GEDE Repository Topic Group: Notes VM 12.2.2019

Editor: Peter Wittenburg

Participating: Maggie Hellström, Ingemar Häggström, Venkata Satagopam, Christophe ?, Damien Boulanger, Peter Wittenburg

A short overview about GEDE was given for the newcomers. Background Info on GEDE can be found here: <https://rd-alliance.org/groups/gede-group-european-data-experts-rda>. The first paper on PIDs from GEDE can be seen here: <https://zenodo.org/record/1116189> .

## Relevance of Repositories for European and National Data infrastructures

Here we just summarise a few statements made:

* repositories are the care takers of scientific data and the ways how this "care taking" is organised is very different in all aspects (storage, data organisation, admin organisation, distribution of functions, etc.)
* an overall strategy for repositories as care takers of scientific data is hardly existing, but is urgently required
* the landscape of repositories is still changing, one cannot speak about final organisational forms although solutions for sustainability get increasingly important
  + in MPG a 50 year institutional guarantee is given for specially marked data
  + the Lund ICOS repository is funded based on limited contracts, copies are stored in other repositories such as EUDAT etc., Swedish authorities invest in changing the role of SNIC and SND to take a broader responsibility for data
  + EISCAT runs an own repository based on limited statements of the involved Research Councils
  + the Luxemburg ELIXIR node is secured by government funding for 10 years, and is part of the wider ELIXIR network and the INTRADEV cluster
  + in France there seems now to be a start on focussing on data infrastructure aspects
* there are many questions about EOSC despite a first good interchange with the cluster projects recently in London
  + it was absolutely unclear how the existing repositories should be integrated and which the benefits should be
  + due to the London meeting[[1]](#footnote-1) it is the first time that the ESFRI repositories are seen as relevant pillars in the EOSC eco-system
  + our paper should help to point to the missing integration of the ESFRIs

## Existing Documents

We went quickly through the two uploaded documents about repositories which are all in the GEDE share and referred to the report of the FAIR Implementation expert group which is also in the GEDE share. A few gaps were identified:

* the sustainability aspect needs to be emphasised
* the special aspects of sensitive data are missing (Venkata will contribute)
* the issue of reduction/deletion should be addressed
* the difference between repository and archives should be explained
* the role of managers/curators/managers needs to be stressed
* the need to support catalogues is relevant to find data

## Paper Skeleton

The skeleton for a possible final paper was discussed and a few points already added.

## Actions

* Maggie will put the first draft skeleton into a Google Doc
* Peter will combine the two basic papers, look for additional aspects in the FAIR group report and take care that missing things will be added, will be done asap but at least until the end of February
* Venkata will add a paragraph on sensitive data and GDPR
* This merged document will also be uploaded as Google doc
* Until 30. March there will be time for everyone to comment on the Google docs
* we should check the status of RDA groups on repositories
* for the final paper we should get some reviews from experts from a few repositorries.

1. Infos about the ESFRI+EOSC workshop on January 30: <https://www.esfri.eu/esfri-events/esfri-ris-eosc-liaison-workshop?qt-event=1#qt-event> [↑](#footnote-ref-1)