

Increasing the efficiency of free software projects through information management

Robert Schuster

Advisor: Christopher Oezbek, Prof. Dr. Lutz Prechelt

Working Group Software Engineering

Freie Universität Berlin

June 2005

Outline

- Introduction to GNU Classpath
- The problem
- The alternative solutions
- The approach: Mediation
- Examples from the experiment
- The Mediation Manual
- Facts and figures
- Lessons learned and conclusion

Outline

- Introduction to GNU Classpath
- The problem
- The alternative solutions
- My approach: Mediation
- Examples from the experiment
- The Mediation Manual
- Facts and figures
- Lessons learned and conclusion

What is GNU Classpath?

- free Java class library
- official GNU project
- for use with multiple JVMs
 - ~15 projects (Kaffe, SableVM, Harmony, JCVm, JAmiga, JamVM, JNode, IKVM.NET, Jamaica, ...)
- founded 1998
- 50 developers over time, ~30 active ones
- ~0.75 Mio SLOCS
- this is where I applied mediation
 - all practical examples are taken from this effort

Outline

- Introduction to GNU Classpath
- The problem
- The alternative solutions
- My approach: Mediation
- Examples from the experiment
- The Mediation Manual
- Facts and figures
- Lessons learned and conclusion

Kinds of F/OSS Projects I

- single person
 - plenty
 - too few developers for communication challenges
 - good: QEmu, cdrwtools
- community based
 - serves a specific need, attracted some developers
 - informal management, no given guidelines
- organised community project
 - part of an organisation's strategy (GNU, Debian, ...)
 - conventions and guidelines given (Apache Mentor)

Kinds of F/OSS Projects II

- company-controlled projects
 - “open-sourced” software (OpenOffice.org, OpenCMS, ...)
 - former copyright holder governs development
 - copyright assignment
 - conventions
 - volunteer contributors are rare

Conditions - Communication

- most projects: mailing list and IRC
- informal manners / conduct
- no one reads everything
 - even worse when traffic is high
- email: publicly archived
- IRC:
 - fast response time
 - is used for socializing, too
 - usually not archived for privacy reasons
- interest:
 - source of problems, understand process

Conditions – Decision Making

- process largely undefined
- social conventions, learn by experience
- problematic:
 - “flame war”
 - “bike shed”
 - no one cares
- outcome not written down (!)
- interest:
 - source of problem

Conditions – Tool Usage

- tools are usually community-born (e.g. CVS)
- preferences
 - “vi!”
 - “emacs!”
 - “joe!”
 - “nano!”
- interest:
 - avoid imposing a certain tool

F/OSS Project's Problem

- outcome of discussions hard to find for non-participating or new developers
 - recurring topics
 - result may not be reached (because of lack of interest)
- concerns of new developers are not considered
- state of independent tasks not publicly visible
 - Where is work going on?
 - What tasks need more support?

Outline

- Introduction to GNU Classpath
- The problem
- The alternative solutions
- My approach: Mediation
- Examples from the experiment
- The Mediation Manual
- Facts and figures
- Lessons learned and conclusion

What others did

- Hipikat
 - find related information from distributed locations
 - Eclipse plugin
 - provides context-sensitive search
 - gathers data from BugZilla, CVS, Newsgroup
- Kerneltraffic
 - website owner monitors severals developer mailing-lists (e.g. Linux Kernel Mailinglist)
 - writes summaries about the threads and publishes them
 - aims to publish news not aimed at project support

What others did ... continued

- Kernelnewbies
 - teaching developers OS kernel development
 - focussed on Linux
 - features mailinglist, IRC channel and newsgroup
 - helps new developers, separated from project

Outline

- Introduction to GNU Classpath
- The problem
- The alternative solutions
- My approach: Mediation
- Examples from the experiment
- The Mediation Manual
- Facts and figures
- Lessons learned and conclusion

Mediation Goals

- lower entry barrier
- provide better overview of project's progress
- support discussions
- teach mediation

Lower entry barrier

- collect newcomer related data
 - design decisions made in the past
 - policies, requirements
 - tool guide
- examples:
 - copyright assignment
 - coding style
 - hacking with Eclipse
 - building from CVS

Enhance overview

- collect
 - who does what?
 - state of affairs
 - where to find more information
- examples:
 - CORBA implementation
 - Eclipse Formatter for Java
 - VM support for certain JNI functions

Support discussions

- identify unsuccessful discussions
 - eg. recurring topics
- collect relevant data
 - explain problem
 - current conclusion
 - links to former discussions
- ask question again
- write down outcome
- examples:
 - click-through licenses
 - support for 1.5 API

Teach mediation

- Why bother?
 - technical side of an issue may be too demanding
 - time
- ideal world: no mediation necessary
 - everyone writes down achievements and plans on their own
- raise consciousness
 - point at existing information
 - teach how to help or update the data
- examples?
 - mostly via IRC (“look at X for topic Z”)
 - 'MediationMissionPage'

Work considerations

- Where does all the information come from?
 - IRC, mailing list, meeting (eg. FOSDEM)
- avoid
 - forcing collaboration
 - imposing usage of specific tool
- discussion obstacles
 - no one participates
 - cannot reach consensus
 - subjectivity

Outline

- Introduction to GNU Classpath
- The problem
- The alternative solutions
- My approach: Mediation
- Examples from the experiment
- The Mediation Manual
- Facts and figures
- Lessons learned and conclusion

Mediation Wiki - Structure

- 3 main parts
 - current topics
 - first steps
 - decisions
- each page: list of issues
- each issue:
 - last change, author, description, references
 - decision: outcome
- description
 - what is the problem about, background information
- outcome
 - what is everybody agreeing on

Mediation Wiki – An example Issue

- “dealing with stub implementations”
- description
 - fake implementation of standard API
 - methods returning null, 0 or false regardless of specification
- outcome:
 - “grep FIXME and start hacking”
 - stub not considered better than missing method
 - must be documented
 - are evil and have to be implemented
- references to 3 separate hot debates on the mailinglist

Mediation on Mailing List I

- recurring question on mailing list: “What are tainted developers allowed to work on?”
- wrote discussion request
 - mentioned existing information (FAQ entry)
 - added what was said in earlier (but incomplete) discussions
 - linked resources as references

Mediation on Mailing List II

- asked related questions
 - “May tainted developers contribute ideas on design decisions (on mailing-list/IRC)?”
 - “May tainted devs fix simple bugs (eg. forgotten imports that cause classpath compilation to fail)”
- definitive answer could be reached through maintainer
- outcome was added to as an issue to the Wiki

Real time mediation on IRC

- inform about updates specific to developer
 - steps in development of character conversion framework
- having a specific information in the Wiki was requested
 - building GNU Classpath from CVS
- pointing to Wiki links during discussions
 - *often*

Outline

- Introduction to GNU Classpath
- The problem
- The alternative solutions
- My approach: Mediation
- Examples from the experiment
- The Mediation Manual
- Facts and figures
- Lessons learned and conclusion

What is the Mediation Manual?

- goal: make mediation applicable for other projects, too
- question and answer style
- explains basics of mediation
- suggests tool
- gives practical examples
- warns about difficulties (eg. subjectivity)

Spreading the word ...

- selected 50 F/OSS projects from sourceforge
 - alpha or beta state
 - at least 3 members
 - age of at least one year
 - at least one release in the last two years
- contacted via mailing list
 - 30 mails reached target without problem
 - 12 mails needed confirmation and passed
 - 8 were rejected

Mediation Manual – Reactions

- suggestions for typographical fixes
- wxGlade: not enough stable members
- NHibernate: liked Wiki idea only
- Syllable: has someone who cares for (system) documentation and usage tips
- PearPC
 - deemed mediation helpful
 - discrepancy between developer and user's knowledge

Mediation Manual - Results

- almost positive reactions
- some project have a need for a mediator
- but no volunteer yet
- at least:
 - some developers have learned about mediation

Outline

- Introduction to GNU Classpath
- The problem
- The alternative solutions
- My approach: Mediation
- Examples from the experiment
- The Mediation Manual
- Facts and figures
- Lessons learned and conclusion

Survey

- interest
 - knowledge about mediation effort
 - evaluation of the practical work
 - self-assessment of developer's participation
 - evaluation of the Wiki and the topics chosen
- 29 questions
 - discrete levels of agreement (strong weak agree/disagree)
 - some free text
- individual invitation sent to every developer
- 11 of them participated in the survey

Survey – Knowledge about Mediation

- basic knowledge (“What is it”)
 - strong reception (81%)
- intermediate knowledge (“How to support”)
 - still strong reception (63%)
- high knowledge (“Doing mediation myself”)
 - rejection prevails (63%)

Survey – Answers

- diverse reactions
 - “It works seamlessly and well, [...]”
 - “I don't know what it is.”
- possible solutions for being informed better
 - “weekly or bi-weekly updates to the mailinglist on what was summarized/added.”
 - “Perhaps some status reports from time to time [...]”

Survey – Results I

- agreement (81%) that mediation helps new developers
- same for long-established but less strong
- clear need to add more respondents for mediation
 - “The active users of GNU classpath, [...]”
 - “I think we could do a better job at engaging the non-technical audience that's willing to help, [...]”

Survey – Results II

- participation by developers is low
 - not written an issue for the Wiki (72%)
 - not edited an existing issue (54%)
 - not answered mediation related questions (63%)
 - not suggested a new topic/issue (72%)
- reasons given by respondents
 - “Too little time and lazyness”
 - “It's not a very high priority for me as yet.”
 - “I'm not sure what the mediator is or why I'd need it.”

Outline

- Introduction to GNU Classpath
- The problem
- The alternative solutions
- My approach: Mediation
- Examples from the experiment
- The Mediation Manual
- Facts and figures
- Lessons learned and conclusion

Lessons learned

- less resistance than expected
 - public announcements did not cause trouble
- Wiki proved to be a practical all-purpose tool
- less discussions on controversial topic than expected
- mediation has limits
 - does not take off the need to learn certain tools
 - some steps have to be experienced (eg. problems when building a compiler)

Conclusion - Advantages

- new developers can learn about aspects of the project that are of interest for them (eg. special policies)
- developers can inform themselves easier about the outcome of former decisions
- important information is not lost but gets filed and written down

Conclusion - Shortcomings

- some developers have not been informed
 - adopt one of the suggested styles of informing developers
- users and other groups should be targetted as well
 - How much time will this consume?

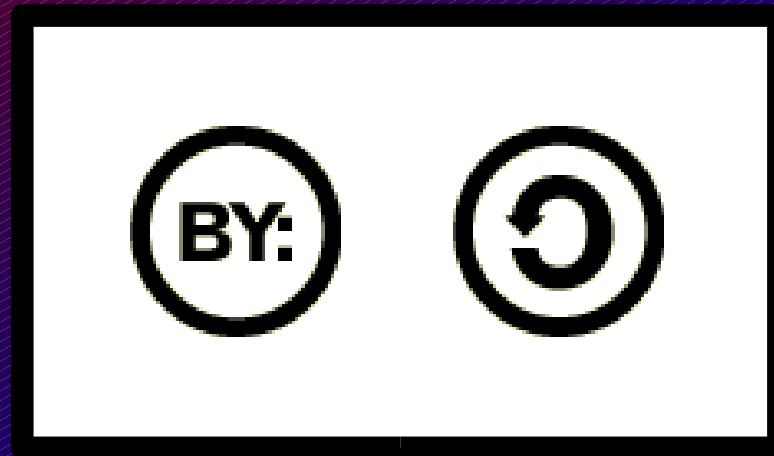
Perspective

- mediation was applied to one project now
 - Does it work for others, too?
- mediation data grows over time
 - Will it stay manageable in the future?
- experiment had a rather short time frame
 - Will mediation stay feasible when applied permanently?

Finished!

- ~~Introduction to GNU Classpath~~
- ~~The problem~~
- ~~The alternative solutions~~
- ~~My approach: Mediation~~
- ~~Examples from the experiment~~
- ~~The Mediation Manual~~
- ~~Facts and figures~~

Some rights preserved ... :-)



This work is licensed under the Creative Commons Attribution-ShareAlike License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-sa/2.0/de/> or send a letter to Creative Commons, 559 Nathan Abbott Way, Stanford, California 94305, USA.