



Partnering with EUDAT – the CLARIN perspective

Dieter Van Uytvanck

CLARIN ERIC

dieter@clarin.eu

3rd EUDAT conference Amsterdam

2014-09-25

CLARIN?

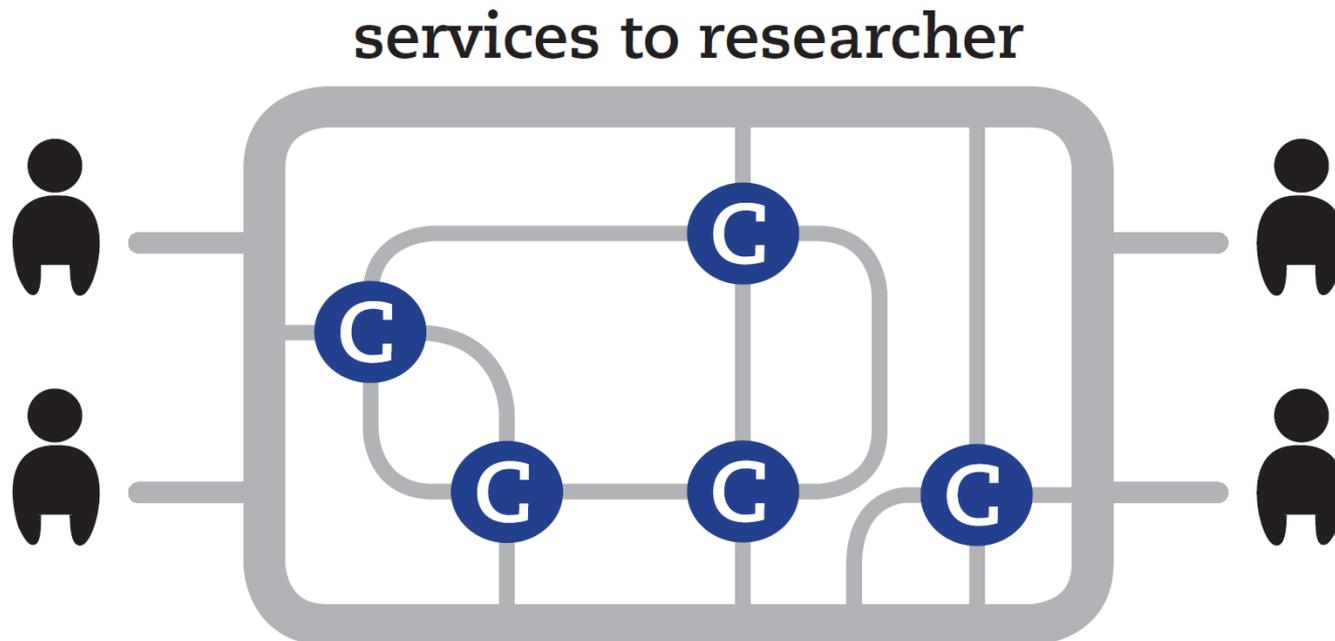


- **Common Language Resources and Technology Infrastructure**
- aims at providing easy and sustainable access for scholars in the humanities and social sciences
 - to **digital language data** (in written, spoken, video or multimodal form)
 - to **advanced tools** to discover, explore, exploit, annotate, analyse or combine them
- More details: www.clarin.eu

CLARIN centres



- A distributed architecture: (http-accessible) files, web applications and web services spread all over Europe
- Some of them password-protected (licenses, privacy, ...)
- User base: also spread over Europe



Organisation CLARIN



- **Members:** countries (AT, DK, CZ, DE, EE, NL, NO, PL) or international organisations (Dutch Language Union)
- **Resources & Services provided** (<http://clarin.eu/services>):
 - Access to language resources (including federated login when needed)
 - Access to language resource processing applications/services
 - Depositing services
 - Metadata catalogue: Virtual Language Observatory
 - “Glue components” like the Virtual Collection Registry (<http://clarin.eu/vcr>)
 - Consulting services
 - (+ a whole set of technical services behind the scenes)

Answers CLARIN



- What are the **agreements** related to access to the RI or the services it provides (e.g. machine time, software, data, support, etc.)?
 - B-centres (= CLARIN centres) providing services to end-users > national consortium agreement
 - A-centres (= CLARIN centres) providing several services to all CLARIN centres > national consortium agreement
 - E-centres (= computing centres) providing services like VMs, monitoring, EPIC/PID services > via an agreement

Answers CLARIN



- What are the current or **planned legal structures and funding models** in place, etc?
 - CLARIN ERIC is a legal entity (funding = contributions from national consortia, H2020 participation possible)
 - Thin top layer, coordinating the national consortia (where the real funding comes from)

Answers CLARIN



- According to the EUDAT CDI concept, what would be the **requirements** from your RI to formally **join and integrate** components of this RI as part of the CDI?
 - Added value (eg by connecting more infrastructure components)
 - Future perspective and stability
 - Technical compatibility

Answers CLARIN



- EUDAT aims to provide storage resources and other related services to the widest numbers of researchers. These resources have a cost and their access should be regulated. In your opinion, what would be the **best model** for accessing these resources:
 - Quality-based: researchers apply for resources which are allocated on the basis of scientific excellence, originality, quality, and feasibility of the applications
 - Quota-based: access is based on quotas determined by e.g. the financial contributions from the CDI partners, or the research programmes agreed with pre-defined users)
 - Market-based: access is granted to anyone against a fee
- Not sure, but market-based at least allows for some planning; difficult to integrate something into CLARIN that each time depends on a scientific assessment committee

Answers CLARIN



- Suppose you are looking for a place to store and take care part of your scientific data for at least 10 years. What conditions/requirements should EUDAT meet to be seen as the **best place to put your data**?
 - Organisational stability
 - Good technical offer (compatible, state of the art)
 - Added value by integration with other services
 - Competitive (price-wise)

Answers CLARIN



- EUDAT is currently a network of independent centers working within a common framework to develop and propose services. At present, contractual agreements (e.g. SLAs) can only be backed by centers as individual legal entities. In your opinion, should EUDAT move towards a **single legal entity**?
 - Difficult to judge this right away – would require an in-depth analysis of pros and cons
 - Experiences with contracting single computing centres are OK, but there are always complexities (e.g. VAT, possibility to have an agreement between ERIC and computing centre)

CLARIN

Common Language Resources and Technology Infrastructure



Thank you for your attention!
