

# Social science for e-science in biodiversity research

## A position paper on behalf of ViBRANT

Prepared for: Digital Social Research: A Forum for Policy and Practice, March 13, 2012

Daphne Duin<sup>1</sup>, David King<sup>2</sup>, Peter van den Besselaar<sup>1</sup>

<sup>1</sup> Dep. of Organization Sciences. & Network Institute, VU-University Amsterdam, Amsterdam, The Netherlands

<sup>2</sup> Department of Computing, The Open University, Milton Keynes, United Kingdom

### Introduction

The application of e-science, the use of information and communications technologies (ICT) to support scientific work, increases the scale, access and diversity of information and collaborations and enables research on complex and interdisciplinary questions. However, smart technologies alone do not make e-science a success. Insights on the user and usage and the understanding of the societal and organizational dimensions contribute to the design, user uptake and sustainable use of the infrastructure. Hence, social science has a significant contribution to make to the development and implementation of e-science facilities. The position we would like to bring forward to this meeting is based on our experience in the e-infrastructure project, called ViBRANT (<http://vbrant.eu>). Below we highlight why and how, according to us, social science contributes to ViBRANT, under what conditions and what issues still need attention.

### E-science project: ViBRANT

ViBRANT is an e-science project with the ambition to increase the number of people actively using e-infrastructures to publish their biodiversity data for use by all. The infrastructure supports mobilization, sharing, reuse and publication of biodiversity data, multiplying the investments already spent by society in the collection of those data and helping to focus the output from the European and global biodiversity research community. ViBRANT is funded under the 7<sup>th</sup> Framework Programme, Research infrastructures. The project consortium consists of an interdisciplinary group of research and technical staff from different research fields, such as: biodiversity research; bioinformatics; computer science and social science. ViBRANT applies a bottom-up, agile development approach, which means that components are developed through close collaboration with the user community, are launched as early as possible, and then their performance is refined in the light of user experience. So understanding the user is crucial for the project's development.

### (e-)social science in ViBRANT

In ViBRANT we identified three specific tasks for social science:

1. Measuring the impact of e-infrastructures. For example, establishing a baseline of researcher activity, and demonstrating how infrastructures add value beyond that baseline. Altmetrics, mainly webmetrics, form a major part of this effort, (eg metrics of content contribution, metrics of citation) are based primarily around digital objects as opposed to traditional published papers. Likewise, log analysis (and determining what want be collected) is also part of this task. Essentially, it is about meta-analysis of user data.
2. Incentivizing user engagement with the system. This is closely connected with 1, but much more focused on the individual user, as opposed to a meta-analysis of users. Again this concerns altmetrics, but also relates to identifying the views of users about an infrastructure, such as what are their primary concerns about engaging with an infrastructure, and what we can do to address these concerns. This task makes important contributions to visualize the uptake of e-science at the level of researchers, research institutions and other stakeholders.
3. Informing functional development activities and prioritization of development activities, particularly via passive intervention with users to collect information on how they use a system. This is task is about developing technical interventions in the system to collect information about what users are doing and

where possible, about the context in which they are working. The meta-analysis of this collected information enables us to prioritize our development activities. Consideration must be given to the data protection issues that surround the collection and processing of this information. For instance a complication for ViBRANT is the different jurisdictions between the UK where the main server is located and Germany where the mirror is based.

### **Conditions for successful (e) social science in ViBRANT**

E-science infrastructures are a fascinating subject for those social scientists interested in the organization of science and in how technology is reshaping scientific practice. According to our experience, the following conditions contribute to the success of this type of research:

- General acknowledgement by the e-science facility and by funding agencies regarding the value of the contributions that social science brings to the design, uptake and sustainability of e-science.
- Support by project management and developers to access user and use data.
- Investment within social science institutions in facilitating storage of large datasets and access to technical expertise to help guarantee data security and other technical support needed to work with large data sets.
- Consideration and expertise on privacy issues and national jurisdiction around the use and access to web data and other ethical questions raised by new forms of data collection and analysis.
- Introduction of team science in social science. E-science is moving quickly and would benefit from a high publication cycle.
- Collaboration with computer science in collecting and analyzing electronic data and large datasets
- Interest among project partners in writing interdisciplinary publications.
- A project management culture with room for individual partners to initiate new collaborations with other partners in the course of the project
- Credits for interdisciplinary research.
- Credits for applied or technical contributions, eg data papers.

### **Challenges to successful (e) social science in ViBRANT**

Thus e-science facilities are an exciting new field of research for social scientists. Nevertheless the type of data and the scale of data also poses challenges to them and their organizations. Hence some of the conditions mentioned above are not yet (sufficiently) fulfilled in our view. In particular, we think the following issues need to be addressed to make social science in e-science flourish:

- General acknowledgement and action by social science institutions to facilitate research on e-data, eg storage of large datasets, other technical support, training in analyzing e-data.
- An encouraging, institutional climate for interdisciplinary research, particularly for collaboration with Computer science. Recognition of publications in outlets outside one's own specialism. This is not only an issue for social sciences but for most disciplines collaborating in ViBRANT.
- Knowledge or access to knowledge regarding privacy issues and different national jurisdictions around web data and other ethical questions raised by new forms of data collection and analysis.
- Culture of team science. Many social scientists and their departments still prefer publications with one or two, at most three, authors.
- Acknowledgement for applied or technical contributions, eg data papers.