

Evaluation Summary Report

Proposal : 309493
Acronym : ILDAP
Program : FP7
Call : FP7-INFRASTRUCTURES-2012-1-INFESO
Funding scheme : Coordination and support action - Coordination - CSA-CA
Duration : 24 months
Activity : INFRA-2012-3.2 - International cooperation with the USA on common e-Infrastructure for scientific data

ILDAP
International Long-term Data and Analysis Preservation

Proposal submitted by :

N°	Proposer name	Country	Total cost (€)	%	Grant requested (€)	%
1	EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH	Switzerland	758,630	52.10	758,630	57.88
2	CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE	France	313,360	21.52	209,559	15.99
3	STIFTUNG DEUTSCHES ELEKTRONEN-SYNCHROTRON DESY	Germany	384,000	26.37	342,400	26.13
	Total		1,455,990	100%	1,310,589	100%

Abstract :

The preservation of scientific data for long-term use and re-analysis has been identified as a key requirement in the field of High Energy Physics and other disciplines such as Astronomy and Astrophysics, as well as Life and Earth Sciences. In collaboration with related projects in the US (in particular in close collaboration with the National Science Foundation and Department of Energy) the proposed project would take the work of the Data Preservation in HEP Study group that defines the physics motivation for long-term data preservation and many of the associated issues, and extend this to cover not only the existing use cases but also consider the needs of the LHC experiments at CERN. This work would ensure the persistent availability of existing data and enable it to be shared between organisations and across national boundaries. Now is the time to define standards for data and meta-data formats and address access and authorization issues for on-going experiments (e.g. those at the LHC) – issues that have historically been addressed only in the final years of a scientific collaboration if at all. In order to perform this work a coordination body would be established that would not only organize workshops devoted to this topic but also address key issues related to long-term data archives, such as infrastructure approaches, the financing models for maintaining these archives, the handling of intellectual property rights both during and after the lifetime of the corresponding scientific collaboration, as well as the required networking of experts both within the HEP domain but also with other disciplines and projects. The results of this work would be made available via Open Access mechanisms and would be actively disseminated at relevant technology-oriented events, such as the IEEE Massive Storage and Technology conference, as well as discipline-focussed meetings, such as the IEEE Nuclear Science Symposium and Medical Imaging Conference and other similar events.

Evaluation :

1. Scientific and/or technological excellence (relevant to the topics addressed by the call) (Threshold 3.0/5 ; Weight 1.00)	Mark:
---	--------------

<p>The proposal focuses on the preservation of data resulting from High Energy Physics (HEP) experiments for a period that goes beyond the actual lifespan of experiments and collaborations. The proposal is narrow, both in terms of scientific discipline (only HEP is covered) and in terms of problem domain (long term preservation), and as such is not well aligned with the call's objective to establish "a EU-US coordination platform aimed at full interoperability of scientific data infrastructure".</p> <p>The problem is convincingly described and the need to address it at an global scale is well-motivated. The combination of data from multiple experiments is a commendable motivation for the ability to reuse data. The proposed prototype validation framework is appealing.</p> <p>The proposal places the physical preservation of HEP data into the context of future re-use of the data. However, it does not provide a convincing description of how these challenges would be addressed during the course of the project.</p> <p>The proposal builds on the results from the DPHEP studies, but does not provide enough insight on what will be done to move beyond and further the DPHEP recommendations. In particular, given the focus of the proposal it is surprising not to find reference to on-going data provenance and workflow preservation initiatives beyond DPHEP. The coordination mechanisms and workplan proposed are sound, building upon existing cooperation between the participants, and establishing a coordination body to define standards for data and meta-data and organize workshops to address issues in the long-term preservation of data. The proposed workpackages logically fit the project's goals, but the descriptions of the deliverables are too generic.</p>	<u>2.50</u>
---	-------------

<p>2. Quality and efficiency of the implementation and the management (Threshold 3.0/5 ; Weight 1.00)</p>	Mark:
<p>The proposed management structure and procedures are appropriate, given the close collaboration already existing between the participants.</p> <p>The proposed boards include representatives only from the EU project partners; this raises questions as to whether the project can function effectively as a EU-US coordination platform.</p> <p>Furthermore, an independent voice in the governance structure – possibly as an additional member of the Project Management Board - could be of value in relating the work of the consortium to the broader data landscape.</p> <p>The proposal does not explicitly describe plans to engage with or assimilate Tier-1 centers or labs.</p> <p>Risk management is basic.</p> <p>The consortium partners and their US counterparts are globally recognized leaders in the HEP field and they have established a strong understanding of the digital preservation problem via their involvement in the DPHEP work.</p> <p>The consortium is small in number of participating organisations, but it would be difficult to envisage a consortium with any higher level of expertise in the area of HEP data preservation. All three are major players in the management and use of HEP data.</p> <p>The allocation of resources is reasonable, but the overall cost is high.</p> <p>The interest of the US counterparts in the proposed project is supported by accompanying letters. However, their effective commitment is weak.</p>	3.00

<p>3. Potential impact through the development, dissemination and use of project results (Threshold 3.0/5 ; Weight 1.00)</p>	Mark:
<p>The impact of this work would be to contribute to the advancement of common data preservation approaches in high energy physics worldwide. However, the proposal does not provide adequate evidence that the results will be usable outside the High-Energy Physics domain. Within this area, the potential impact of the results of this project is substantial in terms of both value to future scientific research and contribution to physics education. A strong case is made in the proposal on the timing of the value of data, i.e. that high value may arise after the experiments producing the data have been completed.</p> <p>The coordination mechanisms necessary to yield long-lasting collaboration are not well defined.</p> <p>The measures for spreading excellence are helpfully divided into four sections: internal to the project (through workshops), external (though the HEP community and published information), through educational outreach and through communication to policy-makers.</p>	4.00

<p>4. Remarks (Threshold 10.0/15)</p>	TOTAL:
	<u>9.50</u>

Does this proposal have ethical issues that need further attention? *(If yes, please complete an ethical issues report form (EIR))*

N

For each criterion under examination, score values indicate the following assessments. Half point scores may be given :

0- *The proposal fails to address the criterion under examination or cannot be judged due to missing or incomplete information*

1- Poor. *The criterion is addressed in an inadequate manner, or there are serious inherent weaknesses.*

2- Fair. *While the proposal broadly addresses the criterion, there are significant weaknesses.*

3- Good. *The proposal addresses the criterion well, although improvements would be necessary.*

4- Very Good. *The proposal addresses the criterion very well, although certain improvements are still possible.*

5- Excellent. *The proposal successfully addresses all relevant aspects of the criterion in question. Any shortcomings are minor.*