

Terms of Reference

For the PLAtform of National eScience/Data Research
Organizations in Europe

PLAN-E

Agreed by the Platform members at the constituting meeting

Amsterdam, September 29, 30 2014

1 Preamble

Collaboration between disciplines, the combination of research and instrumentation –including networks and computational and data resources- has brought about new ways of conducting research. The approach is problem driven, viewed from different disciplinary angles, and usually strongly data intensive. Organizations involved in bridging the gap between the science domain – where the research questions are formulated – and the underlying ICT and e-infrastructural technologies in support of solving those questions, seek to organize themselves in a community platform in support of their activities.

2 Definition¹

eScience is a community of practice, characterized – but not exclusively or limitedly – by the following:

- It is concerned with innovative ways in which ICT can be *applied* to complex scientific or industrial problems;
- It is concerned with the support of multi-disciplinary research, for example through but not limited to cross-type data integration, the managing of structured and unstructured data sets, data-driven research (“Big Data” research) and data analysis;
- It is the application of computer technology to the undertaking of modern scientific investigation, including the preparation, experimentation, data collection, results dissemination, and long-term storage and accessibility of all materials generated through the scientific process²;
- It applies computer algorithms and tools for the interactive specification and maintenance of models and their analysis, visualization and simulation, in order to support scientific *in silico* experiments;
- It is concerned with the optimal use and/or optimization of the use of larger parts of an e-infrastructure for scientific applications for complex scientific problems and/or
- It is concerned with the optimal use and/or optimization of the use of (high-end) computers in scientific applications for complex scientific problems and/or
- It is concerned with addressing scientific usage of computers and/or e-infrastructures in cases where the problems may for example be on the following or a combination of these:
 - o Compute bound
 - o Latency bound
 - o Data streaming bound
 - o Data complexity bound

¹ Whereas the term `science` in UK English strictly refers to the natural sciences or “exact” sciences rather than to the whole domain of academic research areas, in this document and in all subsequent documents of PLAN-E, the term science will be used to refer to all domains in which academic and private sector research is conducted.

² Bohle, S. "What is E-science and How Should it Be Managed?" Nature.com, Spektrum der Wissenschaft (Scientific American), http://www.scilogs.com/scientific_and_medical_libraries/what-is-e-science-and-how-should-it-be-managed/.

- Data size bound;
- In general it is concerned with the *application*, re-use and re-usability of ICT methods, methodologies and tools to support solving complex scientific and/or industrial problems and with *development* where such methods or suitably adapted implementations are lacking;
- It encompasses and advocates strongly advanced visualization and pattern recognition in support of its goals.

3 Background

The traditional way in which research is conducted is changing drastically, due to innovations in ICT, e-infrastructures, and the ease to generate and store massive quantities of data. But whereas a broad overview is required of contemporary knowledge of a research domain to address innovative challenges, the same holds true for knowledge of the rapidly changing ICT and infrastructures world. Linking both worlds is the domain of *escience*. Typical characteristics of *eScience* are the apparent broader applicability of ICT in new scientific domains, the highly improved interconnectivity of systems over networks (internet), the use of newly developed intercommunication languages, such as XML, the new directions in database technologies, the upcoming data-driven research, exemplified by Cern's LHC and SKA, but also by the developments in genomics or, more broadly, the bio-informatics. Since data are becoming increasingly dominant in research, *eScience* is also associated with, but does not coincide with the open data, open source, open access and tools, and data preservation movements and developments. The link is the large dependence of *eScience* on reliable, usable, sharable and accessible data sets. Because the complexity of *eScience* research no longer allows for exact reproducibility (if only because massively parallel and/or distributed computers deliver their results no longer deterministically), one of the bases of conducting science properly, the open publication of (intermediate) data involved in research projects becomes an inevitable requirement.

4 Goals

The principal goal of PLAN-E is to act in support of Enhancing Science: The objective of PLAN-e is to bring together leading influential *e-Science* centres across Europe to help coordinate ongoing innovation in scientific methods and exploitation of infrastructure. The goals of PLAN-E cover all the topics that help promoting the *escience* approach and strengthening the groups and centers conducting *escience*. More in particular, PLAN-E:

- Forms a forum for exchanging knowledge and expertise in the field in order to strengthen the European position in the *eScience* domain;
- Discusses common approaches to *eScience*;
- Communicates about *eScience* and the way it is showing results in all disciplines;
- Represents the European *eScience* scene as embodied by the PLAN-E community externally and internationally in addition to the individual representations from the participating members where applicable. In particular towards the EC in relation to future funding schemes;
- Proposes evaluation criteria for the quality, impact and benefits of *eScience* activities;

- Supports actions towards data stewardship and software availability and sustainability;
- Will take endeavours to stimulate quality and quality ranking of eScience publishing means;
- Facilitates the interaction between its members;
- Will encourage and provide eScience requirements towards improved e-infrastructure provisioning and usage;
- Will communicate best eScience practices regarding the use of e-infrastructures and ICT tools;
- Will strive for the improvement of the skills-level of students and researchers in eScience techniques and stimulate the upgrading of the status of eScience technologists and engineers.

5 PLAN-E

The PLAN-E platform is based on voluntary participation of organizations that share the goals of the platform. In a later stage, if the organization has proven robust, it might be considered to base the organization on formal Letters of Intent. The PLAN-E will consider working with a chosen board with a secretary responsible for the communication. The board will consider drafting a white paper on eScience and its present and future impact on science and the way it is conducted.

In principle the documents produced by PLAN-E are public.