

panel on open access

Carlos Morais Pires
GÉANT and e-Infrastructures
European Commission
Information Society and Media



R&I are priorities of DAE



"The Digital Agenda for Europe outlines policies and actions to maximise the benefit of the digital revolution for all. Supporting **research and innovation** is a key priority of the Agenda, essential if we want to establish a flourishing digital economy."

Neelie Kroes,

Vice-President of the EC, responsible for the Digital Agenda

Data as Infrastructure



http://bit.ly/riding_the_wave

- “Our Vision is a **scientific e-infrastructure** that supports seamless access, use, re-use, and trust of data. In a sense [...] the **data themselves become the infrastructure** – a valuable asset, on which science, technology, the economy and society can advance”.

Report of the High-Level Group on Scientific Data, Oct 2010

“Riding the wave: how can Europe gain from the rising tide of scientific data”

my contribution (1)

- There are generic aspects to Open Access and things that are specific to communities, disciplines, etc...
- The approach of the EC is very ,uch aware of this
- IT/Computer Science are in a privileged position to adopt "new narratives" mechanisms to collaborate, disseminate knowledge, publish papers and data

- Intellectual property is a key issue but we believe that the discussions are getting mature to turn a page in this domain (recent study by KE on approaches in different EU Member States)
- OA is not incompatible with IP
- Most of scientific work is based on prior work and this should come naturally in the efforts to make access wider and affordable
- OA can help lowering unnecessary barriers to scientific discovery

my contribution (2)

- I see the challenges as not being only related with technology...

- My experience with the EU OA pilot is that we need really to work in parallel the policy and the technology

- I see the links between human and machine readable information a key challenge

- Tools for curation thinking about usefulness to others not necessarily specialists in a field...

- An architecture that is open, distributed and open to participation...

- EU is revising the policies to apply in HORIZON 2020

- We are getting a lot of experience with the running pilot

- Be more explicit about the rules

- Make provision for resources

- Data Management Plans

my contribution (3)

- Global standards are not just for interoperability but also for other things such as evaluation of data quality, data classification/ontology and mass storage
- We are working very closely with US
- We are working with G8+8 on these subjects

- There are social and behavioural aspects to be addressed like the clash of cultures between different disciplines, legacy frameworks and the need to rethink organisational models.
- There is a need to train a new generation of data scientists requiring new knowledge about how researchers use and re-use information in different disciplines and countries.
- Data management and governance considerations should be included in the secondary and higher education curricula.

data infrastructures

- Need to discuss further data as “products” and their relation with the demand for value added services

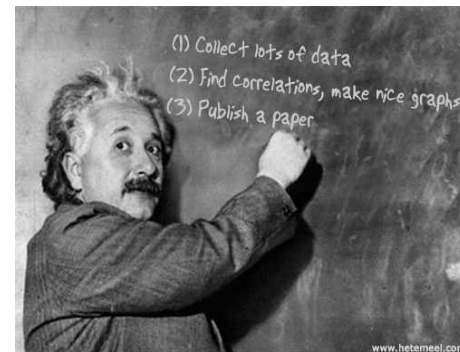
- Global standards are not just for interoperability but also for other things such as evaluation of data quality, data classification/ontology and mass storage

- Not exclusively looking at data produced in the future but also to make existing data available

- Build on existing resources (not reinvent wheels) for cost effectiveness

- There is not enough efforts into curation and that is a huge problem by itself to support data intensive science

Accessing Scientific Data

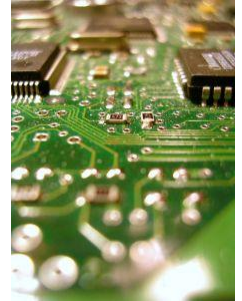


- We don't know how **scholar communication** will adapt to new paradigms **bringing closer human and machine readable information...**
- **Opportunities for innovation** in publishing
 - **Publication + data** = new concept of “enriched publications”
 - **Challenge of quality preservation**
- **What will, in HORIZO2020, follow the FP7 OA Pilot...**
 - OpenAIRE e-Infrastructure
 - Data Management Plans



Data and the discovery process

- underlying computing infrastructures



- Scientific software, models and algorithms embed valuable information and knowledge.
 - software and models for generating, processing and correlating data, software for reproducibility and accuracy of the data, software used for analysis and visualisation...)
- e-Science infrastructures must be able to manage the expected scale of the future data resources.
- Some disciplines in particular will “push the envelope” with respect to what is technologically possible.
- The provenance of scientific data should also be managed through the data infrastructure that should support traceability and reproducibility by recording the derivation history of data.

Thank you!

carlos.morais-pires [at] ec.europa.eu