ScienceSoft
Open Software for Open Science

Alberto Di Meglio, Florida Estrella – CERN
Morris Riedel - FZJ

First Workshop on Maintainable Software Practices in e-Science (SoftwarePractice12)
Chicago - 09/10/2012
Content

• Background and context
• Plans and current activities
• Next steps
EU eInfrastructures

An interoperable Distributed Computing Infrastructure for Europe
EMI in a nutshell

- Collaboration of 4 mw providers (ARC, dCache, gLite, UNICORE)
- 26 partners
- 24M EUR
- 3 years (05/10 – 04/13)
- 55 products
- 2M SLOC
- Open source
EMI Vision and Challenges

Before EMI

3 years

After EMI

Applications Integrators, System Administrators
- Standard interfaces
- Specialized services, professional support and customization

EMI Reference Services

Standards, New technologies (clouds) Users and Infrastructure Requirements

09/10/2012
SoftwarePractice12 - Chicago
Critical issues

• How to ensure technical coordination?
• How to share information across projects and across scientific domains?
• How will the software be supported in the future?
• Are all products sustainable? Are all worth being sustainable?
• Anything that can be exploited commercially? How?
• Is there something else we can use? Is it good? Does it work as it should? Can we contribute to it?
• Who is using the software and how does usage evolve?
• Do users mention the software, acknowledge its contribution to their research?
Initial discussions

• Other projects, communities and SMEs shared the same challenges and concerns or expressed interest in similar discussions

The ScienceSoft Steering Committee (Dec 2011-Feb 2013)
First workshop

• 1st ScienceSoft Workshop – CERN, Feb 2012
  – Original “Steering Committee” and other interested persons, 33 in total
  – Helped streamlining the requirements and possible features

• Overview document:
Roadmap

– **Alpha (March-June 2012)**: refinement of ideas, requirements and features, state-of-the-art

– **Design (July-December 2012)**: initial design and prototyping

– **Concept (January-April 2013)**: proof-of-concept, iterations on features

– **Operation (May 2013 onwards)**: community managed activities

Roadmap: [http://sciencesoft.org/roadmap](http://sciencesoft.org/roadmap)
Features

Catalogues
Software, people, organizations, projects, events, etc.

Processed information
what, where, who, how good/bad, help, licensing, data import/export, notifications, etc.

Marketplace
Communities, services, incubation, crowdsourcing, consultancy, etc.

Identification and citation
Unique identifiers, “softographical” formats, etc.
Sustainability

• Create wider communities around worthwhile software

• Generate revenues by:
  – Enabling the provision of specialized software-related services to well-identified groups
    • *Software engineering, consultancy, specialized technical development, training, etc.*
  – Commercial incubation support
    • *Business plans, legal information, IP, licensing, etc.*
Software Identification & citation

• Improve academic recognition of software development
• Make scientific results fully reproducible
  – Unique IDs, code runs, VMs
• Extend methods from publications and datasets (DOI)
• Agree on a format suitable for citation equivalent to bibliographical formats
State of the art

Applications Database
for the European Grid Infrastructure

The OpenScience Project
Design and prototypes

• Functionality:
  – Web portal
  – Data model, Data discovery and automation
  – Comment and rating system, Subscriptions and notifications
  – Information processing and reporting
  – Marketplace, Software Identification & citation

• Forums on the ScienceSoft portal at

http://sciencesoft.org/forum
First steps (technical)

• Web site at http://sciencesoft.org
  – Based on Drupal (hosted at CERN), features added as modules
  – Easy to contribute features

• Simple data model defined for objects and relationships

• Registration, search, maps

• OpenID-based access, support for Facebook, gmail (not yet public)

• Automation of relationship and data collection being discussed
First steps (strategic)

• Creation of collaborations and networking
  – Existing project and initiatives
  – Decision-makers in international organizations, research centers and funding bodies (EIROForum, EC, FAO)
  – Open source consultancy companies and incubators
Operational Challenges

Governance
Foundation?
Roles, policies
Admin, legal, logistics, etc.

Funding
Public funding
Sponsoring
Paid-services
Voluntary contributions
Next steps

• Finish the prototype according to plan
• Discuss with other projects about common activities and extending/reusing functionality
• Disseminate the results
• 2nd workshop tentatively Feb 2013
  – Governance, funding, future plans
• Based on feedback and participation, stop or continue with operational phase

09/10/2012 SoftwarePractice12 - Chicago