



Contribution ID: 114

Type: **not specified**

ScienceSoft –Open Software for Open Science

Tuesday, 27 March 2012 15:00 (30 minutes)

URL

<http://www.eu-emi.eu>

Impact

As the wide variety of EMI products are used in daily science that should not rely on unreliable EC project funding over the next decade, EMI needs to clearly define and implement a sustainability strategy. This strategy needs to ensure that products currently developed and supported within EMI are continually evolved, supported and maintained after EMI. It is of course clear that the key to sustainability is in the availability of resources to carry on the activities. Resources become available when there is a demand and a will to pay for that demand to be satisfied. The creation of an open software community satisfies the sustainability requirements of establishing a growing base of users and streamlining available funds and resources onto products for which a real demand exists. This process goes also through an investigation of which products have the best chances for further funding or for generating revenues through support and consultancy activities. The impact can be thus summarized in four major benefits to the community. Firstly, it provides a widespread database of information about the software products and their functionality and a channel to advertise the products, collect feedback, and propose additional services to a wider audience, feeding information into search engines and other taxonomy systems. Second, provide a single entry point to accessing source code and binary package from wherever they are stored, easing the tasks of creating distribution for specific communities such as WLCG and EGI (UMD) or PRACE. Third, offer a system to collect and formalize information about the software products, which will help in understanding how the products are used, by whom, and in which contexts. Finally, provide a social networking environment where people can advertise their skills or find the skills they need for their research or development projects. As such ScienceSoft brings benefits and impacts far beyond EMI and the greater scientific community.

Overview (For the conference guide)

The EMI project brings together middleware services from ARC, dCache, gLite and UNICORE in order to provide a harmonized set of interfaces and a streamlined release for the DCI community. Nevertheless, EMI is a project funded by the EC and ends in 2013 leaving it unclear how the harmonized developments will be continued to be maintained. The main objective behind ScienceSoft that will be introduced in this contribution is to create the conditions for the continuing development, support and use of the EMI software products after the end of the EMI project. But EMI will establish ScienceSoft as an incubator for a broad open source community of developers and users of not only the EMI products, but also other software from projects of critical interest for the scientific research communities. This contribution aims to provide an overview of these activities including the value of ScienceSoft to the community, its goals, and its mandate.

Description of the Work

In order to understand the value and work behind ScienceSoft, it is necessary to identify what problems it is trying to address that we outline in our contribution as well. Software developed today by research institutes, universities, research projects, etc. is often stored in local source and binary repositories and available for the duration of a project lifetime only. Subsequently, many cases have been found where important components are distributed from personal Web sites of developers or single institutions. Finding a piece of software based on given functional characteristics is almost impossible. Cases of conflicts are often found between different versions distributed by different people from different places. Source code is even more difficult to locate and access and contributing with patches and fixes, which is a very common activity in the open source world, is traditionally very difficult to do in the research communities. This has been for years a primary complaint from users, but the problem is not limited to EMI. Similar requirements have been expressed by application developers, infrastructure managers and users. Within the HPC community the organization of common repositories of application code and libraries is also a known concern. Another sensitive problem is the general lack of metadata, documentation, or information about the software. Information about who develops, contributes and uses a given program is very difficult to find out and yet the widespread availability of such information would give more visibility and credibility to the software products especially in science that lack marketing mechanisms compared to industry. The establishment of a software rating system will be explored. Other areas of work include software compatibility charts. To sum up, work includes technical implementations, but also sociological implementations via ScienceSoft that we present in this contribution.

Conclusions

ScienceSoft is in the process of being established together with many stakeholders to sustain current DCI technology activities and to set up the foundation for further activities that should become independent from EC funding at large. We can conclude that there are two primary benefits for the scientific user communities depending on whether they need to find software and skills for new or existing activities or to get information and support about software they are already using. The creation of an active software infrastructure is essential to promote the usage and sharing of software and related information to help in both cases. A comprehensive catalogue of software and services is fundamental in finding the right tools or checking whether solutions to problems already exist before committing resources to new initiatives. The possibility of accessing professional profiles of people helps to enable new collaborations making ScienceSoft a 'great chance' for the community.

Primary authors: Dr DI MEGLIO, Alberto (CERN); Mr RIEDEL, MORRIS (JUELICH SUPERCOMPUTING CENTRE)

Co-authors: Prof. OULD-SAADA, Farid (Oslo University); Ms ESTRELLA, Florida (CERN); Dr MATYSKA, Ludek (CESNET); Dr SMIRNOVA, Oxana (Lund University); Dr STEFAN, Peter (NIIF); Prof. BALA, Piotr (UWAR - ICM)

Presenter: Dr DI MEGLIO, Alberto (CERN)

Session Classification: Sustaining the EGI ecosystem: workshop