RDA EU call for action: European Data Economy must focus on interoperability

Brussels, 30 January 2018

Summary of recommendations and Calls for Action

Data localisation restrictions are a too narrow focus for building a Data Economy - instead, interoperability on all levels should be systematically promoted:

Legal interoperability: Removing legislative barriers for data movement and re-use, critical assessment of existing and planned regulation and abolishing it where needed. For example, the scope of the text and data mining (TDM) exception in the proposal for Digital Single Market (DSM) directive on copyright[1] must be widened also to cover commercial use.

Organisational interoperability: Solutions must be developed in collaboration with all stakeholders. There are already good examples of the implementation of federated data infrastructures in Europe[i].

Semantic interoperability: Data must be made understandable for the end-user no matter who and where it is being re-used. Existing semantic tools such as vocabularies, ontologies and enterprise architecture models should be leveraged. RDA Recommendations and Outputs must be adopted and the uptake should be systematically integrated into developing data infrastructures.

Technical interoperability: Making the infrastructures compatible is an elementary building block for sustainable data flows, exchange and re-use between different IT systems and software applications. This must be systematically promoted in parallel and consistently with all other layers of interoperability.

All parallel initiatives under the Digital Single Market umbrella must be developed coherently. Otherwise new barriers for data re-use will be created, which will seriously damage the efforts to build a European Data Economy. Europe needs a coherent legislative framework that supports data re-use.

---

[1] http://ec.europa.eu/transparency/regdoc/rep/1/2016/EN/1-2016-593-EN-F1-1.PDF

research data
sharing without
barriers

Existing work and achievements should be leveraged. The RDA Recommendations and Outputs provide concrete tools to build the social and technical infrastructures that enable open data sharing. Already four RDA Flagship Outputs have been recognised as ICT Technical Specifications by the European Commission, with four more to be officially published in the upcoming months.

The RDA is a neutral forum that enables data interoperability and open data to be shared across barriers through focused Working Groups and Interest Groups, formed of experts from around the world –from academia, industry and government. It is a problem-solving platform for the data science and data management communities, and it should be used to its full potential to promote, discuss and reach consensus on solutions and practical implementation. Furthermore, data management should me recognized as an issue best promoted and regulated through the community of data producers, stewards and end-users.

Data-driven economy

What are the building blocks of a Data Economy? The Research Data Alliance (RDA) was founded in response to a real and urgent need for all research communities to address new ways of conducting research – the data revolution, as some have called it. Suddenly we are in a situation where we have enormous amounts of data, and we have to find ways to work with it, and also to leverage its huge potential. Data is indeed a valuable asset that brings growth, jobs and competitiveness.

Data is ultimately at core of the European Digital Single Market. It is an over-arching, cross-cutting, fundamental element. Be it research data, industrial data, commercial or public data - the possibilities that it brings us are too large to even predict.

What measures are needed to support the European data business? Due to its nature RDA already has a global and cross-disciplinary outreach – it is a grass-roots organization building on the work of volunteers committed to tackle emerging data sharing challenges – it is the perfect place to gather input, increase awareness around key topics and address issues that are critical for both public as well as private sector, acaremia and industry – a single common ground increasingly critical for the European welfare. Europe must act now, and leverage platforms, social structures and infrastructures already in place to avoid fragmentation and adopt a convergent, time and cost efficient approach.

Coherent and supportive legislative basis

The free flow of data is crucial for European competitiveness. Europe needs to urgently create a supportive environment as well as a stable and coherent legislative basis for capitalising on data to promote innovation and growth in Europe. New trends and developments, such as artificial intelligence

and platform economy are dependent on data.

However, regulation is not the solution. Europe should systematically avoid building new barriers for data movement. Thus, the need for new regulation should be critically assessed. Instead, active attempts to reduce and harmonise legislation should be made across the Member States.

Broadening the scope of Data Economy

Removing data location restrictions is one of the measures that has been identified as something that could boost the European Data Economy. However, focusing on data location is not sufficient for creating efficient preconditions for new data-driven business. More focus should be put on ensuring that the data can actually be used. A recent study[ii], commissioned by the EC, shows that data localisation is not one of the key barriers for data sharing and re-use between companies. According to this study, the major obstacles for data sharing are technical, legal, monitoring and skills related. And for data re-use the barriers relate to terms of access, interoperability and cost.

These findings also indicate that investment in data sharing and re-use between companies is regarded as a remarkable source of new business by the companies themselves. According to the study, B2B data sharing as a primary economic activity is expected to grow, and the number of data suppliers and data users may increase in the next five years.

Leveraging existing tools and initiatives

A lot of good and important work has already been done by the EC and member states, and this should be leveraged. Europe should focus on enabling re-use of data between different sectors (commercial, academic and public) and interoperability on all levels: legal, organisational, semantic and technical, identified e.g. in the European Interoperability Framework[iii] developed by the EC. Europe should take advantage of and promote the already reached achievements, such as the RDA Recommendations that have been approved by the EC as ICT Technical Specifications. The European Open Science Cloud and the European Data Infrastructure are major initiatives, with a vision of a research data commons, widely inclusive of all disciplines and Member States. These initiatives will generate as well as require sustainable solutions from the public and private sectors.

It should be ensured, that the development of the Data Economy is coherent and consistent with all ongoing and parallel efforts. Specific focus should be put on semantic interoperability: the ability to exchange and process data in a way that preserves the usability, intelligibility and meaning of data. Putting the user in the driving seat is essential in building a sustainable data management framework, because the end-users will be doing most of the concrete work for consolidating the data economy.

We should not forget, that in developing the data economy, there should be a strong link to the EU

Skills Agenda[iv] in terms of analysing data skills requirements, addressing the gaps and generating solutions. Skills and know-how for data management are urgently needed in Europe. EU should systematically invest in developing relevant skills on all education levels and sectors, including the education of teachers and trainers.

It is absolutely critical that all initiatives under the Digital Single Market strategy are consistent and coherent, in order to avoid a siloed approach where new barriers will certainly be built. One dangerous example is the proposed DSM directive regarding copyright, in which the text and data mining (TDM) exception is narrowed down only to research organisations for the purposes of scientific research. TDM provides considerable benefits for the European economy and research, for example it is a key feature in developing artificial intelligence. European startups need to be supported by a TDM exception that allows for commercial use. If the use of TDM is not fostered and supported in Europe for all sectors, there is a huge risk of not making the most of the brightest minds and of the economic benefits it could bring.

The European Open Science Cloud (EOSC) will place the seamless flow of data, scientific results and knowledge at the heart of the research and innovation process. To make the EOSC a success, Europe should focus on making data FAIR (findable, accessible, interoperable and reusable) within and between scientific disciplines, also promoting data usage between research and industry. EU is investing heavily in building the infrastructure for the EOSC, thus it is of utmost importance to make sure that no legislative barriers exist.

Collaborative actions needed

The RDA is doing its part in creating a sustainable Data Economy, by building the technical and social bridges for data sharing, across borders, across disciplines, and across sectors. It is of utmost importance that we create a way of sharing and re-using data between the commercial, public and research sectors. And this is not only about removing barriers such as data location restrictions. This is about how to ensure data is available for those who need it, and that it is in an understandable format, no matter who is using it and where and in what language they speak. Data should be findable, accessible, interoperable and re-usable for creating innovation, curing rare diseases, finding solutions to grand challenges that require multi-disciplinary approaches etc. But how to actually make it work on a concrete level? This is where we come to the RDA Recommendations and Outputs, the concrete hands-on tools, practices and standards for data movement, that are the basis for putting the plans for creating a data economy into practice.

Just like self-regulation has been the European way in research integrity issues, it should also be our approach of choice when building a Data Economy and ensuring free flow of data. The EC has embraced the learned societies in the form of All European Academies (ALLEA) as a community forum for negotiating and defining what constitutes good research. The European Commission would similarly

support the development of the European Data Economy by further empowering RDA as a key platform for the research data science community and recognising data management as an issue best promoted and regulated through the community of data producers, stewards and end-users. Building trust among a community of responsible and engaged experts is the only mechanism that can offer agility, efficient implementation and legitimisation of measures all at once, in the face of the rapidly transforming ecosystem of digital data.

RDA Europe is calling for a joint effort to create the European Data Economy. All stakeholders need to play the game together, in a coherent way, having a clear and common goal and understanding on how to achieve it.

---

[i] http://www.knowledge-exchange.info/event/federated-rd-infrastructure

[ii] https://ec.europa.eu/digital-single-market/en/news/data-sharing-between-companies-europe-results-study-european-commission-and-way-forward

[iii] https://ec.europa.eu/isa2/eif_en

[iv] http://www.cedefop.europa.eu/en/news-and-press/news/skills-agenda-10-actions-help-equip-people-europe-better-skills