EOSC SYNERGY

LANDSCAPE REPORT

THE NETHERLANDS

EOSC-synergy receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857647

EOSC-SYNERGY

Landscaping Country Report

The Netherlands

Document Identifier:	EOSC-SYNERGY-LANDSCAPE-NL
Date:	2020-10-12
Activity:	WP5
Leader project partner:	CESNET
Responsible project Partner:	DANS-KNAW
Author:	Peter Doorn
Document Status:	DRAFT
Dissemination Level:	
Document PID:	https://doi.org/10.17026/dans-2by-ereu

Abstract:

This landscape analysis report aims to provide an overview of the policies, practices, roadmaps, and strategies around funding, procuring, providing, accessing, and sharing of services and resources in the EOSC scope in The Netherlands. The overview of e-infrastructures, data services, repositories, research infrastructures and related organisations in The Netherlands is based on existing inventories and reports.

Disclaimer and Acknowledgement:

This report does not pretend to be a scientific publication, and it does not strictly abide by academic rules concerning plagiarism. The report freely reuses pieces of text from existing reports and publicly accessible websites with relevant information. However, it does provide references to the sources used, although it does not always use quotation marks or other indicators when one or more sentences are literally used.

Table of Contents

List of Tables

List of Figures

1. Introduction

1.1. Aim and scope of this landscape analysis

This landscape analysis aims to provide an overview of the the Dutch policies, practices, roadmaps and strategies around funding, procuring, providing, accessing and sharing of services and resources in the EOSC scope. Special emphasis will be given to policies and practices in The Netherlands that address the needs of different EOSC user groups (individual researchers, citizen scientists, research projects and collaborations, industries and SMEs) with a focus on the user communities that are already part of the national roadmap for big research facilities.

The work will cover a wide range of policies, including those dealing with generic services for data production, processing and preservation, as well as those for thematic resources and services that target the needs of specific research communities. Transnational access to national resources will also be covered. Existing policies and practices, where applicable, will be related to national and international regulatory frameworks (e.g. handling of privacy-sensitive data in compliance with GDPR).

1.2. Definition/delimitation of "EOSC compliant resources"

The delimitation of "services and resources in the EOSC scope" is not unambiguous. Not only are the rules of participation for inclusion of services and resources in EOSC under development (https://www.eoscsecretariat.eu/working-groups/rules-participation-working-group), there is also no undisputed definition of what to understand under such services and resources. Generally these are regarded as 'facilities' constituting the 'infrastructure' of research, two more terms with various definitions.[1]
Nevertheless, we have to be pragmatic, and the website of the European Commission gives the following: *"Research Infrastructures are facilities that provide resources and services for research communities to conduct research and foster innovation. They can be used beyond research e.g. for education or public services and they may be single-sited, distributed, or virtual.*
They include:

- *major scientific equipment or sets of instruments*
- *collections, archives or scientific data*
- *computing systems and communication networks*
- *any other research and innovation infrastructure of a unique nature which is open to external users"*

[1] See: Francesca Morselli, Research Infrastructures and collaboration in research: A contextualised literature review. Text under development for PhD Dissertation. DANS, 2020.

(https://ec.europa.eu/info/research-and-innovation/strategy/european-research-infrastructures_en; last accessed 8/3/2020).

Overlapping with research infrastructure is the term e-infrastructure, which according to the European Commission, is usually described as: *"a combination of digital technologies (hardware and software), resources (data, services, digital libraries), communications (protocols, access rights and networks), and the people and organisational structures needed to manage them."* (http://erosa.aginfra.eu/definition; last accessed 8/3/2020).

"e-Infrastructures address the needs of European researchers for digital services in terms of networking, computing and data management.

e-Infrastructures provide digital-based services and tools for data- and computing-intensive research in virtual and collaborative environments. They foster the emergence of Open Science, i.e. new working methods based on the shared use of digital tools and resources across different disciplines and technology domains as well as sharing of results and an open way of working together.

e-Infrastructures are key in future development of research infrastructures, as activities go increasingly online and produce vast amounts of data. This support is essential for example to the European Strategy Forum on Research Infrastructures (ESFRI). Furthermore, e-Infrastructures enable and support the circulation of knowledge in Europe online and therefore constitute an essential building block for the European Research Area (ERA).

(https://ec.europa.eu/digital-single-market/en/e-infrastructures; last accessed 8/3/2020).

The e-Infrastructures Reflection Group (e-IRG) uses *"the metaphor of the Commons for the e-Infrastructure resources and related services, which among others refer to networking, computing, storage, data and software, along with digital tools and collaboration opportunities"* (http://e-irg.eu/about; last accessed 8/3/2020).

Digital resources such as data (information) and (software) tools are often stored in digital repositories, which literally are "places where things are stored", and which is hence again a term used to indicate a broad range of "facilities" or "services". In a narrower sense, digital data repositories are places where collections of data(sets/files) are stored and logically organized (i.e., usually described according to some metadata schema), so that they can be found and accessed for reuse. The Science Europe Data Glossary offers a helpful definition and short explanation, see: http://sedataglossary.shoutwiki.com.

The EOSC itself mentions the following categories of "services and resources" on its home page.

Table 1. Services according to type listed in the EOSC marketplace portal

Type of service	**Number**
Compute	38
Data management	87
Networking	12
Processing & Analysis	64
Security & Operations	17
Sharing & Discovery	54

Storage	26
Training & Support	26
All Services	**262**

Source: https://marketplace.eosc-portal.eu/services (last accessed 25/5/2020). This overview is getting more complete as services are added over time.

1.3. Information sources on services, repositories, facilities and infrastructures

For the overview of the services and resources in The Netherlands we used four different lists, inventories and registries. The overview can be seen here: https://docs.google.com/spreadsheets/d/1mvndq1exrgjdVCoaUOkctQ4gqTFPLytK83Oo_QWy9O4/edit?usp=sharing

1.3.1. NWO List of Large-scale Scientific Infrastructures

List compiled for the National Roadmap (version 2018), based on a call for proposals (dated 2016) submitted to the Permanent Committee for Large-Scale Scientific Infrastructure, which was appointed by NWO in July 2015 on behalf of the Ministry of Education, Culture and Science. The ministry's request to the committee was to develop a national strategy for investments in large-scale research facilities. The list of facilities and the Roadmap are regularly updated. In 2019/2020 an update is under way, the information of which is not yet available.

The committee received a total of 164 facilities from 54 different institutions. Of these, a total of 113 have been incorporated into the landscape of existing large-scale research facilities in the Netherlands (see www.onderzoeksfaciliteiten.nl). A number of facilities did not meet the financial lower limit (€10 million) for large-scale research facilities and have therefore not been incorporated into the landscape.

For the first time, the landscape has provided an overview of existing large-scale research facilities, as well as the need for new facilities in the Netherlands. Taken together, the 164 facilities have a combined investment need of more than €3,000 million for the coming five years. Of these, about a hundred facilities have indicated that they would like to be incorporated into the new National Roadmap for Large-Scale Scientific Infrastructure. These infrastructures have an investment need of more than €2,000 million. This amount is indicative and also includes the input of these facilities' own resources and third-party resources.

Nonetheless, the committee believes that the investment need is not commensurate with the National Roadmap's available budget of a total of €200 million for the coming five years. Many infrastructures that are highly important to science in the Netherlands may therefore never see the light of day.

Harmonisation with the field and the evaluation of the Permanent Committee has led to a new National Roadmap consisting of 33 facilities. These are in part individual facilities (16), as was the case with the previous National Roadmap. In addition, the National Roadmap consists of clusters of facilities (17) that are asked to draft a joint investment plan for the entire cluster. This National Roadmap is set to run for a period of 4 years. A number of facilities were not included in the National Roadmap by the Permanent Committee. It mainly concerns facilities that, despite being important to society, have too limited scientific value. Furthermore, a few facilities are not aligned with the strategic frameworks of a scientific field. A number of facilities are insufficiently developed to already be included in the National Roadmap. Even if a facility is on the Roadmap, this does not guarantee (extra) funding. In April 2018, 138 million euros were awarded to ten top research facilities on the Roadmap.

1.3.2. MERIL

The MERIL (Mapping of the European Research Infrastructure Landscape) portal provides access to a database that stores information about openly accessible research infrastructures (RIs) in Europe, across all scientific domains, including the social sciences and humanities (see: https://portal.meril.eu/meril/). The MERIL-2 project has ended and the database is no longer updated, although the portal is still online and searchable. Among other things, MERIL contains information about 1042 infrastructures providing 3015 services in Europe, and 2964 are labeled as "equipment". Of the research infrastructures, 57 organisations have a location in or are coordinated by The Netherlands. Unfortunately, it is not possible in MERIL to select services by country, but on average one infrastructure organisation supplies about three services.
Further exploitation of MERIL data related to research infrastructure services will continue within the newly launched CatRIS project (https://project.catris.eu).

1.3.3. Re3data.org

Re3data has been maintaining a registry of research data repositories since 2012. It indexes and provides extensive information about more than 2450 research data repositories all over the world. According to re3data, there are 56 repositories run by institutions located in The Netherlands (about half of which have co-locations in other countries as well)[2]. Re3data records about 30 characteristics of the repositories, such as access policies, PIDs used, whether the repository is certified, etc. Unfortunately, the standard interface of the web portal does not allow to run metrics on the repositories in a selected country, although via filters it is possible to get an impression of the characteristics of this important group of data services in a country (or for instance, for a disciple; see, for the Netherlands: https://www.re3data.org/search?query=&countries[]=NLD).

As Re3data is based on self-registration, the information must be interpreted with some care. Some self-proclaimed "repositories" refer rather to databases, portals or home pages of

[2] After finalizing the analysis for this report, I found out that two repositories were incorrectly classified as "Dutch". I have notified this to Re3Data, and in the meantime this was corrected. But at the same time, some repositories were added, so that the total stands at 58 on 12/10/2020.

institutions providing some sort of data service, rather than to repositories in the sense of the definition proposed by Science Europe (http://sedataglossary.shoutwiki.com).

1.3.4. Dutch participation in ESFRI Roadmap/Landmarks

A fourth source of information on Dutch services and resources is ESFRI, which maintains a list of projects/proposals on the ESFRI roadmap and a list of ESFRI Landmarks, which are facilities that were on a previous roadmap and that have evolved into more or less sustained organisations (which may or may not take the legal form of an ERIC or European Research Infrastructure Consortium. There are 9 projects and 20 landmarks with a Dutch participation on the 2018 ESFRI List (see: http://roadmap2018.esfri.eu/media/1049/roadmap18-part3.pdf). Moreover, three EIROforum members are also on the ESFRI roadmap, bringing the total of ESFRIs with a Dutch participation at 32.

Figure 1. Map of Large-Scale Scientific Infrastructures in the Netherlands for National Roadmap

Source: http://www.onderzoeksfaciliteiten.nl

1.4. Between projects and (inter-)national institutions (E84)

There is quite some overlap between the four overviews mentioned above, but there are differences as well, which are caused by various factors:

- whether a list is based on self-reporting (such as MERIL and re3data);
- whether being on a list has or may have consequences for funding (such as the national and ESFRI roadmaps);
- whether there is selection involved by an independent committee (usually linked to funding, such as ESFRI or the NWO committee for large scale research facilities);
- what the scope of the inventory is (e.g. repositories in the case of re3data);
- whether the focus is on institutions, projects or services;
- whether the focus is on
- digital infrastructure (repositories, data services, etc.) or on physical facilities (equipment);
- whether the list contains proposals, temporary projects, aspiring or mature infrastructures.

It often happens that institutions provide more than one service, so it may be that in one list we find the institute, and in another one the service(s) they supply. Moreover, services go through a cycle from preparation to construction to implementation, and some proposals may even never see the light of day, although they are still registered in one or more of the lists. And finally, some organisations/services can be missing on certain lists, because they either did not self-register or were not selected for inclusion, because they did not comply with certain criteria (such as the financial lower threshold of 10 million Euros in the case of the Dutch national roadmap -- smaller infrastructure projects fall under other funding instruments). In spite of these deficiencies, we have nothing better than the lists mentioned and we must reckon with the fact they are not perfect. Moreover, the landscape is not fixed, it is changing continuously, so the image we provide reflects the situation at the end of the decade 2010-2020.

Also of interest are the overviews provided by the *Landelijk Coordinatiepunt Research Data Management* (LCRDM). Under the heading “RDM in The Netherlands”, it lists a total of 50 Dutch and 62 international RDM facilities (with a little overlap) in the categories Training, File services, Repositories, Online tools, Working groups, Infrastructures, and Initiatives and organisations (“https://www.lcrdm.nl/en/rdm-in-the-netherlands). In 2017, LCRDM produced an online service catalog of available repositories for Dutch researchers, but that is unfortunately not maintained and will soon be discontinued (https://www.edugroepen.nl/sites/RDM_platform/Dienstencatalogus/SitePages/Dienstencatalogus.aspx).

2. National policies and frameworks for open science support and collaboration

2.1. Formal regulations or publicly available policies

In this section we provide a brief overview of formal regulations or publicly available policies that address the following subjects:

- Open Access to publications
- Research Data Management (RDM)
- Open research data
- Long-term availability of research data
- Compliance of data to the FAIR principles
- Publication of data in a repository
- Publication of data in a certified repository

2.1.1. Open Access to publications

The Dutch government officially adopted an Open Access policy with respect to publications in 2013 (https://www.openaccess.nl/en/in-the-netherlands/what-does-the-government-want). The Dutch government is of the opinion that publicly funded research should be freely accessible. This was the position outlined by State Secretary Sander Dekker in a letter (in Dutch) to the Dutch House of Representatives already in November 2013. He was deliberately opting for the golden route. He aimed to have 60 percent of Dutch academic publications available through open access within five years (2019) and 100 percent within ten years (2024). If not enough progress was made, proposals would follow in 2016 to make open access publication mandatory.

In 2016, the Amsterdam Call for Action on Open Science (https://www.openaccess.nl/sites/www.openaccess.nl/files/documenten/amsterdam-call-for-action-on-open-science.pdf) was drawn up at an Open Science conference organized by the Dutch Presidency of the Council of the European Union on 4 and 5 April 2016 in Amsterdam. The ambition of 100% open access was further strengthened and the date was also adjusted to 100% open access at the end of 2020. The results and actions are formulated in the Amsterdam Call for Action on Open Science. See the summary and comments on the Call for Action (https://www.openaccess.nl/sites/www.openaccess.nl/files/documenten/comments_highlights_on_the_amsterdam_call_for_action_on_open_science.pdf).

The government prioritized the golden route because this was considered the most sustainable in the long term. It was assumed that the publishers' business model would change and that this route would provide the best guarantee for publications to become immediately available. The green route often means lengthy embargo periods.

Although the government policy occasioned considerable debate in academic circles, among publishers, and elsewhere, the main research bodies such as KNAW, NWO and VSNU support the OA policy. In fact, the declarations on Open Access of the early 2000s, in particular the Berlin Declaration on Open Access of 2003, were signed by virtually all Dutch science organisations and universities.

The government also decided to monitor the development of OA in The Netherlands, a task taken up by the VSNU. This is however easier said than done (see: https://www.openaccess.nl/sites/www.openaccess.nl/files/documenten/definitief_definition_framework_oa_vsnu-20150922.pdf). VSNU gave the following estimates for 2016 and 2017, but did not present more recent figures (Table 2):

Table 2. Percentages OA publications 2016 and 2017, Dutch universities by category

"Gold", DOAJ OA	"Hybride" and not DOAJ OA	Only "green" OA	Total OA
13%	20%	9%	42% (2016)
17%	23%	10%	50% (2017)

Source: https://vsnu.nl/en_GB/percentages-open-access-publications-2016-

The OA development can be followed almost in real-time via the NARCIS website maintained by DANS. NARCIS harvests and aggregates the information from the repositories and research information systems of the Dutch Universities, research institutes and Higher Education Institutions (https://www.narcis.nl/metrics/Language/en), containing references and descriptions on over 2 million publications (about 100,000 per year in the second half of the 2010s). The NARCIS figures, based on pragmatic definitions as used by the organisations from which the data is harvested, display the following trend (Table 3):

Table 3. Access to Dutch publications in 34 Academic and Higher Education repositories

Year	Open	Restricted	Embargoed	Closed	Total	Total publications
2016	43%	30%	0%	28%	100%	94786
2017	51%	24%	0%	25%	100%	101622
2018	59%	17%	0%	24%	100%	109840
2019	59%	19%	1%	21%	100%	101694
2020	57%	25%	2%	17%	100%	10668

Source: https://www.narcis.nl/?Language=en Note that the figures for 2019 are not yet complete, and that the figures for 2020 reflect only the situation until March 11 2020 (about 10% of the expected total number of publications).

The figures, which reflect all types of publications and not only articles[3], are quite comparable to those of VSNU. We expect that the percentage of Open Access publications in 2020 will be around 60%, which is close to the original aim of the Dutch Government of 2013. However, the target of 100% OA publications in 2020 was clearly overambitious.

The official OA policy was enlarged into a full-fledged Open Science policy in 2017, when the National Plan Open Science (NPOS) was presented and the National Platform Open Science was created (https://www.openscience.nl/en). The ambitions for 2020 concerned four main topics:

- 100% open access publishing
- Make research data optimally suited for reuse
- Recognising and assessing researchers
- Encouraging and supporting open science

2.1.2. Research Data Management (RDM)

Research Data Management became an issue on the agenda of universities and other research organisations in the early 2010s, especially after some cases of scientific misconduct and data fraud received a lot of public attention (see, e.g. https://en.wikipedia.org/wiki/Diederik_Stapel). Most universities and research organisations, including the Dutch research funding organisations NWO and ZONMW started to develop and successively implement RDM policies. Despite numerous meetings, bodies and platforms for discussion and coordination (such as UKB, LCRDM, NPOS), a variety of policies and requirements emerged, very similar in aim but differing in all kinds of detail. In 2017 NWO pledged to promote the European alignment of RDM policies, a task that was taken up by Science Europe, which published a *Practical Guide to the International Alignment of Research Data Management* in 2018 (https://www.scienceeurope.org/media/jezkhnoo/se_rdm_practical_guide_final.pdf). A growing number of organisations is using these harmonized RDM requirements in their policies.

2.1.3. Open research data

Rather than requiring full and unconditional open access to research data, the policy of most research organisations in The Netherlands is to promote "optimal" access to "FAIR" research data, using the motto "Open if possible, protected when needed". The FAIR data principles, which coincidentally have their cradle in The Netherlands, quickly gained popularity among policy makers, and it is usually required that data comply with these principles. The tendency is that the default access is open, but that restrictions to openness are acceptable on various grounds, especially legal grounds, such as the protection of personal data in the compliance with the GDPR. Access to data has become one of the cornerstones of the Open Science policy, and is usually combined with RDM policies, of which access requirements are a part. The earlier mentioned Amsterdam Call for Action (https://www.openaccess.nl/sites/www.openaccess.nl/files/documenten/amsterdam-call-for-action-on-open-science.pdf) of April 2016 marked the broad acceptance of the Dutch Open Science policy, including access to research data, and it is also one of the projects or action lines of the

[3] Although it is also possible to calculate the metrics on journal articles only.

National Plan/Platform Open Science (NPOS). NPOS seeks to optimize the Dutch Data Infrastructure Landscape in similar ways as EOSC seeks to improve the efficiency and effectiveness of the European landscape (https://www.openscience.nl/en/projects/project-e-exploring-the-dutch-data-landscape).
As we will see later on, compliance with or enforcement of RDM and open/FAIR data policies still leave ample opportunities for improvement.

2.1.4. Long-term availability of research data

Like open data policies, policies with respect to long-term archiving and availability of research data is closely linked to RDM policies. In these policies, usually a distinction is made between the treatment of data during a project and after that project has ended. Researchers who are funded by NWO or ZONMW have to specify in the DMP what is going to happen with the data after the conclusion of a project. There is a clear tendency that depositing the data in a trusted repository is recommended or obligatory, and that the repository complies with a certification standard such as the CoreTrustSeal (https://www.coretrustseal.org, which also has its roots in The Netherlands as it emerged from the Data Seal of Approval). In practice, what will happen with the data often depends on the predominant culture of the research community involved. The data volume is also important. For truly Big Data, the assumption is that the organisations responsible for data creation will somehow keep the data alive or know what is best. Obviously, the need to preserve data that can be reproduced is less than for data that can be measured only once.
Perhaps the awareness of the importance of long-term preservation of research data has the oldest tradition in the social sciences and humanities, where data archives started to emerge from the 1960s onwards. The Steinmetz Foundation (later: the Steinmetz Archive), one of the predecessors of DANS, was for instance founded in 1964.[4]

2.1.5. Compliance of data to the FAIR principles

Compliance with the FAIR principles has quickly become generally accepted by research organisations in The Netherlands, where also one of the GO FAIR International Support and Coordination Offices is located (https://www.go-fair.org/go-fair-initiative/go-fair-offices/). FAIR compliance is usually part of the RDM requirements by research funders and universities. However, both the specification of what it exactly means to comply with the FAIR principles and how to measure that, and how to monitor that compliance, leave much to be desired. The H2020 FAIRsFAIR project, led by DANS, is one of the projects developing FAIR data assessment tools (see: https://fairsfair.eu/).

2.1.6. Publication of data in a repository

The content types mentioned in the self-proclaimed 56 repositories in the Netherlands are the following (note that multiple types are possible; Table 5):

[4] In some domains it is legally required to destroy (personal, medical/health) data after a certain period of time.

Table 5: Repositories in The Netherlands according to content type

Content Types	Repositories
Standard office documents	42
Scientific and statistical data formats	36
Plain text	31
Raw data	26
Images	23
Audiovisual data	19
Structured graphics	17
Archived data	16
Structured text	16
Databases	13
Software applications	7
Configuration data	3
Network Based data	2
Source code	2
Other	17
Total	56

Source: repositories listed in Re3data

The content types that occur most frequently are (textual) documents and (numeric) tabular data, with some kind of image/audiovisual/graphics data as a third important category. Note that there are also a number of repositories storing software or source code.

2.2. Strategies and policies for funding infrastructure services and resources

There are no existing figures on the aggregated amounts of money spent on digital infrastructure, and it is hard to give educated guesses, given the variety of financial streams (local, national, international), and of definitions and specifications of what to include and what not. To give any reliable guestimate would require a more in-depth study.
According to the state budget for 2020, about 55 M€ out of the total budget for scientific education and research (5.2 + 1.2 = 6.4 billion €) is spent on "large-scale research infrastructure", but this covers only a fraction of the total investments and running costs of the many components of the full landscape of facilities and services.[5] In 2018 NWO funded ten top research facilities on the National Roadmap with a total of 138 million euros (distributed over a time period of five years).

[5] Rijksbegroting 2020: VIII Onderwijs, Cultuur en Wetenschap. https://www.rijksoverheid.nl/documenten/begrotingen/2019/09/17/viii-onderwijs-cultuur-en-wetenschap-rijksbegroting-2020

The state budget 2020 further mentions that the Ministry of Education, Culture and Sciences (Dutch abbreviation: OCW) "will also make resources available in 2020 and beyond for investments in scientific research infrastructure. With these resources, a new (biennial) round of the Roadmap Large-scale Scientific Infrastructure will be held by NWO in 2020. In addition, OCW will further develop the investments in the IT infrastructure in the coming year as a result of NWO's ICT report, with additional resources for high-performance computing, networks and data storage, ICT support for data and software intensive research, and making data accessible" (p. 113).

Table 4. Main budget lines for scientific research and research policy, 2018-2024

Amounts x 1000 €	2018	2019	2020	2021	2022	2023	2024
Funding	**1.101.220**	**1.114.385**	**1.049.574**	**1.040.297**	**1.036.661**	**1.036.105**	**1.035.525**
Main Funding	**742.322**	**706.516**	**658.529**	**655.49**	**655.327**	**656.258**	**656.178**
NWO-Law and Law on Higher Education and Scientific Research (WHW)							
– National Organisation for Scientific Research (NWO)	556.834	520.582	521.225	518.388	518.961	518.966	518.886
– Royal Netherlands Academy of Arts and Sciences (KNAW)	89.646	89.527	89.525	89.323	89.132	89.446	89.446
– Royal Library (KB)	95.842	96.407	47.779	47.779	47.234	47.846	47.846
Additional Funding	**358.898**	**407.869**	**391.045**	**384.807**	**381.334**	**379.847**	**379.347**
– NWO Talent Development	160.885	170.885	165.885	165.885	165.885	165.885	165.885
– NWO Applied and Engineering Sciences (STW)	8	8	8	8	8	8	8
– NWO Large-scale research infrastructure	85.38	85.38	55.38	55.38	55.38	55.38	55.38
– National Board of Education Research	28.986	29.957	26.133	21.542	18.069	16.582	16.082
– Polar Research	3.147	3.147	3.147	1.5	1.5	1.5	1.5
– Caribbean Netherlands	2.5	2.5	2.5	2.5	2.5	2.5	2.5
– NWO National Science Agenda (NWA)	70	108	130	130	130	130	130
Subsidies	**22.549**	**27.239**	**25.815**	**26.491**	**26.452**	**26.491**	**26.491**
– Foundation Biodiversity (NLBIF)	550	550	550	550	550	550	550
– Naturalis Biodiversity Center	6.265	6.265	6.265	6.266	6.266	6.266	6.266
– Biomedical Primate Research Centre (BPRC)	9.608	9.608	9.608	9.609	9.609	9.609	9.609
– National Centre for Science and Technology (NCWT) / NEMO Science Museum	3.366	3.366	3.366	3.366	3.366	3.366	3.366

– Netherlands Study Centre for Technology Trends (STT)	221	221	221	221	221	221	221
– Animal Advocacy and Protection (AAP Foundation)	1.032	1.032	1.032	1.032	1.032	1.032	1.032
– National Coordination	1.507	6.197	4.773	5.447	5.408	5.447	5.447
Commissions	**163**	**379**	**340**	**340**	**379**	**340**	**340**
– Commissions	163	379	340	340	379	340	340
Contribution to agencies	**673**	**921**	**842**	**842**	**842**	**842**	**842**
– Education Executive Agency	0	2	2	2	2	2	2
– Netherlands Enterprise Agency	673	919	840	840	840	840	840
Contribution to (inter) national organizations	**92.353**	**95.875**	**95.875**	**95.875**	**95.875**	**95.875**	**95.875**
– Engineering in Medicine and Biology Society (EMBC)	918	941	941	941	941	941	941
– European Molecular Biology Laboratory (EMBL)	5.176	5.227	5.227	5.227	5.227	5.227	5.227
– European Space Agency (ESA)	31.065	31.065	31.065	31.065	31.065	31.065	31.065
– European Organization for Nuclear Research (CERN)	44.199	46.168	46.168	46.168	46.168	46.168	46.168
– European Southern Observatory (ESO)	8.425	9.871	9.902	9.902	9.902	9.902	9.902
– Dutch Language Union / Institute for Dutch Language (NTU/INL)	2.57	2.603	2.572	2.572	2.572	2.572	2.572
Revenues	**504**	**101**	**101**	**101**	**101**	**101**	**101**

Source: Based on Table 16.3 in State Budget 2020, p. 113-114.

According to OECD estimates, about 0,67% of the Dutch GDP goes to public investments in R&D, and about 1.5% of the labour force works in this sector. As can be seen in Table 6, about half of the budget for scientific research and research policy is channeled through NWO. Here we limit the description of funding instruments to those of this research funding organisation, focusing on the major ones for research infrastructure.
However, it is evident from the table that many budget lines will have an infrastructure component. Moreover, other organisations, services and facilities receive their funding through other channels. For example, Statistics Netherlands (CBS, annual budget around 190 M€) is funded by the Ministry of Economic Affairs and Climate Policy, the Royal Netherlands Meteorological Institute (KNMI, annual budget around 60 M€) is an Agency of the Ministry of Infrastructure and Water Management. Like many other state funded organisations, these two agencies contribute significantly to the scientific data infrastructure, but it is unfeasible to attach

financial numbers to the amount to which they do that. Moreover, the universities as well as NWO and KNAW host institutes, which perform infrastructural functions. DANS is an example of such an institute with a budget of about 3.4 M€ per year (excluding projects).
NWO distinguishes about 250 financial instruments divided into six following categories, one of which is "for the realisation and use of large-scale infrastructure". About 45-50 instruments were open for applications at the time of writing this report, some instruments have a temporal character, others return regularly. Although several instruments for the funding of research or scientific collaboration may contain infrastructural components, the instruments on the realisation and use of large-scale infrastructure are most important for the purpose of this landscape analysis. Within this category, there are calls for investments in infrastructure, and subsidies for the use of large facilities (see: https://www.nwo.nl/en/funding/our-funding-instruments).
I*nvestment calls:*

- Investment Grant NWO-Large 2019/2020: Universities and research institutes can make large investments in innovative equipment and data collections with the NWO Large Investment Grant. Investments are financed that are higher than 1.5 M€ (1 M€ for the social sciences and humanities).
- National Roadmap Large-scale Research Infrastructure: The second call for proposals of the National Roadmap Large-Scale Research Infrastructure is intended to enable the 33 (clusters of) research infrastructure on the current National Roadmap to submit applications for funding (minimum grant: 10 M€).
- Scientific infrastructure for astronomy, computer science and mathematics: This programme for medium-sized investments (110,000 to 500,000 €) encourages Dutch universities, research institutes and libraries to invest in innovative equipment or data collections of national or international importance.

In the course of 2020, two new calls are expected for the funding of local (spring 2020) and thematic (autumn 2020) digital competency centres. These calls have not been published yet at the time of writing.

2.2.1. Subsidies for the use of large facilities

- Computing Time on National Computer Facilities: a funding scheme that grants researchers and research groups access to the national advanced computational facilities HPC Cloud, Lisa compute cluster, Cartesius supercomputer (or successor), Data Processing (Grid/Spider) and/or Custom Cloud Solutions.
- The Dutch-Belgian Beamline (DUBBLE) grant enables Dutch and Flemish researchers to carry out experiments at the synchrotron radiation facility in Grenoble.
- Isaac Newton Group of Telescopes: Astronomers in the Netherlands are invited to apply for observation time at the Isaac Newton Group of Telescopes on La Palma (Canary Islands, Spain).
- The Netherlands Centre for Electron Nanoscopy (NeCEN): researchers can request beam time at the NeCEN facilities.

2.2.2. Example of research calls contributing to infrastructure

- Dutch Research Agenda - Theme: Cybersecurity - Towards a secure and trustful digital domain: This call for proposals calls upon knowledge institutions and other interested parties to participate in research and implementation projects involving cyber security and governance and cybersecurity, governance and cryptology.

2.3. Present status with regard to Commission Recommendation (EU) 2018/790 on access to and preservation of scientific information (NI4OS 14_mc)

Commission Recommendation (EU) 2018/790 (which builds on and replaces Recommendation 2012/417/EU) on access to and preservation of scientific information, has attracted relatively little awareness in The Netherlands, especially if we compare it to the attention received by the introduction of GDPR exactly one month later[6]. The GDPR, which has the status of law, completely overshadowed (EU) 2018/790, perhaps because the last one is "only" a recommendation. In Dutch language, hardly any reference to the Recommendation can be found online, save the official text in EU legal document systems such as https://eur-lex.europa.eu/. There is just one Dutch publication online referring to the Recommendation (M. van Eechout, 2018). Recommendation 2018/790 actually consists of 12 recommendations, which are further subdivided into rather detailed instructions for implementation. In Table 7 we provide our evaluation of the current state of affairs in The Netherlands with respect to the Recommendation.

Table 5. Compliance in NL with Commission Recommendation (EU) 2018/790

Nr.	Recommendation text	Compliance level	Remark
1	Member States should set and implement clear policies (as detailed in national action plans) for the dissemination of and open access to scientific publications resulting from publicly funded research.	1	
2	Member States should ensure that research funding institutions responsible for managing public research funding and academic institutions receiving public funding implement the policies and national action plans referred to in point 1 at national level in a coordinated way.	2	Coordination is in progress

[6] Recommendation (EU) 2018/790 is dated 25 April 2018, GDPR came into force on 25 May 2018.

3	Member States should set and implement clear policies (as detailed in national action plans) for the management of research data resulting from publicly funded research, including open access.	1	
4	Member States should ensure that research funding institutions responsible for managing public research funding and academic institutions receiving public funding implement the policies and national action plans referred to in point 3 at national level in a coordinated way.	2	Coordination is in progress
5	Member States should set and implement clear policies (as detailed in national action plans) for reinforcing the preservation and re-use of scientific information (publications, data sets and other research outputs).	2	"Other research outputs", such as software, so far receive little attention
6	Member States should set and implement clear policies (as detailed in national action plans) for further developing infrastructures underpinning the system for access to, preservation, sharing and re-use of scientific information and for promoting their federation within the EOSC.	2	Work is in progress, e.g. in NPOS
7	Member States should ensure synergies among national infrastructures, with the EOSC and other global initiatives by:	2	Work is in progress, e.g. in NPOS
8	Member States should set and implement clear policies (as detailed in national action plans) for the necessary skills and competences of researchers and personnel of academic institutions regarding scientific information.	2	Work is in progress, e.g. in NPOS
9	Member States should set and implement clear policies (as detailed in national action plans) for adjusting, with regards to scientific information, the recruitment and career evaluation system for researchers, the evaluation system for awarding research grants to researchers, and the evaluation systems for research performing institutions.	2	Adjustment is a continuous process, proceeding with varying paces for the policy areas mentioned
10	Member States should participate in multi-stakeholder dialogues on the transition towards open science at national, European and international level on each of the issues addressed in points 1 to 9.	1	The Netherlands is one of the front-runners in this transition

11	Member States should have a national point of reference.	1	For different areas, there are different national focal points, such as NPOS, LCRDM, OpenAIRE NOAD, etc.
12	Member States should inform the Commission 18 months from the publication of this Recommendation in the Official Journal of the European Union, and every two years thereafter, of action taken in response to the elements of this Recommendation.	?	No evidence found of such reporting; perhaps due to low awareness about the Recommendation, and unclarity about who is monitoring which actions

Compliance levels: 1 = full compliance; 2 = partial compliance; ? = unknown

3. EOSC compliant resources

EOSC aims to federate existing resources (data infrastructures) under guidance of a common governance framework, offering a universal entry point via the EOSC portal, although this will not exclude other access channels.
Common specifications and tools to make data FAIR, solutions to ensure legal compliance (in part. GDPR and cybersecurity laws), adoption of existing or new schemes to certify data repositories and service providers as FAIR-compliant, are under development, e.g. in the FAIRsFAIR project mentioned before.
Another EOSC demand will be that access to common core services and to building blocks for developing new, added value services, will be non-discriminatory. Moreover, agreement on possible mechanisms for cost recovery on cross-border access and facilitating joint procurement, integration of services as well as development of new services, has to be reached. This subject will be dealt with in section 4 of this landscaping analysis.
Furthermore, EOSC will have to ensure long-term sustainability of the federating core via the governance framework. As this will only gradually become clear to potential services to be included in the EOSC federation, this is an uncertainty influencing the ambition of such services to become part of EOSC. Finally, EOSC is to identify duplications in services provided and to monitor actual use, to foster economies of scale/scope. The implications of such duplications for individual service providers are not yet known, nor are the advantages of being part of the federated superstructure the EOSC is to provide. All in all, there are no unequivocal yardsticks yet to measure the compliance of services and facilities to be federated into the EOSC.

Despite this situation, many organisations are aware that EOSC is under construction, and several are preparing to become part of it. Such preparation is under way in the five ESFRI clusters:

- Environmental sciences (ENVRI-FAIR)
- Photon and Neutron Sciences (PANOSC)
- Astronomy & Particle Physics (ESCAPE)
- Social Sciences and Humanities (SSHOC)
- Biological & Medical sciences (EOSC-Life)

For this landscaping analysis, we consider the organisations, facilities and services in this overview as potential EOSC components, although it is impossible at this stage to report on their EOSC readiness (and even willingness). In the next sections we report on the characteristics of the services.

3.1. Characteristics of services/resources

3.1.1. Types of services (R5/E1)

We carried out this analysis on the facilities and services in the NWO inventory of Large-Scale Scientific Infrastructures (http://www.onderzoeksfaciliteiten.nl). The classification according to

PILLAR questions R5 and E1 was done by us, where the multiple choice option 4 was modified into “other” in question 4 (and the same option was added to question E1).

Question R5: Does your organisation provide one or more of these services to the research community?
1 - knowledge-based resources such as collections, archives or research data
2 - data and computing systems, and communication networks
3 - major scientific equipment or sets of instruments
4 - other (mainly survey instruments)

Also combinations of answers were possible. The classification was based on the descriptions of the facilities.It is safe to assume that all scientific equipment and instruments (category 3 and 4) produce data, so they would automatically include category 1. However, we only added that category if the data or physical collection was explicitly mentioned in the description of the facility.

Table 6. Types of facilities in the NWO inventory of large-scale scientific infrastructures

Categories	**Frequency**
1 - Knowledge-based resources	10
2 - Data an computing systems	1
3 - Major scientific equipment	75
4 - Other (surveys, etc.)	0
3; 2; 1	2
2; 1	1
4; 1	8
3; 1	10
1; 3	2
3; 1; 2	3
1; 2; 3	1
Total facilities	**113**

Source: NWO inventory of Large-Scale Scientific Infrastructures (http://www.onderzoeksfaciliteiten.nl)

The majority of facilities consist of scientific equipment: 75 out of the total of 113, and even 93 if we include combinations. In this category fall laboratories, telescopes, microscopes, cyclotrons, imaging systems, MRI scanners, simulation devices, etc. Surveys, which we mainly

find in the social sciences (and in some life sciences), may also be regarded as "equipment", but were counted here as "other".
Knowledge-based resources (both digital and physical) occur 10 times as the main purpose of the facility, and an additional 27 times as a combined purpose, so in total 37 times. But of course, *all* facilities produce knowledge bases of some sort.
Data and computing systems (including HPC facilities) occur only once as the single purpose of a facility, but in combination with other functions at least 7 times. Almost all facilities and services are specific for a particular research area or domain.

Question E1: Which service(s) does your organisation provide to the research community?
1 - We offer data infrastructures which store and manage research data (e.g. archive and disseminate data).
2 - We offer high-bandwidth networks which transport research data.
3 - We offer high-performance computing which can be used to process research data.
4 - We offer other services

This question overlaps considerably with Question R5, but makes a finer distinction between data, network and computing services (see Table 7).

Table 7. Types of services in the NWO inventory of large-scale scientific infrastructures

Categories	**Frequency**
1 - Data infrastructures	11
2 - High-bandwidth networks	0
3 - High-performance computing	1
4 - Other	76
3; 2; 1	1
1; 4	3
4; 1	16
4; 1; 3	1
4; 3; 1	1
4; 2; 1; 3	1
1; 2; 3	2
Total facilities	**113**

The "other" category mainly consists of scientific equipment (similar to category 3 in Question R5). There are 11 facilities classified as data infrastructures (+ 25 more in combinations with

other functions). High-bandwidth networks occur only in combinations (4x in total). HPC facilities also occur mainly in combinations, in total 7x.

Here we also see the differences that are directly related to the source of information, and its completeness. For example, Re3Data recorded 55 "data repositories", which is considerably more than the total of 36 in the NWO overview of large-scale facilities. One category not present in the NWO list, is that of generic academic data repositories (not immediately linked to a particular research goal) run by universities and DANS.

3.1.2. Services according to domain (U13/R11/E18)

Questions R11 and E18 in the PILLAR list of questions are identical *"For which scientific discipline(s) does your organisation provide services?"*, so we treat them together. Question U13 is *"What research discipline(s) from your organisation will benefit mostly from EOSC?"*, and we assume that the pattern of answers will be similar to that of R11 and E18, since most organisations provide services to particular disciplines only.

Table 8. Repositories by scientific discipline according to Re3data

PILLAR disciplines	**re3data disciplines**	**Frequency**	**Percentage**
1 - Natural Sciences	Natural Sciences	22	26%
2 - Engineering and Technology	Engineering Sciences	8	9%
3 - Medical and Health Sciences	Life Sciences	24	28%
4 - Agricultural Sciences	Agriculture, Forestry, Horticulture and Veterinary Medicine	4	5%
5 - Social Sciences	Social and Behavioural Sciences	13	15%
6 - Humanities	Humanities	14	16%
88 - Other:	-	0	0%
Total (including 29 classified under >1 discipline)		**85**	**100%**

Source: Re3data, https://www.re3data.org/search?query=&countries[]=NLD

Table 9. Research infrastructures by scientific disciplines according to MERIL

Discipline (in brackets the corresponding PILLAR discipline number)	**Frequency**	**Percentage**
Biological and Medical Sciences (3)	23	26%
Earth and Environmental Sciences (1; 4)	9	10%
Physics, Astronomy, Astrophysics and Mathematics (1)	14	16%

Chemistry and Material Sciences (1; 2?)	11	13%
Information Science and Technology (2)	4	5%
Engineering and Energy (2)	7	8%
Humanities and Arts (6)	10	11%
Social Sciences (5)	10	11%
All disciplines (including combinations)	**88**	**100%**

Source: based on MERIL, https://portal.meril.eu/meril/

The MERIL classification deviates from the PILLAR classification of disciplines, but nevertheless gives an impression of the distribution of service infrastructures over various fields. In both classifications, the life sciences and natural sciences each take about a quarter or more of the infrastructures (Table 10)..

In the NWO landscape inventory, most facilities have multiple disciplinary services, which results in the following distribution (Table 11):

Table 10. Large-scale scientific facilities by discipline according to NWO Landscape inventory

PILLAR Disciplines	NWO Disciplines	Frequency	Percentage
1 - Natural Sciences	Physical Sciences, Earth sciences, Computer science, Physics, Chemistry, Astronomy/ astrophysics, Mathematics	69	35%
2 - Engineering and Technology	Science and Technology, Technology	48	24%
3 - Medical and Health Sciences	Life sciences, Biology, Veterinary medicine, Medicine	57	29%
4 - Agricultural Sciences	(Not separately registered)	-	-
5 - Social Sciences	Area studies, Business administration, Public administration and political science, Communication science, Cultural anthropology, Demography, Economics, Gender studies, Geography/planning, Science of teaching, Development studies, Pedagogics, Psycholog, Law, Sociology	14	7%
6 - Humanities	Archaeology, Digital humanities, Philosophy, History, Art and architecture, Performing arts and	9	5%

	media, Religious studies, Language and literature, Linguistics, History of science		
Total (including multiple disciplines)		**197**	**100%**

Source: Based on NWO inventory of large-scale research facilities, http://www.onderzoeksfaciliteiten.nl

3.1.3. Services to become part of EOSC (E83)

This is a particularly difficult question to answer, as it is as yet not clear to the overwhelming majority of the organisations what "becoming part of EOSC" will mean. Our guess is that, depending on the requirements and conditions, many of the Dutch services, resources, infrastructures and repositories in this landscaping analysis *could* become part of EOSC. We have not asked about the ambitions of the organisations running the facilities, but basing ourselves on discussions during information meetings on the EOSC at conferences and workshops, many Dutch infrastructures take a "wait and see" attitude, although some Dutch organisations have been (or are) active in EOSC preparation or constituting projects (such as EOSC Hub, EOSC Synergy, FAIRsFAIR, etc.). There exists a certain prudence that EOSC might turn into a bureaucratic and technocratic organisation, reducing the autonomy of existing infrastructure organisations, who are asking the basic questions: what is going to be the added value for our scientific community? Moreover, reservations exist that EOSC will be too far removed from the scientists themselves, and that the whole superstructure will become too big to be effective. However, if the cost-benefit ratio will be favourable, no doubt many organisations will join in one way or another.

3.1.4. Expected innovations of services in next five years (E82)

Very difficult to generalize, many infrastructures are constantly innovating, and the possibilities for realization to a large extent depend on funding. Expected innovations are described in proposals, e.g. those submitted for the national roadmap.

3.1.5. Funding of research infrastructures (E19), including "revenues other than funding" (E20)

As described in section 2.2., there are multiple financial mechanisms and streams for funding research infrastructures (RIs). Ultimately, the overwhelming majority of the RIs is funded by the state, first of all by the Ministry of OCW, but other Ministries also fund certain RIs (eg. health, economic affairs, environment, agriculture, etc., depending on the topic).
The majority of the national funds are channeled through the institutions hosting the RIs and (for projects) through the funding organisations NWO and ZONMW. An important additional source is the EU, funding European projects and a number of transnational infrastructures (often temporarily, as for ESFRI).
The number of public-private partnerships in which companies contribute to RIs is limited, and in the Netherlands local authorities (provinces, municipalities) rarely play a funding role for RIs.
In order of importance, the four most important funders can be ranked as follows:

1 - State / ministries
2 - Funding agencies / funding bodies (receiving the funding they distribute from 1)
3 - Research institutions / universities (hosting RIs)
6 - European funds
Of lesser importance are 4 - Regions / towns; 5 - Research communities; 7 - Industries / small and medium-sized enterprises (SMEs).

RIs rarely get their funding through one channel only (E20). A mix of sources of income is the rule rather than the exception. Although detailed information on the contribution of additional sources of income to the total budget is not available, it is safe to assume that there is a lot of variation, often born out of necessity and a lack of stable funding. Except for the main source of stable funding (if any -- there are quite a few infrastructures which struggle to survive with project funding only, a fragile and not very sustainable situation), additional sources of income are primarily externally funded in the form of temporary projects, but also paid services and/or participations in joint services on a subscription or membership basis. They include managed online services (data archiving, repositories, storage, computing), hosting (hardware and services for third parties) and training & consultancy. Selling applications commercially or delivering software as a service is a negligible source of additional income (E201), as much of the academic software is available as open source (or otherwise made freely available), which is often a requirement made by funders if the software is developed in publicly funded projects. For example, the Netherlands eScience Centre (NLeSC) provides a growing Research Software Directory, which aims to promote the impact, exchange and free re-use of research software (https://www.research-software.nl; last accessed on 2/4/2020). DANS and NLeSC jointly present recommendations for promoting sustainable and FAIR software (https://fair-software.nl).

The fragility of the funding of "research infrastructure" in the form of projects has already been noticed multiple times in various documents, including various Science Europe briefing papers and reports (see: https://www.scienceeurope.org/our-priorities/research-infrastructures/ and the references in this landscape analysis).

3.1.6. Role and type of SLA's and potential issues/barriers (E28/E29/E31/E32)

Whether Service Level Agreements (SLAs) are provided depends on the professionalism and maturity of the research infrastructures, and to which customer groups their services are provided. When services are provided on a paid basis, SLAs are usually in place. But many of the services developed in an academic context are provided "as is" without service levels being clearly specified. It has been remarked elsewhere, that even licensing leaves much to be desired in the academic world. Our estimate is that SLAs are available for a minority of the services on offer (E28).
The explanation why this is so is perhaps related to the trust and quality control system in academia, which is rather based on peer review and referencing than on fixed procedures and contracts. SLAs perhaps fit less in the academic culture and may even be perceived as superfluous in a lot of cases.

There is hardly any research infrastructure that does not belong to one or more international organisations. However, such transnational organisations come in a multitude of forms, ranging from temporary coalitions and project consortia to membership organisations such as ERICs. Belonging to such an organisation in itself does not imply that SLAs are implemented. It is our estimation that only in a small number of cases participating in a transnational organisation or federation the international Service Level Agreements (SLAs) or similar contracts are also binding for the national partners (E29). One interesting example of an ERIC actively promoting the implementation of SLAs and other "Contracts, Licenses and Liabilities" among the connected service providers is CESSDA, the ERIC integrating European social science data archives (https://www.cessda.eu/Tools-Services/For-Service-Providers/CESSDA-CDM/Part-1-CRA1-Organisational-Infrastructure/CPA1.2-Contracts-Licenses-and-Liabilities). The organisation has developed a Capability Development Model (CESSDA-CDM), which is the basis upon which an assessment of social science service provision is made, serving as an aid in the improvement of the capabilities of existing and future CESSDA service providers. Although the main emphasis of the model is on social science research data, it is applicable for a wider range of data service organisations (https://www.cessda.eu/Tools-Services/For-Service-Providers/CESSDA-CDM/)

3.2. Data services

3.2.1. Data management, curation and (long-term) preservation

In this section we base ourselves on the information provided by re3data.org, describing 56 repositories in The Netherlands.

3.2.1.1. Documentation and metadata standards (including provenance, metadata languages - E71/E72/E67)

According to re3data, 47 out of the 56 repositories (84%) provide information about the metadata standard used. Dublin Core is the most common standard (29%), followed by DDI (13%) and the DataCite Metadata Schema (11%). All other standards occur 5 times or less.

Table 11. Metadata standards used in 56 repositories in The Netherlands

Metadata standards	Repositories	Percentage
Dublin Core	16	29%
DDI - Data Documentation Initiative	7	13%
DataCite Metadata Schema	6	11%
ISO 19115	5	9%
RDF Data Cube Vocabulary	3	5%
Repository-Developed Metadata Schemas	2	4%

CF Climate and Forecast Metadata Conventions	1	2%
DCAT - Data Catalog Vocabulary	1	2%
Darwin Core	1	2%
FITS - Flexible Image Transport System	1	2%
OAI-ORE - Open Archives Initiative Object Reuse and Exchange	1	2%
PROV	1	2%
SDMX - Statistical Data and Metadata Exchange	1	2%
other	1	2%
none provided	9	16%
Total	47	100%

Source: re3data, https://www.re3data.org/search?query=&countries[]=NLD

All repositories registered in re3data use at least English as a language for metadata, though not always exclusively so. 16 of them (29%) also use Dutch language, and 17 languages are mentioned in between 1 and 3 repositories. One repository reports using 9 different languages.

Table 12. Languages used in 56 repositories in The Netherlands

Repository languages	**Repositories**	**Percentage**
eng (B - English)	56	100%
nld (A - Dutch)	16	29%
fra (C - French)	3	5%
por (F - Other)	3	5%
spa (F - Other)	3	5%
deu (D - German)	2	4%
ara (F - Other)	1	2%
ces (F - Other)	1	2%
dan (F - Other)	1	2%
ell (F - Other)	1	2%
est (F - Other)	1	2%
heb (F - Other)	1	2%
ind (F - Other)	1	2%

ita (E - Italian)	1	2%
pol (F - Other)	1	2%
slv (F - Other)	1	2%
swe (F - Other)	1	2%
tha (F - Other)	1	2%
zho (F - Other)	1	2%
Total	**56**	**100%**

Source: re3data, https://www.re3data.org/search?query=&countries[]=NLD

3.2.1.2. Use of standardized/controlled vocabularies for metadata (E66/E661)

The use of standardized/controlled vocabularies is one of the more problematic elements in the area of making data FAIR, because of three reasons:

- there are innumerable standards at all kinds of levels, so the use of a local or community standard does not necessarily help global standardization across domains;
- controlled vocabularies are typically used for standardizing the categories for specific elements, where the categories are necessarily a limited set (of course an "other" category always offers an escape); it makes a lot of difference for how many (and which) elements controlled vocabularies are used!
- the status of the controlled vocabularies is often unclear, unless there is some normalization/standardization body or dominant industry that recognizes them. Worse, many vocabularies are not professionally maintained, and therefore are not sustainable.

FAIR principles I1, I2 and R1.3 attempt to address these problems, but catalogues, registries and repositories of vocabularies are as yet in their infancy, although several laudable initiatives exist (see, e.g. https://www.rd-alliance.org/groups/vocabulary-services-interest-group.html; Coen et al. 2019; https://bartoc.org/en).

The repositories that are certified and/or that use a metadata standard are also those that employ standardized/controlled vocabularies for metadata.

3.2.1.3. Regulations and policies (E56)

The answers to Question E56 of the PILLAR survey aim to provide an overview on whether service providing organisations (not only repositories) have regulations and policies in place governing the services they provide, concerning the following subjects:

A. Research data management (RDM):
B. Open research data
C. Long-term availability of research data
D. Compliance of data to the FAIR principles
E. Publication of data in a repository
F. Publication of data in a certified repository
G. Other

The Dutch funding organisations NWO and ZonMw, the Academy (KNAW) and all universities require their researchers to submit some form of data management plan (DMP; in some cases a "data management paragraph" in the research plan suffices) for new research projects. DMPs typically include the elements (B) to (F) and more (G).
Whether the service providers, especially the data repositories, have such policies in place is perhaps less relevant, but they often do, if only because they usually belong to a bigger organisation that has such policies.
The implementation and enforcement of the policies is another matter and leaves more to be desired. Moreover, there still is considerable variation in the details of the policies. A small but growing number of organisations is following the Science Europe *Practical Guide to the International Alignment of Research Data Management* (2018).

3.2.1.4. Certified repositories (E73)

Out of the 56 Dutch repositories, 16 (29%) have one or more certificates. 8 repositories comply with the requirements of the Data Seal of Approval (DSA) and 7 with the Core Trust Seal (CTS). Out of this last group, 3 are also CLARIN B Centres, and 1 complies with the German nestor seal DIN 31644.

Table 13. Certified repositories in The Netherlands

Certificates	Repositories	Percentage
DSA	8	14%
CoreTrustSeal	7	13%
CLARIN certificate B	3	5%
DIN 31644	1	2%
RatSWD	1	2%
Repositories with 1+ certificates	16	29%
Total certificates	20	

Source: re3data, https://www.re3data.org/search?query=&countries[]=NLD

3.2.1.5. Compliance with FAIR principles (E63)

Indicators for the practical measurement of and criteria for compliance with the FAIR principles are still being debated and under development (e.g. in the FAIRsFAIR project, see: https://www.fairsfair.eu). Moreover, distinction should be made between compliance at the level of repositories (and other data services) and at the level of data (or other resources) *within* repositories.
As the underlying principles of DSA and CTS are very similar to the FAIR principles, certified repositories already comply with the majority of the FAIR principles. However, several of the FAIR criteria apply to data within a repository, and hence vary from dataset to dataset.

It is expected that the majority of Dutch repositories aim to comply with the FAIR principles, but would welcome practical guidelines for implementation.

3.2.1.6. Machine-readable data catalogues (E57)

Re3data provides several indicators for the presence of machine-readable data catalogues. Obviously, all repositories have some form of a data-access mechanism. One of these is the presence of an API for accessing the data.

Table 14. APIs for access to data in Dutch repositories

API	Repositories	Percentage
OAI-PMH	11	20%
REST	9	16%
FTP	5	9%
NetCDF	4	7%
OpenDAP	2	4%
SPARQL	2	4%
SOAP	1	2%
SWORD	1	2%
other	6	11%
Total	56	73%

Source: re3data, https://www.re3data.org/search?query=&countries[]=NLD

3.2.1.7. Implementation of PIDs and researcher identifiers (E70)

Only slightly over half (54%) of the Dutch repositories use a PID system for persistently referring to data resources, and Handle (23%) and DOI (21%) are clearly the two most used systems, URN and PURL being used only in a handful of repositories.

Table 15. PID Systems used in Dutch repositories

PID systems	Repositories	Percentage
hdl	13	23%
DOI	12	21%
URN	4	7%
PURL	1	2%
none	26	46%

Total	56	100%

Source: re3data, https://www.re3data.org/search?query=&countries[]=NLD

The use of Researcher Identifiers or Author Identifiers (AIDs) is not well registered in re3data. For The Netherlands, the registry provides the following (incomplete) information:

Table 16. Use of Author/Researcher Identifiers in Dutch repositories

AID systems	Repositories
ISNI	1
ORCID	6
none	6

Source: re3data, https://www.re3data.org/search?query=&countries[]=NLD

The Digital Author identifier (DAI), used by DANS and most institutional (university) repositories (next to ISNI and ORCID), is not mentioned in re3data, nor is the ResearcherID provided by Thomson Reuters and used by e.g. the Web of Science. Although we do not know the numbers, ORCID, DAI and ISNI all play an important role in Dutch repositories, while the ResearcherID is also used by publishers. Nevertheless we estimate that AID systems are implemented in a minority of the repositories. That only 6 repositories do not use any AID seems an underestimation of the real situation.

3.2.1.8. Types of data archived (E93)

Re3data provides information on the "content types" in the Dutch repositories, where more than one type is possible. In total, 270 content types are mentioned for the 56 repositories, meaning that on average, a repository has about 5 content types. It is almost certain that the true number of content or data types will be even higher. The DANS archives, for example, contain hundreds of data types and file formats in many versions, although usually they are converted to a smaller number of "preferred" or "archival" formats (see: https://dans.knaw.nl/en/about/services/easy/information-about-depositing-data/before-depositing/file-formats?set_language=en).

Table 17. Content types in Dutch repositories

Content Types	Repositories	Percentage
Standard office documents	42	75%
Scientific and statistical data formats	36	64%
Plain text	31	55%
Raw data	26	46%
Images	23	41%

Audiovisual data	19	34%
Structured graphics	17	30%
other	17	30%
Archived data	16	29%
Structured text	16	29%
Databases	13	23%
Software applications	7	13%
Configuration data	3	5%
Network Based data	2	4%
Source code	2	4%
Total	56	100%

Source: re3data, https://www.re3data.org/search?query=&countries[]=NLD

The table gives an impression of the main content types, although some categories are overlapping. Textual and document data types occur (according to re3data) most frequently, followed by "scientific and statistical data", image data and A/V data.

3.2.1.9. Number of datasets accessible online (E76)

In the PILLAR questionnaire, a dataset is described as "a logical collection of one or more files which is described by a metadata scheme". Different repositories appear to use different ways of grouping resources into logical units, and also use different names to indicate them, such as data collections or dataverses; data belonging to a research project or "study" are alternative ways to distinguish a data set. Some "repositories" registered in re3data are perhaps better characterized as information systems, databases or databanks (containing records) rather than data set storage systems. Some datasets may consist of just one file (from a small spreadsheet to massive tables with billions of measurements), whereas others contain various hierarchical levels with thousands of tables, images, etc. Software code is also a distinct category of resource. This makes counting and comparing the contents of repositories in terms of data sets a virtually impossible task. A partial image can be obtained from the NARCIS portal, aggregating the data resources offered by 22 Dutch organisations (partly corresponding to the repositories of re3data).

Table 18. Numbers of datasets in selected Dutch repositories

Repository	**Number of Datasets**
The Language Archive - MPI	133105
DANS-KNAW	123195

4TU.Centre for Research Data	6095
Wageningen University & Research Centre	1408
SURF Data Repository	696
Utrecht University	640
CentERdata	552
Radboud University Nijmegen	421
University of Amsterdam	336
Netherlands Biodiversity Information Facility - NLBIF	305
Tilburg University	300
Zorggegevens	271
Maastricht University	143
University of Groningen	113
Netherlands eScience Center (software)	104
Erasmus University Rotterdam	66
Leiden University	46
VU University Amsterdam	37
CancerData	34
KNMI	24
NIOZ	11
Trimbos Institute	10
National Library of the Netherlands	6
Total	267918

Source: https://www.narcis.nl (accessed on 5/4/2020); note that most of the University repositories work together in Dataverse.NL, a joint service hosted by DANS. In re3data, Dataverse.NL is registered as a single repository, although some universities using Dataverse.NL also registered their repositories separately.

3.2.1.10. Reuse of datasets, user characteristics and frequency of use (E77, E78)

Reuse statistics are hard to get by. Two recent studies on the reuse of data from the DANS archive EASY shed light on the intensity and variety of data reuse. In a recent article published in the German journal "der Archivar", Doorn (2020) describes the development of the DANS archives and its use from 2007 till 2019. The reuse of datasets has increased in a fairly similar

way as the growth of the content of the archive. In 2007 about 1000 datasets were downloaded for reuse, a number which grew to around 20,000 in 2011. In 2019, the number of downloads grew to about 47,000 datasets.

Table 19. Development of datasets by domains present in DANS EASY archive

Year	Archaeology	Behavioural sciences	Humanities	Life sciences and medicine	Social sciences	Socio-cultural sciences	Total
2007	49	143	74	10	248	555	1079
2008	462	469	276	71	721	2344	4343
2009	4165	656	390	73	853	3032	9169
2010	10874	1278	1168	150	1497	4079	19046
2011	12009	1334	1052	210	1464	4296	20365
2012	20651	1650	2045	299	2022	6349	33016
2013	23762	2240	2197	441	2371	7332	38343
2014*	38547*	2400	3071	499	3003	7491	55011*
2015	22497	2820	3054	607	2849	7888	39715
2016	23617	2315	2547	735	2653	5924	37791
2017	21888	2625	3191	750	3118	6395	37967
2018	23813	2890	3483	914	3009	7411	41520
2019	26068	3320	4298	1352	3175	8875	47088

Note: because datasets can be classified under more than one discipline, the actual total is less than the total of all disciplines; in 2014, one user downloaded all 15,000 archaeological datasets then in the archive, thus inflating the numbers considerably.

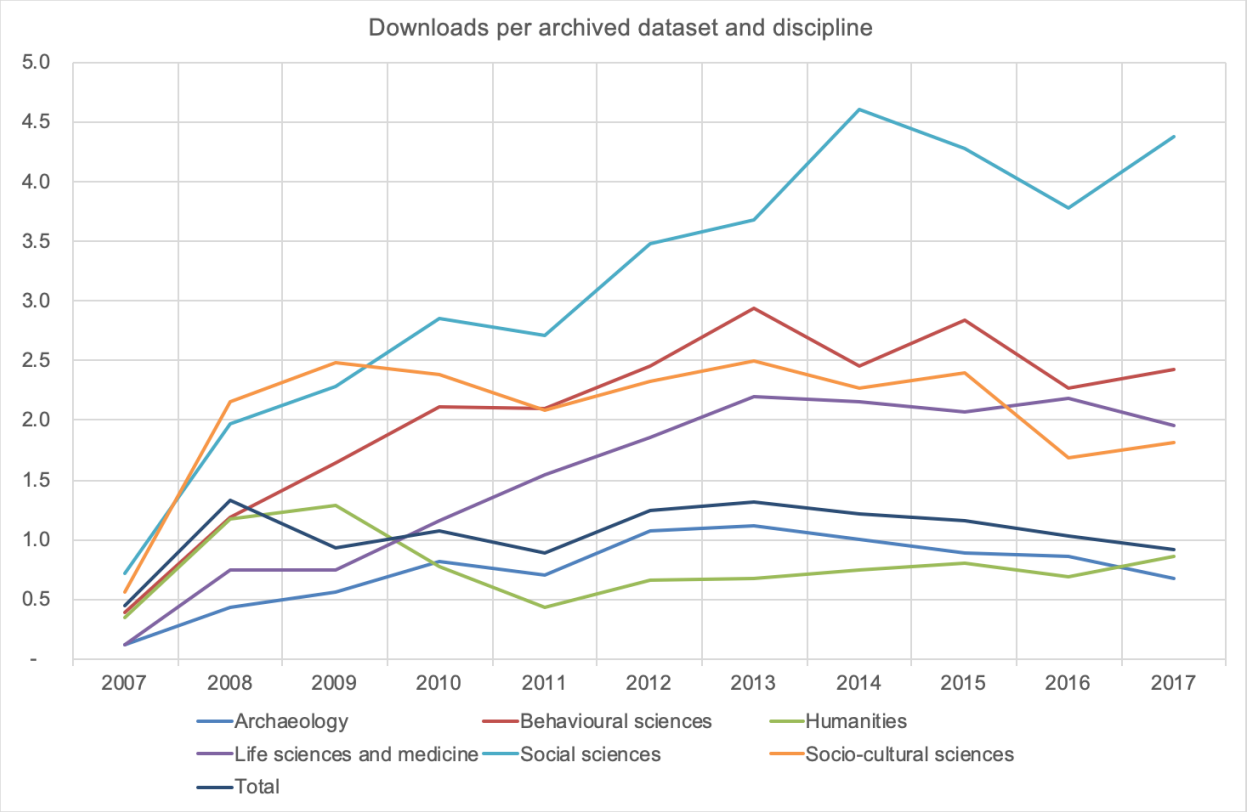

Figure 2. Downloads per archived dataset and discipline from DANS EASY archive

It is instructive to look at the downloads relative to the size of the archive per domain (see Figure). 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 about 1 throughout the investigated period, meaning that every dataset in the archive is downloaded about once every year. Despite the rise of archaeology relative to the social sciences, the data reuse in the social sciences remained clearly on top since 2010, with a reuse ratio of between 3 and 4.5. The social scientists are followed by the groups 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 respectively around 1 and 0.75.

A more qualitative insight in the characteristics of the users and use can be gained from an article by Christine Borgman et al (2019). The authors state "*Our interviews, weblogs, ethnography, and document analyses reveal that a few large contributors provide a steady flow of content, but most are academic researchers who submit data sets infrequently and often restrict access to their files. Consumers are a diverse group that overlaps minimally with contributors. Archivists devote about half their time to aiding contributors with curation processes and half to assisting consumers. Given the diversity and infrequency of usage, human assistance in curation and search remains essential. DANS' knowledge infrastructure encompasses public and private stakeholders who contribute, consume, harvest, and serve their data—many of whom did not exist at the time the DANS collections originated—reinforcing the need for continuous investment in digital data archives as their communities, technologies, and services evolve.*"

3.2.3. Data sharing and access

3.2.3.1. Access policies, restrictions and licenses (E21/E38/E39)

Most repositories in The Netherlands use a variety of access licences, distinguishing between open (80%) and restricted access licences (75%), and less frequently they also contain data that is closed for external access (23%) or where access is embargoed (20%). Note that these percentages are not percentages of the datasets, but of the repositories using these licences.

Table 20. Access licenses in use in Dutch repositories

Data access	Repositories	Percentage
open	45	80%
restricted	42	75%
closed	13	23%
embargoed	11	20%
Total	56	100%

Source: re3data, https://www.re3data.org/search?query=&countries[]=NLD

When access restrictions apply, four types of restrictions are given. Most common is registration by the user, which applies in just over half of the repositories (52%). 5 repositories charge a fee for access, and in 2 cases institutional membership is a requirement. There is a substantial group (41%) making other restrictions. Although we have no detailed information about this, we expect that "for scientific use only" or "non-commercial use only" will be important in this category.

Table 21. Access restrictions applicable in Dutch data repositories

Data access restrictions	Repositories	Percentage
registration	29	52%
feeRequired	5	9%
institutional membership	1	2%
other	23	41%
Total (repositories with restricted access)	42	75%

Source: re3data, https://www.re3data.org/search?query=&countries[]=NLD

3.2.3.2. Impediments to data sharing (E69)

Impediments to data sharing exist, and they appear to vary over time and over disciplines. Requirements by funders and journals play an important role in changing the cultures of data sharing. We have no recent quantitative data on the concerns of depositors. There is a clear

tendency towards a more open sharing of data among researchers, probably related to the encouragement of open science. An impression from this tendency is based on the access categories in the DANS EASY data archive.

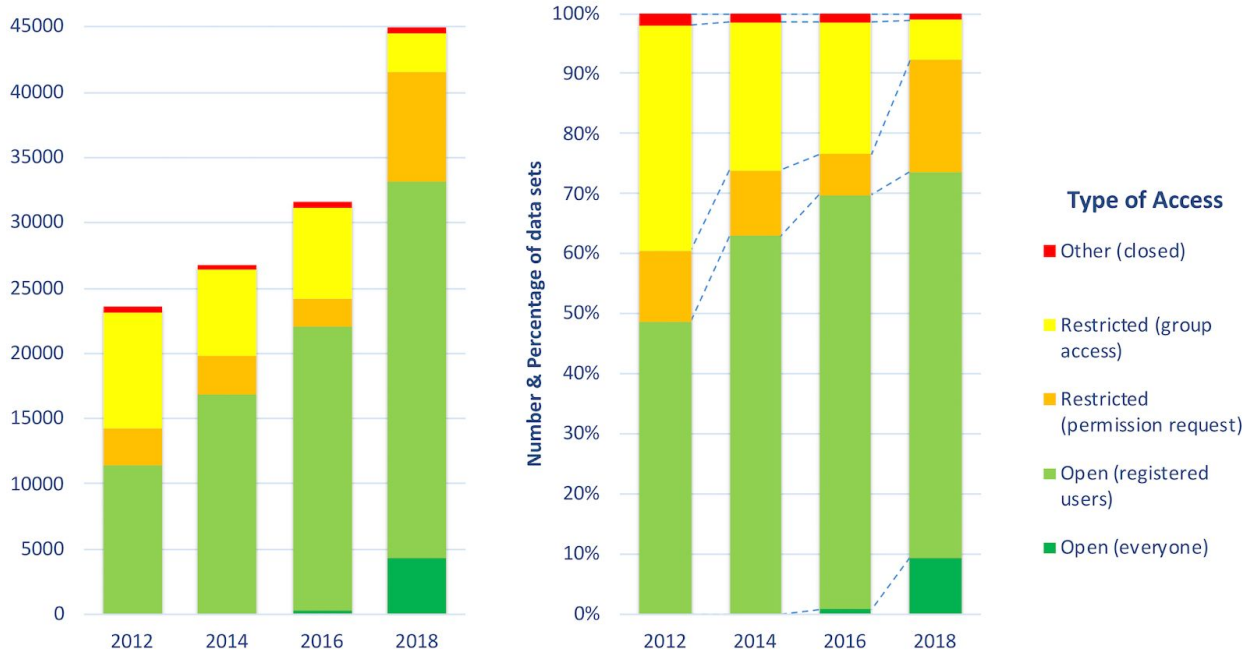

Figure 3. Access categories to datasets in DANS EASY archive, 2012-2018

3.2.3.3. Policies and limits to expansion of access for other groups (E22)

As seen in section 3.2.1., 47 out of 56 Dutch repositories provide information about the metadata schema used. The metadata is accessible to all, be it that some standards are specific for particular communities or disciplines, and may not be compared easily to other standards. The mapping of the various metadata standards to a common standard may pose some challenges, but they are not likely to be fundamental.

3.2.3.4. User charges (E25)

As reported in section 3.2.3, 5 repositories out of 56 charge a fee for access, and in 2 cases institutional membership is a requirement.

3.2.3.5. Data search tools for (meta)data (E64/E65)

Although we have no specifics on this subject, it can be safely assumed that all repositories providing or using a metadata schema (47 out of 56 reported this) or providing an API (see section 3.2.1., question E57) are searchable by some mechanism.

3.2.3.6. Authorization, access control and access federation

The subject of authorization and access gets rather detailed attention in the PILLAR questionnaire, including:

- authentication model used (E40)
- implementation and need for support (E54)
- plans for access federation (E41)
- proxy to eduGAIN (E43)
- management of authorisation information (E44)
- use of REFEDS R&S entity category (E51)

No systematic quantitative information is available on this subject for the Netherlands, but it is certain that repositories are using a variety methods of authorization and access control. It was noted before that 2 repositories require institutional membership for using data. Many organisations in the Dutch academic system make use of SURFconext for authentication and authorization. An overview of the 200 most popular services that are connected to SURFconext is listed here: https://www.surf.nl/en/surfconext-global-access-with-1-set-of-credentials/which-services-can-you-use-with-surfconext.
Some service provider federations (e.g. CLARIN) additionally have an AAI policy in place for their community. Group membership, individual logins, and complete open access also occur. Identity Providers connected to SURFconext who want to use a service from another federation through eduGAIN, do not have to make any technical changes. Service Providers from other federations are connected to the central hub and thus operate according to the hub-and-spoke principle (see: https://wiki.surfnet.nl/display/surfconextdev/Federation+architecture%3A+hub-and-spoke+vs.+mesh)

3.2.3.7. Access to and protection of personal data (E46/E47)

We have no particular information of the number of organisations providing access to (special) personal data, but the repositories in the domains of the humanities, social sciences and biomedical sciences are likely to manage datasets that may contain data about human subjects. The awareness about the consequences of the GDPR for data sharing is much higher among repositories than among researchers, and it can be assumed that the repositories have procedures in place to comply with the GDPR. Anonymization, ensuring that data depositors have informed consent from human subjects (respondents, patients, etc.) for sharing data, and sometimes pseudonymization and encryption are ways of coping.

4. Procurement of and transnational access to services and resources compatible with EOSC

4.1. Transnational access to national resources/services

Most of the infrastructure organisations provide access to the services they offer internationally, and often the metadata and additional documentation are provided in English to make the resources better findable to foreign researchers.Many services and facilities are part of international consortia, or are part of distributed data facilities, as is demonstrated in the table below. And because, as described before, many services are openly accessible, no distinction is made between national or international users.

Table 22. Repositories providing shared services with other (groups of) countries

Countries	Repositories	Percentage
European Union	20	36%
Germany	8	14%
International	8	14%
United States	7	13%
United Kingdom	4	7%
France	3	5%
Sweden	2	4%
Switzerland	2	4%
Australia	1	2%
Belgium	1	2%
Brazil	1	2%
Greece	1	2%
India	1	2%
Indonesia	1	2%
Ireland	1	2%
Italy	1	2%
Japan	1	2%

Korea, Republic of	1	2%
Russian Federation	1	2%
Total repositories	56	100%

Source: re3data, https://www.re3data.org/search?query=&countries[]=NLD

4.2. Potential for harmonization of national procurement policies

4.2.1. Harmonization of joint procurement

The potential for the procurement of big facilities depends on the scale and type of facility. There already exists considerable international cooperation in this area, and it is not easy to generalize about the optimal level of harmonization internationally.
Concerning the type of facility, it makes sense to distinguish generic ones (such as HPC, grids & cloud, mass storage, computer networks) from discipline-specific equipment.
With respect to the first category, the Netherlands quite intensively participates in international bodies and consortia, including EuroHPC, GÉANT, PRACE, EGI, E-IRG, etc.
With respect to the second category, many of the disciplinary facilities are also part of broader networks or consortia.
Whether the procurement of new equipment is a local, regional, national or international endeavour depends on the scale of the investment necessary and on the funding instrument available. The main funding mechanisms and instruments were described in section 2.
International funding for generic and thematic infrastructure often has a matching national component, and usually concerns amounts in the size order of multi-million euros.

4.2.2. Coordinated service provisioning

Over the past decades, an increasing number of services have worked towards international collaboration and coordination, and even integration. This does not mean that all services seamlessly work together. One could rather describe the landscape as a patchwork, but one both with overlaps and with gaps. This is perhaps due to the heterogeneous nature of the science system itself, but also to the competitive funding system, in which independent consortia develop competing proposals. Moreover, not all disciplines develop at the same pace, and facilities useful for one area can not simply be transferred to other ones. The user requirements of theologians for digital services are not the same as those of biologists or particle physicists. At the same time, there are also variations in requirements and developments across countries, as will be apparent from the EOSC-Synergy landscape analyses alone.
In many disciplines, the international harmonization of services started with or got an impetus by the ESFRI-developments of the last 10-15 years. This harmonization has appeared to be a long and difficult process, which is still going on. While leaps forward have been or are being realized, and while such integration processes are progressing, a next leap to integrate across

domains is taking place, and the EOSC as an overall integrator across integrations is aiming to be the superstructure covering the whole cobweb of services. It is not surprising that questions are posed whether EOSC is not too ambitious and what the surplus value will be of the harmonizations within domains that are underway. Care must be taken that the harmonization work within domain-oriented services is not overrun or complicated by yet another layer of coordination.

4.2.3. Coordinated organisation models

Coordination of service provisioning is only possible if organisations providing those services work together in some form of collaborative organisation or consortium. Of the 32 ESFRI projects and landmarks in which The Netherlands participate, 13 have the ERIC status (two more are in phase ERIC step 2), 3 are an AISBL, 3 are EIROforum members, and of the 9 projects the legal status is pending. The remaining two have some other legal form.

4.3. Procuring services/resources

The following national and European regulatory frameworks for integrated procurement for purchasing supplies, resources or services apply (E27/E30). On 1 July 2016 the amended Dutch Public Procurement Act entered into force, which implements the latest EU procurement directives (2014/23/EU, 2014/24/EU and 2014/25/EU). The Public Procurement Act applies to all procedures below and above the EU threshold, which in most situations is currently €214,000.

Dutch public procurement law recognises the general principles of public procurement law (equal treatment, non-discrimination, mutual recognition, proportionality and transparency) and the general principles of Dutch civil law (including pre-contractual good faith).
The Public Procurement Act 2012 (amended per 1 July 2016) applies to all public contracts (in Dutch: https://wetten.overheid.nl/BWBR0032203/2016-07-01).The act is further detailed in the Public Procurement Decree (*Aanbestedingsbesluit*).The Proportionality Guide (*Gids Proportionaliteit*), Works Procurement Regulations 2016 (*Aanbestedingsreglement Werken 2016*) and the European Single Procurement Document / ESPD (*Uniform Europees Aanbestedingsdocument / UEA*) are part of the Public Procurement Decree (in Dutch: https://wetten.overheid.nl/BWBR0032919/2016-07-01).

The Proportionality Guide intends to ensure that all requirements imposed by a contracting authority are proportionate to the object and scope of the public contract (https://www.pianoo.nl/en/legal-framework/proportionality-guide-1st-revision). The Public Procurement Decree states that the Proportionality Guide is to be considered as a mandatory directive. The Proportionality Guide further elaborates on the principle of proportionality and how it should be applied in procurement procedures. Accordingly, the application of the Guide should strengthen the position of small and medium-sized enterprises during tender procedures. Contracting authorities may only deviate from the detailed provisions on proportionality if this is properly motivated in the tender documents.

Nationwide general purchase conditions are ARVODI-2018, ARIV-2018 en ARBIT-2018. These general terms are valid from 15 mei 2018.

The Works Procurement Regulations 2016 describe procedures which are mandatory for procurement for the awarding of (public) works contracts above and below the EU threshold of € 5,350,000 (https://www.pianoo.nl/nl/regelgeving/aanbestedingsreglement-werken-2016-arw-2016).

5. Conclusions

5.1. SWOT analysis

We start these conclusions with a SWOT table, providing an overview of the strengths and weaknesses, opportunities and threats of the Dutch data landscape:

Table 23. SWOT Table

Strengths	Weaknesses
• The Netherlands have a strong position in Europe in the field of open science and FAIR data, as witnessed by the impulse given during the Dutch presidency of the EU in 2016 and by the Dutch contribution to the realization of the FAIR principles. • In 2017, the Dutch government declared open science and open access the standard for scientific research. • There is an active and strong involvement of the Ministry of Education, Culture and Science in promoting open science. • The national research funders (NWO, ZonMw) have included data management in the requirements for grants. • Responsible data management and long-term access to data, are part of the widely supported national code of conduct for scientific integrity. • Almost all knowledge institutions have up-to-date data policies and a reasonably detailed set of data-oriented services for their own researchers. • There is a strong culture of cooperation in the Netherlands, as is evident from the large number of partnerships and consultative bodies (also in large-scale scientific infrastructures) and national ICT (such as SURF) and data service organizations for education and research. • Within the Netherlands we have a large amount of data repositories, generic and discipline specific, of which several are certified. • The DANS institute is the initiator of the international quality mark CoreTrustSeal (formerly Data Seal of Approval) for repositories.	• There is overlap and fragmentation of data-oriented services (and therefore possible inefficiency). Despite the many consultative and coordinating bodies, there is a lack of true cooperation, harmonization measures, and division of tasks. • In spite of existing policies, agreements, protocols and standards, these are not generally accepted and / or implemented. • Most of the research data from recently completed research is located on personal disks or closed network drives of the knowledge institutions and is not FAIR. • Researchers do not always know where they stand, what the rules of the game are (ethical and legal aspects) and what facilities are at their disposal. The diversity in terminologies and research cultures used does not help either. • There is uncertainty among researchers about the handling of company and personal sensitive data during and after their research. Adequate facilities for this are lacking or not known. • The degree of organization within disciplines and domains is very diverse, also depending on available international partnerships (with national focal points) and temporary or structural financing. • Research funding is often on a project basis, while careful data management and infrastructural facilities require longer-term investments. • There is still insufficient recognition and appreciation in the research sector for careful data management, the sharing of data and the associated roles.

	• Data management takes time and also requires skills that are new to many researchers. There is a great shortage of capacity for (support for) careful data management (data stewards).
Opportunities	**Threats**
• The emergence of a federated infrastructure of data competence centers (DCCs) through an incentive scheme for digitization of science. • Collaboration of scientific institutions with private companies and public organizations when it comes to data sharing (initiatives such as the "Amsterdam Data Exchange AMDEX", data science institutes, and public investments in the new "Data Sharing Coalition"). • National cooperation to improve the system of recognition and valuation for data sharing. A relevant position paper states: "Open science and the modernization of the system of recognition and valuation are inextricably linked." • The new Strategy Evaluation Protocol (2021-2026), that forms the basis for scientific evaluations of research units explicitly focuses on the contribution to open science. • The development of the European Open Science Cloud and the new Data Strategy of the European Commission. • The increasing cooperation between the international platforms / networks for research data (e.g., CODATA, GO-FAIR, RDA and WDS => "Data Together"). • A common, national roadmap to solve the experienced bottlenecks, promote national synergy and optimize international connections.	• The separate actions of policy makers, researchers and data support organisations when it comes to open science, data sharing and careful data management. • More and more commercial services are offered on the market that meet the needs of researchers. However, there are risks with regard to ownership of data, dependence on commercial parties, etc. • Too strict laws and regulations (for example regarding copyright and privacy), which discourage the sharing and reuse of research data. • Threats with regard to the security of ICT infrastructures, criminal interest, abuse of research data and cybercrime (e.g. ransomware).

Source and acknowledgement: this is a translated and slightly edited version of the SWOT analysis from chapter 7 "Conclusions and recommendations" of the draft text (consultation version 1 May 2020) by the Project Group "Verkenning Nationaal Datalandschap" in the context of the National Platform Open Science. The author, Melle de Vries, gracefully allowed me to adapt his table for this report.

5.2. Gaps, overlaps and optimization of the landscape

The Dutch data landscape is richly variegated: it counts many services and service providers (50-150, depending on what and how to count: individual services, service providers, legal entities and/or projects), ranging from small to fairly big, and from aspiring projects to well-established institutes.
The downside to this variety, is that the landscape is also characterized by a considerable degree of fragmentation. It is often heard that more coordination is desirable, and there also exists a host of coordinating platforms. Restricting ourselves to national consultative bodies (and also disregarding disciplinary coordinating efforts, there are certainly around a dozen, perhaps more, platforms, networking groups, coordination committees, etc. Apparently, the Dutch data landscape displays a similarity in this respect to the polder landscape, which has a long history of coordination and negotiation that goes back to the Middle Ages (see https://en.wikipedia.org/wiki/Polder_model about the "polder model", a "typically Dutch" way of consensus decision-making).
However, although many want more coordination, few want to *be* coordinated. The "gaps" that are described under the heading "weaknesses" in the SWOT table of the NPOS reconnaissance working group rather have to do with inefficiencies in the system than with a lack of facilities and services as such. Moreover, open science has not yet deeply permeated the research culture, or at least the awareness of its implications varies considerably among research communities. There are indications in the literature that age may be an important factor. Training in FAIR research data management is not yet part of the university curriculum, which explains a lack of skills.

To make the national system of facilities more efficient is a national puzzle, closely connected to optimizing the system internationally, which is the big challenge of the EOSC at the European level. Of course, the science system itself is enormously varied and extremely specialized, which is why it is no surprise that so many facilities exist on the level of domains and (sub)disciplines. Moreover, every Dutch university in the past few years has set up its own data support desk, which often includes running a repository service for its own academic community.

Table 24. Data repositories and RDM support units at Dutch Universities

University	Repository	Data-sets	Webpage RDM Support unit	Webpage Repository
Vrije Universiteit Amsterdam (VU)	DataverseNL	38	https://ub.vu.nl/nl/ub-voor-onderzoekers/research-data-support/index.aspx	https://dataverse.nl/dataverse/vuamsterdam
University of Twente (UT)*	DataverseNL	10	https://www.utwente.nl/en/lisa/service-manuals/product/p885008/researchdatamanagement	https://dataverse.nl/dataverse/twente
Delft University of Technology (TUD)*	DataverseNL	43	https://www.tudelft.nl/en/library/current-topics/research-data-management/	https://dataverse.nl/dataverse/delft

Eindhoven University of Technology (TU/e)*	DataverseNL	4	https://www.tue.nl/en/our-university/library/education-research-support/scientific-publishing/data-coach/	https://dataverse.nl/dataverse/eindhoven
Wageningen University & Research*	DataverseNL	17?	https://www.wur.nl/en/Value-Creation-Cooperation/WDCC/Data-Management-WDCC.htm	No separate Dataverse
Universiteit Utrecht (UU)	DataverseNL	559	https://www.uu.nl/en/research/research-data-management	https://dataverse.nl/dataverse/UU
idem	UU YODA Portal	?		https://i-lab.yoda.uu.nl
Maastricht University (UM)	DataverseNL	145	https://library.maastrichtuniversity.nl/research-support/rdm/	https://dataverse.nl/dataverse/maastricht
Tilburg University	DataverseNL	110	https://blog.uvt.nl/rs/category/research-data-office/?nomobile	https://dataverse.nl/dataverse/tiu
Rijksuniversiteit Groningen (RUG)	DataverseNL	115	https://www.rug.nl/research/research-data-management/	https://dataverse.nl/dataverse/groningen
Universiteit Leiden	DataverseNL	53	https://www.bibliotheek.universiteitleiden.nl/over-ons/centre-for-digital-scholarship	https://dataverse.nl/dataverse/leidenuniversity
Radboud Universiteit Nijmegen	Radboud Repository	483	https://www.ru.nl/rdm/	https://repository.ubn.ru.nl/browse?value=Dataset&type=type
idem	Radboud Donders Repository	1518		https://data.donders.ru.nl/?5
idem	Radboud UMC Digital Research Environment	?		https://www.radboudumc.nl/en/research/radboud-technology-centers/data-stewardship/digital-research-environment
Universiteit van Amsterdam (UvA)	UvA/AUAS figshare	105	https://rdm.uva.nl/en	https://uvaauas.figshare.com
Erasmus University Rotterdam (EUR)		?	https://www.eur.nl/en/library/erasmus-data-service-centre	?
4TU.Researchdata	4TU Data Archive	9240		https://data.4tu.nl

DataverseNL	DataverseNL	1363	https://dataverse.nl

Notes: * The three Technical Universities and the Agricultural University also use a common Data Archive: 4TU.Researchdata; ? unknown due to inaccessibility.

In spite of their official RDM policies and support for FAIR data, it is peculiar that the repositories of three universities are not accessible by the outside world. The number of data sets in the institutional data repositories of most universities is moreover suspiciously low, especially in comparison to the content of the shared data archive of the technical universities or the DANS archive. It is therefore likely that the majority of the data of university researchers is hidden on personal disks or institutional drives, inaccessible to researchers from other institutions.
In spite of good intentions and claims that Open Science is the official policy, we can see a considerable gap between practice and reality in many places. Admittedly, Dutch university researchers also use other data repositories and archives, such as DANS, figshare, Zenodo, Mendeley Data, Dryad, OSF, etc. But estimates are that the majority of the research data in many academic communities is not FAIR and open at all.

The NPOS report proposes another level of national coordination, somewhat comparable to the EOSC at the European level, in order to optimize the system. However, how that coordination layer is to take shape is still under debate. Yet, there appears to be an overkill of coordinating bodies in the Netherlands rather than a lack of them: the university data support units alone have opportunities to exchange information in a staggering seven partnerships and coordinating platforms (Advisory Board DataverseNL, Coordinating SURF Contacts, Implementation Network Digital Competence Centres, LCRDM, Open Science consultation, “Regiegroep” RDM services, University Libraries working group research data), and perhaps there is even more. In the view of the author of this report, there is more talking about coordination than actual cooperation. Even when agreement on common standards or practices is reached, the challenges lay in their implementation. The issue is that the hosting organisations of the facilities described in this report are often tied to particular scientific communities and/or they are part of autonomous institutions, which are free to implement an agreement or not. Local requirements or priorities are an all too easy impediment for enforcing a common ruling, norm or shared service.

Although the discussion is by no means finished, a kind of federated structure seems the best solution, and is proposed in the draft of the NPOS Landscape report. Work is in progress to describe the outline of a “National Data Competence Network”, which might have a “national data coordination and expertise centre” as a central hub. One of the questions that is to be answered is if such a body could represent the country in the EOSC.

A second major issue is that many services and infrastructures depend to a substantial degree on project funding; which seems to be inherent in a science system based on competition for grants. This means that there is a continuous struggle among project proposals, which are often forced by the conditions of the calls to stress innovation over consolidation and continuity. Understandable as it is from the standpoint of research funders focusing on pushing the limits of scientific progress, there is also an element of capital destruction if the funding of proven facilities is discontinued, when an existing service is less innovative than a new proposal.

Perhaps this is also an explanation why the maturity of many scientific facilities is less than it could be, as formalisation (expressed by things such as certification, licensing and SLAs) is less appreciated in funding proposals than scientific status and quality, e.g. expressed by peer reviewed publications. The NPOS report also stresses the need for adaptation of the system of scientific valuation and recognition, and sees this as an opportunity in the SWOT table.

Although software sustainability is mentioned in the NPOS report, it does not return in the SWOT table. Nevertheless, this is clearly an element that has so far received relatively little attention. The Netherlands does not stand alone in this respect: facilities for the FAIR sharing of research software are absent in most countries. Therefore, there is an opportunity for the EOSC to provide solutions for software internationally.

Literature/references

Amsterdam Call for Action on Open Science: a living document reflecting the present state of open science evolution, based on the input of many participating experts and stakeholders of the Amsterdam Conference 'Open Science – From Vision to Action', hosted by the Netherlands' EU Presidency on 4 and 5 April 2016. https://www.government.nl/documents/reports/2016/04/04/amsterdam-call-for-action-on-open-science

Borgman, Christine L., Andrea Scharnhorst & Milena S. Golshan (2019), "Digital data archives as knowledge infrastructures: Mediating data sharing and reuse", in: *JASIST*, 70:8, August 2019, p. 888-904. First published: 24 January 2019. https://doi.org/10.1002/asi.24172

Coen, Gerard et al. Observing trajectories of KOSs Across Space and Time: The DANS KOS Observatory (KOSo). NASKO, [S.l.], p. pp. 35-47, sep. 2019. ISSN 2311-4487. Available at: <https://journals.lib.washington.edu/index.php/nasko/article/view/15619>. Date accessed: 05 Apr. 2020. doi:http://dx.doi.org/10.7152/nasko.v7i1.15619.

Doorn, Peter, "Archiving and Managing Research Data: data services to the domains of the humanities and social sciences and beyond - DANS in the Netherlands", in: *Archivar* 73:1, February 2020, p. 44-50.

Eechoud, M. van (2018), "Bekeert Brussel zich tot artikel 8 Databankenwet? Open Databeleid en de IE-positie van de publieke sector na herziening Richtlijn hergebruik overheidsinformatie", in: *Tijdschrift voor Auteurs-, Media- & Informatierecht* 42:4, 127-133.

e-Infrastructure Reflection Group. National Nodes – Getting organised; how far are we? Implementing e-Infrastructure Commons and the European Open Science Cloud. ISBN: 978-90-823661-6-7. Document adopted on 2019-05-22. http://e-IRG.eu/catalogue/eirg-1006

Integrale aanpak voor digitalisering in de wetenschap: Uitvoeringsplan investeringen digitale onderzoeksinfrastructuur (NWO, Den Haag, oktober 2019) https://www.nwo.nl/documents/enw/uitvoeringsplan-investeringen-digitale-onderzoeksinfrastructuur

Morselli, F. Research Infrastructures and collaboration in research: A contextualised literature review. Text under development for PhD Dissertation. DANS, 2020.

Projectgroep NPOS datalandschap (ed. by Melle de Vries), Eindrapportage – Verkenning nationaal datalandschap (project E). Draft version for consultation: 1 May 2020; Final version: 4 June 2020.

Rijksbegroting 2020: VIII Onderwijs, Cultuur en Wetenschap.
https://www.rijksoverheid.nl/documenten/begrotingen/2019/09/17/viii-onderwijs-cultuur-en-wetenschap-rijksbegroting-2020
https://www.nwo.nl/en/policies/open+science
https://easy.dans.knaw.nl/ui/browse

Science Europe reports:

Long-term Sustainability of Research Infrastructures: Science Europe's Offer to Contribute to Ongoing Efforts (2018)

Briefing Paper on Funding Research Data Management and Related Infrastructures (2016)

Strategic Priorities, Funding and Pan-European Co-operation for Research Infrastructures in Europe (2016)

Science Europe, *Practical Guide to the International Alignment of Research Data Management* in 2018 (https://www.scienceeurope.org/media/jezkhnoo/se_rdm_practical_guide_final.pdf).
Vries, Melle de, Aanzet tot datalandschap; Startnotitie Verkenning nationale datalandschap (project E), v. 1.1 (approved by Steering Group 14 October 2019).

Links to websites of selected organisations:

https://www.knaw.nl/en
https://www.knaw.nl/en/institutes/dans
https://www.clariah.nl/en/
https://www.openscience.nl/en
https://www.nwo.nl/en
https://www.go-fair.org
https://www.surf.nl/en
https://userinfo.surfsara.nl
https://www.esciencecenter.nl
https://www.nikhef.nl/en/
https://www.dtls.nl
https://vsnu.nl/en_GB
https://www.zonmw.nl/en/
https://www.lcrdm.nl/en
https://researchdata.nl/en/
https://odissei-data.nl/en/
https://www.bbmri.nl
https://www.nfu.nl/english/about-the-nfu/