Computational needs for accelerated scientific discovery

User consultation

Contents

Introduction

Scientific research increasingly relies on digital infrastructures. Digital infrastructures enable researchers from around the world to collaborate, share data, and perform complex calculations on large data sets. These infrastructures are not only relevant to the STEM disciplines (science, technology, engineering and mathematics), but increasingly to researchers from a broad range of other scientific backgrounds. Continuous innovation of these infrastructures is important to enable users to benefit from the latest developments in hardware, software and application use. The pace of digitalisation has accelerated in recent years. The adoption of artificial intelligence (AI), for instance, is exceeding expectations in many research disciplines worldwide. As a result, many countries are rethinking their digital infrastructure strategies to enable scientific research and innovation to keep pace.

On behalf of the Ministry of Education, Culture and Science (OCW), NWO intensified its funding for the digital infrastructures for research in 2019 with the Implementation Plan Investments Digital Research Infrastructure[1]. Given the pace of developments, with AI as an emerging system technology, it is time to rethink the Dutch strategy for digital infrastructures for research. This starts with an examination of scientific challenges recognised by Dutch researchers, and their translation into requirements for the digital infrastructure they use.

Please note that the focus in this document is on one specific element of the digital infrastructure for research, namely Tier-1 computing facilities. In addition to high performance computing (HPC), we have also taken high-throughput computing (HTC) into account, which facilitates parallel data processing and is particularly relevant for large-scale scientific instruments in the fields of physics and astronomy. Future considerations of digital infrastructure needs should take into account the whole ecosystem (European, national and local) and include other fundamental elements such as network, data and software facilities. Another crucial aspect is the financial and organisational sustainability of the digital infrastructures for research.

In this document, we describe the expected demand from the scientific community for computing facilities in the coming years. It also considers the impact of AI on this demand, as it will have a huge impact on methodologies and data-driven analyses in a wide range of scientific areas[2]. This science case is based on consultations with a number of principal investigators of Dutch research groups that make extensive use of current computing facilities, and on a questionnaire distributed widely in the scientific community[3], covering a range of approximately 60 scientific disciplines. The following sections describe our general findings. A more detailed account of the scientific ambitions and infrastructure needs of current users can be found in Chapter 4. Further relevant information is provided in the appendices.

---

[1] Uitvoeringsplan investeringen digitale onderzoeksinfrastructuur.pdf (nwo.nl)

[2] For more information on the use and impact of AI in different scientific fields, see Appendix C.

[3] We received responses from 43 researchers from 21 organisations, including universities, research institutions, TDCC and infrastructure partners. More information on the researchers and experts consulted can be found in Appendix A. The questionnaire can be found in Appendix B.

01

Growing demand

The scientific ambitions of all research disciplines in the Netherlands are high, as reflected in the responses of the scientific community to our questionnaire and consultations. Overall, researchers reported a significant increase in demand for capacity, but the percentages varied widely between disciplines. In all cases, computing capacity was considered an urgent matter and essential for continuing top-level scientific research.

The main reasons for this growing demand are the new opportunities offered by digitalisation and data-driven research, the continuous growth of data volumes and the models that researchers are working on, all of which require large-scale data and processing facilities. Digital infrastructures have historically played an important role in the STEM disciplines, but are increasingly doing so for many other scientific fields. This is a result of the availability of larger and more complex data sets, the rise of AI as a research methodology and the adoption of advanced analytical techniques in these fields. This is particularly the case for the Health and Social Sciences & Humanities domains. Broadening the user base of digital infrastructures also leads to different requirements in terms of workflows and user-friendliness.

The increase in data volumes and the growing demand for computing facilities also have implications for (long-term) data storage facilities. Several respondents made explicit reference to this issue. If data production continues to grow at this rate, the demand for (affordable) storage capacity will grow exponentially in the coming years. Another aspect is the specific demand for facilities that are suitable for the processing and storage of sensitive data, which was raised in particular by respondents in the Health and Social Sciences & Humanities domains.

Another driver for the growing demand for computing and data facilities is the participation of Dutch researchers in international large-scale research instruments and infrastructures. Historically, the Netherlands has invested in computing and storage equipment for large-scale research instruments such as the Large Hadron Collider (LHC) at CERN and the Low Frequency Array (LOFAR) distributed radio telescope, securing our position as a strong participant in international research collaborations. Through these engagements, our scientists and institutions contribute significantly to global scientific endeavours, driving innovation and discovery on an international scale. These international collaborations also allow us to retain or attract talent, which tends to boost our economy in a broader sense. With the planned extensions of the LHC to the High Luminosity LHC (HL-LHC) and the upcoming Square Kilometre Array (SKA), the required digital infrastructures will need to grow by at least one order of magnitude. At the same time, large-scale projects have started in the Social Sciences domain with more than a hundred PhD students collecting sensitive data[4], boosting the demand for secure computing and data facilities.

Participation in these international large-scale research instruments and infrastructures is crucial if we want to remain competitive on a scientific level, but also to reinforce the innovation climate in the Netherlands and secure our future national earning capacity. These ambitions have considerable implications for the demand for computing and data capacity. These tools also depend heavily on local and international network infrastructures.

---

[4] For example, the SOCION project, funded by an NWO Summit grant.

The growth in demand is also fuelled by the spectacular rise of AI methods in a wide range of scientific disciplines, opening up new opportunities in data analysis, making existing workflows more efficient through surrogate models and sometimes enabling new approaches to old problems[5]. AI is driving a skyrocketing demand for computing and data facilities, both now and in the near future. AI is also causing a shift from CPU-centric to more GPU-accelerated environments. However, CPUs remain important for specific research areas. At the same time, the application of AI technology is raising questions about knowledge security, privacy and other public values, as well as our strategic digital sovereignty and dependence on dominant companies.

Inseparable from the growing demand for computing and data capacity is the importance of appropriate expert support and training for researchers working with the digital infrastructures. This is relevant for the most demanding 'traditional' users, but also for scientific disciplines that are less experienced in using digital infrastructures. Thus, in addition to the growing demand for computing and data capacity in itself, there is also a growing need for skills development, support and training.

The growing demand for computing and data needs also raises questions about the carbon footprint of digital infrastructures and the possibilities of energy-efficient computing. Energy-efficient computing is identified as one of the strategic focus areas in the National Agenda for Computational Sciences[6]. Current innovation efforts at SURF on the national supercomputer Snellius are already exploring energy-efficient computing and will be continued.

---

[5] More information on the impact of AI on a range of scientific disciplines can be found in Appendix C.

[6] National Agenda Computational Sciences.

02

Current situation

Many respondents to our questionnaire expressed concern about the availability of sufficient computing power and data facilities in the near future, given the growing demand they foresee. Currently, Dutch researchers have access to the national supercomputer Snellius (Tier-1), a general-purpose high-performance computing system. The availability of a Tier-1 facility in the Netherlands makes it possible for researchers to upscale from their local facilities (Tier-2) at universities and research institutions to the national level for larger computations, including expert support. The availability of a Tier-1 facility also enables Dutch researchers to make the leap to international (exascale) facilities (Tier-0). Access to Tier-0 facilities is important for Dutch researchers for the largest computations and for international collaboration. Our recent participation in the LUMI pre-exascale system in Finland and our future participation in Jules Verne in France strengthens our international position.

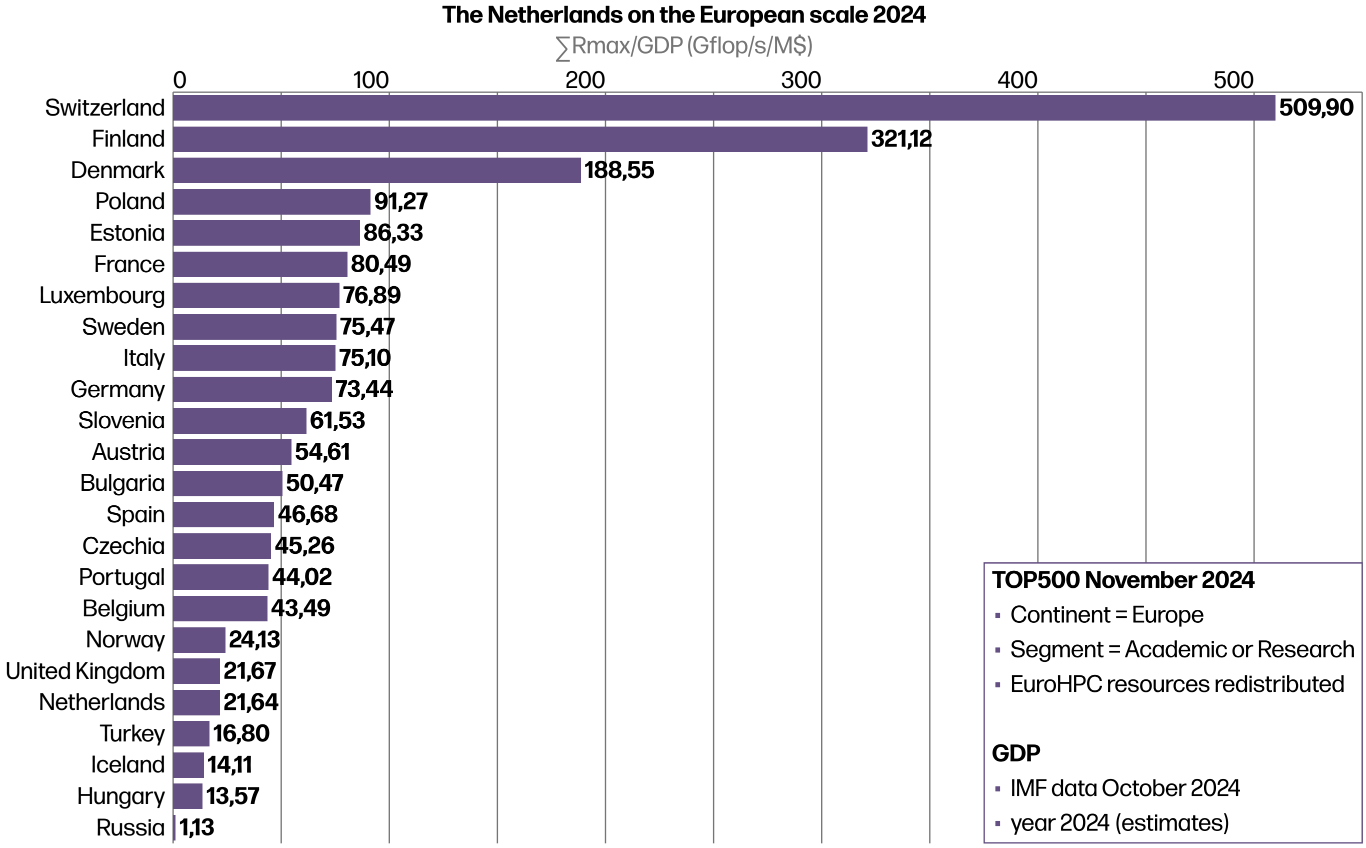
According to most of the respondents, HPC facilities should remain on Dutch soil in the future. The main reason mentioned was the need to process sensitive data in several scientific fields. A number of data providers, such as Statistics Netherlands, the Royal Library and hospitals, do not allow data to be processed abroad. Especially for sensitive data such as personal or business information, these data providers require that the data is stored and processed in a trusted environment in their own country. Digital sovereignty was also mentioned, both in terms of geopolitics (protection of knowledge) and dependence on Big Tech (vendor lock-in risks). Other reasons given for physical HPC facilities in the Netherlands included low barrier access to the infrastructure itself and accessibility of support, performance speed, retention of expertise in the Netherlands and benefits for education, control over design and management of the infrastructure, and technical reasons. For a small number of the respondents, the physical location of the facility should at least be within the European Union or was considered irrelevant.

Some of the respondents expressed concern about structural and adequate funding for HPC facilities, and research infrastructures in general. These concerns reflect the actual international position of the Netherlands in terms of available capacity. The computing capacity available for publicly funded scientific research in the Netherlands is indeed lagging behind compared to other European countries[7]. Many countries are currently expanding their supercomputing facilities. Obvious examples in Europe are France and Germany, but there are other, such as the United Kingdom, Switzerland, Finland, the Czech Republic and Poland, which either already have large new facilities in operation, or are planning to establish some in the short term. The Dutch national supercomputer Snellius, although recently upgraded, is lagging behind, with its successor not scheduled until 2027.

---

[7] See Figure 1.

Figure 1: European country positions based on installed supercomputer peak performance (in Gflop/s) versus gross domestic product in 2024[8].



In recent years, we have seen that both hardware and operating costs have been difficult to predict and have fluctuated greatly due to external factors. Because of the geopolitical situation, energy costs have doubled and high inflation rates have significantly driven up operating costs. As a result, current operating costs can no longer be fully covered by available financing. Hardware prices have also risen, partly due to the focus on AI, which has resulted in huge demand for GPUs. Therefore, to achieve the desired performance growth for a new supercomputer, the level of investment must be increased. While short-term compensation is helpful, a new sustainable financial model is needed.

Another issue affecting the digital infrastructure for research is the current pressure on the job market. The scarcity of (technical) expertise has become more urgent in recent years. This calls for more efficient ways of collaborating and sharing scarce resources.

---

[8] Picture generated via top500.org.

03

What is needed

HPC facilities are a prerequisite for many scientific research disciplines in the Netherlands. Researchers note that access to state-of-the-art digital research infrastructures is necessary to address and realise their scientific ambitions. In their responses, researchers reported a significant increase in demand for computing capacity (both HPC and HTC), (long-term) data storage facilities and appropriate expert support. This applies both to the most demanding 'traditional' users and to scientific disciplines less experienced in the use of digital infrastructures. Dutch participation in international large-scale research instruments also stimulates demand. The emergence of AI technologies further emphasises the need for capacity expansion. Respondents also identified specific requirements, such as facilities suitable for processing and storing sensitive data and user-friendliness.

Furthermore, as research increasingly depends on access to large-scale digital infrastructures, and given other countries' efforts, it will be necessary to strengthen national digital infrastructures for research. All in all, significantly more computing and data capacity will be needed to maintain and strengthen the position of Dutch scientific research at a global level. There should also be enabling elements in the digital infrastructure for research, such as network infrastructures (local, national and international), authorisation and authentication, and more efficient ways of organising the ecosystem. HPC demands should always be considered in the broader perspective of the whole ecosystem of digital infrastructures for research at a local, national and European level.

04

Elaboration of scientific ambitions and infrastructure needs

This section describes in more detail the challenges and needs of eight scientific disciplines that rely heavily on current computing facilities. This elaboration is based on consultations with a number of principal investigators of Dutch research groups and input from several SURF experts. We will briefly discuss the scientific challenges of the disciplines, the implications for computing, data and other elements of the digital infrastructure, future developments, international collaborations and partnerships, and the impact on society.

The following disciplines are considered

- Chemistry and materials sciences;
- Physics (including high-energy physics, as a major user of HTC facilities);
- Technical sciences and engineering;
- Life sciences, computational biology (including genomics) and health (including medical sciences);
- Earth sciences, environment and climate;
- Astronomy (including radio astronomy, as a large user of the SURF high-throughput computing environment);
- Computational sciences (including mathematical sciences and computer science);
- Social sciences and humanities.

4.1 Chemistry and materials sciences

The computing requirements for (theoretical) chemistry and materials sciences are driven by the increasing complexity of simulations, the need for higher accuracy and the integration of advanced computational techniques. In more detail, this holds for various sub-areas in which Dutch researchers are active.

Sub-areas

- *Quantum chemistry calculations:* Accurate quantum chemistry calculations, such as density functional theory (DFT) and post-Hartree-Fock methods;
- *Molecular dynamics simulations:* Large-scale molecular dynamics (MD) simulations, especially those involving long timescales and large systems;
- *Materials modelling:* Simulating the properties of materials, including electronic structure, mechanical properties and thermal behaviour;
- *Nanostructure properties:* Simulating the electronic, mechanical and thermal properties of nanostructures and nanocomposites;
- *Energy materials:* Modelling the properties and behaviour of energy materials, such as batteries, fuel cells and photovoltaics to improve performance and efficiency;
- *Catalysis:* Modelling efficient catalysts for a wide range of applications (e.g. improvement of energy conversion processes, design of new materials);
- *Biomolecular processes:* Simulating the interaction between biomolecules to understand biological function and design new drugs.

Developments

- Extending simulations to increasingly complex, realistic systems;
- Introducing or furthering the use of AI methods, combined with increased use of GPUs, but also the need for appropriate software layers and APIs;
- Increasing interest in the application of quantum computing for specific applications.

Digital infrastructure demands

- Increasing time resolutions from fs to ps to ns[9] for quantum-based simulation models with a moderate number of atoms will require orders of magnitude more computing power;
- Increasing the number of atoms and time resolution to ms in molecular dynamics simulation also requires orders of magnitude more computing power;
- Catalysis simulations are characterised by multiple time scales (from fs to seconds) and also require much more computing power than is currently available;
- Improvement of currently used software, such as Gaussian, VASP, Gromacs, NAMD, etc.;
- Expertise in the efficient use of GPUs;
- Expanded data storage requirements.

International collaboration on large-scale research Instruments

- EMBL (European Molecular Biology Laboratory;
- European XFEL (X-ray Free-Electron Laser);
- CERN (ALICE experiment);
- ESRF (European Synchrotron Radiation Facility);
- European Magnetic Field Laboratory (EMFL);
- European Spallation Source (ESS).

Impact

- More efficient and improved medicine design, shorter drug development cycles and reduced R&D costs;
- Understanding of the structure and behaviour of matter at the molecular and atomic levels. A deep understanding of matter at the molecular and atomic levels is essential to drive innovation, solve global challenges and improve our quality of life;
- Development and characterisation of new materials for applications in electronics, aerospace, medicine, etc.;
- Improved battery performance to accelerate the transition to renewable energy.

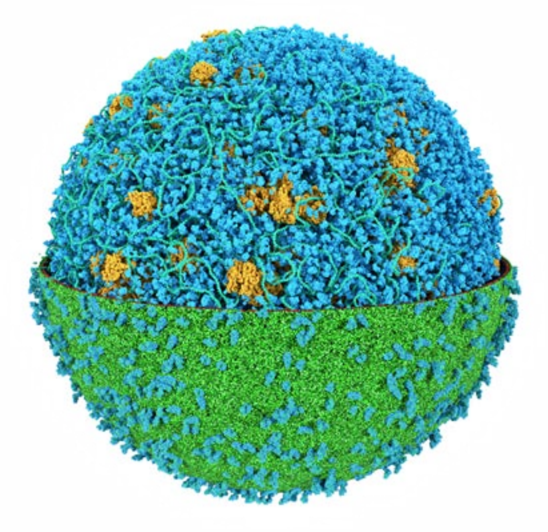
---

[9] fs: femtosecond, ps: picosecond, ns: nanosecond.

Simulation of whole biological cells (Siewert-Jan Marrink, Molecular Dynamics Group, University of Groningen)[10]

The Marrink group is currently using molecular dynamics to simulate whole biological cells. Even the smallest and simplest cells already contain billions of atoms, featuring hundreds of different types of biomolecules (e.g. proteins, DNA, metabolites, lipids). Understanding the complex interplay of these components that underpins the concept of life is incredibly challenging, and large-scale computer simulations play a key role. A next-generation computational infrastructure is essential to meet this challenge and push the boundaries of fundamental knowledge with potential applications in the design of synthetic cells for industrial purposes.

Figure 2: Molecular dynamics calculation of a complete cell: 600 million particles, the world's largest biomolecular simulation. It contains 77000 proteins (blue), single chromosome encoding for 500 genes (cyan), 500 ribosomes (orange) and millions of lipids (green).



4.2 Physics

Computing in physics is driven by a combination of data processing, systems simulation, computational reconstruction and large-scale data storage. Although the scale of the computing challenges differs between the various areas in the physics domain, all have significant computing and storage requirements, driven by the increasing complexity and scale of the experiments. One very demanding sub-area in physics is high-energy physics (HEP), which studies the fundamental particles and forces of nature, typically through experiments using particle accelerators and detectors. The computing and storage requirements for HEP are immense, both because of the volume of data collected (tens of petabytes per year) and because of the need for detailed simulated modelling of the physical processes needed for interpretation. Other areas of physics (e.g. gravitational waves), although smaller in terms of data volume than HEP, also have significant computing and storage requirements. Understanding the structure and interaction of matter at different scales is pushing the envelope for computational modelling of complex systems.

Sub-areas

- *High-energy physics:* analysing experiments in the Large Hadron Collider, exploring the properties of the Higgs particle, explaining the anti-matter asymmetries in the universe and probing new 'beyond-standard-model' physics. This involves massive amounts of data, currently adding tens of petabytes annually;
- *Condensed matter and interfaces:* quantum materials, topological matter, superconductivity, (spin)-proximity interactions and applications in nanoelectronics and spintronic devices;
- *Plasma physics:* advances in plasma physics, including for understanding and controlling the fusion process in the ITER reactor;
- *Magnetohydrodynamics (MHD):* studies of the dynamics of electrically conducting fluids such as plasmas;

---

[10] frontiersin.org/journals/chemistry/articles/10.3389/fchem.2023.1106495/full.

- *Astroparticle physics and gravitational waves:* detecting and analysing high-energy neutrinos and understanding the cosmic sources of neutrinos, detecting the sources of the ripples in space-time in our universe and comparing them to large numbers of simulated cosmic events for interpretation.

Developments

- Accelerated computational co-processing to be incorporated into all stages of the analysis of particle interaction data resulting from LHC and HL-LHC experiments, estimated to increase by 20%-30% per year;
- Combining 'high-throughput' data-intensive computing with high-performance 'fast-interconnect' computing for combined analysis in the same scientific workflow;
- Machine learning inference and AI for analysis, but also the need for appropriate software layers and APIs;
- Entering the exascale era of computing and data, with emphasis on scaling and optimising the performance of application codes.

Digital infrastructure demands

- In HEP, data volumes are expected to increase significantly as current experiments are upgraded and new experiments come online. The High-Luminosity LHC (HL-LHC), set to start in the late 2020s, will produce an order of magnitude more data than the current LHC;
- Solutions must include hierarchical storage systems combining disk and tape storage to manage cost and performance. Distributed computing resources are also crucial;
- Simulations of particle interactions and reconstruction of particle trajectories from detector data are computationally intensive tasks. Advances in detector technology and higher collision rates will require a more powerful computing infrastructure;
- HEP research will necessitate upgrades in network infrastructure (both local and international) to support faster data transfer rates. High-speed optical networks and dedicated research networks such as CERN's LHC Optical Private Network (LHCOPN) will be crucial;
- Gravitational wave analysis requires combining HPC simulations of cosmic events (multi-parameter template simulations in spiralling black holes, neutron star mergers) with AI-accelerated event detection and high-throughput matching of measured data to these templates for interpretation;
- High resolution and small timesteps for modelling plasmas will require orders of magnitude more powerful computing systems.

International collaboration on large-scale research instruments

- CERN (European Organization for Nuclear Research). SURF/Nikhef is one of the worldwide Tier-1 centres in the WLCG collaboration for the LHC;
- ITER (International Thermonuclear Experimental Reactor);
- KM3NeT (Cubic Kilometre Neutrino Telescope);
- LIGO-Virgo-KAGRA Scientific Collaboration (International Gravitational Waves Network);
- Pierre Auger Observatory (Argentina);
- Xenon Dark Matter Project;
- DUNE (Deep Underground Neutrino Experiment).

Impact

- Better understanding of elementary particles and their interactions;
- Grasping the nature of complex composite matter;
- Improved understanding of the behaviour of matter under extreme conditions;
- Better understanding of dark matter and the nature of our universe.

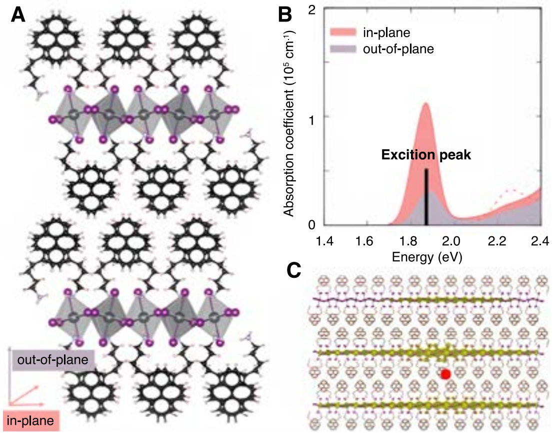
CERN's High Luminosity Large Hadron Collider ESFRI Landmark (David Groep, Nikhef and Maastricht University)

The High Luminosity upgrade will increase the luminosity ('intensity') of the Large Hadron Collider (LHC) by a factor of 10 beyond the LHC's design value. Not only will the detectors then produce raw data at a rate that exceeds by an order of magnitude the volume collected today, but the increased complexity of the data will also increase the processing requirements by several times that factor. Filtering at the source can be handled by using GPU processing, a development that will be extended to globally distributed analysis in the coming few years. The computationally intensive process of finding 'tracks' of particles in what are essentially high-resolution images with dozens of 'overlapping exposures' will combine GPU processing and machine learning inference to deal with this complexity. The key challenge in designing the data processing pipeline will be throughput: with up to 50 gigabytes/second per experiment, even when distributed across 15 sites globally, the platform will require the co-design of networking, storage, processing and science algorithms.

Excitons in low-dimensional metal-halide semiconductors (Linn Leppert, University of Twente)[11]

Excitons are correlated electron-hole pairs that form in semiconductors upon absorption of photons and are key for understanding light conversion and designing materials with tailored properties for a wide range of applications covering photovoltaics, sensing and quantum information. Low-dimensional metal-halide perovskites are an ideal platform for studying excitons due to their tremendous compositional and structural diversity, wide range of tunable properties, facile synthesis and robust stability. Numerical modelling of excitons in such materials from first principles, i.e. without recourse to empirical models, has been challenging thus far due to the system size, complex electronic structures and pronounced structural disorder. However, GPU-accelerated first-principles density functional theory and Green's function-based many-body perturbation theory have recently enabled us to make significant progress in predicting the excitonic properties of low-dimensional perovskites of growing complexity. These calculations provide atomistic insight into the formation, transport and dissociation of excitons, aid the interpretation of experimental results, drive the development of new numerical methods, and inform material design for tailored applications.

Figure 3: A) Crystal structure of the quasi-two-dimensional (pyrene-C4)2PbI4 semiconductor. The in-plane and out-of-plane directions are marked in pink and purple, respectively. B) Optical absorption spectrum indicating the pronounced peak arising from an exciton. C) Probability distribution of the exciton wave function (yellow).



4.3 Technical sciences and engineering

Computational Fluid Dynamics (CFD) and engineering computing requirements are expected to be significant from 2025-2030 due to the increasing complexity of simulations, the need for higher accuracy, and the integration of advanced technologies such as machine learning and real-time analysis.

Sub-areas

- *Aerodynamics:* Detailed analysis of aerodynamic performance, including drag reduction, lift optimisation and noise reduction;
- *Turbulence simulation:* Direct Numerical Simulation (DNS) of fluid flows around aircrafts and vehicles with higher Reynolds numbers than is practical today;
- *Thermal management:* CFD for optimising cooling systems in vehicles, including battery cooling for electric vehicles;

---

[11] L. Leppert, Excitons in Metal-Halide Perovskites from First-Principles Many-Body Perturbation Theory, J. Chem. Phys. 160, 050902 (2024).

- *Multiphase flows:* Dynamics of and interactions between different phases (e.g. gas-liquid, liquid-solid);
- *Energy transition:* Hydrogen production via water and/or CO2 electrolysis;
- *Multi-scale simulations:* Simulations that capture different spatial and time scales involved in specific processes (e.g. transport);
- *Micro- and nanofluidics:* Simulation of the behaviour and flow of fluids (bubbles, drops etc.) through micro- and nanostructures;
- *Wind engineering:* Simulation of wind effects on structures such as buildings, bridges and wind turbines (including turbulence modelling);
- *Water flow:* Modelling water flow in natural and engineered systems, including flood prediction and irrigation systems;
- *Geophysical fluid dynamics:* Melting of glaciers, ocean currents, cloud formation, climate modelling, exchange of CO2 between atmosphere and ocean, internal dynamics of earth and the planets;
- *Structural engineering and design:* Built environment, building physics, real estate, urban design and urban systems.

Developments

- Further use of AI methods, combined with increased GPU usage;
- Algorithm improvement and adaptive mesh refinement implementations.

Digital infrastructure demands

- DNS methods with larger Reynolds numbers will require an order of magnitude more computing power;
- Large memory requirements to cover billions of grid points and achieve higher resolutions;
- Expertise to support the transition from traditional CPU use to GPU and parallel computing.

International collaboration on large-scale research instruments

- VKI (von Karman Institute for Fluid Dynamics); and
- ERCOFTAC (European Research Community on Flow, Turbulence and Combustion);

Impact[12]

- Climate change: Predicting warming and impacts of greenhouse gas emissions on the oceans, sea levels and weather;
- Environment: From water, river and flood management, including pollution and water treatment, to cooling, heating and air management in buildings;
- Energy transition and energy security: From tidal and wind energy, the hydrogen economy and biofuels, to next-generation batteries and high-efficiency transport;
- High-tech manufacturing and high-tech materials, from semiconductors to additive manufacturing, 3D printing and biocompatible production;
- Healthcare: From aerosol distribution of viruses, advanced microfluidic-based diagnostics to personalised treatment and drug production;
- Agriculture and food production: From next-generation indoor vertical farming to food production and processing – areas ideally suited for development in the Netherlands.

Melting rate of icebergs and glaciers (Detlef Lohse, Physics of Fluids, University of Twente)

Understanding the melt rate of icebergs and glaciers is one of the great challenges in environmental fluid dynamics. How do factors such as salinity, temperature and size influence this process? Current models are off by an order of magnitude or more. How do you scale up knowledge from laboratory experiments? The lack of understanding is at a fundamental level. The reason for this is that ice melting is a complex, multi-scale, multi-physics phenomenon. The flow is multiphase and multicomponent, involving phase transitions, multi-way coupling and memory effects. Progress can only be achieved by a concerted effort between experiments, field measurements and numerical simulations, which play a crucial role in trying to understand the ice melting processes and in developing mitigation strategies. In our current project, we aim to achieve a quantitative understanding of melting and dissolution processes in multicomponent, multiphase systems, across scales and at a

---

[12] Source: Flow to the Future: Fluid Dynamics in the Netherlands, J.M. Burgerscentrum, 2023.

fundamental level. To achieve this, we perform numerical simulations for idealised setups, allowing a one-to-one comparison between experiments and numerics/theory.

Renewable energy (Bendiks-Jan Boersma, Mechanical Engineering, Delft University of Technology)

Hydrogen and CO2 electrolysis will play a crucial role in the transition to a renewable society. The products of electrolysis, powered by renewable electricity – hydrogen and hydrocarbons – will serve as the building blocks for renewable fuels and as renewable feedstock for the chemical industry. Electrolysis is a complex physical process that integrates the principles of electrochemistry, thermodynamics and fluid mechanics. While hydrogen electrolysis, and to a lesser extent CO2 electrolysis, are relatively well understood on a small to moderate scale, the primary challenge in the coming decades will be scaling up the electrolysis process from the megawatt to the terawatt scale without excessive material consumption. Achieving this goal will require an immense effort that combines experimental research, numerical simulations and modelling.

4.4 Life sciences, computational biology and health & disease

The computing requirements for health research and medical applications in 2025-2030 are expected to be substantial due to the integration of advanced technologies, increased data volumes and the need for more sophisticated analysis.

Sub-areas

- *Bioinformatics:* Processing large genomic data sets, running complex bioinformatics algorithms and performing simulations of biological systems;
- *Medical imaging:* Image reconstruction, enhancement and analysis resulting from advanced imaging techniques such as MRI, CT and PET;
- *Drug discovery:* Performing molecular simulations, virtual screening and modelling drug interactions that can accelerate the drug discovery process;
- *Computational biomedicine:* Developing data-driven multi-scale models in order to study the human body as a dynamic, multi-scale, complex system;
- *Epidemiology and public health:* Simulating the spread of infectious diseases, evaluating intervention strategies and informing public health policies;
- *Clinical decision support:* AI-driven decision support systems to assist clinicians in diagnosing diseases, planning treatments and predicting patient outcomes.

Developments

- Even greater use of AI techniques, most likely also in clinical situations;
- Use of highly sensitive health data that cannot be anonymised, such as genomic data, provided that privacy is guaranteed;
- EU-wide movement towards the sharing of highly sensitive health data for R&D to speed up innovation;
- The Virtual Human Twin has been identified by the European Commission in December 2023 as a flagship programme at the intersection of health and IT.

Digital infrastructure demands

- Higher spatio-temporal resolution in multi-scale models (from the molecular level to cells, tissues, organs and the whole body) will require orders of magnitude more computing power;
- The absolute need for sensitivity analysis and uncertainty quantification of computational models in the life sciences and medicine will also require orders of magnitude more computing power;
- Increasingly large data sets of genomic and imaging data will require large-scale data storage facilities and fast interconnections between federated storage and supercomputing sites;
- Rapid growth in AI use and user training;
- Hierarchical model of computing resources, ranging from clinical facilities to centralised facilities (edge-computing);
- Large-scale facilities for compliant storage, analysis and the sharing of highly sensitive data sets.

International collaboration on large-scale research instruments

- Health-RI (Dutch Health Research Infrastructure);
- HGP (Human Genome Project);
- EMBL (European Molecular Biology Laboratory);
- EBI (European Bioinformatics Institute);
- ELIXIR (European Life-sciences Infrastructure for Biological Information);
- WeNMR (Worldwide e-Infrastructure for Nuclear Magnetic Resonance Spectroscopy;
- Alzheimers Genetics Hub;
- Project MinE.

Impact

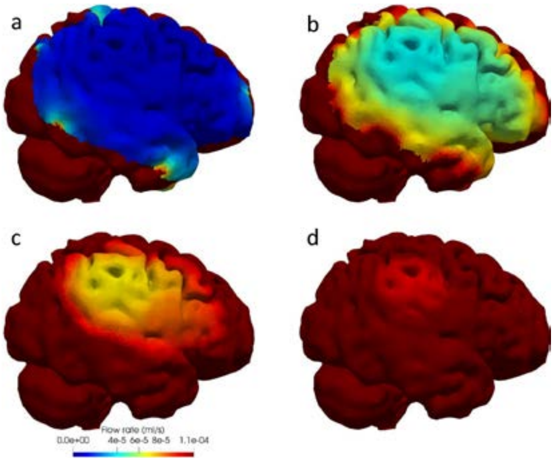
- Better understanding of the physiology and pathophysiology of the human body;
- Better and more accurate drug design;
- Insight into the spread of disease, leading to appropriate interventions;
- Personalised medicine;
- Translational research with applications to clinical decision support or policy support.

Digital twin for the cerebrovascular system (Alfons Hoekstra, University of Amsterdam)

A Digital Twin in Healthcare (DTH) for the cerebrovascular system is being developed at the University of Amsterdam. This is an integrated multi-scale, multi-organ model for cerebral blood, brain perfusion and metabolism, and blood flow and thrombosis along the heart-brain axis. This DTH is applied to acute stroke, both acute ischaemic stroke and haemorrhagic stroke, among others. One example is an in-silico trial for acute ischaemic stroke that simulates the development of infarcted brain tissue after stroke onset, simulates treatment and estimates the clinical endpoint six months after treatment.
In such an in-silico trial, a virtual cohort of a few thousand virtual patients is generated, and for each of these patients a full-blown simulation is carried out, after which the results are compiled in a trial outcome. Each individual virtual patient requires some 8 hours of computing time on 30 cores, so running a single in-silico trial would require 0.25 million CPU hours. Such trials are used to help design more efficient treatments options and clinical trials and to reduce the use of animal testing in preclinical settings.

For more information, see the GEMINI project.[13]

Figure 4: Fractional flow rate change on the pial surface during an AIS for the various collateral scores. Collateral scores: (a) absent, (b) poor, (c) moderate and (d) good. Side view of the brain showing the right hemisphere. The thrombus is located in the right MCA and completely occludes the vessel.[14]



---

[13] The GEMINI project, dth-gemini.eu. C. Miller, R.M. Padmos, M. van der Kolk, T.I. Józsa, N. Samuels, Y. Xue, Stephan J. Payne, Alfons G. Hoekstra, Comput. Biol. Med. 137 (2021) 104802. doi:10.1016/J.COMPBIOMED.2021.104802.

[14] R.M. Padmos, N.A. Terreros, T.I. Józsa, G. Závodszky, H.A. Marquering, C.B.L.M. Majoie, A.G. Hoekstra, Med. Eng. Phys. 91 (2021) 1–11. doi:10.1016/j.medengphy.2021.03.003.

4.5 Earth sciences, environment and climate

The computing requirements for earth system science, and in particular climate research, are expected to be significant by 2025-2030, driven by the increasing complexity of earth system models, the need for higher resolution simulations, and the integration of new data sources and technologies. Also, European initiatives such as Destination Earth, creating a digital twin of earth, will require large digital infrastructures throughout Europe.

Sub-areas

- *Climate modelling:* Developing and optimising advanced earth system models capable of simulating complex interactions between the atmosphere, oceans, land surface, biosphere and ice, including increasing spatial and temporal resolution;
- *Weather prediction:* Enhancing short-term weather prediction models to provide more accurate and timely predictions, including the prediction and simulation of extreme weather events such as hurricanes, heavy precipitation events and heat waves;
- *Climate change projections:* Running long-term earth system model simulations to project future climate change under different greenhouse gas emission pathways, incorporating biogeochemical cycles (e.g. carbon and nitrogen) to study feedback mechanisms, and analysing the potential impacts of climate change on various sectors, including agriculture, water resources and public health;
- *Geodynamical modelling and seismology:* Simulation of phenomena over a huge range of both spatial (metres to thousands of kilometres) and temporal (seconds to billions of years) scales.

Developments

- Incorporation of AI technology to handle and analyse the huge amounts of data;
- Continuous improvement and application of algorithms (physics-based, data-based) in novel hybrid earth system models;
- Streaming of (near)real-time data and their assimilation into models.

Digital infrastructure demands

- Increasing resolution of simulations requires orders of magnitude more computing power; doubling the horizontal resolution in an earth system model typically requires a factor of 10 more core hours;
- The ability to handle increasing amounts of data arising from satellite observations, sensor networks and historical records requires large data storage and retrieval facilities, including real-time processing, both for climate research as well as seismology;
- Expertise for the transition to incorporate data-based (AI) models (for reconstruction, parameterisation and prediction) into earth system models and extended use of GPUs.

International collaboration on large-scale research instruments

- ECMWF (European Center for Medium-Range Weather Forecasts);
- NCAR (National Center for Atmospheric Research);
- LANL (Los Alamos National Laboratory);
- TropOMI (Tropospheric Monitoring Instrument, with KNMI and SRON);
- TANGO (Twin Anthropogenic Greenhouse Gas Observers, with KNMI and SRON);
- Destination Earth (digital twin of earth).

Impact

- Improved early warning systems to minimise damage resulting from extreme weather events and climate tipping points (see example below);
- More reliable climate change projections, offering the possibility to take appropriate long-term measures; and
- Understanding the interactions of various processes that influence the earth system, offering the possibility to simulate the effects of mitigation and adaptation strategies.

Tipping of the Atlantic Meridional Overturning Circulation (Henk Dijkstra, IMAU, University of Utrecht)[15]

The Atlantic Meridional Overturning Circulation (AMOC) plays a crucial role in regulating global and regional climates. Alarming trends show a gradual weakening of the AMOC strength in recent decades. Determining the probability of a future collapse of this large-scale ocean circulation requires the use of complex earth system models, which require massive computing and data handling capabilities. Recent simulations with one of these models have introduced a freshwater forcing into the Atlantic Ocean, causing the AMOC strength to gradually decrease until it collapsed. During the collapse, Europe cooled by about 1°C per decade, with some regions cooling by more than 3°C per decade. Compared to the current global warming rate of 0.2°C per decade, this cooling during a tipping event is unprecedented. Cooler European temperatures have far-reaching consequences, including accelerated warming in other regions and altered precipitation patterns. An abrupt AMOC collapse also leads to a 100 cm rise in European sea levels.

4.6 Astronomy

The computing requirements for astronomy in 2025-2030 are projected to be significant, driven by advances in telescope technology, increased data volumes, sophisticated data analysis techniques and increased HPC. Key considerations for meeting these requirements are outlined below.

Sub-areas

- *Radio astronomy:* Variety of astronomical studies using large-scale research instruments such as LOFAR and the Square Kilometre Array (SKA). It aims to address fundamental questions about the universe, including the formation and evolution of galaxies, dark matter and the nature of gravity. It also covers areas with obvious practical implications, such as space and solar weather;
- *Infrared astronomy:* Understanding of exoplanets, the first billion years of the universe, and many other astrophysical and cosmological phenomena, including data collected by the James Webb Space Telescope;
- *Computational astrophysics:* Emphasis on stellar dynamics, space weather, galaxies and the formation and evolution of star cluster;
- *Gravitational waves:* Analysis of results originating from the LIGO/Virgo inferometers, and possibly followed by the Einstein Telescope.

Developments

- Incorporation of AI technology to handle and analyse the huge amounts of data;
- Continuous improvement and application of algorithms;
- Investigations into energy-accuracy trade-offs;
- Continued efforts to incorporate FAIR data practices.

Digital infrastructure demands

- Research in astronomy is characterised by large amounts of data generated by instruments such as LOFAR, the James Webb Space Telescope and the upcoming SKA. These will generate petabytes to exabytes of data that will need to be transported, stored and processed. This will require orders of magnitude increase in computing power, storage and bandwidth;
- Massive computing requirements to extend large-scale simulations on stars and galaxies;
- Expertise for the transition to incorporating data-based (AI) models and extended use of GPUs;
- Ability to combine and analyse massive data collections in the pursuit of multi-messenger astronomy;
- International collaboration through federated computing centres.

International collaboration on large-scale research instruments

- LOFAR (LOw Frequency aRray);
- SKA (Square Kilometre Array);
- JWST (James Webb Space Telescope);
- ALMA (Atacama Large Millimeter Array);
- EHT (Event Horizon Telescope);
- VLT (Very Large Telescope);
- LIGO/Virgo (Laser Interferometer Gravitational-Wave Observatory/Virgo).

---

[15] Source: Science Advances, 9 February 2024.

Impact

- Better understanding of the nature of the universe;
- Better understanding of space weather conditions that affect technological and communication systems on earth;
- Unique outreach opportunities for public outreach and citizen science.

Massive star cluster formation (Simon Portegies Zwart, Leiden University)

With an international team of astronomers, Simon Portegies Zwart has been working on large-scale simulations, in which he focuses on the formation and evolution of stellar clusters.

The image below shows the results of one of their recent simulations, in which a giant molecular cloud collapses into a stellar cluster. These calculations, performed on the Dutch national supercomputer Snellius, took about 10 million CPU hours. The team is currently working on publishing the results (submitted for publication to Astronomy & Astrophysics (Polak et al., A&A, 2024)).[16]

Figure 5: Public communication of scientific results in cosmology.

This image was taken in April 2024, on a public advertising pillar in New York City. Interesting science communication!

The Square Kilometre Array (SKA) (John Swinbank, Astron, and Raymond Oonk, SURF)

The SKA Observatory[17] is a next-generation radio astronomy facility that will help revolutionise our understanding of the universe and the laws of fundamental physics. The telescopes will be located in the radio-quiet Karoo region of South Africa (SKA_MID) and the Murchison Shire of Western Australia (SKA_LOW). The SKA_MID and SKA_LOW telescopes, taken together, will be capable of producing a stream of 700 PB/year of scientific data products. This is an increase of two orders of magnitude over current facilities, such as LOFAR. Such an extreme volume of data and the associated processing is unprecedented for the astronomical community. To meet the unique challenges posed by this system, a global network of digital infrastructure nodes – termed the SKA Regional Center Network – is being developed. These nodes will be tasked with the delivery and curation of the data, as well as providing a cutting-edge data analysis ecosystem for astronomers.

4.7 Computational sciences

In October 2022, the National Agenda for Computational Sciences was presented to and adopted by the Ministry of Economic Affairs and Climate Policy. The National Agenda sets out a roadmap and focus areas for computational sciences. It defines computational science as an interdisciplinary field that deals with the use of computer modelling and simulation to solve complex problems in various scientific and engineering disciplines.

Indeed, computational sciences are the underlying basis for progress in many, if not all, of the scientific disciplines we have covered so far in this document. Without numerical algorithms, implementation in computer languages, error estimates, performance modelling and optimisation, parallel computing and numerical library development, these scientific disciplines would not have been able to make the progress they have made. The National Agenda focuses

---

[16] Preprint available at: ui.adsabs.harvard.edu/abs/2024arXiv240512286P.

[17] skao.int/en

on themes that cut across and underpin many scientific disciplines. These are:

Focus areas

- *Multiscale modelling and simulation:* Predicting the behaviour of large-scale and complex systems, in which different length and time scales communicate with each other;
- *Data-driven methods:* Distilling good models from large amounts of data to enable insight, design and control/intervention;
- *Ultra-fast computer simulations through machine learning:* Use of this computationally intensive technique to enable extremely accurate and complex calculations;
- *Uncertainty and sensitivity analysis:* Properly assessing the reliability of computer models;
- *Energy-efficient computing:* Reducing the carbon footprint through savings in the computational area itself.

Developments

- Entry into the era of exascale computing, which requires new levels of parallelism in applications for efficient use;
- New quantum algorithms in many fields, e.g. cryptography and materials science, potentially solving problems that are currently intractable for classical computers;
- Advanced AI models, GPU use and algorithms for new processor technologies.

Digital infrastructure demands

- Collaboration between researchers in computation and data-intensive scientific disciplines with computational science experts;
- Availability of high-end computing and storage systems to test new developments;
- Scale-up of applications and algorithms to efficiently use high-end computing systems and exploit large data sets.

International collaboration on large-scale research instruments

- Scientific Large-Scale Infrastructure for Computing / Communication Experimental Studies (SLICES).

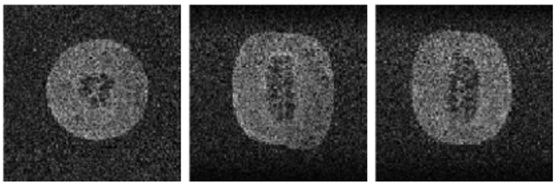
Impact

- Supporting many scientific disciplines through numerical methods and algorithm development;
- Addressing water, food and energy challenges to move towards a sustainable future.

Computational imaging (Martin van Gijzen, Computational Sciences, Delft University of Technology)

Computational imaging is the field of computational science that uses computational tools to create images, improve their quality and extract information. In many applications, images are computed from measurement data. Examples are MR and CT imaging and geo-acoustic inversion. Typically, the computation of the raw image can be formulated as the solution of a least squares problem. Raw images can be of very poor quality and require further processing to be of acceptable quality. Figure 6 gives an example of a three-dimensional scan of a melon taken with a low-cost permanent magnet MRI scanner. It is clear that the image is heavily polluted by noise.

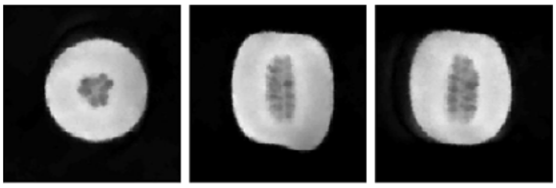
Figure 6: Raw image of a melon.



A quite successful approach to reducing noise in an image is to think of the pixel values as heat, and the noise reduction problem as a heat diffusion problem. The diffusion of heat makes the image smooth. To avoid diffused at the edges, a diffusion coefficient must be used that is small near the edges. Mathematically, this amounts to numerically solving a nonlinear diffusion equation. The two ingredients for computing a good quality image can be combined to an optimisation problem that balances the errors in the least squares problem and the noise reduction problem. To keep the solution time close to real time, special computational techniques must be used that are amenable to parallel computing.

Figure 7 shows the final result of such a computation. The image of the melon has been enhanced to a level that allows us to count the number of seeds.

Figure 7: Noise-filtered image of a melon.



The above example shows the importance of computational science in image processing. To achieve the final result, we used a combination of algorithms from numerical mathematics, optimisation, linear algebra and high-performance computing.

4.8 Social sciences and humanities

The digital research infrastructure in the Social Sciences and Humanities (SSH) domain is lagging behind and often outdated, while the need for high-quality infrastructure is growing as a result of the increasing digitalisation of research materials, the adoption of advanced analytical techniques, and the integration of large and complex data sets.[18] Key considerations for meeting these needs are outlined below.

Sub-areas

- *Humanities research in digital archives:* Analysis of large, already digital archives, such as web archives, email archives and social media to signal and predict social behaviour, increasingly using AI techniques;
- *Linguistics and languages:* Analysis to create a deeper understanding of society and the human cultural record through large-scale processing of digital records of historical texts and spoken and/or written interaction;
- *Literature and the arts:* Analysing texts to understand themes, narrative techniques and cultural contexts;
- *Economic modelling:* Analysing economic systems, predicting market trends and macroeconomic forecasting, supporting policy development and financial planning;
- *Public policy evaluations:* Predicting and evaluating policy decisions using behavioural simulation and agent-based modelling.

Developments

- As more data becomes available in digital form, the demand on storage and computing capacity will grow rapidly. SSH research will become a significant user of high-end computing and data storage facilities;
- Increased use of digital technologies, including artificial intelligence (AI), can make research faster and more efficient and create opportunities for new types of research;
- The growing amount of data available for research and the rise of machine learning are strengthening the opportunities for data science;
- Further use of AI algorithms.

Digital infrastructure demands

- Given their focus on people, their interactions, and the products of their minds and culture, SSH share a need for secure computing facilities to analyse sensitive and copyright-protected data and artefacts;
- User-friendly infrastructure;
- Large-scale storage for extensive data sets, together with services for collecting, sharing, donating and linking sensitive data in compliance with data security and privacy regulations;
- Increasing demand for computing facilities, as a result of the digitalisation of vast amounts of data;
- Coordinated approach to integrating AI techniques into current research;
- Visualisation tools to gain insight into the analysis of large unstructured data sets.

[18] Sectorplannen 2022/2023 SSH: sshraad.nl/sectorplannen-2022-ssh/.

International collaboration on large-scale research instruments

- DARIAH (Digital Research Infrastructure for the Arts and Humanities);
- CLARIN (Common Language Resources and Technology Infrastructure);
- ESS (European Social Survey);
- SHARE (the Survey of Health, Ageing and Retirement in Europe);
- CESSDA (Consortium of European Social Science Data Archives);
- GGP (Generations and Gender Program);
- SoBigData Europe (European open science research infrastructure for social mining);
- CLARIAH (Common Lab Research Infrastructure for the Arts and Humanities);
- NICAS (Netherlands Institute for Conservation+Art+Science+).

Impact

- Better understanding of historical and social developments;
- Helping government and businesses to make better policy decisions;
- The ability to securely and ethically link and analyse a huge range of data such as historical records, textual data, images, survey data and social media data will help SSH research address some of the most pressing issues facing society, such as polarisation, social inequality and behaviour towards environmental change.

ODISSEI secure supercomputing facility (Tom Emery, Erasmus University Rotterdam and director ODISSEI)

At ODISSEI, we have seen how the creation of the ODISSEI secure supercomputing facility has transformed the way social scientists conduct analysis. The high security standards have allowed us to bring a whole new range of complex and exceptionally rich data into computationally powerful environments where they can be integrated into AI research. The creation of a large-scale national AI facility that adheres to these security standards would allow us to train on a far broader and more informative range of data than is currently possible anywhere else in the world. Imagine being able to train an AI model on every single tax return and medical prescription from the last 20 years, and the kinds of problems this could help us solve. This is entirely feasible with a secure, large-scale national AI facility.

Appendix A – Consulted researchers and SURF experts

The following researchers have been consulted:

- Prof. dr. ir. Bendiks Jan Boersma, Delft University of Technology
- Prof. dr. ir. Henk Dijkstra, University of Utrecht
- Dr. T.E. Emery, Erasmus University Rotterdam and director ODISSEI
- Prof. dr. ir. Martin van Gijzen, Delft University of Technology
- Prof. dr. David Groep, Nikhef and University of Maastricht
- Prof. dr. ir. Alfons Hoekstra, University of Amsterdam
- Prof. dr. Linn Leppert, University of Twente
- Prof. dr. Detlef Lohse, University of Twente
- Prof. dr. Siewert-Jan Marrink, University of Groningen
- Prof. dr. Simon Portegies Zwart, Leiden University
- Dr. John Swinbank, ASTRON
- Prof. dr. ir. Kees Vuik, Delft University of Technology

The questionnaire was filled in by 43 respondents from 21 organisations. The following disciplines were covered:

- Artificial intelligence
- Climate research
- Fluid dynamics
- Life sciences
- Chemical engineering
- Material sciences
- Astroparticle physics
- Astrophysics
- Biotechnology
- Earth and environmental sciences
- Health imaging technical research
- Humanities,
- Mechanical engineering
- Medical imaging
- Medicine
- MRI
- Particle physics
- Urban design
- Algorithm/software Development
- Applied mechanics (multi-scale + multi-physics modelling)
- Architectural engineering
- Astronomy
- Atmospheric physics
- Autonomous driving
- Biochemistry
- Bioinformatics
- Building physics and services
- Built environment
- Chemistry
- Condensed matter Physics
- Data science
- Economics
- Energy
- High-energy physics
- HPC
- Law
- Machine learning
- Mass spectrometry-based proteomics and metabolomics
- Mathematics
- Medical sciences
- Numerical analysis

- Pathology
- Physics
- Process technology
- Quantum chemistry
- Scientific computing
- Social sciences
- Structural engineering and design
- Technical domain
- Theology

The following SURF experts were consulted:

- Dr. ir. Axel Berg
- Dr. Irene Bonati
- Dr. Haili Hu
- Daan de Jong M.Sc.
- Dr. Matthieu Laneuville
- Dr. Annette Langedijk
- Walter Lioen M.Sc.
- Dr. ir. Peter Michielse
- Dr. Raymond Oonk

This user consultation was conducted by Ryanne Keltjens (SURF), Naomi Messing (NWO), Peter Michielse (SURF) and Femke Stephan (NWO) under the supervision of the NWO Digitalisation Research Advisory Committee.

Appendix B – Questionnaire

NWO and SURF are currently conducting a general mapping exercise to get a broader overview of the data and computing needs of researchers in HPC, data processing, cloud computing and data storage, in addition to the needs of AI-related research. The purpose of this mapping is to describe the broader landscape for data and computing in the Netherlands via a science case for public scientific research in the Netherlands to the Ministry of Education, Culture and Science (OCW).
This questionnaire will provide a first general view of the data and computing needs for digital infrastructures. The results will be used to strengthen the science case. NWO and SURF are interested in your opinion about the HPC user needs and would like to ask you to fill in the questionnaire below. Answering the following questions will take about 5-10 minutes. The results will be processed anonymously. The respected deadline for submitting the questionnaire is Friday 30 August 2024.
Thank you!

Please find below the questionnaire:

Name and organisation

Name

Organisation

In which scientific discipline do you work?
(E.g. climate research, fluid dynamics, astronomy, high-energy physics, materials science, humanities, ...)

Users

1. Who are the current users of HPC, data processing, HPC cloud computing within your organisation; which digital infrastructural service(s) do they need (HPC, data processing, HPC cloud computing, etc.)?

Curent users *(faculty/research group etc)*	Requires digital infrastructure services *HPC, data processing, HPC cloud, etc.*

2. Do you expect your research group to expand in the coming years, increasing the number of users accessing digital infrastructure?
☐ No ☐ If yes,
please explain and provide a rough indication in percentage of the expected growth in demands for data and computing use over the next 10 years?

User demand for data and computing up to 2035

3. How essential is the growth in demand for data and computing to continue doing top scientific research? Please distinguish between 'nice to have' and 'necessary'.
☐ Nice to have ☐ Necessary Please explain

4. Do you foresee bottlenecks emerging in your organisation(s) in terms of data processing, storage and/or available computing capacity before 2035?
☐ No ☐ Yes, namely

5. What do you consider to be the ideal ratio between CPU and GPU computing capacity?

6. Which trend do you see in the use of/need for of GPUs within your research group/organisation?

7. Do you see a significant increase in the use of AI techniques in your group/organisation's applications?
☐ No ☐ Yes

Other relevant questions

8. In addition to data storage and computing capacity, are there any other aspects of digital infrastructure that are critical to your domain? For example, network needs, etc.
☐ No ☐ Yes, namely

9. Are there any non-infrastructure issues that are critical to the optimal use of data and computing infrastructure (e.g. protection of knowledge, availability of scientific talent, availability of technical support, data and software support, etc.)?
☐ No ☐ Yes, namely

10. Is it important to you that the infrastructure for your research domain is physically located in the Netherlands? If so, why?
☐ No ☐ Yes, namely

11. Is there anything else you would like to mention?

Thank you for your time!

Appendix C - Use and impact of artificial intelligence

Artificial intelligence (AI) is proving to be a transformative technology in a wide range of scientific fields. AI is most promising in scientific areas where large data sets and complex patterns are prevalent, where predictive analytics can significantly enhance decision-making, and where automation can improve efficiency and outcomes. The interdisciplinary nature of AI allows it to contribute to advances in a wide range of scientific fields, driving innovation and discovery at an unprecedented pace.

Below are some of the areas where AI holds the most promise:

Healthcare and medicine

Medical imaging

- **Diagnosis:** AI algorithms can analyse medical images (e.g. X-rays, MRIs, CT scans) with high accuracy, assisting in the early diagnosis of diseases such as cancer, cardiovascular diseases and neurological disorders.
- **Image segmentation:** AI enhances image segmentation, helping to analyse anatomical structures in detail and improve surgical planning.

Genomics and precision medicine

- **Genomic analysis:** AI accelerates the analysis of genomic data, identifying genetic variations associated with disease and aiding in the development of personalised treatments.
- **Drug discovery:** AI models predict the efficacy and safety of new drug compounds, significantly reducing the time and cost of drug development.

Clinical decision support

- **Predictive analytics:** AI systems predict patient outcomes, optimise treatment plans, and provide decision support to clinicians, improving patient care and outcomes.

Environmental science

Climate modelling

- **Weather prediction:** AI improves the accuracy of weather forecasts by integrating vast amounts of data from multiple sources and identifying patterns that traditional models may miss.
- **Climate change projections:** AI models simulate and predict the impacts of climate change on various ecosystems and human systems, aiding in the development of mitigation and adaptation strategies.

Environmental monitoring

- **Pollution tracking:** AI systems monitor and predict air and water quality, track pollution sources, and assess the impact on public health and the environment.
- **Wildlife conservation:** AI analyses data from sensors, cameras and satellites to monitor wildlife populations, track animal movements and combat poaching;

Astronomy and space exploration

Astronomical data analysis

- **Pattern recognition:** AI algorithms identify patterns in vast data sets from telescopes, such as the detection of exoplanets, supernovae and other celestial events.
- **Galaxy classification:** AI systems classify galaxies and other astronomical objects, helping to map the structure of the universe.

Space missions

- **Robotics and automation:** AI enhances the capabilities of robotic explorers on planetary missions, enabling autonomous navigation, data collection and analysis.
- **Satellite imagery:** AI processes and analyses satellite imagery for earth observation, contributing to research in climate science, agriculture and disaster management;

Materials science

Materials discovery

- **Property prediction:** AI models predict the properties of new materials based on their chemical composition and structure, accelerating the discovery of materials with desirable properties.
- **Simulations:** AI enhances molecular dynamics and quantum simulations, providing insights into the behaviour of materials at the atomic and molecular levels.

Manufacturing

- **Process optimisation:** AI optimises manufacturing processes by predicting outcomes, reducing waste and improving quality control.

Physics

Particle physics

- **Data analysis:** AI processes large data sets from particle accelerators such as the Large Hadron Collider, identifying rare events and new particles.
- **Simulation and modelling:** AI improves the accuracy and efficiency of simulations in theoretical and experimental physics.

Quantum physics

- **Quantum computing:** AI aids in the development of quantum algorithms and error correction techniques, advancing the field of quantum computing.

Social sciences and humanities

Behavioural analysis

- **Social media analysis:** AI analyses social media data to understand human behaviour, public opinion and social trends.
- **Economic modelling:** AI models economic systems and predicts market trends, aiding in policy development and financial planning.

Cultural heritage

- **Text analysis:** AI processes and analyses large bodies of text, uncovering patterns and insights in historical documents and literature.
- **Digital preservation:** AI helps in the digitalisation and preservation of cultural heritage artifacts, making them accessible for future research.

Agriculture

Precision agriculture

- **Crop monitoring:** AI systems analyse data from drones and satellites to monitor crop health, predict yields and optimise farming practices.
- **Pest and disease detection:** AI models detect pests and diseases early, enabling timely interventions and reducing crop losses.

Sustainable practices

- **Resource management:** AI optimises the use of water, fertilisers and other resources, promoting sustainable agricultural practices and reducing environmental impact.

Robotics and automation

Autonomous systems

- **Navigation and control:** AI enables autonomous vehicles, drones and robots to navigate complex environments, perform tasks and make decisions in real time.
- **Human-robot interaction:** AI improves human-robot interaction, enhancing collaboration in various fields, from manufacturing to healthcare.