



LIGUE DES BIBLIOTHÈQUES EUROPÉENNES DE RECHERCHE
ASSOCIATION OF EUROPEAN RESEARCH LIBRARIES



DG Connect's proposal for a Framework for Action addressing the area of Research Data e-Infrastructures, March 2013

Responses from OpenAIRE, LIBER and COAR

Key points from our responses:

- The framework is highly relevant and recognizes the changing nature of research support
- Fiches will provide opportunities for libraries and repositories to support open science in a structured way
- It is essential that continuing professional development is supported through engagement with the broader stakeholder community in order to share and develop best practice.

Strengths:

LIBER, COAR and OpenAIRE strongly support the development of an open, interoperable e-infrastructure for scientific data through the engagement of the relevant actors, including libraries and repositories supporting researchers in their scientific endeavor. This framework for action is highly relevant, both for the success of data driven science and also the uptake and impact of open access policies and initiatives. It recognises the changing nature of research and the importance of engaging the whole stakeholder ecosystem in the development and integration of research data e-infrastructures.

LIBER represents an extensive and dynamic network of European research libraries. In our strategy we have identified scholarly communication and research infrastructures as key areas for development and action over the coming years. We wish to strengthen the role of libraries in delivering science to society, and to open up science in a structured and linked way. The seven action areas identified in the fiche will provide libraries with the opportunity to do this. LIBER particularly welcomes the recognition that the research library community had been given in relation to their role in the development of skills and training. Our community has mobilised around this issue in recent years and support at European level to strengthen this activity is urgently needed.

COAR's vision, based on worldwide networked open access digital repositories as key element of the evolving scientific information infrastructure, recognizes the need for the global interoperable data e-infrastructure supporting open science outlined by the Commission. The strategic priorities identified by COAR to implement its vision for the next years are also aligned with the areas for action in the fiches: develop a sustainable open access repositories infrastructure as a key component of the global knowledge infrastructure, support the training of repository managers to adapt their skills and capacities to the new scholarly communication context, promote and define interoperability guidelines and policies and to stimulate the development user added-value services on top of repositories. The EC framework for action, therefore, will be a strong support for the open access repositories community worldwide to

achieve its goals, and will contribute to Europe's leadership at global level.

OpenAIRE, as part of its overarching goal to gather and link together European research output and enable open science through interoperable e-infrastructures for publications and data, considers it vital to equip the next generation of data professionals, users and producers with the necessary infrastructure and skills to deal with data. Much data will originate within the research institution, therefore the library and repositories are well placed to advocate for and support research data sharing. However, for this to happen, the network of research institutions within OpenAIRE need to have the proper foundations in place to allow the sharing, depositing and linking of data.

Identified Gaps

- The group would welcome a more horizontal approach within each of the fiches. Involvement of all relevant stakeholders (for example, researchers, research institutions, universities, libraries, repositories, data centers, publishers etc.) is essential for the supporting of discoverability, navigation and (re)use of research data and literature.
- More attention to research institutions and the long-tail of research is needed. The fiches should give more importance to universities to establish skills and services. In addition the development of a clear workflow between research institutions and data centres should be supported.
- Specific actions addressing research areas which are less far developed regarding digital techniques and data (humanities).
- Storing, managing and preservation should also be a key concern for open access infrastructures, this shouldn't be mutually exclusive. Any work on authentication and authorization will need to take into account developments in digital identifiers.
- Facilitation of online and blended learning (incl. MOOCs) which address both students and practitioners to educate in managing and sharing open data, should be included in the fiches.

Detailed comments on each of the fiches:

Fiche 1: Community support data services

- This fiche has overlooked the role of research institution and IRs. The role of linked data for IRs is such that they should be recognised as a stakeholder amongst semantic data providers.
- The development of standards, ontologies, and metadata for research data will help support cross-disciplinary data sharing and collaboration. It also opens up the opportunity to develop tools and services on top of data centres and repositories.
- This fiche should also focus on supporting and stimulating such service (e.g. usage statistics, text/data mining, research profiles) by drawing use cases and requirements from the communities in question.
- The development of data support data services should consider the relationship to other developing local systems (how this can be done is a real challenge), user-oriented services grown in intensive discussions with the scientists, embedding the repository into services, and differentiating between backend and frontend.

Fiche 2: Infrastructure for open access

- Ensuring that researchers can, on a global scale, have meaningful access to, and exploit, open access content (both via thematic and institutional repositories and data centres) is an essential step in increasing visibility and impact of open access publications. To do this, best practice need to be developed for depositing research data sets that enable linking to relevant documents, and that have high compliance levels driven by appropriate incentives, resources and policies.
- Our vision encompasses supporting the creation of a new, enriched form of scholarly publication that enables the creation and management of relationships between knowledge, claims and data. It is founded on the creation of a knowledge infrastructure that allows the sharing of computationally executable components, such as workflows, computer code and statistical calculations, as scientifically valid content components; and an infrastructure that allows these components to be made openly accessible, reviewed, discussed referenced and attributed.
- More emphasis on incentivisation for data sharing and citation is needed. Support and incentives for the sharing of data and publications should be more effectively developed and implemented in institutions, compatible with national, European and international requirements. We must ensure that research engaging in data sharing can be given recognition for their efforts.

Fiche 3: Storing, managing and preserving research data

- We recognise the need for a coordinated approach to the preservation of research data. From the library and repositories perspective, there is much work to be done in relation to ensuring the sustainability of digital preservation. The development of shared services for the preservation of research data is one solution to helping address this issue. In addition, preservation is not mutually exclusive action, and all activities should have mention of preservation planning.

Fiche 4: Discovery and provenance of research data

- Developing a Digital Identifier e-infrastructure (for digital objects, datasets, etc and authors) is a key issue to develop any interoperable infrastructure. There is a “desperate need” for libraries/repositories to use digital identifiers to be able to create high impact visible services or enable the creation of services on top of repositories.

Fiche 5: Towards global data e-infrastructures

- Increasingly research and data driven science is conducted on a global scale. The growth of open access means that the potential for global collaboration will increase exponentially. The supporting infrastructures for research must also be capable of operating on a global scale in order to realise this potential. We welcome the RDA initiative and the opportunity for libraries and open access repositories, as stakeholders in the global data e-infrastructure, to contribute to and engage in such a forum.

Fiche 6: Authentication & Authorisation

- We support this action, but we would welcome a better balance between the stakeholders being addressed: researchers, research support, libraries and repositories

Fiche 7: Skills and new professions for research data

- The acquisition of skill sets is essential for researchers, research support services, librarians and repository managers. This is particularly important at the institutional level, which needs to link up to the international research community, in relation to standards and good practices, for example via research reports and monitoring. Support and incentives for sharing of data and publications compatible with national, European and international requirements can then be more effectively developed and implemented in institutions,
- While repository work may be philosophically aligned with the work librarians have been carrying out for centuries, the details of repository management – and in particular, the set of services offered around digital content management – are very different and require capacity development and skills training.
- The ODE project¹ has shown that there is clearly demand for libraries to provide support for data management and sharing. There are gaps, which need to be filled, in support for data curation, data management planning, citation etc. The identification of required skills, and development and deployment of curricula and course for data librarians is a crucial step in addressing these gaps.

22 March 22, 2013

¹<http://www.ode-project.eu/>