

Call for Proposals

Innovative eScience Technologies for 'Big Science'

(eTEC-BIG)

Advancing 'Big Science' Research through Innovation in eScience and
Data Science Technologies

2019

A joint call for proposals by The Netherlands eScience Center and SURFsara



Amsterdam, March 2019

Contents

| | | |
|----------|--|-----------|
| 1 | Introduction | 3 |
| 1.1 | Background | 3 |
| 1.2 | Purpose of this call | 4 |
| 1.3 | Available budget | 7 |
| 1.4 | Validity of this call | 7 |
| 2 | Guidelines for applicants | 8 |
| 2.1 | Who can apply? | 8 |
| 2.2 | What can be applied for? | 8 |
| 2.3 | When can applications be submitted? | 10 |
| 2.4 | Preparing an application | 11 |
| 2.5 | Specific conditions | 11 |
| 2.6 | Submitting an application | 13 |
| 3 | Assessment procedure | 14 |
| 3.1 | Procedure | 14 |
| 3.2 | Admissibility and assessment criteria | 15 |
| 4 | Contact details | 17 |
| 4.1 | Contact | 17 |
| 4.2 | Other information | 17 |
| | Appendix A: eScience Center Core Technological Competences | 18 |
| | Appendix B: SURFsara Technological & e-Infrastructure Expertise | 20 |
| | Appendix C: The Research Software Directory (RSD) | 22 |
| | Appendix D: The Dutch National e-Infrastructure | 23 |

1 Introduction

1.1 Background

The Netherlands eScience Center is the Dutch national center of excellence for the development and application of research software to advance academic research.

In many areas of scientific research today, there is an urgent need for methods and tools that help manage and exploit the rapidly increasing wealth of scientific data produced by modern instruments (incl. supercomputers). This urgency is no longer felt solely in headline-grabbing and large-scale initiatives such as the Human Brain Project or the Large Hadron Collider but is rapidly permeating virtually all levels and directions of (experimental) research. Having the methods and tools available to effectively handle large quantities of research data, the process of scientific discovery – from initial idea to scientific breakthrough – is accelerated significantly, and new research opportunities may open up for the very first time. Managing, exploiting, and digitally expanding this wealth of data is a generic challenge and has spawned a new research strategy referred to as (digitally) enhanced science, or eScience.

eScience is an inherently multi-disciplinary pursuit concerned with bridging the gap between the demands of data- and compute-intensive domain research on the one hand, and the capabilities of state-of-the-art ICT on the other – ranging from high-performance computing and networking, to large-scale data management, and novel big data analysis methods. The challenge of eScience is not only to ensure that most value is gained from new research endeavours, but also to exploit the often domain-independent nature of eScience tools. As such, eScience aims to deliver sustainable and discipline overarching solutions, made available for reuse and tailored to problem-specific needs. Also, eScience aims to not reinvent the wheel, but rather to utilize and creatively combine existing methods and tools as much as possible.

The Netherlands eScience Center

The Netherlands eScience Center supports and reinforces multi-disciplinary data- and compute-intensive research in the Netherlands through creative and innovative use of eScience in all its manifestations. To achieve this, the eScience Center fosters research collaborations aiming to accelerate scientific breakthroughs by effectively utilizing advanced ICT, and by making large-scale data analysis possible across multiple disciplines. The eScience Center collaborates closely with various academic and non-academic institutes, commercial technology providers, and universities.

The eScience Center shares its ideas and the tools it develops. Together with a wide range of partner organizations, the eScience Center advances not just the research projects it funds and collaborates in, but the state of academic research in general.

SURFsara is the national center for high-performance computing, data processing and data management to support scientific research with services, expertise and innovation.

SURFsara

SURFsara is the national center for high-performance computing, data processing and data management to support scientific research with services, expertise and innovation. SURFsara creates a bridge between research and advanced ICT. It does so with a passion for scientific research in its DNA and with extensive expertise in its high-performance infrastructure. As such, SURFsara offers a full range of services, expertise, innovation and support in the field of high-performance computing, data services, data processing and analytics, visualization, machine learning and cloud services. SURFsara is part of SURF, the collaborative ICT organization for Dutch education and research.

Core Technological Competences

The eScience Center and SURFsara develop, apply, scout and maintain the ICT technologies, software and services best suited to solving state-of-the-art research questions. The set of core technological competences of the two organizations, detailed in Appendices A and B, is dynamic yet focused. This allows the eScience Center and SURFsara to serve as valuable, even essential, partners in a large variety of research projects dealing with data- and compute-intensive problems.

The technological competences of the eScience Center and SURFsara ensure optimal connectivity between the three expertise areas essential to eScience: 1) algorithmic and methodological domain expertise, 2) technological expertise and software, and 3) e-infrastructure expertise and hardware. While the competences of the eScience Center and SURFsara are overlapping, the eScience Center's expertise is primarily in categories 1 and 2; SURFsara's expertise is primarily in categories 2 and 3.

It is the intention for projects funded by the eScience Center and SURFsara to utilize and exploit these core technological competences in pursuit of their scientific goals. However, the eScience Center and SURFsara are also committed to working with project leaders, strategic associates and partner organizations to develop additional expertise, and to adapt to needs, urgencies and external developments.

In the SURF Open Innovation Lab (SOIL), SURFsara is experimenting in Proof-of-Concepts in collaboration with academia and industry, to develop knowledge and future generation e-infrastructure services and applications.

1.2 Purpose of this call

This call for proposals aims to support research and development of *innovative eScience technologies* and software associated with big data handling, big data analytics and related computational methods, driven by a direct demand from any research area that can be identified broadly by the term '*Big Science*' (see below).

A competitive proposal aims to deliver prototype eScience instruments (e.g. in the form of compute kernels, software libraries, interfaces, algorithms, methods), and will help establish or strengthen research software infrastructures, scientific workflows or service infrastructure, with the aim of enhancing and accelerating breakthroughs in the selected Big Science research area. Each project must focus on one of the 'Technological Research Directions' defined below.

Projects awarded in this call for proposals typically will be led by either:

- a technology-oriented PI from the selected Big Science research area (see 'Big Science' section below), or
- a domain-oriented PI from ICT Science (i.e. a technical discipline such as data science or computer science). *In this case, it is required to include at least one co-applicant & team member from the selected Big Science research area.*

Innovative, reusable eScience Technologies

In the context of this call for proposals, innovative eScience technologies are defined as technological (software) advancements that help enable research activities, or open up research directions, that are not easily pursued using existing methods or tools. The availability of such technologies would impact the daily practice of researchers in the selected Big Science research area, and would contribute significantly to the acceleration of the scientific discovery process therein.

First and foremost, a submitted proposal clearly expresses the research challenges underlying the development of such technologies. The need for the technologies will be substantiated by clear and urgent compute- or data-driven research questions from the selected Big Science research area. For *at least one* of these Big Science-driven research questions a proof-of-concept must be developed and verified.

On top of this, a competitive proposal will explain to what extent, and in what way, the developed technologies and software will have value to other research problems – within the same research domain, or even in other domains. Next to this potential for reuse and generalization, a competitive proposal provides a sustainability section, describing measures taken to ensure the usability and availability of the developed technologies and software beyond the duration of the project itself.

Big Science

In the context of this call, 'Big Science' broadly indicates those scientific research directions 1) whose challenges in terms of complexity and 2) whose needs in terms of data analytics and data management capabilities, as well as processing power, far exceed that of other research endeavors today. Typically, technological developments in Big Science domains are expected to arrive in – and impact – other (non-Big Science) domains only years or even decades later.

Competitive proposals must indicate why and to what extent the selected (domain) research area is compliant with this deliberately non-strict and broad definition of the term Big Science. Note that Big Science may be defined differently in other contexts, e.g. referring to big budgets, big consortia, big laboratories, big machines, etcetera. Research proposals falling under such alternative definitions are certainly welcomed, provided these *also* adhere to the broad definition underlying this call for proposals.

Technological Research Directions

In this call for proposals, three main Technological Research Directions have been selected for which advances in research and development are timely and urgent. Each application must indicate which of the Technological Research Directions the proposed work applies to (most). For all the Technological Research Directions we require proposals to take reproducibility of experiments, and reusability of software and tools, into account. Proposals are evaluated and ranked within their selected Technological Research Direction. This call aims to fund 1 proposal in each direction.

I. Scalable Machine Learning (ML) & AI

Machine Learning offers possibilities to apply dimensional reduction and information extraction to data, without the need for a-priori models to describe the data. ML can be applied to many Big Science data processing, real-time filtering, and simulation activities, for improving classification and selection of data, modeling of experimental effects and data compression, or enhancing numerical simulations, while at the same time reducing execution times.

II. Processing of Streaming Data

In an increasing number of Big Science experiments, unbounded data streams need to be processed to get (near) real-time results or be scaled down for off-line storage. In such cases, smart techniques (e.g. real-time filtering for noise removal, dimension reduction, often combined with subsequent offline data reconstruction) must be applied. These smart techniques, in turn, require detailed mapping to the complex underlying e-Infrastructure, in terms of usability, scalability and performance.

III. Large-scale (Distributed) Data Organization, Management & Semantics

Many Big Science experiments are limited by how efficient ever increasing and more heterogeneous data sets can be accessed and digested. Further complexities arise because data sets often are stored in distributed and hierarchical fashion. Given the trend towards event-based processing away from file-based processing, new models of data organization and management are needed (e.g. FAIR), taking into account e.g. the semantics of the data stored. Although not identical, similar needs arise for organization and management of software, and the reproducibility of experiments.

Open Source/Open Access

This call for proposals specifically targets proposals that aim to develop (standards-based) open source/open access solutions¹, where possible extending or working in concert with the core technological competences of Appendices A and B and made available as part of the Research Software Directory (Appendix C). For all awarded projects and under supervision of the eScience Center an effort is made to enhance development of sustainable and re-usable software solutions. To this end, additional manpower will be made available on top of the awarded grant itself (Section 1.3).

e-Infrastructure

An additional purpose of this call is to improve and expand the link between state-of-the-art scientific research questions and the capabilities of advanced e-Infrastructure

¹ Applicants are asked to endorse and follow the [NLeSC Strategy towards Publishing, Licensing, and IP](#). For alternative agreements, contact the eScience Center before proposal submission.

(e.g. high-performance computing, large-scale data storage, lightpath connectivity; see also Appendix D). For this reason, specific infrastructural support services may be provided by SURFsara or SURFnet to proposals with clear e-Infrastructure needs.

Research Consortium

As stated above, typically the project leader either will be working in (and have leading expertise in) the selected Big Science area, or in ICT Science. Given the nature of many of today's data- and compute-intensive research problems, however, proposals of multidisciplinary teams to achieve their scientific goals are specifically welcomed. Public-private collaborations are positively valued, but the inclusion of industrial partners as part of the research team is not a necessary requirement.

1.3 Available budget

In this call, each grant will consist of two parts: 1) a cash contribution for the employment of local research personnel and other expenses, and 2) in-kind support in the form of skilled eScience Research Engineers and advanced Technology & e-Infrastructure Experts employed by the eScience Center and SURFsara.

The total available budget is as follows: k€ 510 in cash funding, 3.9 PYR² eScience Center in-kind funding, and 3.0 PYR SURFsara in-kind funding. With a total of 3 projects to be awarded, a typical project will receive k€ 170, 1.3 PYR in-kind eScience contribution, and 1.0 PYR in-kind SURFsara contribution. In-kind funding will be in the form of eScience Research Engineers and Technology & e-Infrastructure Experts employed by the eScience Center and SURFsara, whose time is allocated to the project.

On top of this, the eScience Center will contribute 1.0 PYR jointly to the 3 awarded projects, to enhance the development of sustainable and reusable software solutions and to ensure impact beyond the lifetime of the project (and, if possible, beyond the selected Big Science research area). This extra contribution is under supervision of the eScience Center and need not to be included in the project description or budget.

It is the aim of this call to grant one project in each of the Technological Research Directions described in Section 1.2.

1.4 Validity of this call

This call for proposals is valid for the assessment procedure of pre-proposals submitted before the deadline of Thursday **9 May 2019, 14:00 CET**, and full proposals submitted before the deadline of Thursday **29 August 2019, 14:00 CET**.

² PYR = Person Years; 1.0 PYR represents 1680 hours available for the duration of the project.

2 Guidelines for applicants

2.1 Who can apply?

Proposals can be submitted by researchers from any Dutch university or research institute affiliated with NWO or KNAW. Each proposal is to be formally submitted by a single named researcher (the 'principal investigator' or PI) on behalf of a team comprising researchers from one or more institutes. The PI is also the proposed project leader. A copy of the proposal must be submitted to the director or dean of the PI's institute. Proposals are granted only if the PI's institute has been informed of the proposal and accepts the conditions relating to grants awarded in this call.

Further requirements:

The PI must:

- have a PhD;
- hold a contract for at least the duration of the requested project or provide a letter of intent signed by the director or dean of the PI's institute indicating that the PI's employment will be guaranteed for at least the duration of the requested project.

Co-authors of the proposal are encouraged to be part of the research team and to serve as co-applicants of the proposal. This also includes researchers (postdocs) for whom funding is being requested.

The PI will be responsible for scientific progress and reporting. Financial accounting is a shared responsibility of the PI and the Netherlands eScience Center.

It is encouraged to have links with business partners to enlarge the potential for valorization of the project. As such, it is possible for co-applicants to be employed by a non-academic partner.

Further conditions:

In addition to the above requirements, the following conditions hold:

- the PI may submit only one proposal in that capacity in this call;
- the PI may not also submit a (pre-)proposal in that capacity in other calls published by the Netherlands eScience Center in 2019;
- the research team may not submit identical/very similar proposals in this call;
- the PI (and research team) can only submit a full proposal in the second phase of this call if he/she also submitted a pre-proposal in the first phase.

2.2 What can be applied for?

A project grant may be requested for a maximum total in cash budget of **k€ 170**, a minimum in-kind budget of **1.3 PYR** from the eScience Center and a maximum in-kind budget of **1.0 PYR** from SURFsara. The project duration must be between 18 and 36 months.

The available in cash budget is primarily intended to cover expenses for local research personnel. Part of the in cash budget, however, may also be reserved for additional resources and expenditures.

Typically (but not exclusively), the project budget will be as follows:

- a. **1 Postdoc position for the duration of 2 years.** The cost associated with this position includes an individual bench fee of up to k€5 as a contribution to travelling expenses, or other approved expenditure. Budget for local personnel should follow the most recent VSNU agreement³.
- b. **A maximum of k€15** for project-related equipment and software, Open Access publication, and (non-eScience Center / -SURFsara personnel) travel expenses. The costs for equipment and software must equal or exceed k€5, and its necessity must be justified; equipment and software with a purchase price of less than k€5 is deemed part of the standard infrastructure of the local institute and is therefore ineligible for funding. Travel expenses must be incurred by the project team members, with a maximum of k€10. The necessity of travel, equipment and software in support of the project must be justified.

The total of a) + b) may not exceed k€ 170.

- c. **A minimum of 1.3 PYR** in terms of support provided by personnel employed by the eScience Center. Of the total requested PYR, 85% covers all activities performed directly on behalf of the project by one or more eScience Research Engineers, and by the assigned eScience Coordinator who oversees the project at the eScience Center. The remaining 15% comprises activities to the benefit of academic research in general (incl. internal communication, knowledge transfer, and training). In case more than 1.3 PYR is requested, the budget for a) + b) is lowered by k€ 10 per 0.1 PYR.
- d. **A maximum of 1.0 PYR** in terms of support provided by personnel employed by SURFsara. Of the total requested PYR, 95% covers all activities performed directly on behalf of the project by one or more Technology & e-Infrastructure Experts. The remaining 5% comprises activities to the benefit of academic research in general (incl. internal communication, knowledge transfer, and training).

Distribution of the in-kind contribution over the project period will be determined together with the eScience Center and SURFsara at the start of the project.

eScience Research Engineers are scientists employed at the eScience Center who work at the interface of various scientific disciplines and advanced ICT. They will become an integral part of the projected research team focusing on the development

³ See: <http://www.nwo.nl/financiering/hoer-werkt-dat/salaristabellen>. Note that VSNU agreements typically are renewed every year. Therefore, in the pre-proposal the budget is taken as indicative. In the full proposal phase, the budget must follow the latest VSNU agreement precisely.

and implementation of eScience technologies and software. Primarily, they will ensure that the research team will be able to make easy and effective use of the envisioned technological solutions. They will help to interpret the results of the research and to make the delivered eScience tools useable for a broad range of users. Where applicable, they will also co-author research publications together with the research team. The eScience Research Engineers perform their project activities both at the eScience Center in Amsterdam and at the project locations (typically at the institute of the main applicant). In this way, they contribute directly to the project team whilst also providing a direct link to the expertise available throughout the eScience Center and its wider networks.

The eScience Center **Coordinator** is an experienced project manager, responsible for the daily supervision of the Engineer(s) assigned to the project and is – together with the PI – in charge of monitoring progress and the delivery of project results.

SURFsara **Technology & e-Infrastructure Experts** are engineers employed at SURFsara who work on innovation at the interface of advanced e-infrastructure and scientific applications and workflows. They will become an integral part of the projected research team focussing on the engineering of Proof-of-Concepts and prototypes of technological innovations that support the scientific challenges. They will also work on scalability of applications and e-infrastructure components where relevant. Where applicable, they will also co-author research publications together with the research team. The Technology & e-Infrastructure Experts perform their project activities both at SURFsara and the eScience Center in Amsterdam, and at the project locations (typically at the institute of the main applicant). In this way, they contribute directly to the project team whilst also providing a direct link to the expertise available throughout SURFsara and its wider networks.

2.3 When can applications be submitted?

The round consists of two phases:

1. A mandatory pre-proposal phase, in which the main research ideas and projected outcomes are to be outlined. The application form can be obtained from the eScience Center webpage for this call (via www.esciencecenter.nl/funding). The closing date for the submission of pre-proposals is **9 May 2018, 14:00 CET**.
2. The full proposal phase, in which selected applicants are invited by the independent Assessment Committee to submit a detailed application. A full proposal may be submitted only in case the applicants also submitted a pre-proposal in the first phase. The application form can be obtained from the eScience Center webpage for this call (via www.esciencecenter.nl/funding). The closing date for the submission of full proposals is **29 August 2018, 14:00 CET**.

Further information about the procedure, including a timetable, is found in Section 3.1.

2.4 Preparing an application

A pre-proposal has three parts: a fact sheet, the application form, and a list of suggested referees / non-referees. A full proposal has two parts: a fact sheet, and the application form.

- The fact sheet can be completed directly in ISAAC, the electronic application system of NWO (www.isaac.nwo.nl). While completing the fact sheet, you can only make use of ASCII symbols ("plain text"). Therefore, use of (structural) formulas, illustrations, italics etcetera in the fact sheet is not possible. A link to the ISAAC submission system is available on the eScience Center webpage for this call.
- The correct application form must be used for preparing the proposal. The form is obtained from the eScience Center webpage for this Call (via www.esciencecenter.nl/funding). The completed form must be attached to the ISAAC fact sheet as a PDF file.
- A list of possible referees/non-referees should be added to the ISAAC fact sheet, and explicitly not be included in the proposal text itself. For more information, please refer to paragraph 2.6.
- Possible letters of intent from (for example) private partners should be added to the ISAAC fact sheet in a separate PDF-file.
- Applications must be completed in English.
- The layout of the proposal should facilitate its readability. Use a font size of at least 10 points.

2.5 Specific conditions

The specific conditions that are valid for granted proposals are as follows:

- In case the proposal has been submitted to more than one competition, it can be awarded only once;
- Awarded projects must commence within six months of the award date. If the project has not started within that period, the Board of the Netherlands eScience Center and the SURFsara Management Board (Directie) have the right to withdraw the grant.
- In case any components (such as data sets, specialized hardware, etcetera) necessary for starting or continuing the proposed research are not available either at the start of the project or at the date specified in the project workplan, the eScience Center Board and the SURFsara Management Board (Directie) have the right to withdraw the grant.
- In case projected deliverables (i.e. research papers, software tools, data sets, or otherwise) have not been realized at the approximate date specified in the project workplan, the eScience Center Board and the SURFsara Management Board (Directie) have the right to withdraw the grant.

- **FAIR:** Researchers funded within this call for proposals must use the FAIR principles with respect to the sharing of data⁴.

Open Access

All scientific publications resulting from research that is funded by grants derived from this call for proposals are to be immediately (at the time of publication) freely accessible worldwide (Open Access). There are several ways for researchers to publish Open Access. A detailed explanation regarding Open Access can be found on www.nwo.nl/openscience-en.

Data Management / Software Sustainability

Responsible data management and high-quality software are part of good research. The eScience Center expects data and software that emerge from publicly funded research to become freely and sustainably available, as much as possible, for reuse by other researchers. Furthermore, the eScience Center aims to raise awareness among researchers of the importance of responsible data management and software sustainability, amongst other to enhance correctness and reproducibility of scientific results. Proposals should therefore satisfy the data management and software sustainability protocols of the eScience Center. Both protocols consist of two steps:

1. Data management section / software sustainability section

The data management and software sustainability sections are part of the research proposal. Researchers should answer several questions about data management and software sustainability within their intended research project. Therefore, before the research starts the researcher will be asked to think about how the data collected must be ordered and categorized, and how the research software created will be licensed and published, so that these can be made freely available. Measures will often need to be taken during data and software production to make long-term storage, dissemination and re-use possible, also after the project has finished. Researchers can state which research data and software they consider relevant for storage, publication and reuse.

2. Data management plan / software sustainability plan

After a proposal has been awarded, the researcher should elaborate the data management and software sustainability *sections* into data management and software sustainability *plans*. Both plans should be provided to the eScience Center within a maximum of 4 months after the project has started. The eScience Center will approve the plans as quickly as possible. Approval of the data management and software sustainability plans is a condition for disbursement of the funding. The plans can be adjusted during the research. The eScience Center requests PIs to use one of several *certified data management templates* (see: www.lcrdm.nl), to best match the details of the awarded project and/or any specific requirements of the PI's research institute. Further information on data management and software sustainability are available on the eScience Center website (www.esciencecenter.nl/funding).

⁴ The FAIR Data Principles, see <https://www.force11.org/group/fairgroup/fairprinciples>

The NWO Regulation on Granting applies insofar as it does not deviate from the specific conditions in this call for proposals.

2.6 Submitting an application

- A proposal can be submitted only via NWO's electronic application system ISAAC. Applications submitted otherwise will not be admitted to the selection procedure. The PI is obliged to submit the proposal via his/her personal ISAAC account. If the PI does not yet have an ISAAC account, this should be created at least one day before the submission deadline (www.isaac.nwo.nl). Possible registration problems may then still be resolved on time. If the PI already has a personal ISAAC account, the creation of a new account is not needed. For technical questions, please contact the ISAAC helpdesk (see Section 4.1.2).
- The proposal must be submitted as a PDF document and should arrive no later than the deadline set in Section 2.3.
- In the submission process in ISAAC, you will be requested to provide additional information. Please take this into account with regards to the set deadline.
- Applicants are requested to suggest up to three international referees. Please give their full names, email address and web address. You should not propose anyone with whom you have recently collaborated or with whom you intend to collaborate in the near future, whether as co-authors or in other forms of joint undertaking. Only referees who are not directly involved in the research project and research team to which your application refers can be considered. Moreover, suggested referees must not currently hold an appointment in the Netherlands.
- It is possible (but not mandatory) to give the names of up to three people who should NOT be approached as referee. In the interest of confidentiality, these names should not be included in the application itself but be provided in ISAAC separately (before the deadline set in Section 2.3).
- The proposal summary provided in ISAAC, and the summary for non-experts, may be used for publication purposes, should your application be granted.
- In accordance with the agreement between NWO and the Association of Dutch Universities (VSNU), applicants should inform their employing institute of the submission by sending a copy of the application to the scientific director or dean of the institute or department. It is therefore assumed that the employing institute or university is informed of, and accepts, this Call's granting conditions.

3 Assessment procedure

3.1 Procedure

Information event

To further inform interested applicants of the specific aims of this call for proposals, the mission and approach of both the eScience Center and SURFsara, the role and expertise of the eScience Research Engineers and the SURFsara Technology & e-infrastructure Experts, the software technologies implemented and applied by the eScience Center and SURFsara, and the specific capabilities of the Dutch National e-Infrastructure, an information event will be organized at Amsterdam Science Park on **9 April 2019**. Registration is required via www.esciencecenter.nl. Presence of at least one team member at the event is highly recommended, but not mandatory.

Pre-proposal

The first evaluation of all submitted pre-proposals is carried out by eScience experts and technology leads at both the eScience Center and SURFsara. This evaluation covers only the eScience and technology criteria outlined in Section 3.2 (eScience and technology state-of-the-art, sustainability, re-use potential and lateral impact). In this phase, the scientific quality, novelty and impact is explicitly not evaluated. The results of this evaluation are handed over to the independent Assessment Committee (AC; see Section 4.2) as input for their assessment of the pre-proposals.

Based on the pre-proposals and the results of the above evaluation, and using the criteria outlined in Section 3.2, the Assessment Committee reviews the proposals, prioritizes the proposals within each *Technological Research Direction* (see Section 1.1), and advises the eScience Center and SURFsara on the highest ranked proposals that should be worked out into a full proposal.

Approximately eight weeks after the pre-proposal submission deadline, applicants will receive the AC advice, stating whether they are invited to proceed with a full proposal. Based on the AC advice, the eScience Center and SURFsara may suggest that applicants intending to work on closely related subjects submit a joint proposal.

Mid-procedure meeting

Applicants invited to submit a full proposal are also invited for a personal meeting with eScience Center and SURFsara experts. In the meeting, applicants are given advice on how to best incorporate any advice of the Assessment Committee, how to exploit the competences of the eScience Center and SURFsara in full, and how to best cover all review criteria.

Applicants with a negative AC advice are not invited for a mid-procedure meeting. The eScience Center and SURFsara have capacity to provide this support for the invited finalists only.

eScience Center and SURFsara employees are not allowed to write any part of the proposal, or to serve as co-applicant.

Full proposal

All submitted full proposals are sent to international independent peer-reviewers. The review reports of the peer-reviewers are sent to the applicant, who will be given the opportunity to write a rebuttal. In a meeting, the AC will discuss all proposals using the submitted proposals, the peer-reviews and the rebuttals from the applicants. Proposals will be assessed following the criteria explained in Section 3.2.

Only in case a negative evaluation was given by the eScience Center and SURFsara in the pre-proposal phase, the eScience Center and SURFsara may give further advice to the Assessment Committee based on the full proposal, again only covering the eScience and technology criteria. For all other proposals, no further advice will be given to the Assessment Committee. In case the second evaluation by the eScience Center and SURFsara is negative as well, the eScience Center and SURFsara have the right to provide a binding advice to the AC, i.e.: to not grant the proposal.

Applications are assessed within the Technological Research Direction selected by the proposers at the time of submission. Based on the prioritization and ranking of applications in each of the three Technological Research Directions defined in this call, the AC will compose a recommendation for granting and rejection to the Board of the Netherlands eScience Center and the SURFsara Management Board (Directie). The AC aims to recommend one proposal for funding in each of the Technological Research Directions in this call.

Granting Decision

The Board of the Netherlands eScience Center and the SURFsara Management Board (Directie) decide on the awarded grants, based on the AC recommendations.

Timetable

| | |
|-------------------------------|---|
| <i>9 April 2019</i> | Information event at Amsterdam Science Park |
| <i>9 May 2019</i> | Deadline pre-proposals |
| <i>Beginning of July 2019</i> | Announcement of results of pre-proposal round |
| <i>29 August 2019</i> | Deadline full proposals |
| <i>End of October 2018</i> | Applicants receive reviewer's comments and are given the opportunity to respond |
| <i>Mid-November 2018</i> | AC evaluation and prioritization |
| <i>Mid-December 2018</i> | Applicants informed of final decision |

3.2 Admissibility and assessment criteria

3.2.1 Formal admissibility of applications

A proposal will be assessed only when all of the following conditions have been met:

- the proposal has been submitted by a researcher at a recognized institution (see Section 2.1);
- the proposal is consistent with the purpose of the call (see Section 1.2);
- the proposal was submitted online via ISAAC;
- the proposal was submitted before the deadline;
- the proposal meets the conditions and requirements of this call for proposals.

Once officially declared admissible, the proposal will be processed. Proposals with serious errors or omissions may be disqualified.

3.2.2 Assessment of contents

Proposals will be assessed by the referees and by the Assessment Committee based on the criteria below:

Scientific quality (25%)

- the proposed research should be at the forefront of the state-of-the-art of technological developments, also at an international level;
- the research team should be of the highest quality, and – if possible – already be recognized as representative (and authoritative) with respect to the proposed research questions, direction, and long-term vision.

Scientific novelty and impact (25%)

- the proposed research should be novel and represent a number of essential steps towards the long term aim of solving a specific, major scientific challenge;
- the proposed research should potentially change the modus operandi of research practice within the 'Big Science' discipline area of focus, in terms of broadness, scale, speed of result-delivery, or otherwise;
- the proposed research can be expected to lead to one or more significant results (either from a technological perspective, or in terms of a significant advancement within the 'Big Science' discipline area of focus);
- the proposal must indicate which efforts are made to promote the results of the project (publications, demonstrations, workshops, training, etcetera).

eScience and technology state-of-the-art (25%)

- the eScience and data science technologies (e.g. software for data analytics, data management, efficient computing, etcetera) researched and developed should be state-of-the-art and appropriate, meaning that no alternative (proven) technologies exist that would serve better in solving the (new) 'Big Science' research questions of focus or lead to more significant scientific breakthroughs;
- the research team should show awareness of the state-of-the-art of alternative (relevant) technologies; PIs can first contact the eScience Center and SURFsara, if needed.

Lateral impact, re-use and sustainability (25%)

- the proposal must indicate how the proposed solutions will find use beyond the proposed work itself, preferably across disciplines, also after finalization of the project;
- the proposed solutions and (software) deliverables must be open source/open access and permit use and/or interpretation by other researchers;
- the proposal must indicate how the project will build further collaborations, in science, industry, or both; inclusion of concrete letters of intent from such foreseen partners will be valued positively, but is not required;
- the proposal must indicate how long-term maintenance and sustainability of project results will be secured and managed.

4 Contact details

4.1 Contact

4.1.1 Specific questions about this call

If you have specific questions about this call for proposals and the assessment procedure, please contact:

Dr. Barbara van der Sar-Reumer, Program Coordinator NWO
Tel.: +31 (0)70 349 4602
Email: e-science@nwo.nl

For questions about the Netherlands eScience Center, or the eScience requirements for this call, please contact:

Dr. Frank J. Seinstra, Program Director Netherlands eScience Center
Tel.: +31 (0)20 460 4770
Email: dtec-call@esciencecenter.nl

For questions about SURFsara, or SURFsara's contribution to this call, please contact:

Dr. Axel Berg, Manager SURF Open Innovation Lab (SOIL) at SURFsara
Tel.: +31 (0)20 800 1300
Email: axel.berg@surfsara.nl

4.1.2 Technical questions about the electronic application system ISAAC

For technical questions about the use of ISAAC, please contact the ISAAC helpdesk. Applicants are requested to read the ISAAC manual before consulting the helpdesk.

The ISAAC helpdesk is available from Monday to Friday from 10.00 to 17.00 hours on +31 (0)20 346 7179. You can also send your questions to isaac.helpdesk@nwo.nl. You will receive a reply within two working days.

4.2 Other information

4.2.1 Members of the Assessment Committee

A separate Assessment Committee (AC) will be set up for the evaluation of the proposals submitted in this call. The AC will consist of independent eScience and domain experts. An expert is not allowed to be part of the Assessment Committee in case a member of his/her research group submits a proposal in this call.

Appendix A:

eScience Center Core Technological Competences

The Netherlands eScience Center is the Dutch national center of excellence for the development and application of research software to advance academic research. We contribute to research projects in at least two important ways:

1. We continuously scout the international spectrum of research software; we have a broad overview of relevant software solutions and a detailed understanding of how to apply these in a broad range of research disciplines;
2. We have expertise to extend and build high-quality, sustainable, and reusable research software using modern software development techniques and standards.

Our core competence is the *creation* and *application* of research software. What software is already available? When and how can we apply this software? Can we extend already existing software? How do we build new software, if needed? In the process of extending and building software we apply high standards of software quality, and put significant effort into testing, documentation and packaging.

In addition to this core competence, we focus our efforts in three expertise areas: Optimized Data Handling, Big Data Analytics, and Efficient Computing. Together, these cover a large part of the spectrum of required software and expertise in research projects. Below each of these expertise areas is outlined further.

Optimized Data Handling

This expertise area includes a.o.:

- FAIR data
- streaming data
- databases
- linked data
- data fusion

Storing, accessing and sharing voluminous and rapidly generated data

Data are generated at increasing speed and abundance due to the miniaturization and parallelization of experiments, the deployment of sensors and the digitization of experimental practices. From radio telescopes to social media, the development and application of methods to store, access and share large volumes of rapidly generated data are becoming universally important.

At the eScience Center, we have expert knowledge on handling large volumes of data (using both traditional databases and their NoSQL alternatives), processing streaming data (as produced by sensors such as radio telescopes), and linked data (typically used to add meaning to text data). In addition, we have ample experience in sharing data according to FAIR-principles (i.e.: making data Findable, Accessible, Interoperable, and Reusable).

Big Data Analytics

This expertise area includes a.o.:

- machine learning
- natural language processing
- search
- computer vision
- visualization

Identifying patterns and relationships

From data to information to knowledge to insight. Current research challenges demand robust and reliable methods to identify the patterns and relationships contained in, but also obscured by, large amounts of data.

eScience approaches can enable researchers to recognize sources of relevant information, prepare raw data, use statistical tools, extract and search for meaningful information, recognize potential problems and make visualizations to communicate their findings.

With the application of statistics and state-of-the-art machine learning techniques at its core, the use of data-analytics and visualization are generic requirements for many scientists. Combining 'big data' with theory and conceptual models enables scientists to structure the wealth of data and provide skillful forecasts.

Efficient Computing

This expertise area includes a.o.:

- high-performance, distributed, and energy-aware computing
- efficient algorithms
- scalability
- ease-of-use

Optimizing for performance

As the ambition and data volumes of researchers grow, processing requirements grow accordingly. To keep up with the sizes of the data and models, software must be optimized for performance (resulting in more processing power per computer) and/or scalability (allowing more computers to share the processing load).

By applying state-of-the-art technologies such as GPUs, a significant performance increase can be achieved, while simultaneously reducing energy requirements. This requires expert knowledge, however, as GPUs are very hard to program.

Often, research data are stored in multiple locations and are too large to gather in a single place. In such cases, it may be necessary to move computing to the data, and not vice versa. Such a distributed computing solution requires specialized software to organize which computation runs where. For the user, such techniques are a means to an end, and must be made transparent to not get in the way of the research itself.

Appendix B:

SURFsara Technological & e-Infrastructure Expertise

SURFsara is the national center for high-performance computing, data processing and data management to support scientific research with services, expertise and innovation. SURFsara creates a bridge between research and advanced ICT. We do so with a passion for scientific research in our DNA and with extensive expertise in its high-performance infrastructure. SURFsara offers a full range of services, expertise, innovation and support in the field of high-performance computing, data services, data processing and analytics, visualization, machine learning and cloud services.

We contribute to research projects in a number of ways:

- In direct collaboration with researchers and/or research communities, we provide expertise in support and projects to develop scalable e-Infrastructure/application solutions for simulations, data processing and/or data management;
- Through our SURF Open Innovation Lab (SOIL), we collaborate in early innovation and experiments with research institutions and industry, with the purpose to obtain more knowledge about new technologies as well as work-flows and applications. This concerns testing concepts and proof-of-concepts in areas like quantum computing, machine learning, digital data exchanges, energy efficient computing, etc. The knowledge, expertise and facilities developed within the SURF Open Innovation Lab can be used by SURF, SURF institutions and partners for technology, product and service innovation.

Our core competence is scouting, building and providing optimal ICT-infrastructure tools to support scientific research. With optimal we mean easy to use, scalable, and high performance. We understand the ICT challenges in scientific research and know about new ICT technologies and how to apply those in supporting research. We do this always in close collaboration with the researchers and research communities themselves.

SURFsara is organized in three competence and service areas, where innovation, service development and service provision take place: High-Performance Computing, Data Processing and Data Management. Below each of these competence areas is outlined further.

High-Performance Computing

This expertise area includes a.o.:

- HPC system & processor architectures and setup (incl. accelerators)
- performance optimization and parallelization of (HPC-) applications
- application porting
- scalable machine learning
- scientific visualization, augmented reality and virtual reality

We have expertise in single core optimization (making applications faster per core), parallelization and support in making applications scalable. To take full advantage of

the greatly increased performance of contemporary HPC systems, in-depth knowledge of the underlying hardware and programming models is required. At SURFsara, we have experience with a large number of supercomputer architectures. This experience enables us to ensure applications make the best possible use of the power of HPC systems. This also includes the development of specific software for specific research activities.

We also have expertise in development and improving visualization software. This included developing large-scale scientific visualizations and high-resolution visualizations using tiled-panel displays. We have expertise with remote visualization with remote rendering clusters and with streaming visualizations via a network. We also developed visualization applications based on Google Earth, Google Maps, VTK, OpenGL, etc.

Data Processing

This expertise area includes a.o.:

- modern cloud technologies & services (including e.g. containers)
- grid technology
- large scale data processing
- streaming data processing
- IoT platforms
- scalable data analytics

Choosing the right frameworks or tools for big data analysis is not easy. Projects can start with small amounts of data, but ultimately often have to deal with unforeseen workloads. We have expertise to making the right decisions from the start, as well as in setting up a scalable and cost-effective solution from SURF or a commercial cloud infrastructure. We have a wide range of expertise, from streaming analysis to (no)SQL data storage, from container-based solutions to serverless architectures etc.

Data Management

This expertise area includes a.o.:

- research data management platforms
- large scale data storage
- long-term data preservation
- FAIR data
- policy-based data management
- persistent identifiers

We have many years of experience in developing and managing data storage infrastructures for research communities. We know the challenges of long-term data retention, including data management tools. We also have expertise in on how to set up data infrastructures and specific data management platforms optimally.

Appendix C:

The Research Software Directory (RSD)

The Research Software Directory (RSD)⁵ is the eScience Center's primary facility for open, sustainable and re-usable research software, expertise, and eScience research. First and foremost, the RSD contains *research software*. In part, this software constitutes results of the collaborations between the eScience Center and its project partners. Other parts of the RSD are formed by software that is developed in-house at the eScience Center, and by software developed by external parties to which the eScience Center has made significant contributions. The eScience Research Engineers contribute to the RSD by generalizing and inserting the technologies they develop in the projects in which they are partnering as a research team member.

Apart from the software itself, the RSD contains *supporting material* associated with the actual tools, applications, scientific workflows, algorithms and libraries. This material can take the form of documentation, best practice guides, tutorials, training material, papers, demos, blog posts, etcetera. This collection of supporting material grows as software is re-used in other projects. In this way, the software in the RSD is presented in its *research context*. This context helps researchers to quickly judge if a certain piece of software is relevant to their particular problem, if others in their field are using it, how to get started with the software, and whom to contact for questions. This improves the findability of software and promotes its re-use.

While the RSD is the primary facility for managing and disseminating software created in the eScience Center's project portfolio, all aspects of the RSD can be applied in a broader context than just a single project. As such, the RSD supports multiple research efforts, an entire research discipline, and even multiple disciplines. The RSD explicitly aims to promote the exchange and re-use of knowledge and best practices and to prevent fragmentation and duplication of research software.

Serving research communities

It must be stressed that the technological developments undertaken by and with the eScience Center are not aimed at realizing benefits for the eScience Center itself. All developments are in support of the scientific goals of the research project, with the additional aim to also serve other research communities as much as possible, now and in the future.

⁵ See also: www.esciencecenter.nl/expertise/ and <https://www.research-software.nl/>.

Appendix D:

In this call, all applicants are asked to indicate the project's e-Infrastructure needs, in terms of compute hours, data storage capacity, lightpath connectivity, or otherwise. A 'use-or-explain' policy will be applied, meaning that

- projects *without* e-Infrastructure needs are asked to give a brief explanation;
- projects with clear e-Infrastructure needs are expected to select the hardware resources and services as part of the Dutch National e-Infrastructure as first option, and to indicate the expected extent of use;
- projects with clear e-Infrastructure needs that aim to use international (e.g. PRACE, EUDAT, EGI, XSEDE, etcetera) or commercial (e.g. web, cloud, etcetera) hardware and services instead are required to give a brief explanation.

The use of the Dutch National e-Infrastructure is not a requirement, nor is it a formal review criterion. However, in all cases in which the Dutch National e-Infrastructure is not used, a justification should be provided.

The Dutch National e-Infrastructure

In this call, the Dutch National e-Infrastructure is defined as follows:

all publicly-funded hardware resources (e.g. compute, data, visualization, networking, etcetera) and directly connected support services (people, software), set up and maintained with the aim to support publicly-funded research in the Netherlands, and made available to either all or a selected subset of all researchers from a.o. Dutch universities and research institutes affiliated with NWO or KNAW.

The definition distinguishes between hardware resources and services available to all researchers in the Netherlands (Category I), and those made available to a selected subset (Category II). The Category I e-Infrastructure, outlined below, is formed by the hardware resources and services provided and maintained by SURFsara, SURFnet, DANS, and – in part – also by Nikhef and RUG-CIT.

The Category II e-Infrastructure is formed by all other hardware resources and services that are accessible to a selected group of researchers following thematic or geographic criteria. Examples of such infrastructures include the Distributed ASCI Supercomputer (DAS) and the many stand-alone local facilities at various universities (e.g. the Peregrine cluster (RUG-CIT), the GPFS data storage facilities (Target), the WUR HPC Cluster (Wageningen), etcetera).

Overview: Category I e-Infrastructure

While it is impossible to provide a complete overview of all resources part of the Dutch National e-Infrastructure in this call text, the following provides entrance points to the major Category I e-Infrastructure resources and services. For more information, it is advised to contact the organizations and institutes responsible for these resources directly, in particular SURF: <https://www.surf.nl/en/contact.html>.

Compute Resources and Services

- Cartesius: National supercomputer for maximum performance
<https://userinfo.surfsara.nl/systems/cartesius>
- HPC Cloud: Complete control over your own computing infrastructure
<https://www.surf.nl/en/services-and-products/hpc-cloud/index.html>
- Grid: Distributed computing system for fast processing of large data sets
<https://www.surf.nl/en/services-and-products/grid/index.html>
- Hadoop cluster: Big data processing and analysis
<https://www.surf.nl/en/services-and-products/big-data-services/index.html>

Data Resources and Services

- Research Drive: Collaboratively share and save large volumes of research data
<https://www.surf.nl/en/services-and-products/research-drive/index.html>
- Data Archive: Secure long-term storage of research data on tape
<https://www.surf.nl/en/services-and-products/data-archive/index.html>
- DataverseNL: Store, share and register research data online
<https://www.dans.knaw.nl/en/about/services/DataverseNL>
- EASY: Online archiving, depositing and downloading of research data
<https://www.dans.knaw.nl/en/about/services/easy>

Networking Resources and Services

- Lightpaths (SURFlichtpaden): Ultra-fast and high capacity connectivity
<https://www.surf.nl/en/services-and-products/surflichtpaden/index.html>

Cloud and Collaborative Resources and Services

- SURFconext: Online collaboration and services in a single environment
<https://www.surf.nl/en/services-and-products/surfconext/index.html>
- SURFdrive; Personal and secure cloud storage, synchronization, sharing
<https://www.surf.nl/en/services-and-products/surfdrive/surfdrive.html>

Visualization Resources and Services

- Remote visualization: Visualize large datasets on your desktop
<https://www.surf.nl/en/services-and-products/visualisation/index.html>
- Collaboratorium: Sophisticated presentation and visualization aids
<https://www.surf.nl/en/themes/research/overview-services-for-research/big-data-analytics-and-visualisation-processing-data/index.html>

SURF

For a complete overview of all Category I services provided by SURF, see:
<https://www.surf.nl/en/services-and-products>

DANS

For a complete overview of all Category I services provided by DANS, see:
<http://www.dans.knaw.nl/en/about/services>