tr>
<tr>
<td>(including research grants)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>72</td>
<td>136</td>
<td>104</td>
<td>73</td>
<td>257</td>
<td>140</td>
</tr>
<tr>
<td><strong>Total funding</strong></td>
<td>3.838</td>
<td>4.411</td>
<td>4.536</td>
<td>4.798</td>
<td>4.493</td>
<td>4.737</td>
</tr>
<tr>
<td>Expenditure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other costs</td>
<td>1.008</td>
<td>1.020</td>
<td>1.053</td>
<td>1.436</td>
<td>1.022</td>
<td>1.001</td>
</tr>
<tr>
<td><strong>Total expenditure</strong></td>
<td>3.919</td>
<td>4.482</td>
<td>4.752</td>
<td>4.976</td>
<td>4.110</td>
<td>4.264</td>
</tr>
<tr>
<td><strong>Total result</strong></td>
<td>-81</td>
<td>-71</td>
<td>-216</td>
<td>-178</td>
<td>384</td>
<td>474</td>
</tr>
</tbody>
</table>
### Appendix 6: D3d – PhD candidates

<table>
<thead>
<tr>
<th>Starting year</th>
<th>Enrolment male/female</th>
<th>Total</th>
<th>Graduated in year 4 / earlier</th>
<th>Not yet finished</th>
<th>Discontinued</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012</td>
<td>M</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2017</td>
<td>F</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
# Appendix 7: List of abbreviations and definitions

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4TU.Centre for Research Data</td>
<td>Also known as 4TU.ResearchData; an initiative of the three Dutch technical universities and joined by Wageningen University and Research Centre</td>
</tr>
<tr>
<td>ADS</td>
<td>Archeology Data Service</td>
</tr>
<tr>
<td>ANDS</td>
<td>Australian National Data Service</td>
</tr>
<tr>
<td>ARIADNE</td>
<td>Advanced Research Infrastructure for Archaeological Dataset Networking</td>
</tr>
<tr>
<td>BagIt</td>
<td>Hierarchical file packaging format designed to support disk-based storage and network transfer of arbitrary digital content</td>
</tr>
<tr>
<td>CBS</td>
<td>Statistics Netherlands/ Centraal Bureau voor de Statistiek</td>
</tr>
<tr>
<td>CEDAR</td>
<td>Dutch historical census data on the Semantic Web; Census data open linked</td>
</tr>
<tr>
<td>CERIF</td>
<td>Common European Research Information Format</td>
</tr>
<tr>
<td>CERN</td>
<td>European Organization for Nuclear Research/ Conseil Européen pour la Recherche Nucléaire</td>
</tr>
<tr>
<td>CESSDA (ERIC)</td>
<td>Consortium of European Social Science Data Archives</td>
</tr>
<tr>
<td>CLARIAH</td>
<td>Common Lab Research Infrastructure for the Arts and Humanities</td>
</tr>
<tr>
<td>CLARIAH-PLUS</td>
<td>Research Infrastructure; proposal submitted for the NWO Roadmap 2017</td>
</tr>
<tr>
<td>CLARIN</td>
<td>Common Language Resources and technology Infrastructure</td>
</tr>
<tr>
<td>DANS</td>
<td>Data Archiving and Networked Services</td>
</tr>
<tr>
<td>DARIAH</td>
<td>Digital Research Infrastructure for the Arts and Humanities</td>
</tr>
<tr>
<td>DARIAH CIO</td>
<td>DARIAH’s Chief Integration Officer; see also DARIAH</td>
</tr>
<tr>
<td>DataCite</td>
<td>Global non-profit organisation that provides persistent identifiers (DOIs) for research data</td>
</tr>
<tr>
<td>DataTag</td>
<td>An informative label from a controlled vocabulary that can be applied to a dataset. It carries distinct semantics, summarizing sufficient conditions for a specific set of automated actions over the data. A dataset is labelled with a tag on the basis of a systematic interrogation of a data controller, conducted using a specified set of survey questions, and inferential rules for tag assignment. Each label formally corresponds to a set of assertions regarding permissible or impermissible actions over the dataset.</td>
</tr>
<tr>
<td>DataverseNL</td>
<td>Open source software application to share, cite and archive data</td>
</tr>
<tr>
<td>DCC</td>
<td>Digital Curation Centre</td>
</tr>
<tr>
<td>DCCD</td>
<td>Digital Collaboratory for Cultural Dendrochronology</td>
</tr>
<tr>
<td>DH</td>
<td>Digital Humanities</td>
</tr>
<tr>
<td>DIN 31644</td>
<td>Deutsches Institut für Normung, Information and documentation - Criteria for trustworthy digital archives</td>
</tr>
<tr>
<td>DMP</td>
<td>Data Management Plan</td>
</tr>
<tr>
<td>DRYAD</td>
<td>Nonprofit repository for data underlying the international scientific and medical literature</td>
</tr>
<tr>
<td>DSA</td>
<td>Data Seal of Approval</td>
</tr>
<tr>
<td>DTL</td>
<td>Dutch Techcentre for the Life Sciences</td>
</tr>
<tr>
<td>EASY</td>
<td>Electronic Archiving System</td>
</tr>
</tbody>
</table>
EC European Commission
EGI DARIAH CC Agreement between a group of EGI resource centres and the DARIAH Competence Centre
EHRI 2 European Holocaust Research Infrastructure, second phase
EOSC European Open Science Cloud
ERIC European Research Infrastructure Consortium
ESFRI The European Strategy Forum on Research Infrastructures
ESFRI Landmark Research Infrastructures that were implemented or reached the implementation phase under the Roadmap and that are established as major elements of competitiveness of the ERA
ESFRI Roadmap Identifies vital new European RIs for the next 10-20 years
EU European Union
EUDAT2020 European Association for Databases in Education and Training; programme for 2017-2020
EU Horizon 2020 Horizon 2020 is the financial instrument implementing the Innovation Union, a Europe 2020 flagship initiative aimed at securing Europe's global competitiveness
EuroCRIS Current Research Information System, international organization
FAIR Findable, Accessible, Interoperable, Reusable
Fedora (3) Open source Linux-based operating system (version 3)
Figshare Online digital repository where researchers can preserve and share their research outputs; portfolio company of Digital Science, operated by Macmillan Publishers
FOBO (model) Front office/back office model
GDPR General Data Protection Regulation
GESIS Leibniz-Institute for the Social Sciences, the largest infrastructure institution for the Social Sciences in Germany/ Gesellschaft Sozialwissenschaftlicher Infrastructureinrichtungen
IBG The Netherlands Institute of Sound and Vision/ Nederlands Instituut voor Beeld en Geluid
ICPSR Inter-university Consortium for Political and Social Research
ICSU-WDS International Council for Science - World Data System (see also former WDS)
IMPACT-EV Evaluating the Impact and Outcomes of EU SSH Research
INRIA French Institute for Research in Computer Science and Automation/ Institut national de recherche en informatique et en automatique
KB National Library of the Netherlands/ Koninklijke Bibliotheek
KNAW The Royal Netherlands Academy of Arts and Sciences/ Koninklijke Nederlandse Academie van Wetenschappen
LCRDM  National Coordination Point Research Data Management/ Landelijk Coördinatiepunt research data management
LOD     Linked Open Data
M3      Molecule, Man, Society/Molecuul, Mens, Maatschappij
Mendeley Desktop and web program produced by Elsevier for managing and sharing research papers, discovering research data and collaborating online
Meresco Open Source Heritage Software. Under the Meresco name, a number of Dutch public institutions initiated the development and sharing of innovative quality software components with respect to metadata exchange and search
MoU     Memorandum of Understanding
NA      National Archives of the Netherlands/ Nationaal Archief
NARCIS  National Academic Research and Collaborations Information System
NCDD    Dutch Digital Preservation Coalition/ Nationale Coalitie Digitale Duurzaamheid
NDE     Digital Heritage Network/ Netwerk Digitaal Erfgoed
nestorSeal See also DIN 31644; based on the German DIN 31644-standard, the SEAL consists of a complex and comprehensive set of 34 criteria that an archive has to comply with.
NLeSC   Netherlands eScience Center
NLGis   Netherlands Geographic Information System (project)
NSD     Norwegian Centre for Research Data/ Norsk senter for forskningsdata
NWO     The Netherlands Organisation for Scientific Research/ Nederlandse Organisatie voor Wetenschappelijk Onderzoek
ODISSEI Open Data Infrastructure for Social Science and Economic Innovations
OpenAIRE 2020 Open Access Infrastructure for Research in Europe
PARTHENOS Pooling Activities, Resources & Tools for Heritage E-research
PBO     Periodiek Bestuurlijk Overleg (periodical meeting between the DANS and KNAW board)
PRELIDA  Preserving Linked Data
R&I     Research & Innovation
RCE     Cultural Heritage Agency of the Netherlands/ Rijksdienst voor het Cultureel Erfgoed
RDA     Research Data Alliance
RDJ     Research Data Journal for the Humanities and the Social Sciences
RDNL    Research Data Netherlands
Re3Data Registry of Research Data Repositories
Re-SEARCH Contextual Search for Scientific Research Data
RI      Research Infrastructure
RIVM    National Institute for Public Health and the Environment/ Rijksinstituut voor Volksgezondheid en Milieu
SCP     The Netherlands Institute for Social Research/ Sociaal en Cultureel Planbureau
SHA     Software Heritage Archive
SHEBANQ (project)  System for HEBrew Text: ANnotations for Queries and Markup
SMART (targets)  Specific, Measurable, Assignable and Time-related
SURFsara  The collaborative ICT organisation for Dutch education and research
SWORD 2  Simple Web-service Offering Repository Deposit, version 2; protocol for depositing content from one location to another
(SWOT)  (Strengths, Weaknesses, Opportunities and Threats)
SYNVAR  Tracing Syntactic Diversity in Biblical Hebrew Texts
TDR  Trustworthy Digital Repository
UKDA  United Kingdom Data Archive
VSNU  Association of Universities in the Netherlands/Vereniging van Universiteiten
VU  VU University Amsterdam / Vrije Universiteit
WDS  World Data System (see also ICSU-WDS)
WODC  Research and Documentation Centre of the Ministry of Justice/ Wetenschappelijk Onderzoek- en Documentatiecentrum
WTMC  Netherlands Graduate Research School of Science, Technology and Modern Culture
Zenodo  Research data repository. It was created by OpenAIRE and CERN to provide a place for researchers to deposit datasets.
Appendix 8: A selection of national and international projects carried out by DANS in the past few years, and some recent data/research infrastructure proposals

A selection of national projects:
- With RCE (Cultural Heritage Agency of the Netherlands): Digital Collaboration for Cultural Dendrochronology (DCCD)
- With VU University Amsterdam: Tracing Syntactic Diversity in Biblical Hebrew Texts (SYNVAR)
- With NCDD (Dutch Digital Preservation Coalition)/NDE (Digital Heritage Network): Software Sustainability
- With NCDD/NDE: Certifying digital archives
- With CLARIAH partners: Development of technical infrastructure

A selection of international/European projects:
- OpenAIRE 2020: Open Access Infrastructure
- IMPACT-EV: Evaluating the Impact and Outcomes of EU SSH Research
- EUDAT2020: Research Data Services, Expertise & Technology Solutions
- EHRI2: European Holocaust Research Infrastructure
- EOSC pilot: preparing the European Open Science Cloud

A selection of recent proposals for national data/research infrastructure roadmaps:
- NWO Roadmap: ODISSEI and CLARIAH-Plus
- KNWA agenda for future large-scale research facilities: M3-proposal (Molecule, Man, Society)
- NWO advice for the national digital infrastructure for scientific research: DANS is extensively mentioned as an indispensable part of the chain (in connection with SURFsara and the NLeSC).
- National Science Agenda: proposals are being formulated
- DANS is core partner of several international ESFRI Landmarks (such as DARIAH and CESSDA) and member of several consortia with the ambition to become ESFRI Roadmap projects.

36 See also our list of abbreviations and definitions
37 See our projects page for more information: https://dans.knaw.nl/en/projects/projects
Appendix 9: Other relevant documents

**Previous formal external assessment (2010/2011):**
- Comments by the DANS Board (Dutch): http://knaw.nl/shared/resources/actueel/bestanden/Reactie_directie_DANS_op_evaluatieapport.pdf
- Comments by the academic committee DANS (Dutch): http://knaw.nl/shared/resources/actueel/bestanden/Reactie_WAR_op_evaluatieapport.pdf
- Academy's position on the DANS evaluation (Dutch): http://knaw.nl/shared/resources/actueel/bestanden/KNAW-standpunt_evaluatie_DANS.pdf

**Previous mid-term assessment (2014; in Dutch only):**
https://drive.google.com/drive/folders/0B8Ero1DLtjA3MEyEyknFY1hRbzA?usp=sharing
- Reactie KNAW en NWO op advies van den Herik: https://drive.google.com/file/d/0B8Ero1DLtjA3eNuaUZXSFGFWaDA/view?usp=sharing

**DANS Strategic documents:**