Manifesto

The New Global Data Generation

Simplifying and Guaranteeing
Access and Sharing
in e-Science Scenarios

Background

Most researchers during their careers will access large research infrastructures. They are becoming increasingly vital for large international research projects and range from single large facilities to the linking together of massive data sets and tools in ways that transcend individual disciplines. In Europe, ESFRI exists for cases where the infrastructures transcend national boundaries or are beyond the means of individual Member States. At all levels the amount of data created, maintained and shared by these infrastructures has been growing by orders of magnitude over a very short period of time, and the management of this data deluge has become a critical issue if the sheer volume is not to overwhelm researchers. Despite important initiatives to improve the training of people to manage these infrastructures (RAMIRI) and to find suitable legal structures (e.g. the ERIC), there are still major challenges if they are to deliver their full potential.

Sustainability

A frequent problem encountered by such infrastructures is not the initial capital commitment but rather the ongoing running costs, and the costs of curating data over long periods after the initial work has been undertaken, and the research project funding has ended. Initial investments from the European Commission in preparatory phase infrastructure projects have been extremely successful, and it is very promising to see the first ERICs being set up, with significant contributions from the member states. However, this might not be sufficient in the long run. Research infrastructures reflect research communities, which are usually trans-national initiatives. This means that there is a need to have funding streams from the member states and from the European commission to guarantee long-term sustainability and to maintain the trans-national momentum. Yet we cannot see that there is a general agreement about a shared funding model for the future.

We urgently request the European Commission and the member states to come to the necessary agreements about shared funding models for sustainable research infrastructures in order to ensure European competitiveness.

Simplifying the Rights Situation

The legal and ethical minefields of copyright, licensing, codes of conduct and use of user credentials are particularly difficult problems for those involved in building common data repositories and giving access to them. In the spirit of the "fifth freedom" of the European Union, for the free circulation of and access to knowledge, harmonized solutions are required urgently. The current lack of clarity and harmonization will inevitably lead to researchers inadvertently doing things of dubious legality through ignorance, or having to compromise for fear of crossing a legal boundary. A key issue concerns privacy rights of individuals. These need to be harmonized and clarified so that all researchers and citizens can be assured their data will not be misused, but also so that an exaggerated perception of risk in this area does not prevent the secure exchange of data.

We urgently request that the stakeholders in the European Commission, in the member states, in the research infrastructures and institutions, and in relevant industrial and commercial sectors sit together and agree on harmonized and simplified licenses and templates that can be used to facilitate access by researchers to data.

Simplifying Distributed Authentication and Authorization

Many key research infrastructure services require the reliable identification of individuals, their group affiliations and entitlements to promote convenient access based on single-sign-on principles. The situation with respect to identification for authentication and authorization in Europe and in some member states is very fragmented and incoherent. In collaboration with GÉANT/eduGain and the national NRENs (managing national identity federations), infrastructure projects are working to overcome remaining technical problems, but we are faced with some organizational and legal aspects which hamper progress. Many identity providers are hesitating or refusing to provide the required attributes to allow us to identify users and to make decisions about authorization to access resources, including normally non-sensitive information such as names and email addresses. Yet a number of countries are not prepared to participate in a pan-European trust federation and for others it takes far too much time to sign formal agreements in establishing a trust relation with infrastructures as service providers. In addition, we still lack some generic technical components for establishing smooth authentication and authorization, for example in the area of identity delegation in service-oriented architectures.

Although new initiatives such as creating a Code of Conduct for trusted service providers are in discussion, research infrastructures need to invest in setting up ad hoc community-based identity provider solutions to not lose momentum, but this means large investments to serve the thousands of researchers who require access to the services, and could lead to divergent and wasteful reproduction of services.

We urgently request that the stakeholders in the European Commission, in the member states and institutions ensure that the necessary authentication infrastructure is in place at the institutional and national levels and policies across the continent are being harmonized.

Summary

We urge policy-makers to facilitate agreement, harmonization and investment in:

- shared funding models for trans-national infrastructures
- simplified licenses to facilitate access by researchers to data
- access and identity management at national and pan-European levels