1 Introduction

1.1 About this document

1.1.1 This is the final report of the discovery phase of Jisc’s co-design consultation. It has been prepared by Dr Max Hammond under contract to Jisc, during the period February-May 2017.

1.2 Background

The NGRE Co-design challenge

1.2.1 As part of Jisc’s co-design consultation, Jisc identified six potential challenges that face people working in education and research. The Next Generation Research Environments challenge was one of these six and stated: “We think it’s time to define the future of research environments and determine how such environments can support the current and future needs of researchers.”

1.2.2 Researchers continue to push the boundaries of research, throughout the research lifecycle, using the latest tools and technologies to support their work. Jisc asked the community whether the time has come to define what a next-generation research environment should look like and if modular solutions might satisfy the increasing demands of researchers.

The discovery phase

1.2.3 Following an initial round of voting, and solicitation of comments, this report was commissioned to investigate the issue in greater detail. Everyone who contributed to the first phase was invited to take part in an interview to expand on their perspectives. A further set of expert stakeholders were contacted, and further desk research was undertaken, in particular to understand the current VRE landscape.

1.2.4 This report is not a comprehensive review of the VREs, although references are given to other work that addresses this field in more detail. This report instead describes what we believe defines the next generation of research environments, and has specific recommendations to Jisc regarding how to support the development of these environments.

2 From VRE to NGRE

2.1 Introduction

2.1.1 The idea of a Virtual Research Environment (VRE, also known as a Science Gateway or Virtual Laboratory) is not new: Jisc and others have been conducting work in this area for
over a decade. Many VREs are now mature, but there is no single definition of what a VRE is. Jisc’s original definition of VREs was broad, albeit aspirational.¹

A VRE helps researchers from all disciplines to work collaboratively by managing the increasingly complex range of tasks involved in carrying out research.

2.1.2 Current views are more specific, and range from a narrow scope:²

web-based enterprise information systems that provide scientists with customized and easy access to community-specific data collections, computational tools, and collaborative services on e-Infrastructures

to a broad one:³

an electronic web-based environment for a) access to data, tools, resources; b) cooperation or collaboration with other researchers at the same or different institutions; c) cooperation at the intra- and inter-institutional levels; or d) preserving or taking care of data and other outputs.

2.2 The VRE landscape

2.2.1 A review of the impact of science gateways (which we see as equivalent to VREs) is currently being developed by colleagues elsewhere⁴ – rather than pre-empt that work, this section will highlight key points for understanding the environment.

Diverse landscape, close to the researcher

2.2.2 There is a diverse landscape of VREs, many of which serve one discrete community of researchers. As many VREs are closely bound to one infrastructure (often with a limited set of clearly defined workflows), it often makes little sense to attempt to expand them to broader groups of researchers (but see below). These VREs largely exist as self-contained systems where research is actually undertaken, and are thus very strongly connected to that research.

2.2.3 This model has worked well for enabling access to tools within a community, but is inefficient when VREs are developed independently (without drivers to re-use existing components, it has often been seen as easier to develop new), is challenging when

---

¹ It is probably fair to say that whereas the outputs of the VRE programmes were significant, they did not deliver this vision – they were largely discipline-specific, and typically interfaces to e-Infrastructure. In many ways, the work of Jisc and others developed the understanding which led to the more specific, and more limited definitions of a VRE.


⁴ Michelle Barker et al., ‘The Global Impact of Science Gateways, Virtual Research Environments and Virtual Laboratories’ Personal communication of draft publication.
multidisciplinary research is undertaken, and is largely ineffective when diverse research communities must be rapidly brought together to address high-priority current issues.

**Links to infrastructure, not institution**

2.2.4 Most VREs are linked to e-Infrastructure; they typically exist to allow researchers to utilise complex computational (and now data management and analysis) tools through a portal. This position, as the point of contact between researchers and infrastructure, has provided some significant benefits to researchers, making available shared tools that would otherwise be beyond the limits of affordability (or of skills) of individual researchers.

2.2.5 VREs are typically (but certainly not always) independent from institutional systems used by researchers. They use their own user authentication and authorisation processes, and transfer of data between the e-Infrastructure and other compute and storage resources operated by the researchers is a manual process operated by the researcher – it is not intrinsic to the VRE.

**Generalisation and science clouds**

2.2.6 Although VRE development is still largely focused on the very specific systems that address one class of research problem, there is a clear trend toward developing interoperability, and towards platforms and frameworks that make interoperability and interworking easier to establish (see section 2.4).

2.2.7 Technologies such as virtualisation and “cloud” have now become established in research communities, and are providing the tools to make the provision of more generalised research infrastructures practicable: when the underlying platforms are no longer tied to specific workflows or even institutions, it is far easier to develop models for sharing them amongst diverse groups of researchers.

2.2.8 Paired with abstraction of hardware and platforms, the ongoing increase in the maturity of approaches for managing research data (for example, through the FAIR principles\(^5\)) is leading to generalisation of the technologies for handling and analysing the outputs.

**Sustainability is a key issue**

2.2.9 One of the concerns raised by interviewees for this review is that VREs are often unsustainable. They are linked closely to one research group, research project, domain, or infrastructure, and when this element changes or changes focus, the VRE becomes deprecated or unmaintained. In many cases, this may not matter – if the infrastructure that the VRE was designed to support is decommissioned, it is clearly unnecessary to maintain a portal to it. However, many VREs have had to solve the same or similar problems, and the learning from prior work is not effectively taken forward, in particular to those developing VREs in other domains.

2.2.10 As VREs become larger and more powerful, and in particular more general, the inefficiencies generated by this way of working are likely to increase. Issues of sustainability will become increasingly important – who maintains a VRE once the initial development work (and contingent funding) is finished?

This is related to broader questions of software sustainability and reproducibility in science, on which we will not expand here.

### 2.3 What is the “next generation”?

Interviewees consulted for this discovery phase held a wide range of perspectives as what they believed the next generation of research environment would mean to them. These fall broadly into four classes:

- **Administrative**: an environment that supports the research process, but not the research itself
- **A virtual lab**: an environment that supports the conduct of research directly. This is closely aligned with the VRE definitions above.
- **A portal/dashboard**: bringing together information from or access to other systems
- **A set of standards**: to allow systems to interoperate with each other.

Clearly, these are not orthogonal – a portal could contain administrative and/or research systems, and standards are necessary to underpin everything else. The key element that is different to definitions of VREs is the large proportion of interviewees who saw research administration as being integral in an NGRE.

**The scope of an NGRE**

Put another way, interviewees broadly saw NGREs as being either a more-capable VRE (focused primarily on the execution of research, the collection and sharing of data), or having a much larger scope, and covering the entire research lifecycle. The broader scope is in many ways equivalent to combining the capabilities of a complete CRIS system with a complete VRE and a collaborative authoring platform.

![Figure 2-1: research lifecycle](image)

We believe that implicit within this broader scope is the factor that distinguishes between the current generation of tools and the next generation. Whereas current tools may fulfil parts of this lifecycle, integrations are often *ad hoc*, with disconnects at important stages. NGREs will support international, interdisciplinary collaboration as core business, including the technical, legal and process designs required to do so. This requires integrations of:

- concepts;
- people (identification, authentication and authorisation);
tools (including compute and storage resources, software and workflows);
- datasets;
- outputs.

2.3.5 How Jisc might support progress toward an NGRE is discussed in section 3 below.

2.4 **Steps toward an NGRE**

2.4.1 We do not believe that the NGRE will be a normative system – many interviewees made the point strongly that each researcher creates their own research environment; “my research environment is the internet” was a common statement. Clearly, elements of their environment may be institutional or collaboration-specific tools, including VREs – but other parts are not.

**Standards**

2.4.2 Integration of the diverse range of data and systems that comprise a research environment requires standards. These standards may be formal, informal, open, closed, commercial, voluntary, ad-hoc, or any other structure. For example, a proprietary API to a university finance system is a standard no less than an IEEE technical standard, although the approach taken to developing these standards, and the approach required to managing dependencies on these standards will also vary.

2.4.3 Given the open and non-normative structures that we believe will represent NGREs, an open and inclusive approach to standards should be maintained: rather than considering a limited set of standards that are used within a set research stack, the focus should continue to be on identifying areas where a lack of standardisation is leading to specific challenges to specific research processes, and considering approaches to alleviating these.

**Platforms**

2.4.4 One step toward an NGRE is the creation of platforms, which could allow more efficient use of effort which is currently spent on specific VREs. We view a platform in this context as a system or set of systems that enables generalisation of some elements of VREs, for example to facilitate researchers from different fields sharing resources and/or collaborating on multi-disciplinary research.

2.4.5 We believe that platforms represent sensible and pragmatic approaches to reducing the cost and increasing the benefit of VREs, and are one step on the path to the next-generation environments. Platforms can themselves be scoped at a wide range of levels of generality, and we believe that the most benefit is likely to be found in pragmatic connections between the VREs used either i) for closely-related fields of research, or ii) where there is an established requirement to collaborate with another field. The broader and more diverse the group of researchers to whom a platform is targeted, the less likely we believe it to be that it will be successful.
Examples of platforms include ArcticConnect\(^6\) and GenAP\(^7\), both funded by CANARIE, EVER-EST, funded by the European Commission\(^8\), and a range of “virtual laboratories” funded by Nectar\(^9\). A current EC funding programme specifically encourages modular development of VREs – essentially the creation of platforms\(^10\). Included within this is the VRE4EIC project\(^11\), which has ambitions to establish a reference VRE implementation that is applicable to multi-disciplinary research. It is, however, specifically designed to integrate extant ESFRI VREs, rather than the kind of open approach that a Framework would encourage.

Two further examples of platforms with somewhat different goals are the European Open Science Cloud\(^12\) which is broad in scope, reaching from underpinning infrastructure through to data reuse, and the NIH Data Commons\(^13\) which is discipline-specific and tightly focused on data sharing and reuse.

**Frameworks**

Frameworks represent the next logical step toward abstraction from VREs. Where a platform integrates a set of tools, a framework provides the tools to integrate tools, and thereby supports the development of multiple platforms\(^14\).

A successful framework will attract the development of integrations with both research-specific systems and those external systems that are part of the research lifecycle, but only over time. It is necessary to support a certain number of “core” services to make the framework attractive as a starting point.

Developing a framework in abstract is likely to prove highly challenging. Until one (or more) such frameworks become well established, it seems more likely that frameworks will grow out of domain-specific platforms, rather than platforms being built from frameworks. As the maturity and scope of frameworks increases, it is likely that starting with one will be an obvious design choice when developing domain-specific platforms.

A comprehensive framework is probably the entity that comes closest to truly representing the technology of an NGRE, although it is neither a necessary nor a sufficient prerequisite! Effective frameworks will dramatically reduce the effort in creating research environments – but it is possible to do so without one, and simply using a framework does not lead to the kind of NGRE we envisage here.

\(^14\) The Jisc VRE programme funded frameworks, but these do not match the definition of frameworks that we use here. The frameworks funded under that programme abstracted specific infrastructures within specific domains, rather than providing a generic framework for abstraction and integration, as discussed here.
The only activity that we are aware of that has so far taken this approach is the Open Science Framework, developed by the Centre for Open Science. This is ambitious in approach, although currently limited in capability, and not yet mature.

3 Recommendations for Jisc

3.1 Introduction

3.1.1 There are two perspectives on what researchers want to achieve, and thereby where Jisc could consider providing support or services. One is a simple reduction in the effort required in managing research. This is the end-to-end lifecycle view, where the whole process hangs together better than it does at the moment. The second is to make collaborating on research easier than it is at present. The key challenges here appear to be managing authentication and authorisation decisions for inter-institutional (and often international) collaborations. The VREs and platforms can do this within their own domains, but it makes interdisciplinary research hard. Sharing arbitrary data in managed, secure ways is a key demand.

3.2 Recommendations

3.2.1 This section sets out recommendations to Jisc, based on our analysis of the landscape and direction of travel. This analysis is external – it is not driven by Jisc’s own commercial decisions, which will of course influence decisions about what, if any, action to take on these recommendations.

3.2.2 The idea of NGREs is well-aligned with Jisc’s interests, mission, and other activities. However, the nature of NGREs is likely to be significantly more varied than the other activities that make up Jisc’s R&D and service portfolio. As such:

**Recommendation 1**: Jisc should not attempt to develop any kind of “NGRE as a service”. The clear message from interviewees is that NGREs will be extremely diverse, and made up of components operated by a very wide range of actors, and connected with each other in very varied ways. This does not lend itself to a service provision model.

3.2.3 Although the idea of developing platforms to integrate research environments is appealing, it is our view that these platforms are intrinsically and closely linked to the research communities that develop and use them. Jisc’s approach to co-design and service development does not lend itself well to the development of these kinds of systems.

**Recommendation 2**: Jisc should not attempt to develop domain-specific research platforms. This is best undertaken by the research community itself, and is best supported by a model that funds these researchers.

---

3.2.4 The platforms that are currently being developed or in use are almost exclusively focused on specific research domains, and are typically tied to specific research infrastructures. Cross-domain infrastructures are currently immature, and the scale required for these platforms is large. Jisc by itself does not have the necessary reach nor scope to develop such a platform effectively, but there are potential opportunities for Jisc to influence and support the development of these platforms, for the benefit of its customers.

3.2.5 International collaboration is essential for Jisc in this space, whereas trying to build these components in their own right is not. Jisc should continue to work with those developing (or funding the development of) these tools. Jisc should itself build services and the components necessary for those services where they support specific requirements from its customers, with consideration of international connections from the very outset: UK-specific solutions simply aren’t going to be effective for most elements of the NGRE.

Recommendation 3: Jisc should not attempt to develop a cross-domain research platform as a service. Rather, it should continue to seek opportunities to engage with these platforms as they are developed elsewhere, to promote the adoption of concepts and standards that are beneficial to Jisc’s customers (see also recommendations 6-8).

Recommendation 4: Jisc should continue to support, where appropriate, the development of the underpinning standards and identifiers that are necessary to develop platforms and frameworks.

Recommendation 5: Jisc should consider the development of research environments – and the fit of its own services within research environments – on a global scale. International collaboration in research is the norm, and systems must be designed to support this. Jisc should continue to be involved in international activities that share experience, or support the development or harmonisation of approaches internationally.

3.2.6 It is our view that the clearest opportunities for Jisc to support the development of NGREs are from the bottom up: all the services that Jisc provides are part of the research environment, and Jisc must ensure that these services are designed to integrate into broad and diverse global research environments.

3.2.7 Jisc has developed significant experience with university administrative systems through the Research Data Shared Service activity. Although many CRIS systems purport to have capabilities in this area, satisfaction appears to be generally low. Jisc may be able to use its experience to support the integration of the research and research administration domains.

Recommendation 6: Jisc should consider interactions with research platforms and frameworks as part of the design of all its services. Jisc’s services are part of the research environment, and by ensuring that these can be easily integrated into diverse VREs, e-Infrastructures, and platforms, they will be able to form key parts of the vision of NGREs.

Recommendation 7: Jisc should actively consider the integrations of its developmental Research Data Shared Service with other elements of the research environment. In particular, there appears to be demand for closer integrations between active research data and archival research data.

Recommendation 8: Jisc should seek opportunities to use the “University of Jisc” test environment to test or develop integrations between research environments and
administrative systems. Jisc should monitor whether any research platform or framework should become part of the UoJ.

4 Bibliography


5 List of consultees

The following individuals generously supported this review through interviews and other input. They were asked to provide their input in a personal capacity, so institutional affiliations are not given below.

- Michelle Barker
- Anthony Basiel
- Jane Belger
- Fiona Bradley
- Chris Brown
- John Casey
- Matthew Dovey
- Beth Hall
- Alex Hardisty
- Jim Hensman
- Scott Henwood
- Ray Kent
- Natalie Meyers
- Frank Norman
- John Shemilt
- Mark Toole
- John Tuck
- Martin Turner
- Scott Wilson