

Eco System of Services (on Data)

Peter Wittenburg
Max-Planck-Institute for Psycholinguistics
Nijmegen, NL

Why this topic?

- this is what we often find in proposals, ideas etc.:
 - “we are supported by initiative X”
 - “would you please support our proposal A”
 - no ref to ongoing work - ignorance or reinventing the wheel?
 - etc.
- is this still acceptable?
 - since 2006/8/12 ESFRI RIs
 - since many years e-Infrastructures
 - have large institutions at EU level
 - there must be increasingly more services at various levels (otherwise we are doing things completely wrong)
- we need statements such as
 - we will use service A or tool B from initiative X (-> SLA)
 - neither X nor Y offer service A, therefore we build A

eco system
of services

Are there examples?

- oh yes – partly from my domain view
 - we need to set up our own repository infrastructure for humanities resources
 - don't we have yet offers from CLARIN and DARIAH?
 - we need to set up our own AAI solutions
 - don't we have yet a smoothly functioning system in EU?
 - we need to set up our own language processing infrastructure
 - don't we yet have yet some frameworks within CLARIN?
 - we need to set up our own machine for Big Data
 - don't we have yet an offer by PRACE and EGI?
 - etc.
- it may be that the existing services are not appropriate but then come up with a convincing argument
- **AND: system only works if the offers are explicit**

What is the problem?

- of course we need
 - visibility of services – how to find (also anticipating cross-walks between disciplines)
 - explicit statements about terms of usage (who, when, restrictions, costs, etc.)
 - explicit statement of continuation of service (if only 3 years then say so)
- otherwise we will not be able to save money
- excellent recent example: ZENODO by OpenAIRE + CERN (it is European! – of course some remaining questions)
- EUDAT also ready to launch services (but 3 years issue)

What did we (I) try?

- wanted to use this e-IRG meeting to send out a message and see what kind of reactions we will get
- somehow we (I) completely failed – is it bad? (not at all)
- many reasons
 - who are you – why should we care – it means some work so who has sufficient authority?
 - we have so much to offer – just look at our web-site(s) well – what an argument – and now?
 - what should we describe – which services do you mean that's a good point – but need to start and learn
 - what would be the consequence – can't handle requests exactly – that is the point
 - don't know what the life-time will be exactly – need solutions – otherwise many **Potyomkin villages**

Is it a trivial problem?

- no it is not a trivial problem
 - different types of services
(storing/archiving, computational services, HPC service, web-services vs. web-applications, etc.)
 - large variety across disciplines
 - how to describe them properly + simple enough
- we are not the first trying that
 - think of UDDI etc.
 - CLARIN has a specific MD description for Web-services
- in NL an agreement to start creating a light-weight wiki type with some rough categorization
- obviously we need ONE portal maintained by neutral and trusted people (brokers)
 - who can do it in Europe?

How to go on?

- should we start it or better not?
- who has the authority - ESFRI, EC, e-IRG, etc.?
- why not form a WG (use RDA framework) to specify a registry?
- who takes care of the continuous effort?
(EUDAT + EGI?)

- obviously must start simple first to have low threshold

- what is your opinion?
- will come back to this at the end of the meeting

What today and tomorrow?

- decided to select a few speakers to show the spectrum of services
- today
 - Nuria Bel (Barcelona) – Web-Services Infrastructure
 - Bernard Bel (Aix en Provence) - Archiving services
- tomorrow
 - Jon Ison (EBI) – Broad Service offer
 - Giuseppe Fiamini (Cineca) – data services
 - Daan Broeder (Nijmegen) – language services