* **Task 7.2 Metrics and cost for the Photon and Neutron community EOSC (M9-M36) (Lead: CERIC-ERIC; Participants: ESRF, ILL, XFEL.EU, ESS, ELI)**

From the GA

* Objectives:  Study - even by using advanced methodologies - of the cost per partner for maintaining the infrastructure required for providing FAIR data (archiving, data services etc.) and explore different scenarios for financing the long term costs.
* **Task 7.2 Metrics and cost for the Photon and Neutron community EOSC (M9-M36) (Lead: CERIC-ERIC; Participants: ESRF, ILL, XFEL.EU, ESS, ELI)**
* Analysis and development of metrics for the evaluation of costs and added value of the services provided to the community. This clearly depends and connects to the developed data policies and on the overall architectural choices for the Photon and Neutron community EOSC.

**Scoping the evaluation of costs**

1. It is proposed that WP7 collects the full cost of setting up and maintaining the infrastructure required for providing FAIR data
2. It is proposed that we collect the individual costs for every single partner, and present the final results as an interval.
3. For some of the partners, the costs for setting up the infrastructure will refer to past actions. It may not be possible to retrieve the exact amounts (or to update them). We must find some compromise for dealing with expenses made in different years.
4. For some other partners (ELI, CERIC), the costs for setting up the infrastructure will refer to future actions, so they will be an estimation based on quotations and a priori assignment of resources.
5. All facilities have a different level of FAIR data maturity, therefore past and future efforts depend on this.
6. Do we need to distinguish between what is “ordinary” data management by the facilities and what id an additional cost due to the EOSC?
7. It is proposed that we follow the RDA FAIR Data Maturity Model to determine the services we must take into account.
8. .. other ideas?
9. It is proposed that WP7 collects the full cost of setting up and maintaining the infrastructure required for providing FAIR data

Reasons for yes: this is an extremely useful exercise for all the community. It will evidence what some facilities have already invested and (if auditable) this may also be used as a part of the cost to be reimbursed by H2020/HE. It will be useful for facilities that are not compliant with the FAIR principles to some extent, to estimate the costs to become compliant. Finally, it will be useful to estimate/define the cost of the maintenance.

Reasons for no: Complex exercise

1. It is proposed that we collect the individual costs for every single partner, and present the final results as an interval.

Probably, costs will vary from one partner to the other (e.g salaries). On this basis, costs may not be comparable. However, there may be value in showing that the cost of services may change from country to country, and to identify the main sources of these differences. The alternative is to define together a list of services and provide a single cost, based on quotations, for a given context/geographical area.

1. For some of the partners, the costs for setting up the infrastructure will refer to past actions. It may not be possible to retrieve the exact amounts (or to update them). We must find some compromise for dealing with expenses made in different years.

This is related to the methodology for collecting/calculating costs. In this regard we have different options:

* To use actual costs, no matter when they were incurred
* To use actual costs, updating them with a well established parameter (e.g. inflation)
* To use current costs, based on updated quotations for similar services.

1. For some other partners (ELI, CERIC), the costs for setting up the infrastructure will refer to future actions, so they will be an estimation based on quotations and a priori assignment of resources.

No other way to obtain these costs other than based on quotations and estimation of the necessary manpower, e.g. based on previous experiences by other partners, so this knowledge may turn useful.

1. All facilities have a different level of FAIR data maturity, therefore past and future efforts depend on this.

The goal of PaNOSC is to make FAIR data a reality in all partners, meaning we all need to develop or improve our infrastructure to some extent. The incurred costs will represent the starting of PaNOSC. We could assume that at the end of the project, we will all provide FAIR data, so all the resources received for the (technical?) WPs are a part of the cost. We may decide that training is a necessary service and include it. It is hard at this stage to assume that there will be costs sustained to build the infrastructure after PaNOSC.

1. Do we need to distinguish between what is “ordinary” data management by the facilities and what id an additional cost due to the EOSC?

An initial proposal was to separate costs that are strictly related to the facility from costs that are held only because of the EOSC. There is a grey area here so this division will be based on a convention on how we decide to attribute the costs. We can continue and decide to drop this idea, if too difficult to agree on the criteria. But we need to define the criteria for what is ordinary and what is additional.

1. It is proposed that we follow the RDA FAIR Data Maturity Model to determine the services we must take into account.

“Findability, Accessibility, Interoperability and Reusability –the FAIR principles –intend to define a minimal set of related but independent and separable guiding principles and practices that enable both machines and humans to find, access, interoperate and re-use research data and metadata.”

e.g Data is identified by a persistent identifier: what does this require?