Note on GEDE Directions

Peter

This is an internal note for the co-chairs discussions. First I describe what has been said during the meeting and what we know from elsewhere. Then we should come to an evaluation of the suggestions.

# 1. What has been said

## 1.1 Changed Situation

* Data driven science in Europe is moving fast mainly driven by domain-oriented approaches with global inclusiveness (global labs). This includes methodological aspects (AI, ML, etc.) as well as infrastructural solutions (repositories, interoperability, etc.). The complexity introduced by the size of these domains has already achieved a scale which is hardly to manage. COVID is accelerating towards machine actionable solutions and there is a general trend to federative solutions.
* Infrastructures’ role is to let us do better science and to offer solutions increasing efficiency and reliability. This happens in a time where competition is increasing.
* Much time (70-80%) is still spent on data wrangling. Urgent need to do training and to educate data stewards who will do the transition towards FAIR. Automation will be done were possible.
* The cultural change towards sharing is in progress and can’t be stopped, but is slow and hampered still by many blockages. There is an increasing believe that deep changes will take time and require patience. Trust, an improved “trust infrastructure” technologies and incentives will be required to improve sharing. Technology will not be able to cover all trust aspects. Science is changing in its core and some are afraid of these changes.
* RDA EU has founded now a LE and can act to get and administer funds and organise the European community. Reaching out to not so well represented communities and countries is one goal. Another one is to help organising discussion processes such as for example done by GEDE.
* RDA, CODATA, GOFAIR, WDS joined forces to “Data Together” (TG) and could play a strong role internationally in achieving convergence.
* EOSC Governance has been set up including first boards and working groups. WG discussions were too often dominated by funding interests and not so much by independent experts. Soon an EOSC LE will be setup to replace the existing governance, yet it is too early to anticipate future structures and the partnership program with the EC will also influence how EOSC will evolve.
* EOSC does not exist, all claims to be “EOSC” compliant are intentionally misleading. “EOSC portal” is a disaster.
* Should separate between infrastructures for basic services (CPU, storage, bandwidth) from the EOSC type infrastructure addressing interoperability and reuse. EOSC infrastructure can be seen as a thin federation layer bringing together existing infrastructures based on FAIR princinples.
* Clusters are limited time projects, the ESFRIs/ERICs, large national projects and research organisations are meant to be persistent and thus the carriers of the “real” work. Many do not feel represented by the clusters.
* The differences between short-term solutions and long-term thinking have been clearly understood now due to recent discussions about funding options.

## 1.2 Character of GEDE

* GEDE stands for agreement finding by independent experts on deepest level. It acts neutrally, since no direct funding interests are involved. It is an open space for discussions.
* Due to its mission it focussed on technical topics, but also run surveys always addressing the “real world of practitioners” in large RI that are working on solutions.
* The GEDE experts are anchored in their RIs and thus act in both directions.
* GEDE is a European initiative, but started integrating international experts especially in its DO/FDO activities that have a global impact.
* GEDE established a domain of trust over the years and as voluntary organisation has shown to be able to act fast.
* GEDE created until now documents worked out by the “real world of practitioners” on relevant topics mainly for the practitioners, i.e. it did not make so much PR.
* GEDE work until now is voluntary and thus has its limitations.

## 1.3 Suggestions for GEDE

Many suggestions have been made about gaps in the decision structure and how GEDE could change. Here we focus on GEDE although there is a narrow relation with RDA and RDA EU[[1]](#footnote-1).

* GEDE should formally be operating under RDA EU and RDA EU should become member of EOSC. This would ensure that GEDE can contribute to EOSC as well.
* Whatever changes GEDE should maintain its major characteristics and should build on what it has already demonstrated.
* Changes in the tasks of GEDE would require an optimised organisation.

**Scope**

* GEDE should not only address technological issues but also cultural ones. GEDE needs to come up with suggestions for policy makers as well and not just remain in “internal, academic” discussions
* GEDE should address the leadership, silverbacks and other important stakeholders not convinced yet. We should address the hierarchies in the RO to change the evaluation criteria to push the cultural change
* GEDE should help RDA to reach out to less-involved disciplines and countries.
* As one concrete action GEDE should produce a paper on the 10 most urgent things to be done now. This should be discussed and endorsed by DT to have global impact.
* GEDE should continue to work on technological topics of interest but consider its limited resources. FDO work needs to be continued, but other topics are important.

**Sounding Board**

* There are currently too many voices from data science. We lack a respected authority that comes up with evaluations that includes the “real world of practitioners” acting in an independent fashion.
* GEDE could become an authority for consultancy and give input to EOSC based on its neutrality and independence. This can be described as an “auditing” or “recognised sounding board” function.
* GEDE therefore should organise an “infrastructure council” with clearly nominated experts.

# 2. What could be done

1. More of GEDE outputs should be endorsed by RDA. [↑](#footnote-ref-1)