Excellent research requires excellent infrastructure

Advisory report on the national digital infrastructure for scientific research

June 2017

Executive Summary

NWO was asked by the Cabinet to provide advice about the resources needed to ensure that the digital infrastructure meets the needs of all research communities, remains at a world-class level, and is capable of keeping up with developments in computer and data science. The 'ICT Subcommittee' was installed by NWO's Permanent Committee for Large-Scale Scientific Infrastructure to provide background and advice about the requirements for a digital infrastructure that meet the ambitions set out in the 2016 National Roadmap Large-Scale Scientific Infrastructure. (See Chapter 1). The ICT Subcommittee was given three tasks, as follows:

1. Provide an investment plan for the national digital infrastructure, within the current budget, including recommendations for how this could be evaluated.
2. Provide advice on what additional resources are needed to ensure that the Dutch national digital infrastructure is at a comparable level to that of other advanced knowledge economies.
3. Advise the Permanent Committee regarding criteria to be used in the evaluation of proposals for the National Roadmap.

In the Netherlands, as in all advanced knowledge societies, digital infrastructure is essential for all scientific and scholarly research, not only for storing and analysing large volumes of data but also for generating data with simulations, for visualisations, and for engaging in real-time collaboration with researchers around the world. The Netherlands enjoys a leading position in the provision of digital infrastructure for research, and for broader societal and economic use. This fundamental infrastructure on which other research infrastructures depend is often taken for granted by researchers and policy makers, even though it is impossible in the 21st century to conduct high-quality research without high-quality and robust digital infrastructure. Given the massive explosion of data in recent years, and the dependence on digital technologies by researchers in all disciplines, the demands on the digital infrastructure and those providing it continue to grow. Other developments on the horizon, such as the European Open Science Cloud and the Internet of Things, will lead to even further infrastructural requirements. (See Chapter 2)

Digital infrastructure includes five elements: Networks, Computing facilities; Data, eScience and Authorisation and authentication (cybersecurity). All of these elements include the support of highly skilled professionals. Three organisations are key to providing digital infrastructure at the Dutch national level: SURF, the Netherlands eScience Center, and DANS. But universities, public research organisations, and discipline-specific collaborations also have their own important facilities, and equipment and software may also be provided by commercial organisations. The annual turnover for SURF alone is approximately 75 million euros, which includes not only the costs for infrastructural innovation (the focus of this report) but also the costs for providing the higher education and the basic research infrastructure.

This report focuses on the resources needed for ensuring that the growing demands of data- and compute-intensive research can be met, that the digital infrastructure of the Netherlands is of an internationally competitive level and therefore contributes to the current high level of scientific research in the Netherlands can at least be maintained, and that Dutch-based researchers continue to have the opportunity to participate in European and global collaborative projects. (See Appendix 4.1)

Task 1. Provide an investment plan for the national digital infrastructure, within the current budget, including recommendations for how this could be evaluated.

Conclusion. The current resources for innovating the national digital infrastructure (21 million euros per annum) are inadequate for meeting both the growing demands from researchers, and for incorporating the continued developments in digital capabilities. If there is inadequate structural support, then all stakeholders involved in providing infrastructure will continue to react on an ad hoc basis in order to meet users' needs. If this situation continues, it will result in sub-optimal infrastructure provision, leading to lower research quality, increased 'brain drain', fragmentation of provision, and risks to data security and privacy. (See Chapter 3.1) The question of evaluation is thus no longer relevant here, and is addressed in Recommendation 2 below.

Task 2. Provide advice on what additional resources are needed to ensure that the Dutch national digital infrastructure is at a comparable level to that of other advanced knowledge economies.

Recommendation 1. Additional structural funding of 27 million euros per annum (for a total of 50 million euros for innovation) is needed to meet the growing needs of Dutch researchers, and to ensure that the digital infrastructure remains at a world-class level. The additional funding includes (i) 10 million euros identified in earlier advice from *ICTRegie* (2008), plus (ii) 17 million euros to meet the growing demands and ambitions of the Dutch research community, as set out in the National Research Agenda, the National Plan for Open Science (OCW, 2017), and the National Roadmap (NWO, 2016). To date, (i) has been financed on an ad hoc basis, but it must become part of the ongoing, structural funding. (See Chapter 3.2)

Recommendation 2. Draw upon the expertise of the national organisations, SURF, the Netherlands eScience Center and DANS, to ensure optimum use is made of both the digital infrastructure and additional investments. Coordinate the existing evaluations of these national organisations to ensure that review of infrastructural provision is adequately addressed, and to ensure that the additional structural funding is invested in such a way as to maintain the integrity and quality of the national digital research infrastructure. (See Chapter 3.2)

Task 3. Advise the Permanent Committee regarding criteria to be used in the evaluation of proposals for the National Roadmap.

Recommendation 3. Researchers must demonstrate awareness of the material and operational costs of digital infrastructures, and where appropriate include additional costs in their proposals for large-scale infrastructure and research projects. Both domain-specific and, where possible, cross-domain collaboration should be encouraged, not only to reduce costs but also to stimulate the emergence of standards and the joint development of software and tools, and to facilitate learning from shared experience. (See Chapter 3.3)

Additional recommendations:

- Strengthen dialogue and coordination between all parties providing and using the national digital infrastructure. (Chapter 2.3)
- Increase the synergy between fundamental computer and data science, and digital infrastructure. (Chapter 3.4)
- Stimulate local support from ICT departments and libraries to assist researchers in understanding their digital needs and options. (Chapter 3.4)
- Contribute to greater digital literacy of all researchers, and stimulate the educational and career opportunities for data scientists. (Chapter 3.4)

Table of contents

1 Introduction

Digital infrastructure is a special part of science: as long as it functions smoothly it remains practically invisible, but if it fails it is enormously disruptive. Confidence in the infrastructure is so high that its presence and correct functioning are often taken for granted, and only noticed when it *fails*. Due to the exponential growth in the use of and dependence on research software, high-speed networks and (big) data in all academic disciplines, demands on the current digital infrastructure have grown massively and rapidly. Developments in the field of cloud computing, open science and open data, security and privacy also impose ever-higher demands on the national digital infrastructure. Permanent, structural investment in a reliable, sustainable and broadly accessible high-quality infrastructure and infrastructure support is therefore essential. Without a solid infrastructure of networks, computer facilities, data services, research software and support mechanisms, the scientists of today and tomorrow will hardly be able to continue their work. The recent Deloitte report, *Dutch Digital Infrastructure 2016: Enabling the Digital Economy and Society* has furthermore revealed how the interaction between an excellent digital infrastructure and a broad spectrum of commercial activities has blossomed over the past few years. Any shortfall in investments in the infrastructure will therefore have harmful consequences for both the Dutch economy and society as a whole. Other countries are moving forward rapidly, and as a result we are at risk of losing talented researchers, as well as failing to attract new researchers in the future.

Against the background of this rapid and disruptive process of digitalisation and increasing competition, the Advisory Council for Science, Technology and Innovation (Dutch abbreviation: AWTI) advised the government in 2015 in its report *Klaar voor de Toekomst* (Ready for the Future) to increase investment in a world-class digital infrastructure for excellent research in the Netherlands. In response, the government called upon NWO to map out the needs of scientists in terms of the digital infrastructure. The question of whether sufficient resources are available to maintain our position in the international lead group would also need to be considered. In response, the ICT Subcommittee (see annex 4.7 for its composition) was appointed by the Permanent Committee for Large-Scale Scientific Infrastructure (Dutch abbreviation: PC-GWI). The ICT Subcommittee was called upon to 'provide advice on the optimum deployment of the available resources for the ICT infrastructure' (hereafter 'digital infrastructure', and in other countries sometimes referred to as 'e-infrastructure' or 'cyber-infrastructure'). Considering the harmonisation between investments by SURF in the digital infrastructure and those by the NWO in large-scale research facilities was highlighted.

The PC-GWI assigned the ICT Subcommittee the following three tasks:

1. Provide an investment plan for the national digital infrastructure, within the current budget, including recommendations for how this could be evaluated.
2. Provide advice on what additional resources are needed to ensure that the Dutch national digital infrastructure is at a comparable level to that of other advanced knowledge economies.
3. Advise the Permanent Committee regarding criteria to be used in the evaluation of proposals for the National Roadmap.

Researchers' own ambitions, the increasing digitalisation of society and the economy all call for digital infrastructures that require investments far exceeding the financial capacity of single knowledge institutions or of individual users. In the current situation, every Dutch researcher requires immediate access to a wide range of reliable services for communication and collaboration, and to process, analyse, store and share data. A national infrastructure makes all these activities possible by providing an integrated array of facilities that can be offered at lower costs as a consequence of cooperation and economies of scale. By combining needs via the national digital infrastructures, specific advanced services in the Netherlands can be financed efficiently and in a timely manner. Government support is needed in order to maintain the solidarity model, and to benefit all researchers. At an early stage, major investments tend to benefit a relatively small group of advanced users, and more widespread adoption across the disciplines only takes place at a later stage. The relatively small group of pioneers is not able to cover the costs of new and inherently expensive infrastructure individually. However, delaying investments until all research organisations are is willing to invest would quickly place the Netherlands at a disadvantage.

On 13 December 2016, the PC-GWI presented the National Roadmap for Large-Scale Scientific Infrastructure. This Roadmap provides a strategic framework for large-scale research infrastructure for the next five years. Digital infrastructure is playing an increasingly important role in research, and as such provides a foundation for leading research. It is essential that Dutch researchers have access to an advanced digital infrastructure. Ensuring that access requires follow-up investments in a national digital infrastructure that also enjoys a high-level international reputation. Establishing such an infrastructure will also help to meet and fulfil the ambitions of the National Plan for Open Science (OCW, 2107) and the National Research Agenda, the top sectors and the National Roadmap. In the Netherlands, SURF, DANS and the Netherlands eScience Center are the organisations that have been given a national task in respect of the digital infrastructure, and operate alongside more local initiatives such as the 4TU.Centre for Research Data. SURF has a coordinating role for the national digital infrastructure. Local and domain-specific facilities are generally financed locally or on a project basis. As a consequence, they are typically neither nationally accessible, nor suitable for multiple scientific domains. It is essential that in the future, harmonisation and collaboration between these initiatives and the national initiatives be intensified.

2 Science and the digital infrastructure

2.1 Trends in science

The past few years have seen an exponential rise in the volume of data produced for research in practically every scientific discipline. This rise is caused by the possibilities offered by new technologies, for example those employed within large-scale research facilities, advanced software, more powerful computing facilities and sensor networks. In addition, ever-growing volumes of data are available for research that does not relate directly to large-scale research facilities. Examples include the large volumes of text and audio-visual material from social media and online news sites used by political scientists and media researchers. In healthcare research, increasing volumes of data are generated by wearable consumer devices, including smartphones and other portable smart devices.

The majority of large-scale research facilities on the National Roadmap Large-Scale Scientific Infrastructure (2016) involve substantial ICT and data components. The facilities and collaborative ventures between scientists are constantly seeking the boundaries of technical capabilities, and can only be implemented if high-quality digital infrastructures are available. Projects of this type effectively illustrate the emerging developments in practically every domain, in which similar developments in terms of explosive data growth are undeniably taking place. The shift towards more (international) cooperation and the working methods employed by new generations of researchers demand infrastructure and related services for all researchers.

Dependent on the digital infrastructure

MinE is an international project for tracing the genetic causes of ALS, a deadly muscle and nerve disease. For this research, material is required from at least 15,000 patients and 7,500 control subjects. Following the sequencing (reading out) of a single DNA sample, a file is generated of between 75 and 100 gigabytes. If you multiply that by 22,500, the outcome is some 2 million gigabytes of data. That in turn equates to a stack of hard drives as high as the Cathedral Tower in Utrecht.

Physicists from all around the world work together at CERN. Using large detectors, they study tens of millions of collisions between protons – every second – and the particles that are generated as a result. A worldwide network of computers (in a grid) filters, records and processes the data from these collisions. Each year, millions of gigabytes of data are generated.

Astronomers are currently elaborating the building plans for the Square Kilometre Array (SKA). It is now already clear that this, the largest telescope in the world, will also lead to enormous data challenges. Imagine the total volume of data on the Internet right across the planet. That same volume of data, a billion gigabytes a day, will be continuously received, as soon as the SKA is activated.

The growing need for digital infrastructure across the full academic spectrum is reflected in the ePLAN study *Infrastructuur Duurzaam op Maat* (Sustainable, Tailor-made Infrastructure) report, for example. This report, based on a series of workshops and a questionnaire sent out to more than 1000 researchers from all academic disciplines, reveals that there are serious concerns about the lack of investment in the Dutch digital infrastructure. Above all, researchers expect a huge increase in demand for *data & cloud services*, more support and assistance in their daily data-intensive research work.

This heavy dependency between scientific developments and digital infrastructure has led the European Commission to deliver an additional boost to this ecosystem. With its plan for a European Open Science Cloud, the Commission eventually hopes to realise a FAIR research infrastructure for data and services (FAIR - Findable, Accessible, Interoperable, Reusable). One essential requirement for establishing such an infrastructure is effective dialogue between scientific domains and providers of digital infrastructure, services and eScience.

Open Science is an approach to the scientific process based on more intensive cooperation and knowledge sharing through the use of digital technologies and new cooperation tools. As a result, research data, methods and publications will become transparent, reusable and more accessible for scientists, business and industry and society.

In February 2017, the Ministry of Education, Culture and Science presented the ambitions of the Netherlands in the National Plan Open Science. This Plan concludes that support is required in every phase of the research process, from infrastructure services to training, from hands-on support to the development of research software. The plan also supports the ambition for a consistent system to ensure FAIR access to research data. To support these developments, the GO-FAIR initiative is focusing on the development of powerful national and domain-specific *open science cloud nodes*, the goal of which is to bring about a cultural sea change through training and the implementation of FAIR data and FAIR services. To enable data-intensive research and to ensure the success of such initiatives as the European Open Science Cloud, the National Plan Open Science and GO-FAIR, it will be essential to considerably expand and upgrade the current infrastructure. In these and other future major scientific projects, the technological challenges facing scientists are set to become even more complex and more extensive. In addition, a growing number of individual scientists from a whole raft of scientific disciplines will experience an ever-growing need for high-quality hands-on support in data analysis and software applications.

Cloud computing

The development of cloud computing has led to numerous new digital services from commercial providers, but has not eradicated the need for national digital infrastructures and services. After all, researchers constantly seek out the boundaries of technical capacities, thereby creating a need for service, support and expertise that is not (yet) offered by the market, or at least not with the right conditions.

This applies to the high-end demand for High Performance Computing (HPC), for example, for which scientists clearly have a greater need for capacity than what is currently offered by the market. HPC cloud computing is a development which, in some sense, is still in its infancy. If we consider the TOP500, the upper rankings are occupied by national supercomputers, the supercomputers of meteorological and other institutes, including research institutes, and some government computers. The only companies high in the TOP500 are not providers of HPC cloud computing but users who need their own supercomputer for their own business processes (e.g. Total, Petroleum Geo-services).

There are different forms of cloud: public cloud, private cloud, community cloud and hybrid cloud. In cases where the generic functionality of the public cloud (equal for all customers) is sufficient, on behalf of its members, SURF enters into contracts with commercial providers. In those contracts, in addition to the price paid for services, agreements on privacy, security, user authentication and intellectual property are certainly of equal importance. SURF shares these experiences on a European level, since economies of scale are useful in improving the services on offer and ensuring better dialogue with the cloud suppliers. This has recently led to deals with Amazon and Microsoft.

In situations where privacy and confidentiality are very important, one option is a private cloud. Using a private cloud, the user has complete control over data, security and service quality. The applications are made available within a private, virtual organisation. Governments, businesses, institutions and large-scale European research infrastructure have their own private cloud, which may be located locally or provided by a cloud supplier.

The community cloud combines the confidentiality of a private cloud with the size advantages of a public cloud. Users from multiple organisations work on the same infrastructure. This is only possible if they trust one another sufficiently and impose similar requirements. At the request of its members, SURF is developing a variety of community cloud services for research and education, to serve the specific needs of users. In that process, SURF offers guarantees about the location of the data, confidentiality and privacy. This is important in the case of personal details or confidential medical research data, where there is a demand for guarantees that the data is stored in a particular country.

A relatively new development is the hybrid cloud, which offers combined use of various cloud types. Together with a number of early adopters, SURF has developed a new service with this in mind, in which SURF acts as a broker and manager on behalf of the user institutions. All the institutions need to do is to monitor and manage their use, thereby freeing up more time for ICT departments to concentrate on their primary tasks: supporting education and research.

2.2 Infrastructure as an essential requirement

A broadly accessible, high-quality national digital infrastructure is an essential requirement for efficiently and effectively facilitating competitive science. The urgency of this need was recently emphasised by the AWTI. In the report *Houd de basis gezond (Maintain a healthy basis)*, the digital infrastructure is identified as one of the five priorities for additional investments in research and innovation. The AWTI has warned that the Netherlands is at risk of lagging behind in competitiveness and earning capacity, and in the worst case, falling behind the international average.

Concerning the role of the digital infrastructure, the AWTI argues that:

> "The ICT research infrastructure forms the backbone of any 21st century knowledge and innovation system. Practically all meaningful scientific and technological developments are dependent on ICT. Communication and cooperation, knowledge development (computing power, advanced algorithms and big data) and scientific progress are more than ever dependent on a high-quality ICT infrastructure."

By increasing investment, the essential capacity expansion can be achieved, the growing volume of research data can be well exploited, and the necessary hardware and software support can be provided to allow these tools to be used for research.

The increased size and central nature of the role attributed to such themes as digitalisation, big data and ICT are also reflected in a number of other recent publications. Among others, VSNU (*Digitale Samenleving/Digital Society*), VNO-NCW (*Investeren in een digitale kwantumsprong/Investing in a digital quantum leap*) and Nederland-ICT (*Groei door Digitalisering/Growth through digitalsation*) have all called for more focus on these themes. Digitalisation is the overarching development that is capable of helping every sector to move forward, and in that process, high-quality digital infrastructures are an indispensable basic element. In a joint statement issued in the spring of 2016, the ICT Ministers of the G7 once again emphasised the importance of digital infrastructures for research and education. The National Research Agenda (Dutch abbreviation - NWA), and in particular the Big Data Route, provide further evidence that the importance of data (including big data), and as a consequence the importance of digital infrastructures and specialist support, are constantly growing.

The majority of research facilities on the NWO and KNAW National Roadmaps (both published in 2016) call for high-quality technological skills and support. These facilities are often embedded in European joint ventures such as ESFRI (European Strategy Forum on Research Infrastructures). International collaboration can only be implemented if a digital infrastructure is available.

The European Open Science Cloud mentioned earlier and the role of the Dutch stakeholders in this initiative call for good coordination and further strengthening of the digital infrastructure. Although the term Open Science Cloud is still very much in a state of development, a number of elements are now clearly defined in respect of the needs of scientists: extensive ease of use, integration, single sign-on, reproducibility of analyses and methods, transparency of data locations, training, suitable services and experts capable of supporting the researchers. Against that background, the success of the cloud will be determined by the robustness of the system, the obstacles perceived by researchers and, of course, the volume and quality of the resources that make up the cloud. Realising the vision of the European Commission, and indeed of the Netherlands on a national level, will require a series of expansions and improvements to the existing digital infrastructure, including harmonisation and consultation with users. The European Commission has issued an indication of the necessary investments, to the tune of 6.7 billion euros, the majority of which will have to be provided by the Member States. The European Open Science Cloud will, for example, have to be supported by a European Data Infrastructure of federated national networks and computing and storage facilities.

The Netherlands currently occupies an excellent starting position to play an active role in these international developments: the public digital infrastructures are well organised, with heavily engaged users. In that sense, the Netherlands acts as a role model for many countries. Nonetheless, it remains essential to make a major step forward in terms of innovation and coordination, in order to meet the growing demand for services to scientists.

2.3 The national digital infrastructure

The national digital infrastructure consists of the facilities and services that are in principle available to all scientists, irrespective of the scientific domain in which they work. As such, they form the foundations for the already referred to research activities and for the implementation of the National Research Agenda (NWA) and the Roadmap. These foundations need to be more than sufficient to support scientific efforts across the board. No European funds are available for national digital infrastructures. Each country is required to provide for its own national networks, computing facilities, data storage and support. In addition, to permit participation in international infrastructure projects, countries must contribute their own facilities and financial resources, including network connections or a national supercomputer.

The Subcommittee has examined digital infrastructures and services that play a national and/or international role. The figure below represents the five different elements of the digital infrastructure. The organisations that offer non-domain-specific digital infrastructures at a national level are SURF, DANS and the Netherlands eScience Center. While SURF is active in every element of the infrastructure, the eScience Center focuses mainly on research software and advanced research support, while DANS offers services in the field of data storage and data archiving. Annex 4.1 offers a more detailed description of the various organisations. Under the auspices of SURF, RUG-CIT and Nikhef offer services on a national level, via a federated model. The RUG-CIT, Nikhef and SURF balance investments and the deployment and availability of grid and cloud services for their scientific user communities. Chapter 3 provides more substantive information about the digital infrastructure.

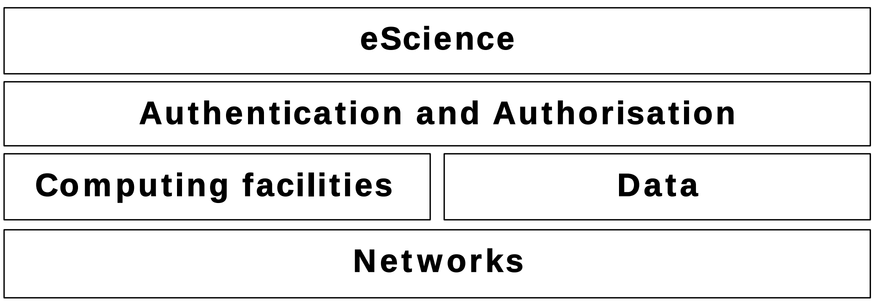

Figure 1: Elements of the digital infrastructure.

Alongside these national facilities, there are numerous local, regional and domain-specific infrastructures and services, such as campus networks, data centres (e.g. 4TU.Centre for Research Data) and research support departments, as well as domain-specific infrastructures. These are not considered in this advisory document, because they fulfil a local or domain-specific need, or are financed locally. This also applies to the KNMI, MARIN and NLR, which offer considerable computing facilities in-house, but do not provide national services to other researchers. These facilities are often also financed with public funds. The recommendations provided here focus on these facilities in respect of overall management for guaranteeing the optimum use of those resources.

For that same reason, the data infrastructures of Roadmap projects such as BBMRI, ELIXIR-NL and ODISSEI are not considered in this advisory document: these infrastructures are organised on a domain-specific basis, and as such do not serve as part of the foundation for all areas of science. At the same time, in the call for proposals for the Roadmap, the effective use of existing national digital infrastructures is urged.

This advisory document calls for more intensive dialogue between the future research facilities and the national digital infrastructure, with a view to increasing the effective use of that infrastructure. Similar to the European situation in ESFRI, this document proposes that the future research facilities on the National Roadmap should describe and budget for the necessary use of digital infrastructure facilities, as well as indicating how various aspects can be implemented, in the development and implementation of the facilities. In that connection, it is essential to achieve a balance between local support and domain-specific customisation, and the efficiency and effectiveness of national and international facilities.

Organisation of the national digital infrastructures in the Netherlands

The national digital infrastructures are to a large extent organised within the SURF cooperative, in which universities, universities of applied sciences, university medical centres, research institutions and growing numbers of senior secondary vocational education (Dutch abbreviation - MBO) institutions work together on an administrative, policy-specific and operational level, towards ICT innovation and digital infrastructures. In 2009, the government embraced the recommendations from *ICTRegie* issued in 2008, to place all digital infrastructures for scientific research under the auspices of SURF, thereby placing direct responsibility for strategy and innovation in the hands of the users. This led to a merger between SARA and SURF, and the absorption of the activities of NCF in SURF.

The Netherlands eScience Center was established in 2011 by SURF and NWO. It bridges the gap between the ambitions of scientists and the increasingly complex digital infrastructure and ICT, including knowledge and skills from computer and data science. The eScience Center supports researchers by developing high-quality research software, for example, in close collaboration with the research groups involved.

DANS is an initiative of KNAW and NWO. It is the national data archive and centre of expertise for access to research data and scientific information. DANS promotes permanent access to digital research data and encourages researchers to archive and reuse data in a sustainable manner. DANS also grants access to thousands of scientific datasets, publications and other research information in the Netherlands, and is responsible for research into sustainable access to digital information and software sustainability.

3 Recommendations

3.1 Investment plan within the current budget

Provide a sustainable investment plan for the scientific digital infrastructure within the current budget, for carrying out scientific research for the next five years, including recommendations for how this could be evaluated.

Description of the current situation

The current structural budget to the tune of 21 million euros for the national digital infrastructure is entrusted to SURF and the Netherlands eScience Center. There is also separate financing from NWO and KNAW, for DANS. Figure 1 outlines the elements of the digital infrastructure for which these organisations are responsible.

The two most capital-intensive elements of the digital infrastructure are the network and the (*high-performance*) computing infrastructure. The underlying principle in all choices on investments and the continued development of services is that on the basis of market developments, SURF judges which knowledge, products and services will be developed in-house, and which should be tendered to the market.

- On average, the investments required in a network amount to at least 10m euros per annum (Figure 2). This conclusion is based on SURF's extensive experience with advanced networks. This amount includes renewal of the knowledge base, the testing of new concepts, the implementation of pilots with advanced users and the establishment of reserves for future purchase. These costs are estimated at approximately 6 million euros (of the total of 10 million) per annum. Migration, operating costs and ad-hoc expansions are financed on the basis of charges applied to the affiliated education and research institutes.

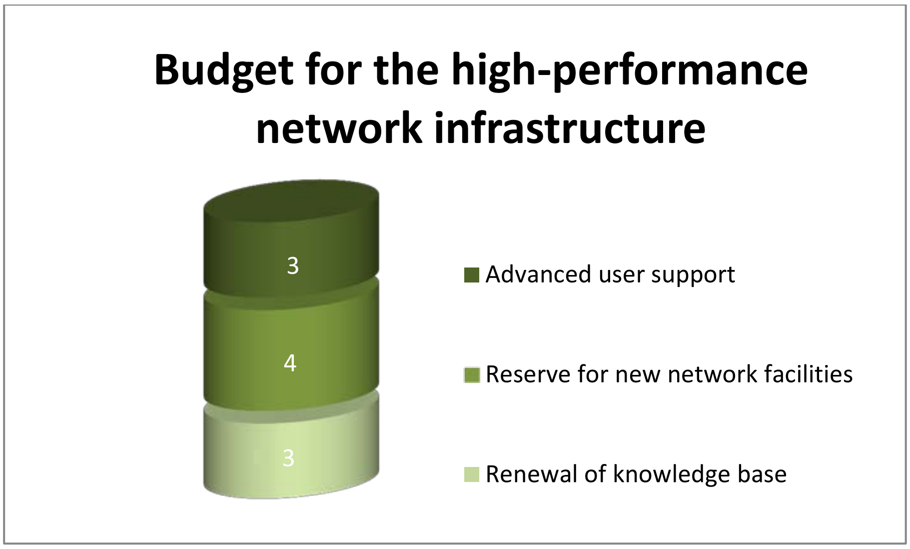

Figure 2: Budget for the high-performance network infrastructure (Amounts in €m)

- Including the requests for user support, the investments required for the current *high-performance* computing infrastructure amount to approximately 18m euros per annum (see also Chapter 4.1.2 for an overview of the various computing facilities and their types of users). This amount includes renewal of the knowledge base, the testing of new concepts (such as *deep learning*), the implementation of pilots and the establishment of reserves for the purchase of a supercomputer comparable to existing and other computing facilities (see Chapter 4.1.2), and the operating costs for keeping the system up and running (including power supply and accommodation). In practice, there is huge demand for advanced and community support. This growing demand for support, in particular for computing and data facilities, means that the originally calculated budget for support is exceeded by 4m euros per annum. Examples of intensive user support are LOFAR, TROPOMI, MinE, BBMRI and Virgo-LIGO, whereby the majority of the budget requirements are charged to SURF. On the other hand, support for users with less advanced needs but who still require HPC facilities is growing exponentially, and already exceeds the amount originally budgeted. As a consequence, SURF is currently not able to set aside sufficient reserves for a new supercomputer. Figure 3 provides an overview of the current costs for high performance computing.

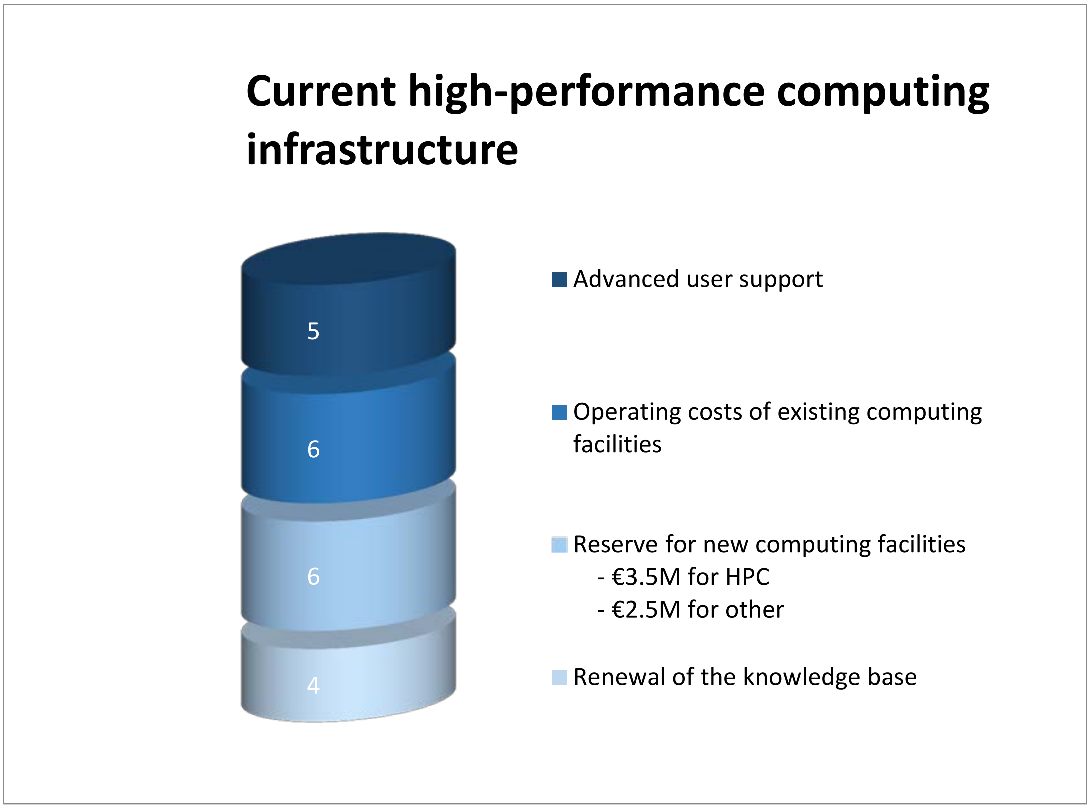

Figure 3: Current high performance computing infrastructure (Amounts in €m)

As well as these two facilities, investments are also being made in knowledge and know-how in the field of cybersecurity, including secure and user-friendly methods for accessing services. The importance of a secure (learning and) working environment, to which access can be gained in a secure (*two-factor authentication*) and low-threshold (*single sign-on*) manner has grown. The

complexity of *virtual research environments* is also set to increase, and with it the importance of a high-level AAI (authentication and authorisation infrastructure – see also Chapter 4.1.4).
Added to this, over the past few years, we have noted the arrival of advanced data storage, *curation* and *analytics*. Investments are currently at around the 3 million euros mark, but demand is clearly growing.

Over and above these capital-intensive facilities, 2011 saw the establishment of the Netherlands eScience Center, which focuses on the element of eScience in Figure 1. For this purpose, roughly 6m euros have been reserved from structural funds. The total budget for financing the current digital infrastructure is 37(18+10+3+6) million euros (Figure 4).

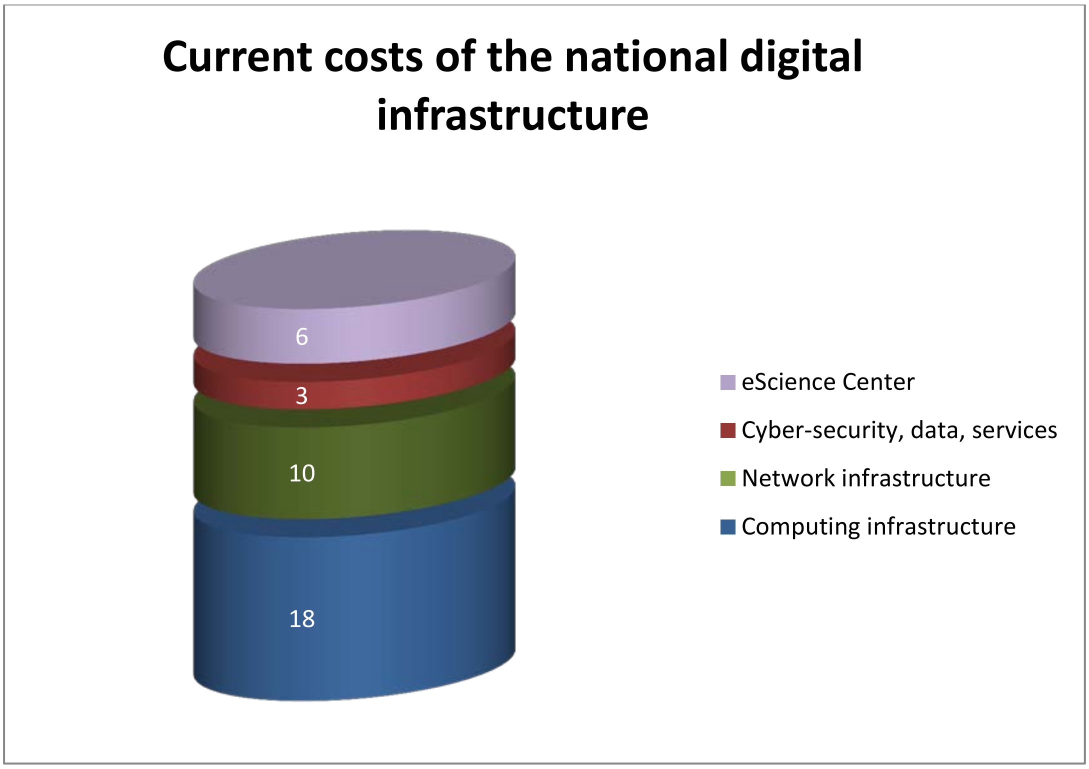

Figure 4: Current costs of the national digital infrastructure (Amounts in € million)

Figure 5 shows the breakdown of the current budget. As described above, there is a shortfall of 4million euros per annum, for covering current costs.

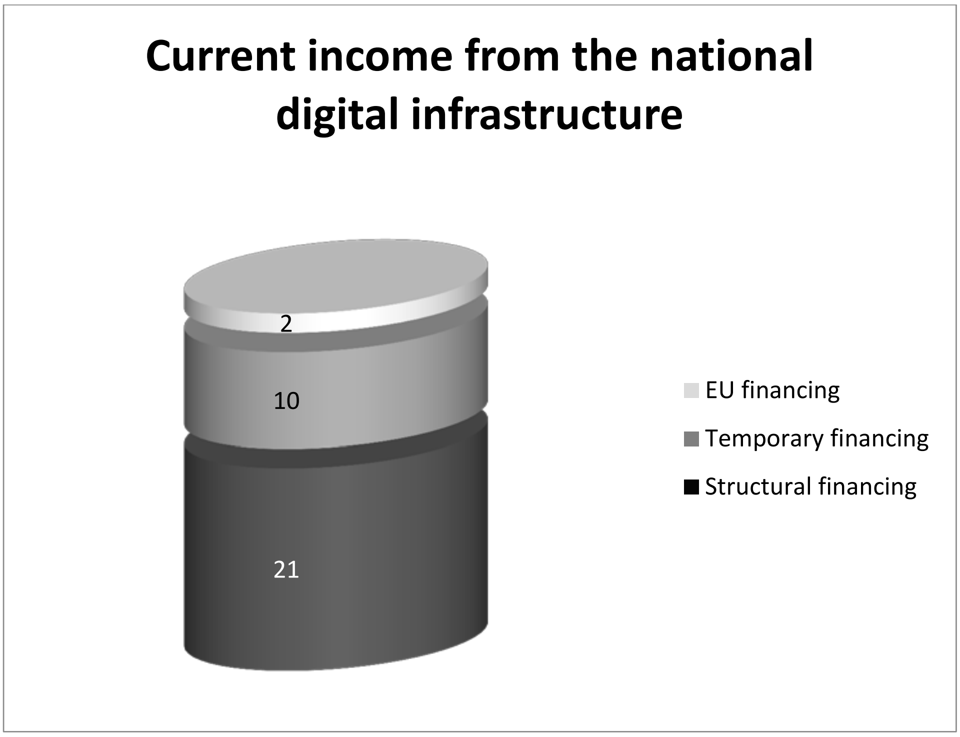

Figure 5: Current income national digital infrastructure (Amounts in € million)

In developing a long-term investment plan, SURF and the eScience Center are confronted with (i) the fact that it is currently unusual for researchers to budget and pay for advanced digital services (and eScience), and (ii) the growing demand for support. These factors have led to pressure on and indeed a shortfall in the reserves for the computing and data infrastructure, and a low success rate (10 percent) at the eScience Center. By means of this advisory document – which runs parallel to the international recommendations issued in the framework of the ESFRI roadmap, for example – the Subcommittee wishes to address this first challenge (see i) in respect of large-scale scientific infrastructures. At present, SURF supports the various existing infrastructures from its own funds.

The financing of the digital infrastructure in the Netherlands is insufficient to cover current costs. In addition, one third of the available budget is realised via ad-hoc, temporary financing. The financing of DANS itself is under pressure due to the growing demand for services in data and archiving. There is only very limited capacity for matching European subsidies.

The Subcommittee therefore concludes that providing a robust and sustainable digital infrastructure for competitive science research with international ambitions is not possible within the current budget.

Consequences of failing to increase the budgetary capacity

The explosive growth and exponential rise in the complexity of and dependency on data (including big data), analytics and research software in all scientific disciplines call for substantial interventions in the national digital infrastructure. The current level of investment is insufficient to enable the Netherlands to keep pace with the rapid technological developments of today and tomorrow. If the Netherlands wishes to remain a leading international scientific player, major investments in innovation, capacity, support, security and reliability are unavoidable (see Chapter 3.2).

Over the past few years, we have seen a huge rise in the demand for data services, research software and high-quality scientific support. This applies all the more to scientific domains that have recently become heavily digitalised (e.g. through the availability of digital sources and advanced analysis software) but which have no fundamental tradition of ICT applications within their field. The significant demand for a national infrastructure with solid support mechanisms has been identified in all domains of scientific and scholarly research, as revealed in the 2016 report based on the national ePLAN survey among more than 1000 Dutch researchers, for example. Furthermore, the past few years have seen major developments at the interface between ICT and science, which at present are insufficiently supported or at least not to the full extent, by the current scientific digital infrastructure, such as virtualisation, the Internet of Things, the cloud, authentication and authorisation, cybersecurity, data management, software sustainability and open science.

Perhaps the most important development is the massive rise in the volume of data in science – which was not in fact foreseen, even just a few years ago. As the growth of data has outpaced the development of the capabilities of storing that data cost-efficiently, we are clearly facing major challenges. In most cases, the data in question is big data: data which by nature is complex, fast and vast, and which cannot be employed for scientific research without thorough filtering and analysis. For global climate data alone, even very conservative estimates suggest that the volume of data will be multiplied by at least seven times, in the period 2015-2030. In a period of less than five years, the volume of data from the LOFAR radio telescope has grown by a factor of 25: from less than one petabyte to 25 petabytes. On the basis of estimates by research scientists, Compute Canada (2016) has calculated that over the next five years, demand for computing power will rise by a factor of seven, and that there will be a need for fifteen times the current storage capacity (see Annex 4.4).

The Subcommittee recommends freeing up additional resources, since there are no other sources of financing available. Since the phasing out of the Economic Structure Enhancing Fund (Dutch abbreviation - FES), there are no more calls for short-term or temporary funding, and European financing for a national digital infrastructure is not possible. Because it would be undesirable for a domain-independent digital infrastructure for *all* scientists to be in competition with the projects of users, the national digital infrastructure has not been included in the Roadmap 2016.

Although the capacity of ICT hardware (per euro) is rising, this in no way means that the digital infrastructure can manage with less funding. In the same way that each new generation of laptops or smartphones is not cheaper, but does offer more capacity and more capabilities per euro, so the costs of high-end infrastructure are also not set to fall.

And although capacity is rising, the number of users is rising at an even faster pace; nor are the costs for short- and long-term intensive support and software development falling.
Maintaining the current level of investments will not only have far-reaching consequences for individual scientists, but will also result in a number of financial and substantive risks for science as a whole. These are summarised below.

Loss of quality

The quality of much scientific research depends to a great extent on the quality and capacity of the digital infrastructure. Insufficient investment at a national level means network speeds and computing capacity are outperformed by foreign parties. This effectively represents an implicit choice no longer to provide support to science at an internationally competitive level. It would, for example, mean that the Netherlands would be forced to surrender its position as an international leader in data-intensive fields of science such as physics and astronomy.

Level of investment in Finland, Switzerland and Sweden

If we look at the levels of investment in other European Member States, the threat of the Netherlands lagging behind is once again confirmed. In Finland, for example, almost twice as much is invested (adjusted for the size of the population) in CSC, responsible for the Finnish digital infrastructure. The most notable feature of investments in Finland is that in relative terms, much higher levels of investment are made in services and support for computing facilities and computational science. The gap with Switzerland is also large. Investments in high-performance computing alone (no 8. In the world, and highest European ranked) are around 8-10 times higher than in the Netherlands and Finland. Spread over a period of 4 years, government contributions in the field of computing alone amount to around 43 million euros per annum. It is also useful to consider the investment plans in Sweden, where over the next few years, around 20 million euros per annum will be spent on high-performance computing. This amount is almost equal to the recommended 21 million euros per annum for high-performance computing in the Netherlands. At present, the budget in the Netherlands is 16 million euros, of which, it must be remembered, 5 million is sourced from temporary funds.

Isolation

Participation in major international projects and infrastructure will no longer be possible. Due to the insufficient investments in networks and computing facilities, the matching and participation requirements for European projects and *Research Infrastructures* can no longer be satisfied. As a result, for example, the Netherlands would no longer be able to gain access to Europe’s fastest supercomputer within PRACE, the international Research Infrastructure for Supercomputing. Two examples of the current use of PRACE facilities are simulation of the interaction between proteins and new medicines, and simulations of extreme weather conditions in the future climate. There will also be insufficient capacity to offer the combined and basic support infrastructure to scientists and universities wishing to participate in European and national developments in the field of Open Science. In more general terms, it will become clear that international collaborative scientific ventures could be put at risk due to the lack of sufficient added value from the Netherlands.

Quality loss & isolation in practice

Consequences of insufficient investments for ...

Climate research

The Netherlands is susceptible to climate change. Above all, the threat of too much and too little water is of clear relevance. To be able to issue statements on climate development, complex computer models are commonly used. It is essential to use those models to represent details of the atmosphere, oceans and planetary ice. Among other things, climate development is dependent on cloud formations and ocean currents. The first exascale computers are due to be operational in around 2023. These computers are up to 1000 times more powerful than the next generation of the national supercomputer. They will make it possible to simulate the climate in more detail. The American Department of Energy is already taking the lead with models that simulate clouds in detail, and in Europe too investments are being made. All these models generate huge volumes of data, in addition to the data from satellites, ground observations and many other sources.

Analysing all that data is a major challenge which will result in new infrastructures that integrate data storage and the computers used for the relevant calculations. STFC has already invested in an innovative infrastructure of this kind, in the United Kingdom. There is already a supercomputer at the UK Met Office that is used exclusively for weather forecasting and climate research, which is ten times faster than the Cartesius operated by SURFsara. In Germany, (at the DKRZ in Hamburg), there is a supercomputer (Mistral) used exclusively for climate research, that is a factor of four times larger/faster than the Cartesius. Current investments in the Netherlands are lagging behind those of our major neighbours and the US. As a direct consequence, Dutch academic parties no longer have access to the best infrastructure, while participation in competitive international research consortia is coming under threat.[1]

Consequences of insufficient investments for ...

Astronomy

Astronomy is the branch of science that answers the most fundamental human questions: “where do we come from?” and “are we alone?”. In *absolute* figures, astronomy in the Netherlands currently performs in the world’s top three. Through national cooperation, therefore, the Dutch astronomers are punching well above their class on a global scale. To make this possible, a complete chain of telescope development, data analysis and large-scale simulation is essential.

The world’s largest telescopes, costing more than 1 billion euros, are designed and built in European and global collaborative projects. So long as the Netherlands maintains a leading scientific position, this generates major orders for Dutch industry. Such a leadership position is dependent on technological innovation and performance given that image processing for modern telescopes is extremely demanding. For example, in 2022, the Square Kilometre Array (SKA) will only be able to *see* after the aerials are combined by an Exaflops supercomputer. Missions and telescopes such as Euclid, JWST, SKA and the ELT will then generate on average 100 PB of new data every year, for Dutch astronomers, who will then use hundreds of Petaflops of capacity computing, to analyse the data thus generated.

---

[1] https://www.hpcwire.com/2016/09/07/exascale-computing-project-awards-39-8m-22-projects/
http://www.metoffice.gov.uk/news/releases/archive/2014/new-hpc
https://www.primavera-h2020.eu/
http://cordis.europa.eu/project/rcn/197542_en.html
http://www.jasmin.ac.uk/

Our understanding of the universe from which we emerged will only be truly complete if our best simulations eventually reproduce and explain the observations made. Only if Dutch scientists have access to exascale capability computing will they maintain their lead over their international competitors. Failure to invest structurally in the digital infrastructure could represent a direct threat to the leading global position of Dutch astronomy.

Consequences of insufficient investments for ...

Life sciences

In an international context, the life sciences are focused on the development of knowledge that directly influences daily life, from healthcare to nutrition, and from agriculture to biodiversity. Dutch life scientists are amongst the world leaders, but their advances are increasingly disrupted by ICT limitations. Within the field there is therefore an urgent need for a higher-quality ICT infrastructure. Further fragmentation of investments in the e-infrastructure could prove utterly disastrous for the life sciences.

Within the life sciences, the volume and variety of datasets are growing exponentially. Researchers are already producing exabytes of imaging and omics data (in particular in the fields of genomics and metabolomics) every year, and the overall volume of data is doubling every six months. Furthermore, these data are generated simultaneously at a whole raft of different locations. As a consequence, globally there is huge demand for high-quality capacity for data storage, data exchange and data analysis.

One additional challenge facing human life sciences research is the privacy-sensitive character of much of the data. This fact calls for high-performance, distributed analytics (otherwise known as 'distributed learning'). Within this process, sub analyses are first carried out locally, in a large number of distributed data sources, prior to combination at meta level, to allow conclusions to be drawn. What is needed is an infrastructure made up of large, central computing and storage capacity (at present provided by SURF) and a heavily interconnected network with local HPC capacity that, to an increasing extent, will be managed by the larger research institutes, backed up by the knowledge and skills needed to deploy the data for the life sciences (as currently carried out by eScience). An approach of this kind will require major adjustments in terms of software and security, areas in which the Netherlands is currently a ground breaker.

Elsewhere, with ever-growing intensity, the life sciences are using data from other domains such as earth sciences, social sciences and nutrition. It is of key importance that data systems from such divergent domains be able to communicate by concentrating on international data principles that ensure that data, standards and systems remain Findable, Accessible, Interoperable and Reusable (FAIR) for both people and machines. This in turn calls for cooperation between all parties in the Dutch research field, close coordination, and harmonisation between local and national ICT investments. Only in this way will the Netherlands be able to continue to play a leading role in future life science research.

Consequences of insufficient investments for...

Social sciences & the humanities

Thanks to the emergence of the data-driven society, both the sources and methods for research have undergone fundamental change over the past ten years. In the field of humanities, much cultural heritage material has been digitised, while huge volumes of 'born digital' material are made available on a daily basis. The social sciences have new methods for organising surveys (through the use of sensors) and have access to a wealth of data generated via social media.

Roadmap projects such as CLARIAH (PLUS) and ODISSEI promote the spread of innovation in these domains, while making use of and simultaneously expanding the existing infrastructure. Investments are needed in two adjacent areas in order to provide a constantly growing and structurally financed infrastructure (the basis) and to guarantee capacity, to secure the sustainable continuation of temporary and project-financed services (tools, data).

Investments in infrastructure are necessary because big data research itself is demanding ever-greater computing capacity, a trend that is set to continue. Many social science and humanities research projects today are dependent on HPC computing. Recent examples are the SHICO project that makes it possible to monitor cultural concepts over time, and Golden Agents, a project that helps historians to understand the origins of the innovative capacity in the Golden Age, by combining a variety of different sources.

A lack of structural investment in the digital infrastructure will certainly hinder these projects or even lead to their results not being secured. That in turn will prevent further work being carried out, on the basis of historical data. In that case, Dutch humanities and social sciences researchers will lose ground in comparison with their foreign competitors.

Brain drain

The attractiveness of the Netherlands as a location for researchers, technology start-ups and ground-breaking ICT businesses will rapidly disappear if investments continue to fall behind. Data-intensive scientific researchers, data scientists and innovative knowledge will disappear to the countries around us, where substantially more is being invested in digital infrastructures for research and education.

Fragmentation

The absence of national investments will lead to individual research groups building their own solutions, which as a consequence will be fragmented, non-interoperable, less cost-efficient and far less powerful. The absence of national supervision and coordination may result in the build-up of unused capacity at various locations.

Privacy and security

Insufficient investments in reliable and secure networks and data facilities will bring about a fall in confidence in data security and the privacy protection of personal details, for example in the field of medical research. Because there is insufficient trust among users that commercial cloud computing and storage are able to offer sufficient guarantees of the security of research data, the absence of a public national infrastructure could have a significant adverse effect on scientific innovation.
This will affect research particularly in the medical and social sciences which involve the processing of huge volumes of personal details.

ICT innovation

By not allowing investments in the digital infrastructure to continue to grow, or by even failing to purchase new digital infrastructure such as HPC capacity, the level of ICT research and the number of ICT researchers (including many data scientists) will inevitably fall. Excellent infrastructures work like magnets for top talent (including top talent from abroad). The disappearance of human capital in the Netherlands in this rapidly growing field could very quickly result in a shortfall in terms of the availability of knowledge and talent, which could be extremely difficult to make up in the future.

The question of an evaluation mechanism is addressed in Chapter 3.2.

3.2 Digital infrastructure amongst the international leaders

Produce advice on the necessary additions to the first task for raising/maintaining the Dutch digital infrastructure at a level comparable to that of other leading knowledge economies

For the Netherlands as a leading knowledge economy, a high-quality digital infrastructure is of essential importance. Dutch institutions have the ambition to:

- successfully serve a broad and diverse group of advanced users in accordance with their needs. This in turn requires continued investment in excellent facilities, developed in consultation with users. Together with local institutions, a broad set of facilities will be created for the (advanced) users of the digital infrastructure. This is not a static process – hence the need for continued investment.
- be a strategic partner for the leading players in the market, to enable them to supply products at attractive prices at an early stage, for experimentation and pilot studies by advanced users. This can only be achieved by an infrastructure comparable to that offered by other leading players. This combination of excellent knowledge and excellent facilities means that SURF will continue to be viewed as a high-level strategic discussion partner.

Realising these ambitions calls for an excellent and continuously innovating knowledge base. Given the lightning pace of these developments, the infrastructure itself has to be regularly renewed (every four to five years) on the basis of the requirements presented by the scientific community. Finally, advanced support is needed to assist the scientists in using these facilities.

In order to fulfil these ambitions, the Subcommittee recommends a total level of investment of 50m euros per annum, for innovation in the Dutch digital infrastructure. The Subcommittee expects that 4 million euros can be generated via the resources from the National Roadmap and EU-funded projects. The Subcommittee therefore recommends generating additional investments of 25 million euros per annum. Of this amount, 10 million euros is intended for converting current temporary budgets into structural ones, and 4m euros for topping up the shortfall in the operation of the existing supercomputer. The remainder of the budget, to the tune of 11 million euros, will be utilised to realise additional ambitions in the provision of support to researchers, investment in e-science and research software, and delivering an additional boost in the field of high-performance computing.

	Current budget				Additional financing based on current recommendations	Desired scope ICT infra
	Basic	Temporary NL government	EU projects	Total current		
Networks	5	4	1	10	1	11
Cybersecurity (AAI), Data	2	1	-	3	3	6
Computing facilities	8	5	1	14	8	22
eScience Center	6	-	-	6	5	11
Total	**21**	**10**	**2**	**33**	**17**	**50**

Table 1. Current budget and recommended additional funding (in € million/annum)
Note. The amounts under the heading 'Current budget' do not include the current DANS budget. DANS is included in the additional recommended financing

4 mllion euros/annum extra for Networks, Cybersecurity and Data Services

This amount is needed to raise structural financing to a level at which the Dutch ambitions as outlined in the Roadmap and the National Research Agenda are facilitated. At present there is an urgent structural shortfall, as we demonstrated above. We have also seen the rates of data growth and digitalisation far outpace the predictions when previous recommendations were issued. Excellent ICT facilities are an absolute precondition for successful participation in international research infrastructures, both in terms of gaining access to these types of infrastructure abroad, and attracting them to locate in the Netherlands. This investment combines a mixture of network and computing facilities, data storage and the necessary support for large-scale projects not listed in the Roadmap. Matching will also be required for numerous national and European projects, to allow an infrastructural contribution to be made. This Chapter of the recommendations also ties in with another Dutch ambition: establishing a digital main port as the basis for an attractive establishment climate (see Chapter 2.2).

8 million euros/annum extra for bringing Dutch high-performance computing into line with other leading knowledge economies in Europe

The answer to question 1 (see Chapter 3.1) shows that the widespread calls for support are detracting from the amounts reserved for a new HPC cluster. Additional investments will also be required in high-performance computing (HPC) in order to keep pace with the countries around the Netherlands. Neighbouring countries (Germany, France, Switzerland) are already working towards exascale computers (see EuroHPC, etc.), that offer scientists undreamed-of possibilities in tackling major challenges to society (better climate simulations for issues relating to global warming, for example) and carrying out existing calculations more cost effectively. Dutch researchers can only gain access to the large-scale European computing facilities if the Netherlands also establishes national computing facilities on a scale typically equivalent to 10 percent of Europe's largest computing facilities. Scientists must also be able to test their models on national systems, before making use of large-scale European facilities. This was indeed one of the conclusions of the ePLAN survey.

There is also growing demand for a broader range of computing facilities. Large-scale simulations in astronomy, climate studies and material studies, for example, call for different forms of computing capacity than the large-scale analysis of data in genomics. The availability of greater variety in terms of infrastructure is desperately needed. Use can of course be made of commercial cloud solutions. In the long term, these advanced computing facilities will also require additional user support. With that in mind, funding will initially have to be deployed to finance the purchase of a new supercomputer, to ensure that the Netherlands remains on a par with surrounding countries. A share of the budget (approximately 2 million euros per annum) will then have to be spent to fund the expected growth in user support. Eventually, a budget of this scale will enable the Netherlands to establish an annual reserve of approximately 8 million euros per annum for replacement of the high-performance computing facilities, 7 million euros per annum for user support, 6 million euros per annum for operating costs and 1 million euros per annum for innovation in the knowledge base.

5 M euros/annum extra for eScience and research software

The Subcommittee recommends investing 5 million euros/annum extra in eScience and research software. This funding is essential to meet the growing demand by users for specialist services and support of this kind. Although only a small proportion of Dutch researchers are conversant with the capabilities of the eScience Center (approximately 20% according to ePLAN), the limited funding means that only 10% of all requests received are granted by the eScience Center. This percentage is expected to fall even further, given the massive growth of data and software-intensive research in all scientific disciplines. As a consequence, scientists themselves will have to spend a large proportion of their valuable and limited research time on developing and employing software, and using the digital infrastructure. This will not only cause delays for the researchers, but will also take place at the expense of the quality and sustainability of the software thus developed. With that in mind, we urgently recommend making more funding available to support scientists in the future development of specialist research software, which can then be actively made available to the national scientific community.

The Subcommittee also recommends investing in the development and maintenance of a more sustainable infrastructure for research software. At present, the sole focus in the Netherlands is on establishing a sustainable and general infrastructure for networks (SURF), computing capacity (SURF) and data & archiving (SURF and DANS). The Research Software Directory (previously known as the eScience Technology Platform) represents a first step towards an infrastructure for software. Via this platform, software developed from individual projects is made available according to the principles of Open Science, cooperation between individual scientists is encouraged, and fragmentation (whereby the same software is concurrently developed at different places) is limited. Additional investments in the field of eScience will also provide the essential shoring up of the ties between Dutch ICT research (computer and data science) and the national infrastructure on the one hand, and data-intensive scientific research on the other. This will in turn stimulate the use and utilisation of the full capacity of the digital infrastructure, while at the same time facilitating more intensive multidisciplinary cooperation between ICT research and other scientific domains.

Confirm uniformity and focus in policy

Based on the recommendations from *ICTRegie*, SURF has been given control over the digital infrastructures for scientific research, as a result of which immediate responsibility for strategy and innovation has increasingly been placed in the hands of the users.

The Subcommittee therefore recommends that similar supervision be provided for public investments in ICT and digital infrastructures for research and education. To make this possible, official discussions on the management process should be organised between the Ministries of Education, Culture and Science, Defence, Economic Affairs, Infrastructure and the Environment, and Public Health, Welfare and Sport, which involve SURF. These Ministries are all involved in investments in digital Infrastructures for research and education, for example for the KNMI, MARIN, and NLR, as well as for the national library (Dutch abbreviation – KB)) and university libraries. In these discussion sessions, the Ministries should be required to keep one another informed of proposed investments, and investigate potential opportunities for synergy.

The Subcommittee also recommends that when investing in large-scale research facilities, the NWO should ascertain whether these facilities are able to make use of national digital infrastructures, and how any such investments in digital infrastructures could also be utilised by a wider group of users, in a federated structure. See also Chapter 3.3.

Develop a vision on new themes and manage its implementation

There are a number of new themes which first call for the development of a vision in the field of research and innovation. Under the overall management of SURF, within the current and requested budget restrictions, it is possible to specify the deployment and innovation of the national digital infrastructure. In line with the ePLAN survey, it is proposed that ePLAN should asses the needs as expressed by domain researchers, on an annual basis. New technological developments, such as developments of sensor networks and the Internet of Things, are examples of requirements in terms of new facilities and services. The National Plan Open Science is another good recent example of the effectiveness of surveys of this kind. The Subcommittee calls for special attention for the further elaboration of the National Plan for Open Science.

The parallel development of comparable services is ineffective. Avoiding this situation calls for cooperation between the various organisations, research facilities and providers of ICT infrastructure services involved. The Subcommittee recommends that the relevant scenarios be investigated. A number of ambitions outlined in Open Science can be achieved within the current recommendations, in particular by developing a broader variety of facilities and introducing newly developed services and support.

Evaluation mechanism: tie in with existing evaluations

The Subcommittee recommends harmonising the existing evaluation mechanisms for SURF, the eScience Center and DANS, to allow an adequate investigation and evaluation of the availability of digital infrastructure for science in the Netherlands, as a whole. This process will also ensure that the effectiveness of additional investments in the national infrastructure can be successfully monitored.

Multi-year financing for SURF is allocated on the basis of a multi-year plan approved by the affiliated institutions, in which investments, innovation, service development, service provision and support are all described. A mid-term review is held halfway through the plan period by an evaluation committee, which includes international members. The outcomes are shared with the members of SURF, including NWO.

Multi-year financing is allocated to the eScience Center on the basis of a collaboration agreement between NWO and SURF. The eScience Center itself is evaluated externally every six years by an independent evaluation committee, and the results are subsequently shared with SURF and NWO. DANS is evaluated according to the standard evaluation protocol operated by NWO, KNAW and VNSU (SEP), with an independent external evaluation every six years.

3.3 Recommendations for the National Roadmap

Advise the Permanent Committee on the formulation of the criteria for ICT in evaluating proposals for the National Roadmap.

The Roadmap was published in 2016, and issued a number of sound recommendations in respect of ICT (in particular in Chapters 4.3 & 4.6 of the Roadmap). Nonetheless, the Subcommittee wishes to issue a number of additional recommendations for evaluating ICT in applications for the National Roadmap.

Use the e-Needs questions from ESFRI in the call for the National Roadmap

In the upcoming call for the Roadmap, the Subcommittee recommends using the ESFRI-2018 questions (e-Needs). The use of the e-Needs questionnaire will guarantee uniformity in applications issued by researchers both in the Netherlands and in Europe. In addition, the answers will offer an insight into the relationship between project infrastructures and national or international digital infrastructures, as well as revealing possible opportunities for synergy.

Ensure the presence of sufficient ICT expertise in the Roadmap assessment committee

The Subcommittee supports the recommendation from the Permanent Committee to ensure that in assembling the assessment committee, sufficient expertise is available to allow assessment of the ICT component. ESFRI has opted to call upon members of the e-IRG to assess the European calls. In line with this decision, the Subcommittee recommends that in the Netherlands, the necessary expertise be provided by SURF, DANS and the eScience Center.

Also in other funding calls, request information about the need for digital infrastructure

The Subcommittee considers asking researchers responding to the Roadmap call about their use of digital infrastructure and digital support as a valuable first step. This initiative should be further extended to other national and European calls for research funding. In this way, the calls will raise awareness among researchers in all domains of the possibilities offered by digital infrastructures and support, in addition to which these infrastructures could be better utilised.

Encourage the federated use of digital infrastructures

The national digital infrastructures that are available are not always sufficient to meet the requirements of users. In such cases, if the evaluation committee considers it reasonable for applicants to develop an individual or additional digital infrastructure for their own domain or cluster, it is essential that immediate consideration be given to the question whether this infrastructure could in fact also have a broader scope. The Subcommittee therefore recommends encouraging the federation of digital infrastructures subject to the management of SURF. One current example is Nikhef, which alongside SURF investments, is currently being used to harmonise the use and availability of grid services and cloud services for scientific user communities.

Encourage cooperation between domain-specific infrastructures

A variety of large-scale national research facilities have a solid infrastructural data component, including BBMRI and ELIXIR-NL (Life Sciences), CLARIAH (humanities), ICOS (climate) and ODISSEI (social sciences and economics). The Subcommittee recommends encouraging co-operation between research facilities of this kind, and other research facilities, to promote co-operation between them and with the national service providers, for developing tools and software, standards and policy. In this way, knowledge will also be shared outside their own cluster, while work can be jointly undertaken on the establishment of standards. Finally, it would be useful if various research facilities were to become involved in the GO-FAIR and Research Data Netherlands (RDNL) initiatives.

3.4 Additional recommendations

Because the digital infrastructures are not 'islands' and because more work is needed to fulfil the ambitions of the National Research Agenda and the Roadmap, the Subcommittee calls for attention for other policies that go beyond the immediate sphere of influence of the parties that provide the national digital infrastructures and support.

Strengthen cohesion between ICT research and the digital infrastructure

The Subcommittee recommends strengthening the cohesion between ICT research and the digital infrastructure. As a digital main port, the Netherlands recognises the importance of knowledge and innovation in the field of digital infrastructures. With that in mind, within scientific programming, more focus should be placed on subjects that generate new insights or techniques, for the national digital infrastructure. The eScience Center should be able to fulfil a role in adopting new methods and techniques, and testing their usability for national digital infrastructures, in collaboration with ICT researchers. Finally, it is important that usable concepts and techniques are implemented by the providers of national digital infrastructures.

Organise local research support for researchers

There is growing demand among researchers for digital support and expertise in their research process, certainly in those disciplines where handling software and large volumes of data are new. It is neither scalable nor practical to organise this form of support exclusively at national level, because in many cases all that is required is short-term, hands-on support for not very complex challenges. The Subcommittee therefore recommends that in the framework of their vision 'The Digital Society', the VSNU encourages Dutch universities to connect or expand their ICT departments and/or libraries into research support departments.

These departments can then serve as the first point of contact for researchers, as well as representing links between themselves and the national support and national digital infrastructures.

Train data scientists, data stewards and digital world citizens

The digitalisation of science calls for employees trained in digital skills. Shortages of ICT professionals place us at a disadvantage. The Subcommittee therefore recommends giving digital skills a more prominent position at every level of education, also in scientific education. Make data and programming skills an integral part of all science disciplines and teaching programmes, in the same way that statistics is a fundamental element.

Train sufficient numbers of data scientists and data stewards: a huge demand is emerging in every sector for data scientists and stewards, in business, in government and in science itself. The European High-Level Expert Group for the European Open Science Cloud has estimated that Europe needs around half a million data stewards. The Subcommittee recommends that in the framework of their vision 'The Digital Society', the VSNU should consider how universities can offer data stewards attractive career prospects. This recommendation ties in well with the recommendations from Platform Education 2032, and the KNAW in 2012.

4 Annexes

4.1 Overview of the national digital infrastructure

SURF, DANS (a KNAW and NWO initiative) and the eScience Center (SURF and NWO initiative) are the organisations with a national task in respect of the digital infrastructure. They offer a range of infrastructural services, as outlined below: a description of the infrastructure, its value for researchers and a brief analysis of the developments required over the next few years to offer good facility support to researchers.

Although there are local and domain-specific networks and computing facilities in the Netherlands, they are financed either locally or on a project basis, and in most cases are not nationally accessible or are unsuitable for use in multiple scientific domains. As a result, they are only referred to in passing. International infrastructures are described briefly because participation in such infrastructures requires a national contribution of resources, and because they are essential for international cooperation by researchers.

4.1.1 Networks

A network links together the biobanks of BBMRI-NL, transports data from CERN to computing and datacentres around the world, offers researchers access to the distributed data in CLARIAH, enables researchers to cooperate online, and connects cloud services (both commercial and non-commercial) to the institutions. Commercial networks in many cases fail to fulfil the requirements of the majority of data-intensive scientists in terms of quality, speed, security and capacity.

Organisation

Access to networks is achieved locally via the campus networks of the institutions. These are linked to the national network operated by SURF. In turn SURF works alongside European and other international partners to facilitate worldwide cooperation between researchers. SURF also operates alongside European (GÉANT) and other international partners (NRENS) to enable researchers to cooperate worldwide.

Development of a fixed network

Among new network generations, capacity has for years been the key factor. Although still important, it is starting to emerge that the basis for the current Internet is itself not designed for the changes in use. Researchers no longer store large volumes of data in their own data centres; instead, they transmit them to combined databases and the cloud, via public connections. Private networks offer a solution for large volumes of data or for data in respect of which security is important; such as medical research data. This can be achieved by establishing rapid and secure light paths between a large-scale research facility and a data or computing centre.

Against that background, it is just as important that network services be made more advanced and more flexibly by ensuring the availability of new intelligent techniques (virtual and programmable networks) in building blocks.

In this way, service providers and users are themselves able to create new applications on the thus offered building blocks. Compare this situation to a smartphone, whereby third parties build applications on the functionalities offered by the smartphone manufacturer. By enabling creative ideas to be realised, innovation is accelerated and an ecosystem is established for providers and customers in which the needs of researchers can be responded to better and faster.

Development of wireless networks

The end points of the network are becoming increasingly wireless. This certainly applies to users who are seeking to gain access to the network via wireless systems. The high demands placed by users on bandwidth and quality of the wireless network mean that institutions have an ever-growing need for advanced knowledge and support in the roll-out of these networks.

In many cases, the instruments themselves are also becoming mobile. Take sensor networks, for example, the Internet of Things. Some sensors transmit data once a day, and others must remain permanently alert to the environment such as those in autonomous vehicles. Examples of sensor networks in science are networks for fine particulate measurement, earthquake sensors and smart cities. We can expect exponential growth in sensor networks, accompanied by high demand from researchers in terms of reliability, security, bandwidth, coverage, etc.

4.1.2 Computing facilities

For simulations and for processing and analysing data, researchers need computer facilities. Climate models and genetic models are processed using large computer clusters, but linguists, literary scholars and cartographers are also increasingly relying on advanced computer facilities.

For their calculations, many scientists use national computer facilities, such as those offered by SURF. However, because not every application requires the same type of computing system, the national computer infrastructure provided by SURF is made up of different types of computing capability. The current systems on offer, use types and user types are reproduced below in two diagrams.

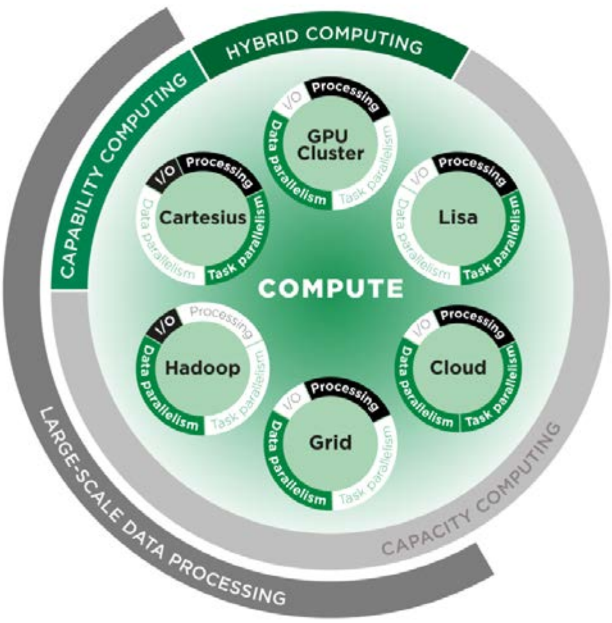

Figure 1 SURF computing portfolio

Typical e-infrastructure usage

Infrastructure	Typical usage	Typical users	Use cases
Cartesius	• Capability computing • MPI applications	• Research groups • Consortia	• Climate modeling • Fluid dynamics • Galaxy simulation
LISA	• Job farming • MPI applications	• Research groups • Individual researchers	• GWAS studies • Magnetic field research
Grid	• Job farming • I/O heavy applications	• Consortia	• LHC experiments • LOFAR • NGS data analysis
HPC Cloud	• 3rd party PaaS • Courses • High Memory apps • Microsoft Windows (& other licensed software)	• Individual researchers • Consortia	• Galaxy platform • AMC compute course • De novo sequence alignment • Stock exchange ticker data
Hadoop	• Pattern recognition / data mining • Web crawls • Information retrieval	• Individual researchers	• Twitter mining • TwiNL • SETI analysis

Figure 2 Use of the SURF computing portfolio

Organisation

Under the auspices of SURF, the digital infrastructure is working alongside partners within research institutions on a federated infrastructure. For example, Nikhef, RUG-CIT and SURF together are harmonising their investments in and the use and availability of grid and cloud services, on behalf of their scientific user communities. Within a number of large-scale research facilities with exceptional data production, HPC clusters are an integral part of the facility, used for carrying out initial data selection (LOFAR and LHC, for example). Following that initial selection, data are then transported around the world, so that researchers can use their national computing facilities in their research work involving that data.

A number of public institutions operate a smaller HPC facility, including the KNMI (weather forecasting), MARIN (hydrodynamic simulations) and NLR. These facilities offer capacity on a national scale. Within the Netherlands, the DAS-5 (Distributed ASCI Supercomputer 5) is available within the field of ICT research for testing new theories and concepts. DAS 5 operates as a flexible research infrastructure.

The development of computing

Given the huge growth in data volumes and the diversity and complexity of data types, the demand for advanced computer facilities and services is constantly expanding. This not only applies to supercomputers but for all types of computing services.

Computing capacity must continue to grow, and at least keep pace with international developments, in order to be able to offer Dutch researchers access to even larger European supercomputers. In that process, sustainability is becoming increasingly important. Computing facilities are becoming increasingly connected to data storage to facilitate data-intensive research. The demand for specialist software and statistics packages is also growing, since they allow the analysis of large, complex datasets.

4.1.3 Data storage and data archiving

Researchers wish to store their raw and processed research data, even when it comes to large, diverse and complex datasets. After all, the data must be reusable for further research or for checking research results. Certainly in the case of data from costly or long-term research programme it is extremely desirable that the same data be available for reuse by other researchers.

A challenge for data services is that although data must be stored securely, at the same time it must be accessible to researchers. There are data services at different levels: close to the storage facilities, specifically for optimum, secure and rapid access; services at a high level of aggregation and domain-specific services.

Organisation

In addition to local data services at individual institutions, there are a number of players that offer data services at a national and regional level. SURF and DANS are national providers while the 4TU.Centre for Research Data provides the service on behalf of the four Dutch universities of technology.

DANS encourages sustainable access to digital research data and encourages researchers to sustainability archive their data, for reuse. DANS also offers access to thousands of scientific datasets, publications and other research information, in the Netherlands. SURF offers a range of services in the field of big data, including secure and long-term storage, simple storage and data sharing, big data analysis and coding, to improve data retrievability.

DANS, SURF and the 4TU.Centre for Research Data work together within Research Data Netherlands (RDNL) and the National Coordination Point for Research Data Management. They each operate so-called back offices, whereby the front offices (for example the university libraries or university ICT departments) maintain direct contact with the researchers. There are also domain-specific data services, provided by large-scale research facilities, for example.

Development of data services

Although the sharing of data is not yet common practice among researchers, given the political desire for open data (an extension of open access to publications), it is merely a question of time before the financial backers of research programmes start imposing open data requirements. The explosive growth of data volume is also generally recognised, particularly with the advent of numerous sensor networks. The ePLAN survey reveals that scientists expect exponential growth in their demand for data resources. DANS is increasingly working on behalf of all disciplines that demonstrate a demand for data services.

There is a clear need to improve the storage of, access to and sharing of data. The GO-FAIR initiative, aimed to assist researchers to make data Findable, Accessible, Interoperable and Re-usable (FAIR) is clearly important. Data, first and foremost big and open data, represent a challenge that crosses all domains, and affects all research infrastructures. Without the software for reading, processing or interpreting data, the data soon lose their meaning. It is against that background that DANS is increasingly focusing on software sustainability.

4.1.4 Authentication and authorisation infrastructure (AAI) (or 'cybersecurity')

Researchers are co-operating more and more frequently on both a national and international level, and are sharing research facilities and data. To determine whether a specific researcher is authorised to access data, it is important to be able to ascertain his or her identity. The correctness of the identity must be reliably determined, while at the same time guaranteeing privacy, particularly in respect of access to internal or external online services.

Organisation

To make this possible, SURF has developed an authentication and authorisation infrastructure (AAI) which is also available to lecturing staff and students. Researchers log in with a single account and via single sign-on facilities have secure access to services that are made available by national and international partners and providers (both commercial and non-commercial), while limiting the exchange of personal details to an absolute minimum. More than 130 cloud services are already linked to SURFconext, including services for online co-operation, storage, video and research. Examples are Google Apps, N@tschool and Kluwer Navigator.

Development of the AAI

For sensitive data, such as eHRM, student information systems and applications with patent-sensitive research data or patient information subject to privacy restrictions, various institutions are increasingly making use of cloud services. However, these services require more stringent forms of authentication than simply user name/password security, with a view to limiting the risk of data breaches and other security incidents.

At present, a researcher's account is still linked to a particular institution. Work is currently underway on providing researchers in the future with a reliable personal identity. This will offer researchers greater freedom, and allow them to make use of services, independent of their specific institution, enabling them to simply transfer their accounts if they move to another institution. User privacy is another area that continues to demand close attention, as is the question of guaranteeing the reliability of a person's identity.

There is a need for an international standard in respect of AAI. Traditional methods for confirming an identity through the use of certificates are not scalable for large communities. Consequently, AAI solutions are more and more common in large-scale research infrastructures. As a consequence, digital infrastructures often work closely together with these communities, for testing, harmonising and implementing existing solutions.

4.1.5 eScience

The transition to data-intensive and computing-intensive research means that researchers must be able to make optimum use of advanced research software. Whereas in the past many researchers were able to carry out their data processing and analysis work with relatively simple local hardware and software, the requirements have grown enormously over the past few years. It is becoming increasingly simple for research software to gain remote access to computing facilities. The capabilities of cloud-based computing and data services have also expanded hugely.

Organisation

To better utilise the digital infrastructure and the insights gained from computer and data science, the NWO and SURF established the Netherlands eScience Center, in 2011. The eScience Center serves as a link between researchers and the existing digital infrastructure, and assists researchers in developing research software for the processing and analysis of large and complex datasets. Researchers simply submit their project applications to the eScience Center. If approved, they receive (a combination of) in-kind support from eScience Research Engineers and in-cash contributions.

On a national scale, the eScience Center coordinates ePLAN, the platform for eScience/data research organisations. On a European scale, this same service is provided via PLAN-E, the platform that brings together Europe's eScience Centers.

Development of eScience and research software

Practically every discipline is facing ever more complex challenges in terms of the analysis of large volumes of texts, photographs, sensor data, climate data, satellite data, social media data, databases or survey data. The effective use of the digital infrastructure combined with the latest insights from computer and data science are key contributing factors to scientific and scholarly impact. The growing demand for support services must be met.

At present, research software is mainly developed on a project basis within the eScience Center, in close collaboration with specific users. However, there is a clear need to provide a structural facility that creates access to, maintains and continues the development of research software. The Research Software Directory was initiated in 2016 with exactly that in mind. The goal is for this online platform to become a national digital software infrastructure capable of supplying researchers with guides, guidelines and other eScience and software documentation.

4.1.6 Services and Skills

Support and expertise, human capital and service provision are integral elements of the digital infrastructures referred to above. Without human resources capable of offering the necessary knowledge and scientific experience of management, consultancy and practical application of ICT, scientists will be unable to make use of the existing infrastructure.

Organisation

Support and expertise are provided by all organisations involved in digital infrastructures, service provision for and support to researchers.

The services are provided at local level by universities and research institutions and at national level by the providers of national digital infrastructures and (eScience) services such as DANS, SURF and eScience Center.

SURF offers support to researchers in selecting the ideal digital tools, and organises training programmes on a variety of subjects such as current ICT themes and hands-on system training. In providing this particular service, SURF works alongside the eScience Center and NIKHEF, in the Support4Research programme.

Whereas SURF tends to focus on short-term advice programmes, the eScience Center supplies intensive and high-quality eScience support, such as scientific support and advice on the use of the digital infrastructure and the development of research software. In collaboration with SURF, the eScience Center also provides training to scientists in the field of programming, data analysis and the use of computing facilities.

DANS in turn provides training and consultancy services, and is responsible for research into sustainable access to digital information. Researchers in every stage of their research work can contact the 4TU.Centre for Research Data for advice and support on data management. This ranges from drawing up a data management plan and facilitating a data lab through to the eventual depositing of data in an archive.

Development
The Open Science Cloud is due to take up a prominent position in the world of research. There is growing demand among researchers for support in open research, open notebooks, open source, open data, open access and data stewardship. This support to researchers is still very much in the development stage, also in terms of the use of the digital infrastructure and the provision of research software. The challenge lies in making the support scalable and enabling experiences accrued in one discipline to be used in others. The clear key elements are that the focus remains on the researcher and that any approach developed is multidisciplinary.

4.1.7 International digital infrastructures

Although international digital infrastructures are not part of the national infrastructure, we have included a brief description of the most important European infrastructures, because the opposite clearly is the case: national infrastructures are part of a larger international infrastructure, which is essential for facilitating international cooperation between researchers. National organisations are required to make a contribution, which as such is an integral element of the total budget available to the national digital infrastructure.

GÉANT
GÉANT manages a European backbone between the 41 national research networks (NRENs). Within GÉANT, the NRENs also collaborate on innovation, standardisation, harmonisation, AAI and new services. SURF is a project partner. SURF contributes to the network operating costs according to a model based on the purchase of bandwidth, and Gross Domestic Product.

PRACE
PRACE offers competing European researchers access to the largest European HPC systems (Tier-0). At present there are only a few Tier-0 systems: in Germany, France, Italy, Spain and Switzerland. Researchers can only gain access to these systems if their own country contributes national computing facilities on at least tier-1 scale to the project, and their national computing centre makes a financial contribution according to a model based on historical use and Gross Domestic Product. The contribution from SURF involves the national supercomputer.

EGI
EGI offers advanced computer services to scientists, multinational projects and research infrastructures. EGI is a product of the grid infrastructure that was built for processing the data from the LHC experiments. EGI is a federation of more than 300 data and computing centres spread across 22 countries in Europe and the rest of the world. In the Netherlands, SURF, Nikhef and the RUG-CIT are all part of the EGI facility. EGI offers services in the field of (federated) high-throughput computing, storage and cloud. The head offices of EGI are based in Amsterdam.

EUDAT
EUDAT offers generic data services, storage and support for data, to a variety of research communities and for individual researchers, via its 35 European partners. The data are stored in fifteen European countries, close to Europe's most powerful supercomputers. EUDAT provides access and storage, from informal data sharing through to long-term archiving, identification and findability of both long-tail and big data, throughout the entire lifecycle of the research data in question. SURF, DANS and KNMI are partners in EUDAT.

4.2 National digital infrastructure survey

To gain a clear picture of the current and future needs of scientists in terms of digital infrastructure, the national platform for eScience and data science organisations conducted a survey at the end of 2015 among more than 1000 researchers (university lecturers, senior university lecturers and professors) spread across all disciplines, 242 responses were received from researchers who replied on behalf of their study group. In addition, four targeted workshops were organised in 2016 among researchers representing Life Sciences, Humanities, Physics and the Environment.

The conclusions most relevant to these recommendations appear below.

- The developments in the research domain are characterised by a manageable number of keywords: above all more data, but also computing, specifically also more truly *large-scale* computing, more image processing and visualisation, more requests for (support in) data analysis and concerns about security and privacy;
- There is already much multidisciplinary cooperation, with a clear focus on international cooperation;
- The rise in the importance and volume of data for research covers all disciplines and all types of data;
- There is a real expectation of exponential growth in demand for data resources (into the exabytes domain) and data services;
- Major concerns were expressed about the general shortfall in the level of investments in the national ICT infrastructure in relation to the massive growth in demand for all types of resources, and in comparison with the infrastructure available to competitors elsewhere in the world. The absence of reserve funding for future investments, for example in a new national supercomputer, is a further cause for concern;
- The backlog in terms of a national HPC facility as compared with the international lead group should not be allowed to exceed a factor of 10 for researchers, if they hope to keep up with the competition. This is also necessary in order to maintain (the right to) access to European facilities such as PRACE, the European infrastructure for supercomputers;
- Of the majority of respondents who announced that they would be computing more, one third indicated that the work involved is memory-intensive;
- In addition to a supercomputer, there is also demand for a robust second layer of cluster facilities, provided centrally or coordinated within the institutions;
- Generally speaking, the availability of a national cloud service or a European-based cloud service would be appreciated, in connection with the expectation that privacy and security can be better regulated in this way than via commercial services;
- There is a clear call for improvement in the ease of use of the ICT infrastructure and access to the infrastructure;
- Support by such parties as SURF and the eScience Center is identified as a valuable and necessary facility for the efficient use of the ICT infrastructure in research;
- More than half of all respondents indicated that they had never heard of the terms stewardship, research data management and/or software sustainability. Nonetheless, 90% (!) of respondents identified these subjects as of considerable importance; important enough to warrant national policy (80%).

4.3 Growth of scientific data

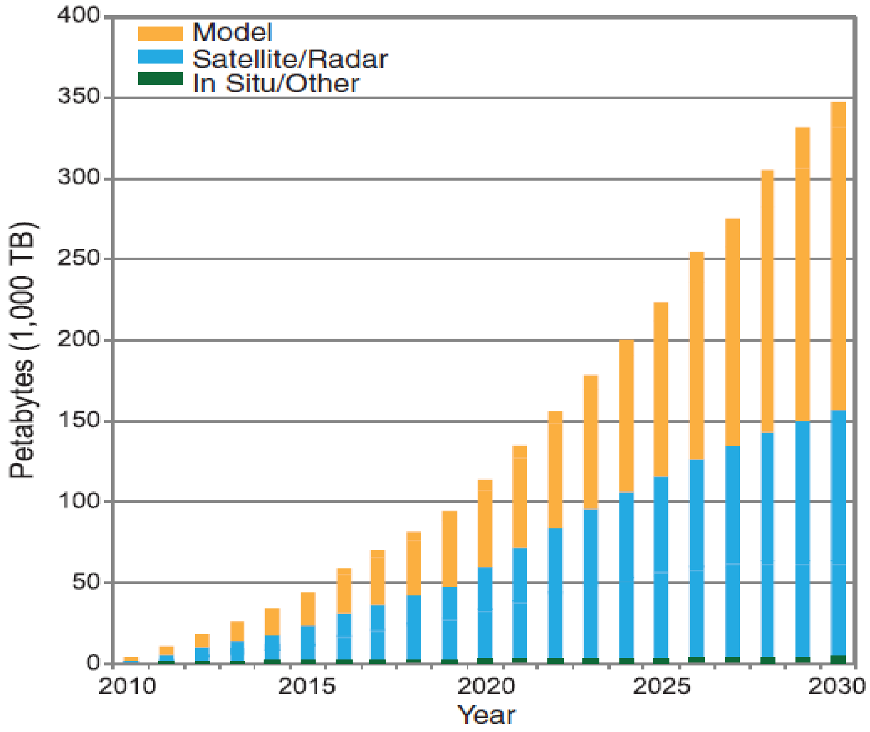

Figure 3 Worldwide scale of climate data (Science, 2011)

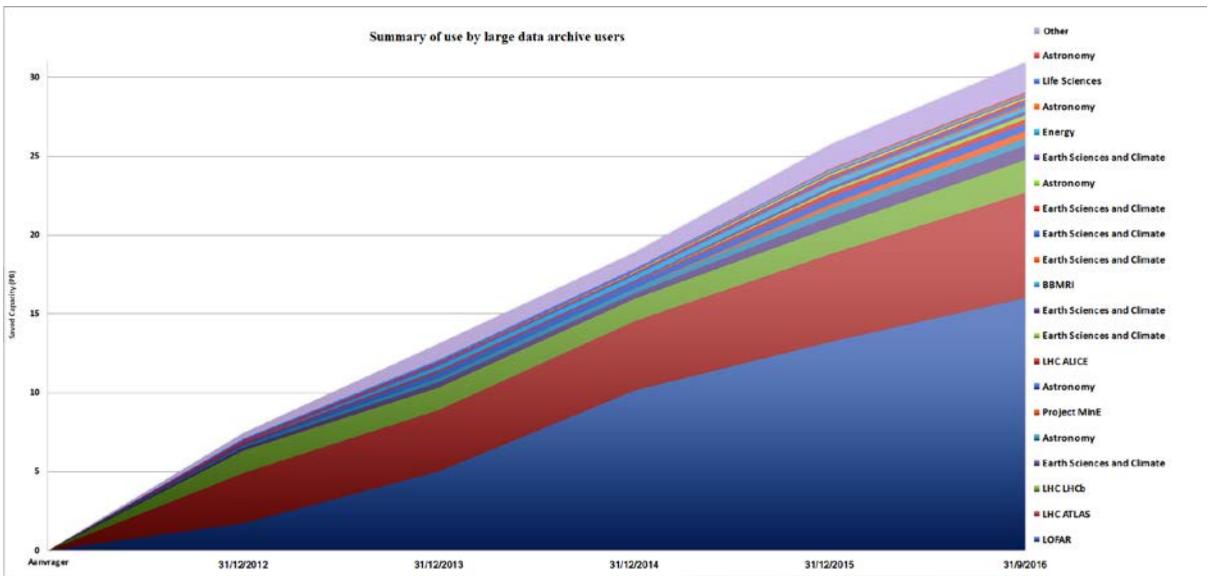

Figure 4 SURFSara Central Data Infrastructure. Summary of use by large data archive users

4.4 E-Needs questions in ESFRI submission format 2018

A) E-NEEDS questions for new proposals

4.1. Outline the Data Management Plan (DMP) and data access policy of the RI. If applicable, describe how data would become accessible to the public: (maximum 3000 characters with spacing)

4.2. Describe and quantify what e-infrastructure services - e.g. resources for storage, computing, networking, tools for data management, security, access, remote analysis, etc. - your RI will need: (maximum 2000 characters with spacing)

4.3. Describe how the e-infrastructure services needed by your RI will be implemented, specifying the potential need of external e-infrastructure resources and the relations to external e-infrastructures: (maximum 2000 characters with spacing)

4.4. Describe how the RI will contribute to the development of the European and global e-infrastructure landscape at all levels (institutional, regional, national, international) - including e.g. the e-infrastructure commons and the European Open Science Cloud (EOSC): (maximum 2000 characters with spacing)

B) E-NEEDS questions for 2008 & 2010 PROJECT MONITORING

4.1. Summarise the data management, data access and data security policies of your RI - including procedures for ensuring the sustainability of data – and clarify to what degree they are approved: (maximum 3000 characters with spacing)

4.2. Describe the technical design and operations of the e-infrastructure services -e.g. resources for storage, computing, networking, tools for data management, security, access, remote analysis, etc. - for your RI and clarify to what degree it is approved *and implemented*: (maximum 2000 characters with spacing)

4.3. Describe the use of external e-infrastructure resources and the relations to external e-infrastructures - including agreements with external partners delivering core e-infrastructure services for your RI – and clarify to what degree this use is approved within your consortium: (maximum 2000 characters with spacing)

4.4 Describe how the RI contributes to the development of the European and global e-infrastructure landscape at all levels (institutional, regional, national, international) - including e.g. the e-infrastructure commons and the European Open Science Cloud (EOSC): (maximum 2000 characters with spacing)

4.5 If available, upload the approved Data Management Plan (DMP): (upload with limit 1 MB)

C) E-NEEDS questions for LANDMARKS

4.1. Summarise the data management, data access and data security policies of your RI - including procedures for ensuring the sustainability of data – and clarify to what degree they are approved *and implemented*: (maximum 3000 characters with spacing)

4.2. Describe the technical design and operations of the e-infrastructure services -e.g. resources for storage, computing, networking, tools for data management, security, access, remote analysis, etc. - for your RI and clarify to what degree it is approved *and implemented*: (maximum 2000 characters with spacing)

4.3. Describe the use of external e-infrastructure resources and the relations to external e-infrastructures - including agreements with external partners delivering core e-infrastructure services for your RI – and clarify to what degree you have *signed agreements*: (maximum 2000 characters with spacing)

4.4. Describe how the RI contributes to the development of the European and global e-infrastructure landscape at all levels (institutional, regional, national, international) - including e.g. the e-infrastructure commons and the European Open Science Cloud (EOSC): (maximum 2000 characters with spacing)

4.5 Upload the approved Data Management Plan (DMP): (upload with limit 1 MB)

4.5 Glossary

BiG Grid self-evaluation
In connection with the final evaluation of the Big Grid project, a self-evaluation was undertaken according to the evaluation framework from the *Impuls Grootschalige Onderzoeksfaciliteiten* (Large-Scale Research Facilities Impulse).

Compute Canada
Compute Canada facilitates the acceleration of research and innovation by offering state-of-the-art computing systems, storage and software to researchers. Compute Canada provides Canadian researchers and their partners in other sectors with high-end computing facilities and support. For more information: https://www.computecanada.ca/home/

European Open Science Cloud
The European Open Science Cloud aims to realise an integrated European infrastructure including access, management and financing mechanisms that broadens optimum reuse of research data and facilitates open science in general. For more information:
http://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud

GO-FAIR initiative
GO-FAIR is an international initiative coordinated by the Netherlands that supports a revolutionary transformation of Open, data-intensive science and innovation, with a worldwide social impact. The aim is to bring about a scalable implementation environment to create an Internet of FAIR data and services, where analysis tools can be combined with relevant data, because both the tools and the data are Findable, Accessible, Interoperable and Reusable (FAIR).

***ICTRegie* recommendations (2008)**
Under the auspices of NWO, *ICTRegie* issued recommendations in 2008, at the request of the Ministries of Education, Culture and Science and Economic Affairs, on the digital infrastructure in the Netherlands. The recommendations were twofold: *ICTRegie* issued recommendations for a new governance system whereby SURF was given control over the digital infrastructures. In addition, *ICTRegie* issued recommendations for investments in the digital infrastructure.

National Coordination Point Research Data Management (Dutch abbreviation - LCRDM)
The National Coordination Point Research Data Management is responsible for the national coordination of research data management policy, thereby facilitating a national approach to research data management in the Netherlands. Crucial to this approach is close consultation with teaching and research institutions. As a result, the Netherlands will be able to develop and implement research data management policy efficiently and effectively. SURF fulfils this coordinating role while the VSNU research and valorisation steering committee serves as a sounding board.

Together with the institutions, five issues have been identified to be tackled by the National Coordination Point in its working committees: facilities and data infrastructure, legal aspects and authority, financing, research support and advice and awareness of the advantages.
The LCRDM working committees, each staffed by experts from the field, will move these issues forward to support the development of national research data management policy. This is carried out in close and considerate consultation with all relevant stakeholders and research communities, and in line with the latest developments at European Level. For more information: www.lcrdm.nl

National Plan Open Science (NPOS)
The National Plan Open Science is a national plan that describes the ambitions of the Netherlands, current activities and possible future steps for open science. By signing the National Plan Open Science, all stakeholders will express their intention to move forward together, in achieving the open science ambitions.

Research Data Netherlands (RDNL)
Research Data Netherlands is a collaborative venture between the service providers 4TU.Centre for Research Data, DANS and SURF. This coalition, that is also open to membership by other parties, brings together the three data archives in the field of sustainable data archiving. For more information: www.researchdata.nl

Taskforce NWO (2010)
The Taskforce NWO, established in 2010, was a follow-up to the recommendations from *ICTRegie*. On the one hand, the plan described by the Taskforce was a more sober version of the recommendations from *ICTRegie*, while on the other hand it contained a proposal on how the plans of *ICTRegie* for the infrastructure, as supported by the Dutch government, could be financed. NWO, the Ministry of Education, Culture and Science, the Ministry of Economic Affairs and SURF were all part of the Taskforce.

4.6 Background literature

AWTI (2015). *Klaar voor de toekomst. Naar een brede strategie voor ICT*.

AWTI (2016). *Houd de basis gezond. Prioriteiten voor extra investeringen in onderzoek en innovatie*.

Commission High-Level Expert Group on the European Open Science Cloud (2016). *Realising the European Open Science Cloud.*

Compute Canada (2016). *Investing in advanced research computing resources*. https://www.computecanada.ca/research/investing-in-advanced-research-computing-resources-is-essential-to-competitive-science-and-innovation-in-canada/

Deloitte (2016). *Dutch Digital Infrastructure 2016: Enabling the Digital Economy and Society.*

ePLAN (2016). *Infrastructuur duurzaam op maat. Een analyse van de vraag naar ICT-infrastructuur en services in Nederland.*

ICTRegie (2008). *Verkenning naar de gewenste ICTinfrastructuur voor het onderzoek in Nederland*.

Kenniscoalitie (2015). *De Nationale Wetenschapsagenda: Vragen, verbindingen, vergezichten*.

KNAW (2016). *KNAW-Agenda grootschalige onderzoeksfaciliteiten*.

KNAW (2012). *Digitale geletterdheid in het voortgezet onderwijs.*

National Academies of Sciences, Engineering, and Medicine (2016). *Future Directions for NSF Advanced Computing Infrastructure to Support U.S. Science and Engineering in 2017-2020.*

Nederland ICT (2016). *Groei door digitalisering*.

NL next level (2016). *Investeren in een digitale kwantumsprong*.

NWO (2016). *Nationale Roadmap Grootschalige Wetenschappelijke Infrastructuur*.

OCW (2017). *National Plan Open Science*.

Platform Onderwijs 2032 (2016). *Ons onderwijs 2032. Eindadvies.*

Staatssekretariat für Bildung, Forschung und Innovation SBFI (2015). *Schweizer Roadmap für Forschungsinfrastrukturen im Hinblick auf die BFI-Botschaft 2017-2020*.

Taskforce Financiering Wetenschappelijke ICT-Infrastructuur (2010). *Hoe de nationale ICT-infrastructuur voor research en hoger onderwijs te financieren.*

VSNU (2016). *De Digitale Samenleving. Nederland en zijn universiteiten: internationale pioniers in mensgerichte informatietechnologie*.

4.7 Composition of the ICT infrastructure advisory subcommittee

Chair
Sally Wyatt (Maastricht University and Huygens ING, KNAW)

Subcommittee members
Patrick van Beers (personal title)

Henk Dijkstra (Utrecht University)

Erik Fledderus (SURF)

Bernard de Geus (Virtual Laboratory Plant Breeding)

Wilco Hazeleger (Netherlands eScience Center)

José de Kruif (Utrecht University)

Joeri van Leeuwen (ASTRON)

Andrea Scharnhorst (DANS)

Ana-Lucia Varbanescu (University of Amsterdam)

Jacob de Vlieg (Eindhoven University of Technology)

Acknowledgements
The Advisory Committee is very grateful to Tom Bakker (eScience Center) and Ilse Koning (SURF) for their extensive work in preparing this report. We would also like to thank Yvette Tuin (NWO) for her support of the activities of the Advisory Committee.