In brief

Online research environments in 2027

The vision of SURF

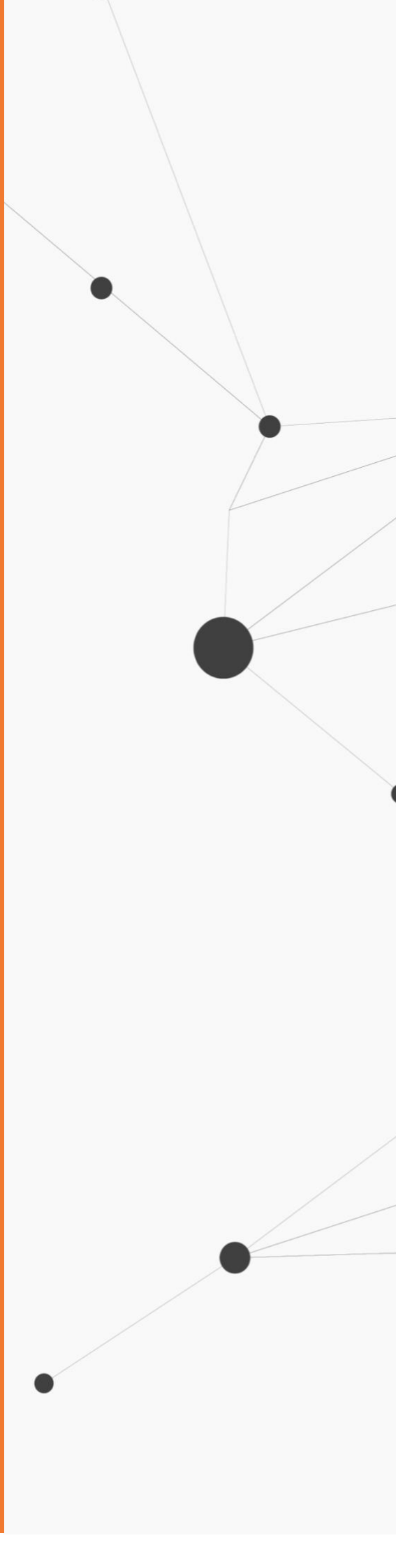


Table of contents

Introduction

Growing need

Collaboration is an essential factor for success in research. Rarely do researchers publish on their own, and research is developing in a very multidisciplinary and cross-institutional way. Researchers from the same institution often work together, but they're also increasingly working with researchers from other scientific institutions or companies, both inside and outside the Netherlands. Project consortia in which various institutions collaborate on a research area benefit from online research environments.

Over the past few decades, data analysis has become increasingly prominent in scientific research due to digitalisation. Better and faster technology means that data sets have grown in size and data processing has become increasingly faster and more complex. This has become particularly evident with the rise of machine learning (ML) and other artificial intelligence (AI) techniques. But particularly in the field of data analysis, researchers are not yet collaborating much, partly because facilities and infrastructures are not set up for this. The challenges to collaboration with online research environments include facilitating legal frameworks for data sharing and achieving the desired interoperability between institutions.

Working together on a vision

Online research environments (Virtual Research Environments, VREs) are intended, among other things, for collaboration between researchers from *different* institutions, such as project consortia and research communities. This makes them different from many other IT applications, for example in terms of data security, interface standardisation and applicability of software licences.

It is precisely because VREs must enable collaboration across institutions that it is important for institutions to think together about the direction in which these environments should evolve. That is why SURF – the cooperative in which research institutions naturally already find each other – has worked together with members on a vision for the development of VREs up to 2027.

The vision on VRE for 2027 described here is the result of discussions and ideas from stakeholders in the SURF community. They discussed what online research environments in the Netherlands could look like, striking a realistic balance between what is desirable and what is achievable. This balance is important because the wishes and ambitions of different (types of) members of the SURF community vary considerably. A more detailed report on this can be found in the vision document 'Online

Onderzoeksomgevingen in 2027’ (in Dutch) on the SURF website: https://www.SURF.nl/online-onderzoeksomgevings-in-2027.

Starting point for discussion

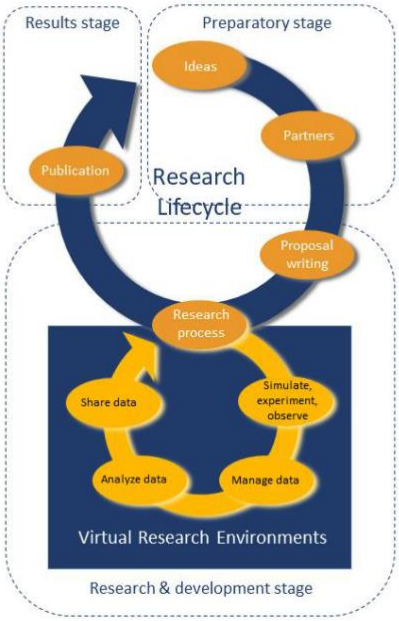
This vision document aims to provide a starting point for further discussions, for example in meetings in your own institution about VREs, and in meetings with colleagues from other institutions. It is important to constantly assess (changing) needs and insights.

1 Vision on VRE

What is a VRE?

Before we discuss the vision on online research environments, it is important to outline what we mean by a VRE. What is the purpose of VREs and how are they structured?

A VRE is, of course. embedded in the research process. The figure shows the context of a VRE in the Research Lifecycle (source: researchgate.net) as described in the VRE Reference Architecture[1].



A Virtual Research Environment (VRE) is a working environment that enables researchers from the same or different institutions to work on research online, together or alone. A VRE provides the tools and data researchers need for their research.

A VRE consists of online workspaces. A workspace can be set up with the procedures, tools, computing power and storage capacity that researchers from one or more research institutions need to conduct their research.

[1] See the document 'Further Development of a VRE Reference Architecture', 13 February 2020

Typical tools are those for importing, preparing, integrating, simulating, annotating, analysing, managing and sharing data.

A workspace is specific to a research team: the available tools and data and the service level offered meet the wishes of the researchers in that team.

The collection of workspaces is made available on a VRE platform as a service set up per institution or organisation. Portals are part of the platforms; they form digital entry points to the workspaces.

The vision in brief

We articulate the vision on online research environments as follows:

> *By 2027, we envisage a VRE ecosystem of portals and platforms that comply with national and international standards so that researchers can easily collaborate in their workspaces with good access to data and software in repositories. This will require national alignment and coordination of the ecosystem, including developing a VRE target architecture with members, encouraging federated collaboration and concluding suitable software licences. SURF is prepared and available to take on this coordinating role.*
>
> *The VRE ecosystem must be organised in such a way that innovation of parts remains possible, with a process for integrating new services into the architecture and the possibility of sharing new workspaces developed as open source projects.*

Partly integrated ecosystem

In the SURF vision, we aim for a partly integrated ecosystem in which the VRE platforms play an important role. This role is balanced with that of the VRE portals, each of which acts as a representative of users in their own domain. Those organisations offering VRE portals make agreements with the VRE platforms on principles for effective functioning of the workspaces, such as ease of access, compliance with FAIR principles, security, connectivity, scalability and suitability for automation.

In order to achieve interoperability between the VRE platforms and portals, it is essential to establish national control of standards and interfaces between parts of the VRE platforms. This is primarily desirable at the technical infrastructure level (such as linking authentication and authorisation facilities), but also at the level of protocols and standards for interfaces between data, computing and software facilities. Standards and links are managed by representatives of research communities and members of the SURF community. A reference architecture with agreements on standards and links determines authority through a governance structure.

In the ecosystem, it is possible to make VRE workspaces managed by one institution available to researchers at another institution. Such federated working is assured by links in operating facilities, such as budget allocation and accounting. As mentioned in the HOSA,[2] an orchestration function ensures the connection of individual components in order to properly regulate traffic between them in terms of supply and demand.

This vision is a compromise between 'fully non-integrated', where digital collaboration is difficult due to restricted access and interchangeability, and 'fully integrated', where there is little room for individual interpretation and innovation of the whole.

2 SURF's role and tasks

SURF's role in the national ecosystem

The starting points for the roles of institutions and research communities are laid out in the document 'Samenwerkingsmodellen voor VRE-waardeketens' (Collaboration Models for VRE Value Chains')[3]

Institutions and research communities are in charge of their VRE portal or platform, including first-line user support. The number of support tasks that researchers want to carry out themselves differs. Some want to keep a lot in their own hands, while others want to outsource a lot and limit themselves to the role of process manager for purchased functionalities. Institutions that have a DCC see a major role for the DCC in the institution's portal. The DCC can act as first-line support for the VRE users of the institution itself, for example as a data steward or local developer of custom solutions.

Members of the SURF community agree that SURF has a role to play in the national ecosystem. This role consists of two main tasks:

- Coordination of the ecosystem with regard to the national management and brokerage role
- Development and management of interoperable parts of the national part of the ecosystem infrastructure according to a defined architecture

[2] Hoger Onderwijs Sector Architectuur (Higher Education Sector Architecture) (HOSA) | SURF.nl

[3] See document 'Samenwerkingsmodellen voor VRE-waardeketens' (Collaboration Models for VRE Value Chains'), 12 November 2020

Most research communities and members see a role for SURF as the coordinator of the national aspects of a VRE ecosystem in particular (as part of a steering committee, i.e. together with members of the SURF community).

SURF's tasks in the national ecosystem

In addition to SURF's current tasks in service provision and innovation, SURF will take on the following tasks specifically for VRE:

- Coordinating and making (national) agreements on architecture standards (such as HOSA) within the VRE ecosystem.
- Managing a national architecture in which VRE standards and interfaces are defined that align with international standards such as EOSC.
- Coordinating and shaping (mostly through partners) innovation of the national ecosystem.
- Developing or implementing licensing models that are cross-institutional and valid within the ecosystem.
- SURF can ensure favourable terms of use by jointly purchasing Microsoft Azure, Google Cloud or Amazon AWS.
- SURF can coordinate, publish and manage a service catalogue of the VRE components, platforms and workspaces in the national ecosystem, including qualification of aspects such as ease of use, required licences, etc. Certainly in the case of federated working, it is necessary for one party to ensure the findability of components developed by institutions; but this role could at least partly be assigned to the research communities and institutions.
- SURF is asked to encourage institutions, alliances and research communities to purchase, manage, adapt and make existing open source and commercial products and services available on a national basis.

3 Challenges

Different workspaces

There are many different user requirements and solutions within the VRE ecosystem. Some of the key reasons for this are:

- Research communities and research institutions are purely focused on (scientific) research and often have very advanced users who have an in-depth understanding of how their tools work and can intervene where necessary.
- In addition to researchers, universities of applied sciences also have many lecturers and students in practice-based research as users of online research environments. These groups mainly need

a simple workspace that 'just works'. User support is very important here.
- Research often involves collaboration with companies. A solution must therefore also be offered for collaboration with companies.
- The requirements for VRE workspaces vary depending on the type of data being worked with. For example, health research puts more stringent requirements on the processing of patient data.

It is therefore important to check the national ecosystem carefully to see whether all the different target groups are being served with workspaces that match their use profiles.

Collaboration with the market

It is a challenge to continue ensuring coherence and interoperability in the VRE ecosystem. Agreements will be reached mutually on this, depending on market developments and the demand of researchers. The use and management of facilities in the ecosystem requires governance for selection in terms of quality, usability and functionalities such as security.

Some members believe that researchers should make maximum use of open source products, supplemented by their own national products and services (such as SURF Research Cloud). Another group believes that we should not build too many functionalities, because the market (including big tech) will probably be able to deliver them more cheaply and reliably. This group is committed to making good, binding agreements with commercial suppliers with regard to compliance with the requirements of public values. This focuses primarily on security and privacy requirements and on making agreements to minimise vendor lock-in.

In the discussions on a partly integrated ecosystem, this issue will need to be discussed extensively and intensively.

Critical success factors

- Standardisation between many different solutions to ensure interoperability
- The innovative power of the ecosystem for online research environments, such as how long it takes to adapt services to a changed reality
- A suitable licensing model for cross-institutional use of software and an appropriate model for data ownership and data sharing dynamics
- Feasibility of financing the workspaces.

4 Next steps

SURF will continue to discuss the subject of online research environments with its members in the coming years.

Architecture

Architecture is an important means of translating the described vision into a technical and functional implementation. To this end, a target architecture is outlined using the previously mentioned VRE reference architecture and HOSA.

Roadmaps

SURF collaborates with members in innovation zones on a number of complex innovation issues. The ambitions of the zones are defined in roadmaps. The roadmaps are ratified by the SURF Members' Council. Developments related to VRE are also reflected in these roadmaps. The following topics will be on the agenda in the near future:

- How do we solve the licensing issue with regard to contracts?
- What interfaces will we create to connect VRE components?
- Which components will be supplied by institutions and which by external suppliers?
- What agreements will we make regarding the import and export of (privacy-sensitive) data?