

Interoperability and data lock-in in forges

Sébastien Campion

INRIA

9th November, Fossa 2010, Grenoble



Overview

Definitions and Uses

2 Interoperability, Coclico project



Overview

Definitions and Uses

2 Interoperability, Coclico project



What kind of data?

- Source code (sources and dependencies)
- Documentation (installation guide, manual)
- Bug trackers
- Lifecycle (Roadmap, TODO, requirements for enhancements)
- People (community, authors, mailing lists, ...)



What usage?

- A example use case, a segfault or an error appears during the first uses.
 - Search on google.com
 - Find an answer in the bug tracker
 - Apply the patch

All data other than source code are also fundamental in a software developpement process

- Academic practices and lifecycle
 - free software
 - launch a startup
 - transfer a software to a community, change forge hosting
 - old papers and experimentations need archiving process

Must be free and mobile



The data lock-in problem

What happens if:

- your forge service is down, hosting unavailable ?
- Google Code decides to change their policies ?
- Oracle shuts down kenai.com ?
- you want to change your forge ?

A shared problem, "Three Systemic Problems with Open-Source Hosting Site" October 8th, 2009, Eric S. Raymond http://esr.ibiblio.org/?p=1282



Coclico WP2: Interoperability and data exchange

We establish a proposal for an interoperability standard for exchange of data between software development forges.

Two steps:

- On't reinvent the wheel, use native format from forge tools
 - svndump, mailman archives, ...
 - trackers: ForgePlucker (Google Code, Berlios, Trac, Savane, Sourceforge, FusionForge)
- Organise and federate
 - Toolbox and framework integration with semantic index
 - semantic web tools for automatic translation



Work Done: First step

A small use case: export/import from FusionForge.

- Lots of contributions to ForgePlucker to enable export of trackers, people, permissions, news, forums, file release systems
- Patch FusionForge API to enable data import operations
- Export process create an archive project.tar.gz

What have we learned?

- Forge APIs are not sufficient, how to backdate an artifact ?
- Data is structured but without semantics.

Demo at the end https://forge.projet-coclico.org/projects/wp2/



In progress, and future

In Progress

- Refactoring and code cleaning
- Pushing patch on the mainstream
- Driving and integrating each data extraction process into a single toolbox

Future

- Choose an other Forge (Codendi, Trac, ...)
- Promote idea and software to build and maintain community
- Integrate and use web semantic tools to improve data conversion process
- Production use on INRIAGForge



Demo: export/archive trackers, forums, documents, files from INRIAGForge



Demo: import data in a new project

