1 Overview of the EVOp Project

There is an emerging and urgent need for new approaches to environmental challenges in the broad context of sustainability. Scientists, businesses and policymakers are asking questions that are far more interdisciplinary than in the past. Unfortunately, an unexpected outcome of the explosion of data and associated information is the growing disconnect between and within the supply of scientific knowledge, and the demand for that knowledge from the private and government sectors. NERC commissioned the Environmental Virtual Observatory pilot project to explore the question:

> "Is there a way of providing the 'wiring' to help people access the resources they need, be they a scientist, policy maker, industrial body, regulator or member of the public?"

1.1 The hypothesis

The hypothesis to be explored was that novel cloud computing technologies could be exploited to increase accessibility in a data-intensive world to "filter" and integrate this information to manageable levels as well as provide visualization and presentation services to make it easier to gain creative insights and build collaborations. This has been called the 4th paradigm (Gray, 2007; Hey et al, 2009). The ultimate aim was to make NERC science more efficient, effective and transparent. The transparency issue has been highlighted in a recent Royal Society report as requiring urgent attention to increase public confidence in the process of translation of scientific evidence through to policy making (The Royal Society, 2012).

1.2 Beyond data

A second hypothesis was to test if there is value going 'beyond data' to include models which are effectively a synthesis of current understanding and one of the main tools NERC scientists use to integrate complex data, upscale and make projections under future scenarios. Within the terrestrial and freshwater communities many models address environmental questions concerning soil and water quality, flood and drought risk, and ecosystem structure and function (e.g. Defra recently identified ca. 60 models currently used in diffuse pollution modelling alone). These models simulate complex physical, chemical and biological process interactions. There is a challenge however in gaining access to these models, linking them together to deliver more holistic outputs and objectively testing to the level needed by end-users who need to make policy or management decisions based on their outputs. This leads to the second question to be tested:

> "How can we create a culture of more open and rigorous testing and evaluation of the current models necessary to improve process understanding, process representation in models and thus model forecast accuracy?"

Two main challenges currently limit this ongoing model development; the spatial/temporal limitations of our observational capacity and the lack of integration of data, models and visualisation tools across the air-land-water domains. However, recent advances in technical methods allow for detection of real-time changes in biogeochemical, hydrological and ecological functioning (e.g. molecular markers, isotopic and spectroscopic approaches, land and space-based observational techniques). Given a platform where these observations could be explored, accessed and integrated with models across domains, a fast and more informed analysis of system change would emerge, leading to tools which identify options for immediate, targeted and thus more cost-effective management interventions.

The EVOp project therefore needed to develop a platform for both data and models to answer these two questions. The approach taken was to represent data and models as services in a secure cyber-infrastructure. In the long term there would be a need to ensure it possessed a robust architecture, standards, and global access linking private and public clouds, GRID and HPC environments where appropriate. Such a platform would directly address the two challenges, leading to improved exploitation of NERC data and models, and a more integrated response to urgent environmental challenges. The alignment of science supply and demand in the context of continuing scientific uncertainty will depend on seeking out new relationships, overcoming language and cultural barriers to enable collaboration, and merging models and data to evaluate scenarios.

The £2 million pilot Environment Virtual Observatory pilot (EVOp) project was commissioned by NERC in January 2011 to test these two hypotheses and identify opportunities and challenges that might lead to a potentially far greater investment by NERC in partnership with stakeholders into the future.

1.3 Building the EVOp team

NERC recognized the project required a community-led cyber-infrastructure development and new

approaches to scientific workflows that describe, compose, model and execute ensembles of data, models, tools and visualisations on distributed resources with global access.

The EVOp team therefore required a mix of IT and computer specialists, a test research community of scientists and a potential future end-user community drawn from government, regulators and industry.

A sandpit was organized which brought together a community from which a single large project consortium representing 12 institutions emerged incorporating a mix of IT specialists and soil-water scientists (see online Annex), later this would be supplemented by additional involvement from individuals from other organisations Soil-water science was proposed by NERC as an ideal community to test this approach as it is experienced in cross-disciplinary working , has a good range of well-tested models and is well-linked to a range of end-users tackling significant environmental challenges such as flooding, diffuse pollution and climate change.

The consortium elected a leadership team to represent the different communities needed to deliver the pilot project spanning IT specialism (Robert Gurney), basic research (Bridget Emmett) and industry needs (Adrian McDonald). A Project Advisory Group was established covering a wide range of potential end-users including representatives from government, industry, regulators, policy makers and funders (see online Annex).

1.4 Project Structure

A major objective of the consortium from the beginning was to ensure that the work programme was science and end-user led underpinned by a robust exploration of the available IT technologies.

Five principal areas of activity were identified, some requiring narrow and deep testing of the approach across different operational scales (i.e. data and model application from local to global scale) whilst other areas needed a broad and shallow exploration across a range of challenges (i.e. IPR and data security). The five areas were:

i. legal and security issues associated with security of data handling and consumption ensuring the EVOp can scale rapidly without compromising the integrity of data;
ii. data licensing, platform hosting and licensing, and use of the platform taking into account the range of data initiatives including data.gov.uk;
iii. options for cloud infrastructure with security-by-design inbuilt;
iv. development of standards and inter-operability;
v. case studies focused on soil-water process understanding and management at three scales (local, national and global).

Six work packages were developed involving overlapping teams to cover these areas of activity:

i. WP1 Leadership and management (covering legal, security, outreach issues, future funding and responsibility for commissioning and delivery of the global exemplar in Year 2);
ii. WP2 Cloud infrastructure (data licensing, platform hosting, cloud infrastructure, inter-operability, standards);
iii. WP3 Modelling in the cloud;
iv. WP4 Local Exemplar;
v. WP5 National Exemplar (hydrological and biogeochemical exemplars);
vi. WP6 International engagement.

1.5 Deliverables

The high level deliverables reinforced the pilot status and the expectations placed on the project. They were selected to facilitate advances in understanding about the issues and feasibility of delivering such a service to the community. The development process was therefore experimental and although some useful and interesting science emerged and technology solutions generated, the resulting service was not envisaged to become operational as an outcome of the pilot project. In recognition of this status, emphasis was placed on identifying future ways to progress the EVO concept. Furthermore, throughout the project, the team endeavoured to raise awareness about the EVO beyond those communities already familiar with the potential capability provided by IT. The deliverables were:

i. A tested web service providing web based environmental models and data across a limited number of exemplars at a variety of scales. The nominal scales used are referred to as local (catchments in the range 10 to 200 km^2), national with some sub-national division and global.
ii. An identification and analysis of barriers and opportunities revealed in the development of deliverable (i). The barriers considered range from technical feasibility, ownership, governance and financial through to the to the legal and liability issues. Opportunities range from improved science (both in new questions and new solutions) because of the better data and modelling synthesis and the resulting communication through improved public awareness to financial opportunities and market leadership.
iii. A skilled and engaged community is, in part, a direct outcome of deliverable (ii). In a pilot project, engagement is more easily achieved than the 'skilling' up of many groups of stakeholders because the product is not created until a considerable way through a project. The skills development is therefore not seen as a technical training skill but as a higher-level conceptual and vision framework - the skill to recognise the opportunity and potential and to contribute to the

moulding of the future information framework. Within the team, translation across the science/IT interface inevitably resulted in improved skill sets.

iv. Funding and partners in place. The final deliverable is the positioning of the EVOp for further development. To strengthen partnerships through membership of important cross disciplinary alliances such as LWEC, contributions and organisation of national and international conferences and the partnership with industry and across research agendas.

1.6 End-user engagement and the use of storyboards

A major innovation of the project as a whole was to use storyboards to ensure the exemplars were grounded in real questions/challenges by end-users. Potential end-users were considered to cross a wide range of communities from governments to public, and industry to regulators. Initial scoping of likely questions from a range of stakeholders are indicated in Figure 1.1, a full list is provided in the online Annex.; these were developed with, and approved by, our Project Advisory Group. Single specific questions and storyboards were then developed for the local, national and global exemplars. These are summarised in Figure 1.2, with the fully annotated storyboards provided in the online Annex. A requirement of these storyboards and exemplars was to test particular aspects of the data IPR, model operability in the cloud and critically different elements of the cyber infrastructure as indicated in Figure 1.3 .

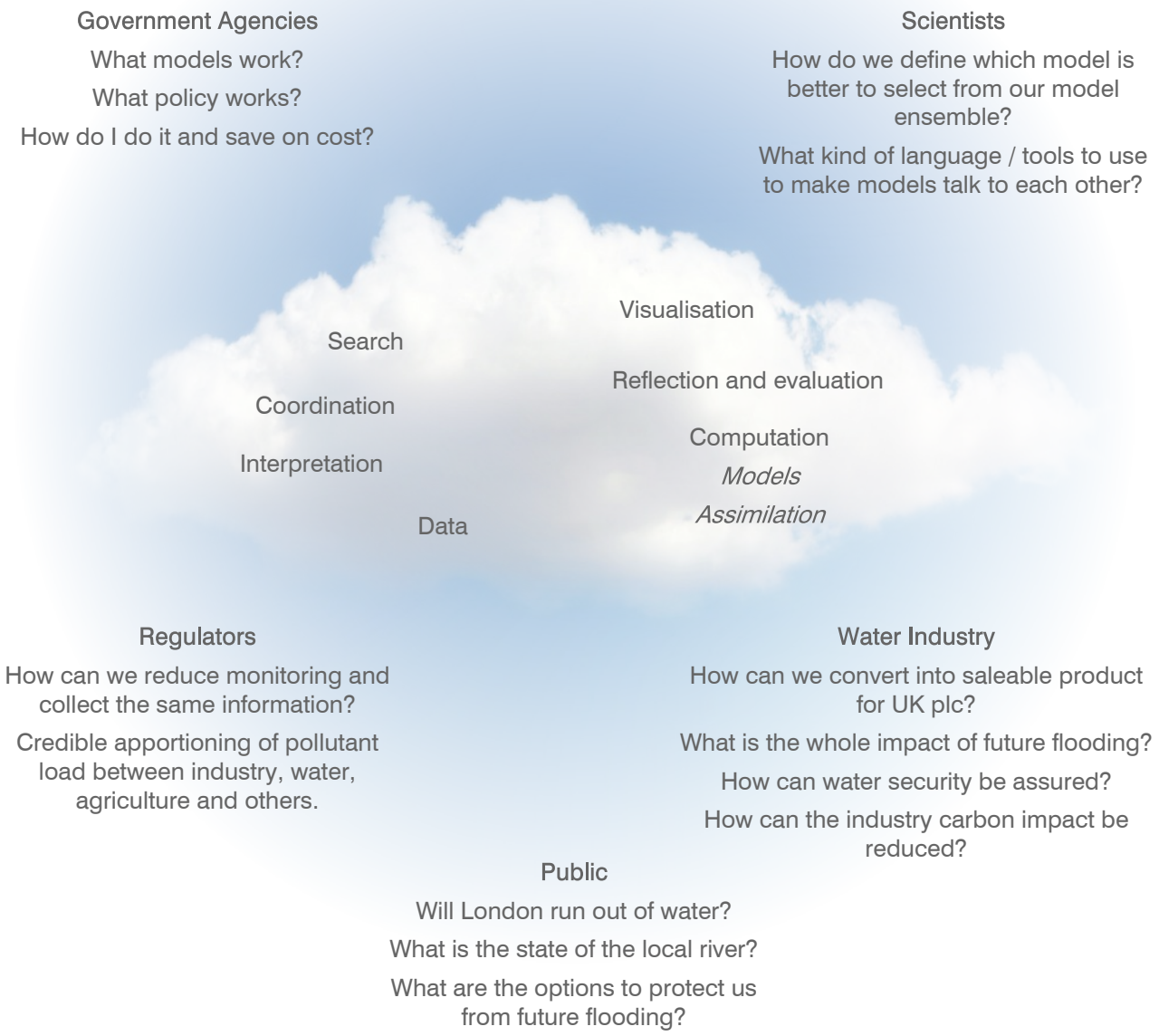

Figure 1.1 Exemplar questions provided by a range of EVO stakeholders.

	Local	National hydrology	National Biogeochemistry	Global
User	Farmers and local stakeholders; insurers	Power companies primarily, but other alerted stakeholders associated with abstractions	Government departments and agencies	NERC scientists; DECC
Issue	Flood risk	Drought	Water Framework Directive compliance and OSPAR reporting	Uncertainty in GCMs and soil C
End product	Local catchment flow and water data and user-friendly interface for modelling tools to forecast flood risk	Selection tool for modelling water resource; Alert tool for industry as to when weather conditions may threaten energy production due to drought	Modelling tool to quantify nutrient fluxes for a range of catchments and marine water bodies at different scales	GCM analogue tool linked to impact assessment model to quantify global regions where uncertainty in change of soil C are greatest
Unique data needs OS / NERC / EA / Met	Live and historical flow data; Local web cams	Live EA data; Access to forecast products		Global databases and driving variables
Science demonstrated	Linking sensors, data and visual data together; Land management decision support tool; Preparedness assessment tool	Multiple model application for hydrology; Uncertainty; Model selection tool; An ensemble of coupled predictive capability	Multiple model application for biogeochemistry; Uncertainty; Regional modelling framework; Model selection tool; Data rich to data poor catchment	Sensitivity to climate change impact models to uncertainy between different GCMs and speed in assessing different emission 'storylines'
EVO 'added value'	Access to data and models; Integration of functionality	National security; Alerts for end-users; Old and new data view and model forecasts for multiple sites; An exemplar for a real time security management issue for important national infrastructure for multiple pressures	Online scenario testing; Ensemble macronutrient modelling; Scaling and multiple metrics	Access to climate change impacts assessment tool with increased capacity and faster working
Future potential	Adaptive modelling' for local conditions; Ask an expert blog; Preparedness tool	Approach shows data/model cloud resources for identifying forecast alerts for any environmental threshold	Reverse engineering' tools to find solutions; Dynamic coupling with hydrological modelling	Linking to other land-atmosphere models; Addition of valuation tool

Figure 1.2 Mapping of exemplars to users and issues.

Framework properties		Local hydrology (Eden)	National hydrology	National biogeochmeistry	Global
Essential infrastructure properties	Everything as a service	O	O	O	O
	Sharing of everything	O	O	O	O
	Openness and interoperability	O	O	O	O
	Transparency	O	O	O	O
	Ease of use by different communities	O	O	O	O
The cloud as a utility	Alternative business models	o	O	O	
	Elasticity	o	O		O
	The managed cloud	o	o	o	
Enhanced management	Tailored management models		O		
	Multi-cloud management	•	•	•	
Web 2.0 techniques	Supporting mashups	•	•	•	
	Supporting workflows	•	•	•	
	Service discovery	o	•	•	
	Enhanced discovery	o			
Systems of systems	Combining with ubiquitous computing	o	•		
	Combining with mobility	•	•	•	

Figure 1.3 Date security.

O = major effort, o = some effort, • = left for full EVO
O = completed, O = required

1.7 Data security

Concerns over data security were identified as a critical issue for many potential providers. Within the EVOp project a scoping exercise was undertaken through a workshop approach led by the data security industry. A report was delivered to the EVOp leadership team to help inform NERC of future needs, should a fully operational EVO be commissioned. The EVOp Cyber Security Advisory Board proposed six key areas that would require further consideration should the EVO concept be commercialised. Details are contained in the report provided in the online Annex, and can be summarized as:

i. Striking the right balance (confidentiality, integrity and availability): A clear understanding of the core principles and their corresponding priorities is important in the 'Security by Design' approach. Additional core principles that should be considered include non-repudiation, authentication and privacy. It is also clear that the importance of each core principal will vary based on security standards, classifications and the target audiences for each model, tool and data set.

ii. Cloud Security: Considerations associated with cloud security fall into two areas. These are related to cloud providers and customers (i.e. the EVO). In the continuously evolving domain of cloud security, the Cyber Security Advisory Board will provide important guidance to ensure the right cloud providers are chosen and all cloud security issues are considered.

iii. Data Protection Methods (encryption and other): Perimeter or layer protection methods (i.e. firewalls and IDS/IPS) are common focal points in the protection of data. Whilst these methods serve their place, an additional consideration for the EVO is security of the systems that will store and process the most sensitive data. Implementation of appropriate encryption methods may be a suitable method for ensuring that data is protected within the data centre(s), not just at the perimeter. Understanding the ideal encryption methods and how and when to implement them is vital in maintaining the high performance of EVO virtual modelling and the transmission of live data (such as river levels, real-time temperatures, etc.).

iv. Application Security: Whenever an EVO application interacts with suppliers and customers (end users), both the data and the application itself must be protected. As the EVO is expected to allow end-users to create data workflows to modify how data is analysed, it will be important to keep an audit trail of user activities and alert on behaviour which falls outside the norm. Just as important is the need to have an audit trail for all activities undertaken on any database which stores critical EVO information.

v. EVOp Portal Security: The EVOp portal will aggregate content from all of the EVOp systems providing a means for users to explore a variety of data sources and execute simulations and models in the cloud. It is vital that portal security ensures that only an authorised user can generate requests to the applications server(s).

vi. EVOp Security Resource Considerations: A core team will need to be assigned with responsibility for developing a Security Policy Framework and driving cyber security as a business as usual function.

While the six key areas cover the key cyber security considerations such as legal implications and the interaction between these topics; alignment to current security standards and their robustness within a rapidly evolving sector, the legal implications on security, and domain specific; security considerations are also in need of further investigation.

1.8 Legal Issues

A preliminary report on legal issues was prepared by The London Institute of Space Policy and Law (ISPL) and Edwards Wildman as an output of an EVOp-convened workshop. The white paper is split into three parts:

i. Data Collection and Licensing Issues: There will be a number of issues to consider in respect of the collection of the underlying data to be populated in the EVO platform. These include issues arising out of the use of third party intellectual property rights or underlying data that is otherwise in the public domain.

ii. The EVOp Platform: There will be a number of issues associated with the hosting of the EVO platform itself and of the development of new tools and applications. This includes issues arising out of the use of cloud computing technology.

iii. The use of the EVOp Platform and Licensing Issues: There will be a number of issues associated with the use of the EVOp platform, including the onward licensing of the output data and the EVOp platform's potential liability for any reliance placed on such output data.

The legal implications for development of the EVO were considered in parallel to the cyber security considerations. The interdependency of these two areas was acknowledged and further consideration on how to remain joined-up would be required if EVO were to become operational.

1.9 Outreach and community building

One of the deliverables of the EVOp project was to help develop an engaged and educated community. Members of the team were enthusiastic in this endeavour and participated in a wide range of activities including briefings, workshops, summer schools and conferences.

1.9.1 Briefings and conferences

Members of the team were active in promoting the EVOp vision and key outputs within different forums. A summary of interaction with key stakeholders and

events where EVOp featured are provided in the online Annex.

1.9.2 Training

Training was achieved in several ways throughout the course of the project:

i. Seminar on cloud computing: Representatives from WP2 organised a seminar for EVOp team members and staff at Lancaster University during the early stages of the project to provide a foundation level of knowledge and consequently, a baseline for communication.

ii. Involvement of early career scientists: Fourteen Post Doctoral Research Assistants and Research Assistants were employed across science and technology work packages.

iii. Summer School: The leadership team initiated a summer school at Istituto Veneto in Venice sponsored by NERC. The focus of the 2011 summer school was closely aligned to the EVOp and the six EVOp PDRAs that attended received an overview of cloud computing and the type of models that have been used already in environmental sciences in cloud services. Students with no previous advanced computing experience were able to design a working web service in three days, or carry out new work in data assimilation.

These activities demonstrate the EVOp, or an allied concept, could be rolled out to the science community, and also the operational community, without a very high training barrier. In particular, the summer school illustrated that it would be possible to develop a training course, possibly on-line or through webinars, to deliver the necessary training quickly and without considerable expense for scientists to appreciate the potential of the approach. It must be noted however that professional trained computer scientists are essential to design and implement the actual cyberinfrastructure. Indeed, one of the main issues during the EVOp project was the over-reliance by scientists on too few computer specialists. Any future related activities would need to correct this and provide and a more balanced distribution of skills within the team.

1.9.3 Project Advisory Group

The project benefited from the input from an Advisory Group which met on four occasions (approximately every six months). Membership of the Project Advisory Group was diverse and incorporated representatives from the water and IT industry, regulatory bodies, government, and academia (see online Annex); they were extremely supportive of the potential and need for an EVOp approach to gain greater efficiency, effectiveness and transparency of NERC science.

1.9.4 Alignment with other UK national initiatives

Awareness of aligned initiatives was essential in identifying how a future fully operational EVO could co-deliver more efficient, accessible and transparent data and models. We identified the following aligned initiatives and held various briefing/meetings with key participants including NERC Theme Leaders and Principal Investigators:

- Data: NERC has invested National Capability funding in an array of Data Centres which ensure secure long-term storage and access to data resources. Any future EVO platform would need to work together with the Data Centres and develop a clear interface to these as well as other initiatives such as data.gov.uk, the UK locator programme and those ongoing in the Met Office and Ordnance Survey. This would increase uptake of current investment and enhance ongoing work on international standards and data access. Testing of links to some Data Centres were explored within the Pilot and links and discussions with other data providers will inevitably be ongoing in future EVO-related initiatives.
- Science: A range of research investments are in place which cross the air, land, water, geological and climate communities providing resources and researchers actively pursuing the integration of NERC science. Of particular relevance to the EVO activities are the Programmes which seek to bring together the biogeochemical-hydrological and ecological communities (e.g. Macronutrients, BESS, Changing Water Cycles and Network of Sensors). These programmes all seek to improve the science which underpins the sustainable exploitation of our natural capital therefore providing a wealth of resources to underpin future EVO activities.
- Modelling: NERC's Integrated Environmental Modelling Initiative which emerged from NERC's proposed Modelling Strategy seeks to identify benefits of improved model integration and data exchange tools. Running concurrently to EVOp, a web portal, or 'Experimental Zone' was initiated under the NERC PURE programme to facilitate the sharing and fusion of models and data from a variety of different sources information to practitioners in environmental risk management.
- Tools: The Environmental Science to Services Partnership (ESSP) is a joint initiative by NERC, the Ordnance Survey and the Met Office to develop new products from current data and knowledge to improve uptake and impact of environmental science. Thus there is input by Defra, the EA and NERC capabilities through CEH and BGS.

1.9.5 Alignment with other international initiatives

In addition to the array of data, standards and climate initiatives ongoing in the European and global arena the following have made specific links to the EVOp team and indicated their interest in joint future collaborations:

- NSF EarthCube: Developing a framework to create and manage knowledge in geosciences to understand and predict the Earth system. There are opportunities to move forward in the short-term with joint legal workshops where EarthCube can

learn from our current knowledge and in turn, NSF can fund future joint workshops together with joint SAVI grants.

- NSF Neon: A continental-scale research platform for understanding and forecasting the impacts of climate change, land use change and invasive species, on ecological processes and on interactions of the biosphere with the geosphere, hydrosphere and atmosphere. NEON will collect data from 106 sites, over the next 30 years, from an investment of $434M.
- Knowledge Systems for Sustainable Landscape Management Initiative: An initiative to develop more sustainable integrated land management solutions building upon existing data collections and analysis tools, with investments for the development and deployment of new tools. Partners include USAID, Oak Ridge National Lab, NASA, CGIAR, World Bank, NASA, CSIRO.
- Belmont Forum: High level group of the world's major and emerging funders of global environmental change research and international science councils. NERC International team facilitated participation in activities to define future research agenda.

1.9.6 EVOp International Conference

AAn international conference to explore how new information technologies, and in particular cloud computing, could be used in the environmental sector was organised by the EVOp with support from five learned societies (British Ecological Society; British Hydrological Society; Royal Geographical Society; Royal Meteorological Society; and The British Society of Soil Science). The conference, "Harnessing Emerging IT Technology for Environmental Science - A 2020 Vision" was held at the Royal Geographical Society in London on 16 May 2012 . The event brought together a diverse range of topics in the form of presentations, interactive demonstrations and posters (see online Annex).

From an early stage in the project, the need for an EVO-led event to provide a dedicated platform where the ambitions and outputs of the pilot could be showcased was apparent. Whilst the team would be active in communicating different aspects of the project within the various academic forums relevant to their expertise, there would be no single academic conference that could adequately represent all areas of interest. It was envisaged that an EVO-led event would be an opportunity where the future horizon for environment science, as influenced by technology, could be explored in addition to raising awareness about the EVOp project.

Although unique at that time in its ambition to combine data, models and tools into a single platform, the EVOp has drawn upon knowledge and links with other international science and technology programmes and initiatives active in generating new standards and new ways of working. Through the conference, the pilot EVO played a role in promoting the need for groups from disparate disciplines to come together and address the following questions:

- How can advances in IT help to solve or ameliorate major environmental issues?
- What are the practical barriers hindering and the opportunities to encourage integration between IT, research and user communities?
- What approaches, individuals and institutions appear to constitute the cutting edge of this IT - environment integration?

Over 160 people attended the event which included 36 posters/demonstrations on relevant topics and sixteen presentations. Each session offered a different perspective, aimed to inform and provoke discussion on the potential links between the environmental science and IT communities both in the UK and elsewhere. Lord Selborne provided the keynote address in which he outlined government policy and the programme landscape. Other speakers included representatives from the EU and UK government, IT industry, academic and research sectors:

ARUP Australia, British Ecological Society, CEH, Defra, Dtex, European Commission, Geoscience, Google, Manchester University, Met Office, Microsoft, University of Cambridge, University of Colorado, and Willis Group Holdings.

The full conference programme is available within the online Annex. Recordings have been prepared by the Environmental Sustainability Knowledge Transfer Network and can be viewed on the ESKTN TV YouTube channel. Website, EVO portals and publications

1.9.7 EVOp website

A website - www.evo-uk.org - was established early on to advertise the activities and ambitions of the project and to bring in national and international partners.

Initial work to scope out a potential landing page and the navigation routes end-users may select was undertaken by the team with design direction provide by website consultants INTRO. A key issue identified was the different ways in which end-users liked to access resources. An attempt to accommodate these different preferences is reflected in a landing page which has options to explore the EVO resources by location, topic and data. For the more experienced user, there is also provision of direct access to models, tools, apps and the workflow library (Figure 1.4).

The current design provides an adequate mechanism for demonstrating the pilot web service to EVO end-users; it was however noted that further work would be needed to develop different portals for different communities should the EVO ever be operationalised. A recommendation from the EVO pilot is that any future related initiative should focus more resources on this aspect and include a separate activity for planning and end-user testing of the portal interface.

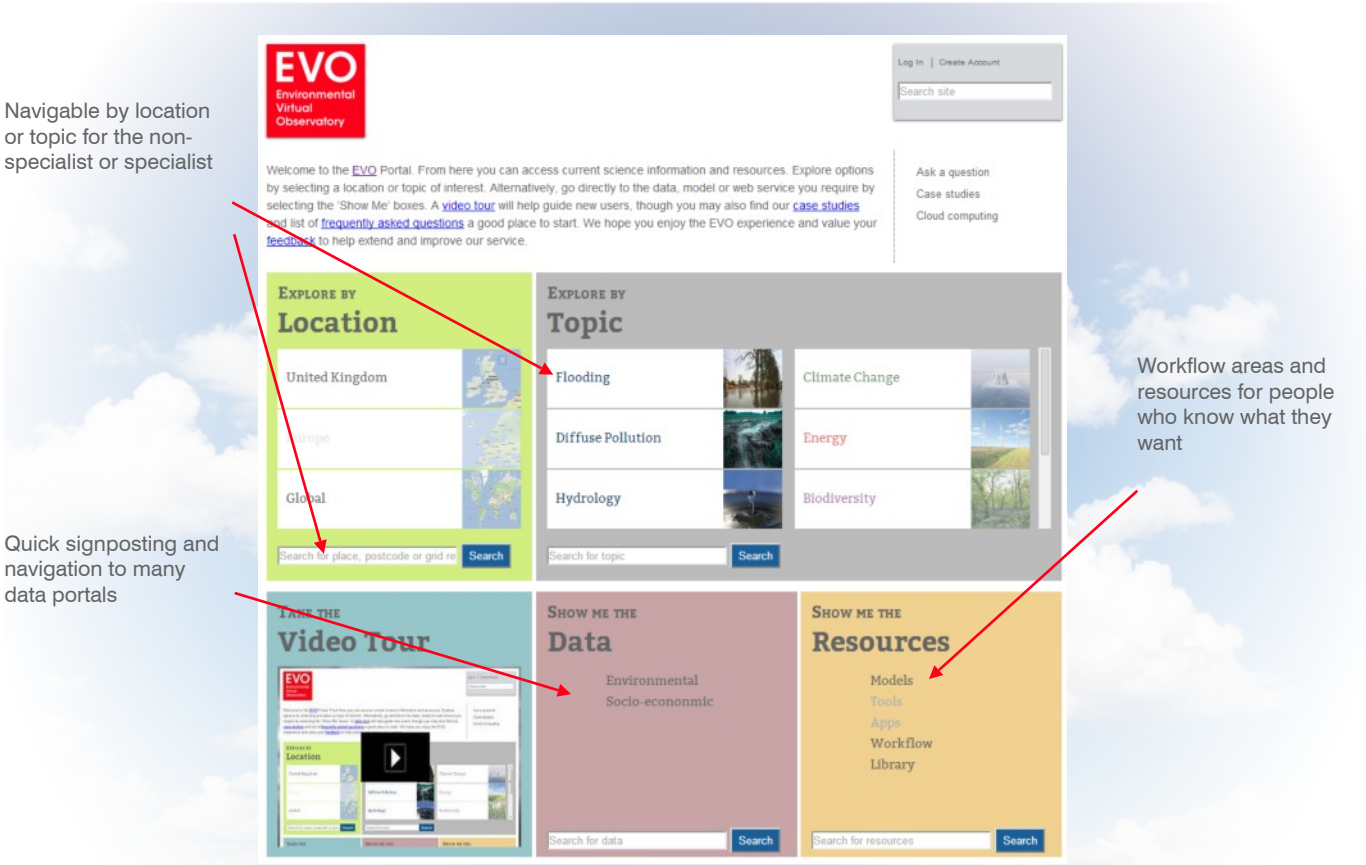

Figure 1.4 Design of the EVO portal.

1.10 The EVOp legacy

As one of the key aims of the project was to provide information to the wider community on opportunities and barriers in the EVOp approach, a series of legacy publications are provided within the online Annex. Videos of the local, national and global teams talking through the exemplars are also available on the EVOp website along with two-page flyers covering each of the work packages. An international publication summarising the overall project is also available (International Innovation paper - online link). This publication is an open access publication designed to communicate worldwide environmental research and development. The publication is distributed to over 30'000 stakeholder readers at all levels in the government, policy, research and related health stakeholder sectors and communicates the impact and relevance of both fundamental and applied research in the field.

Finally, EVOp has played a role in raising awareness and contributed to establishing an engaged and educated community that are becoming increasingly more open and ready to exploit the advantages new technology holds and benefit from future initiatives. Thus, perhaps one of the most important legacies of the project are the funding opportunities which have emerged from NERC and others.

1.11 Funding landscape and partners

1.11.1 NERC

By early 2012, it became evident that to make the EVO operational would require mixed funding to underpin the complex range of activities spanning: capital, research and private sector investment along with international and stakeholder collaboration. NERC have opted to continue a range of activities under the Environment Information Initiative.

- Environmental Big Data Investments: JASMIN, CEMS, Environmental Research Workbench, Environmental Big Data Capital Call.
- Innovation Activities: NERC Environmental Data: short projects to consider applications, products and services as a precursor to the joint NERC/TSB call, Solving Business Problems with Environmental Data".
- Belmont Forum: e-Infrastructure Knowledge Hub programme).

Several of these initiatives have adopted approaches that utilise and build upon knowledge acquired during the EVOp explicitly and thus the legacy and value of NERC's early investment was well founded.

1.11.2 Agencies, government and industry

A forum was convened by NERC to discuss the needs and opportunities for demonstrating tangible economic

and societal benefits from NERC's research investments as a whole. This included representatives from Defra, the EA, the National River Trust, British Water, Thames Water, the Welsh Government, the Scottish Government, Natural England, UKWIR, the Water Consultancy Community, SEPA, the Environmental Sustainability KTN, the NERC Water KE Programme, NERC Science Management and the EVOp project. A range of activities were proposed to better integrate the UK modelling resource and provide a forum to promote exchange between the developers and users of environmental models. Key resources to realise this ambition were to operationalise the outcomes of research programmes such as Macronutrient Cycles, Biodiversity & Ecosystem Service Sustainability, Changing Water Cycles and Environmental Virtual Observatory. This would serve not only to inform public policy and regulation but enhance business performance and practice, with consequent benefits for both UK plc and the environment.

Specific funding opportunities which have recently emerged that seek a cloud based, or EVOp-type approach, include:

- A call for by the EA for requiring cloud-based data, exploration and visualisation approaches for delivering the Water Framework Directive.
- Commissioning of an integrated data and modeling platform for diffuse pollution management and ecosystem services by Defra, the EA, NERC and Scottish Government from a consortium led by CEH, one of the EVOp leadership team.

1.12 Lessons learned, and opportunities and barriers identified

These are outlined in various parts of this final report but in summary:

- Consider the needs of end users and maintain this focus throughout development. Perhaps the most successful part of the EVO project was its strong focus on end users. The advice from the Project Advisory Group to use storyboards to articulate the needs of the user informed both the overall appearance and development of the portal, as well as informing the IT requirements. This approach is now embedded at the heart of other recent funding calls. The AGILE approach was also influential in the project design and encouraged exposure of the developing system to stakeholder input. This aspect is often the most used and communicated part of the project.
- The concept of deploying data, models and tools as services in the cloud was demonstrated to be an effective way forward. A mix of commercial and private cloud providers is likely to be the optimum solution. The cost of commercial providers is likely to be prohibitive for large modeling applications. There is much to be learned from approaches adopted by other communities e.g. the astronomy community and bioinformatics community (e.g. CloudBioLimux) which could accelerate the progress of the environmental community in this field.
- There is a continuing need for international collaboration in the development of standards and inter-operability. The EVOp project effectively borrowed from existing activities in this area. The Belmont Forum activities should ensure progress is made in the area in the immediate future.
- It is critical to ensure an appropriate balance and mix of science/IT skills in any future initiatives. The EVOp despite trying to ensure this from the outset had insufficient IT / computer science postdoctoral researchers. There was a pinch-point between the enthusiastic and engaged catchment scientists with many ideas and the capacity to make these operational within the cloud.
- Incorporate a small number of carefully selected exemplars in the early stages of system development. This was a major success of the EVO project to both provide cohesion to the diverse and large team and facilitate outreach to a broad section of the potential end-user community. The pace of development is inevitably enhanced by utilising established science for the exemplars.
- The culture change. One of the areas of the EVOp project of most interest to end-users was the potential for a change in research culture. Could the community move towards a more open and transparent way of working? Would they be willing to web-enable their models and encourage independent testing to gain greater trust by the end-user community. The potential for the EVO-like service to effectively act as a 'market place' for models; potentially only those made freely available and web-enabled in the EVO or cloud would gain traction in the end-user community was appealing to some end-users as they sought ways to move forward to develop more tested and integrated modelling approaches to tackle the complex environmental challenges they face.

1.13 Conclusions

Although not operational, the EVOp portal has demonstrated the value of integrating fragmented and widely-distributed public and private sector data, expert knowledge, modelling tools and visualisation services. It has illustrated how cloud computing improves the efficiency, speed and effective use of such resources. Valuable lessons have been learned and key issues identified for future investigation. Out of all the impacts, perhaps most important of all is the support and enthusiasm it engendered from its stakeholder groups, the potential for expansion to other science areas, industry applications and benefits arising from social and economic data were evident.

From the outset, the EVOp was purposefully ambitious. It sought to demonstrate new capability and ways of working, to expand the knowledge base and sign post the way for the future. In achieving these aims the EVOp has undoubtedly had a bearing on current

funding and approaches being taken in this arena. In the context of research programmes, the outputs and impact of this pilot are significant considering the modest investment and timeframe for implementation.

1.14 References

Gray J. 2007 Talk given by Jim Gray to the NRC-CSTB in Mountain View, CA, on January 11, 2007

Hey T, Tansley S, Tolle, K (Eds). 2009 The Fourth Paradigm Data-Intensive Scientific Discovery. Microsoft Research

The Royal Society, 2012, Science as an open enterprise. The Royal Society Policy Centre.

Related links

NSF EarthCube

http://www.earthcube.org

Earth System Science Partnership (ESSP)

http://www.essp.org

Environment Information Initiative

http://www.nerc.ac.uk/research/capability/

Environmental Sustainability Knowledge Transfer Network

https://connect.innovateuk.org/web/sustainability ktn

ESKTN TV YouTube channel

http://www.youtube.com/playlist?list=PLE862DA58 53C19459

Natural Environment Research Council (NERC)

http://www.nerc.ac.uk

London Institute of Space Policy and Law (ISPL)

http://www.space-institute.org