

Alambic: An Open-Source Platform for Software Engineering Data Management

The Case of Embedded Software Development



Boris Baldassari



boris.baldassari@castalia.solutions



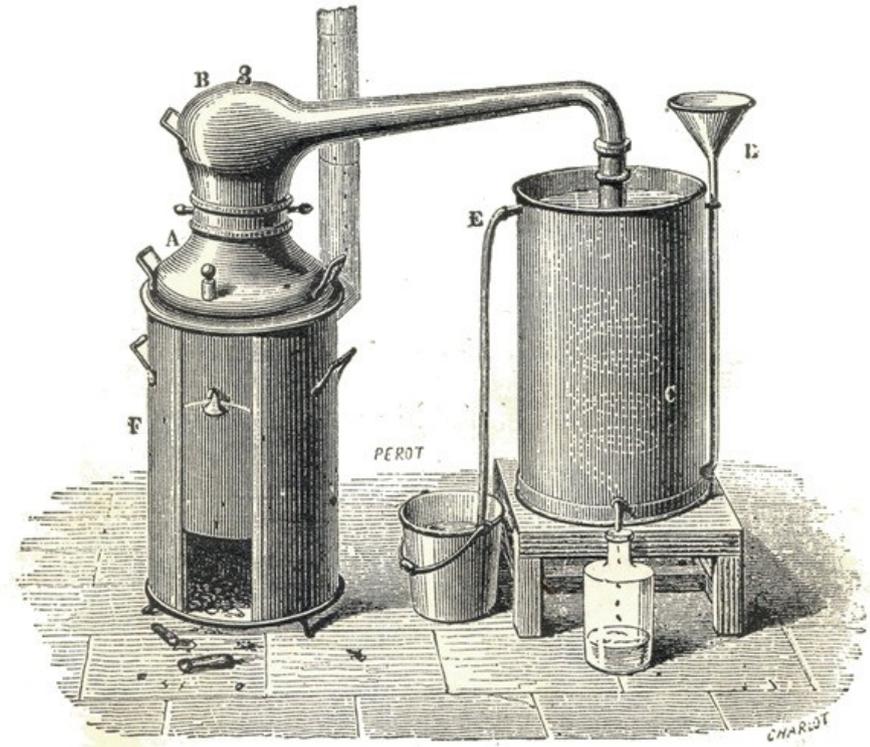
<http://castalia.solutions>

Check the full article! →



Plan

- Software quality, open data, and Alambic
- Alambic
 - Introduction: why Alambic
 - Architecture
 - Plugins: MetricsGrimoire, Hudson CI, PMD, Stack Overflow
- Use Cases
 - The PolarSys dashboard
 - The Eclipse forge
- Alambic 3.0 – one step forward
- Conclusion & Perspectives



Software quality, open data and Alambic

Software quality has different meanings for different contexts :

- Open-source projects (e.g. includes community attributes),
- Embedded software (e.g. stresses reliability and testability),
- Desktop software (e.g. stresses usability and maintainability).

► Defining a **quality model** helps people realise what quality is in their context, and how to measure and improve it.

Transparency (open data) is a key feature of software development.

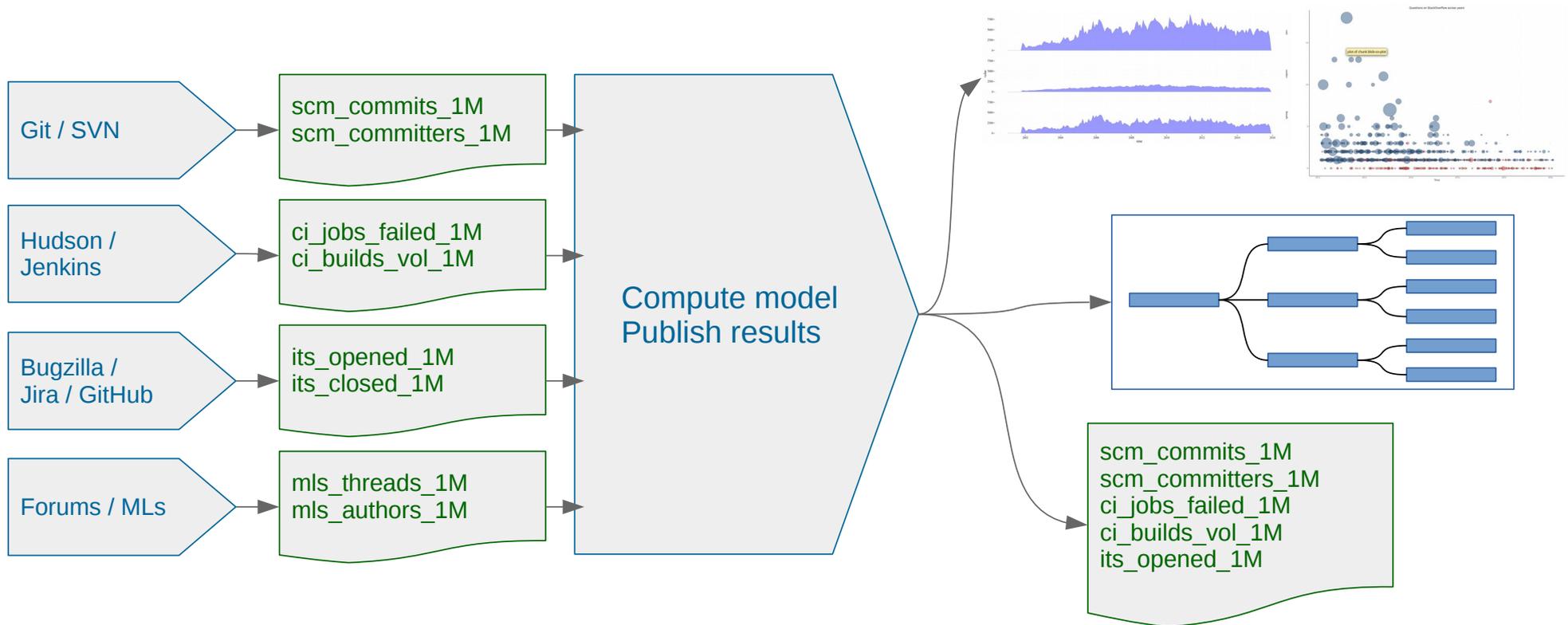
- For open-source projects,
- For proprietary projects – Agile methodology explicitly promotes it.

Alambic – Introduction

Why Alambic?

- A fairly big amount of data is generated along the development process, including requirements, code, issues, builds, delivery, support, and more...
- Alambic automatically retrieves, curates, and organises information from these data sources and presents it to the user in a unified vision:
 - Dedicated plugins can run checks, actions, and reports.
 - Visualisations show various aspects of the project behaviour.
 - Recommendations are proposed to improve practices and quality.
- All data is made available in a neat format (JSON, CSV) so data enthusiasts can easily play with it, and use external tools for reporting, visualisation, data mining, post-processing..

Alambic – Architecture



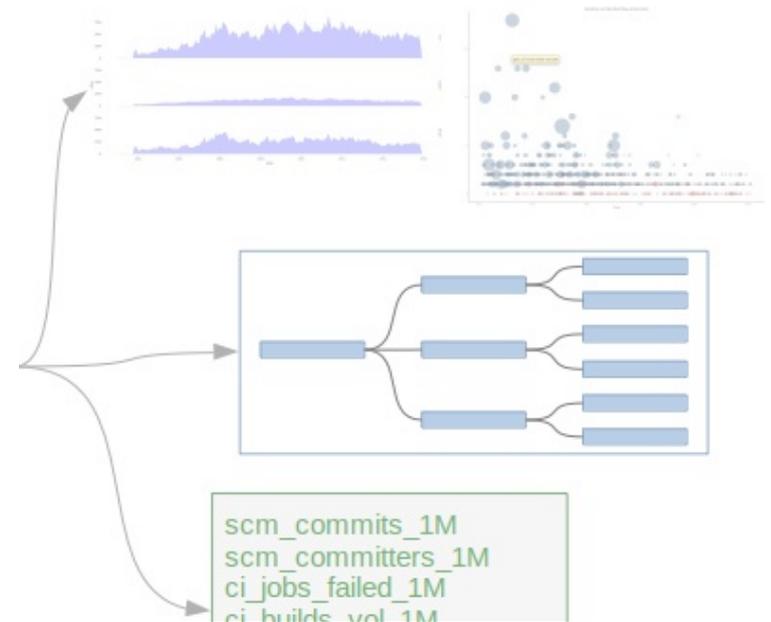
Alambic outputs

Alambic provides:

- A dashboard to browse all information online.
- Downloads for metrics, attributes, recommendations (JSON, CSV).
- Exportable figures: PNG/SVG, HTML snippets, interactive plots.
- Visualisations: quality model, data, history, specific data sources analysis.

Alambic 3.0

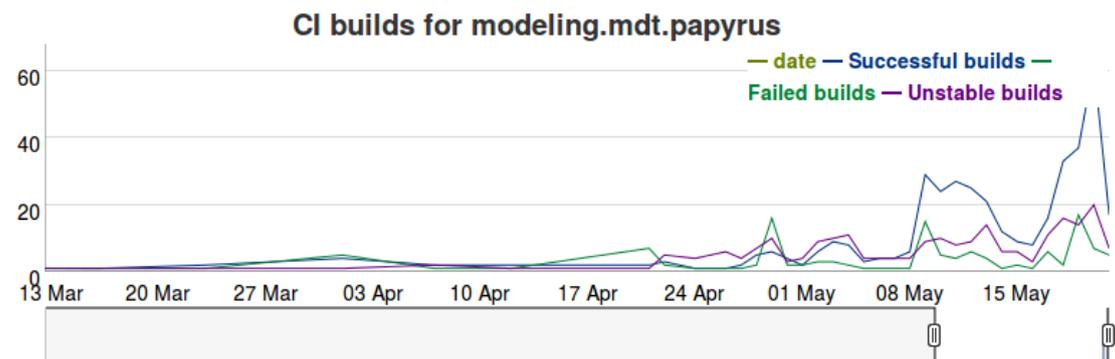
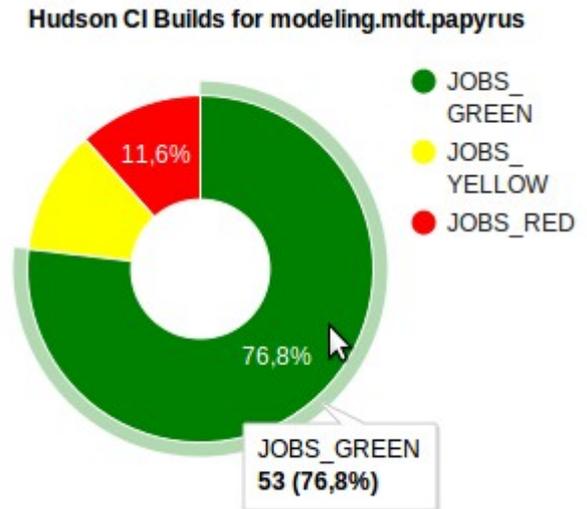
- PDF reports.
- Plotly graphics for online editing.



Alambic plugins – Hudson CI engine

This plugin connects to a Hudson CI instance and provides

- Metrics
number of jobs, jobs failed/successful/unstable,
builds failure rate..
- Figures
builds history, jobs results
- Data
builds, jobs (JSON, CSV)
- Recommendations.



Alambic plugins: Eclipse ITS, SCM, MLS

A set of plugins to retrieve information from the Eclipse dashboard [EDA].

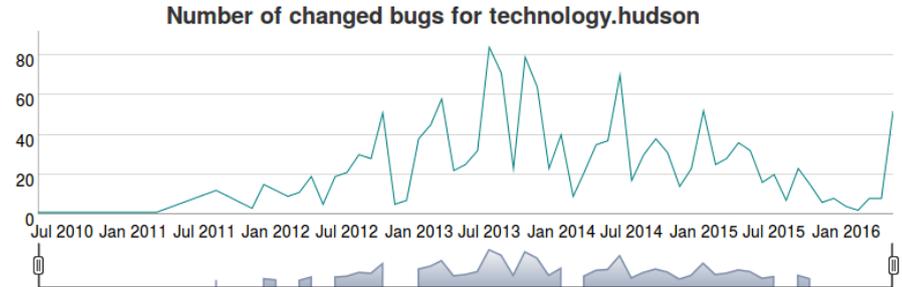
Community metrics are computed using Bitergia's MetricsGrimoire [BMG], a generic tool to analyse software development repositories and tools metadata.

- **Issue Tracking System (ITS)**

- Metrics: its_closed_7, its_closed_30, its_closed_365, its_bugs_closed, its_bugs_changed, ...
- Figures, visualisation
- Recommendations

[EDA] <http://dashboard.eclipse.org>

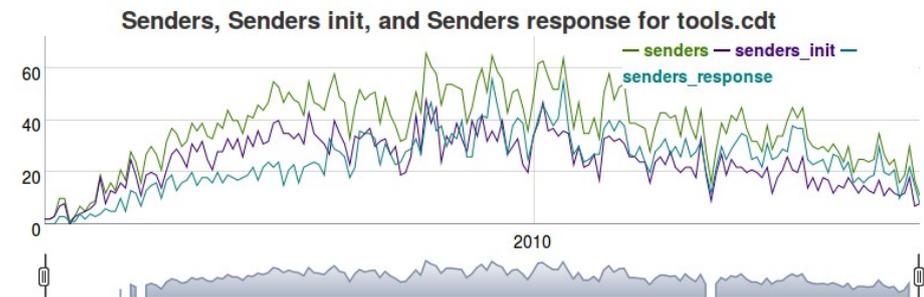
[BMG] <https://metricsgrimoire.github.io>



Alambic plugins: Eclipse ITS, SCM, MLS

A set of plugins for community metrics.

- **Software Configuration Management (SCM)**
 - Metrics: scm_commits_7, scm_commits_30, scm_commits_365, scm_authors_7, scm_committers, scm_files, ...
 - Figures, visualisation
 - Recommendations
- **Mailing Lists (MLS)**
 - Metrics: mls_senders_7, mls_senders_30, mls_senders_365, mls_sent_7, mls_threads, mls_senders_response, ...
 - Figures, visualisation
 - Recommendations

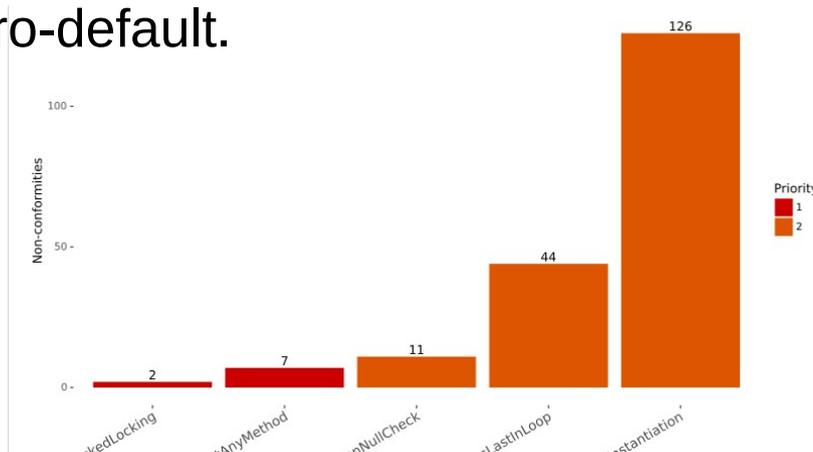
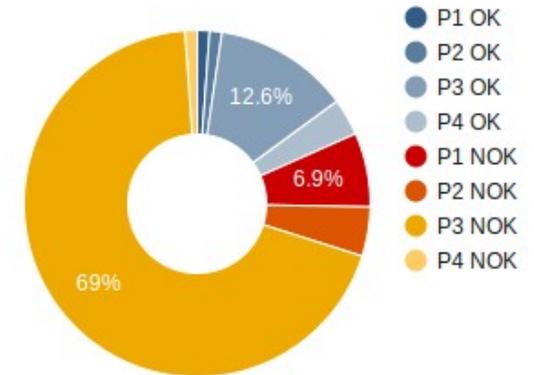


Alambic plugins – PMD Analysis & Configuration

A plugin to help users understand the output of PMD and configure the tool.

- PMD Analysis
 - Get a summary of violations found by PMD
 - Identify the most important violations to resolve.
- PMD Configuration:
 - Use only most important rules, with fewer violations.
 - Target *conscious* zero-default.

Rules checked



Mnemonic
DoubleCheckedLocking
AbstractClassWithoutAnyMethod
BrokenNullCheck
AvoidBranchingStatementAsLastInLoop
BooleanInstantiation

High-priority rules with a low number of violations

Use Cases: The PolarSys dashboard

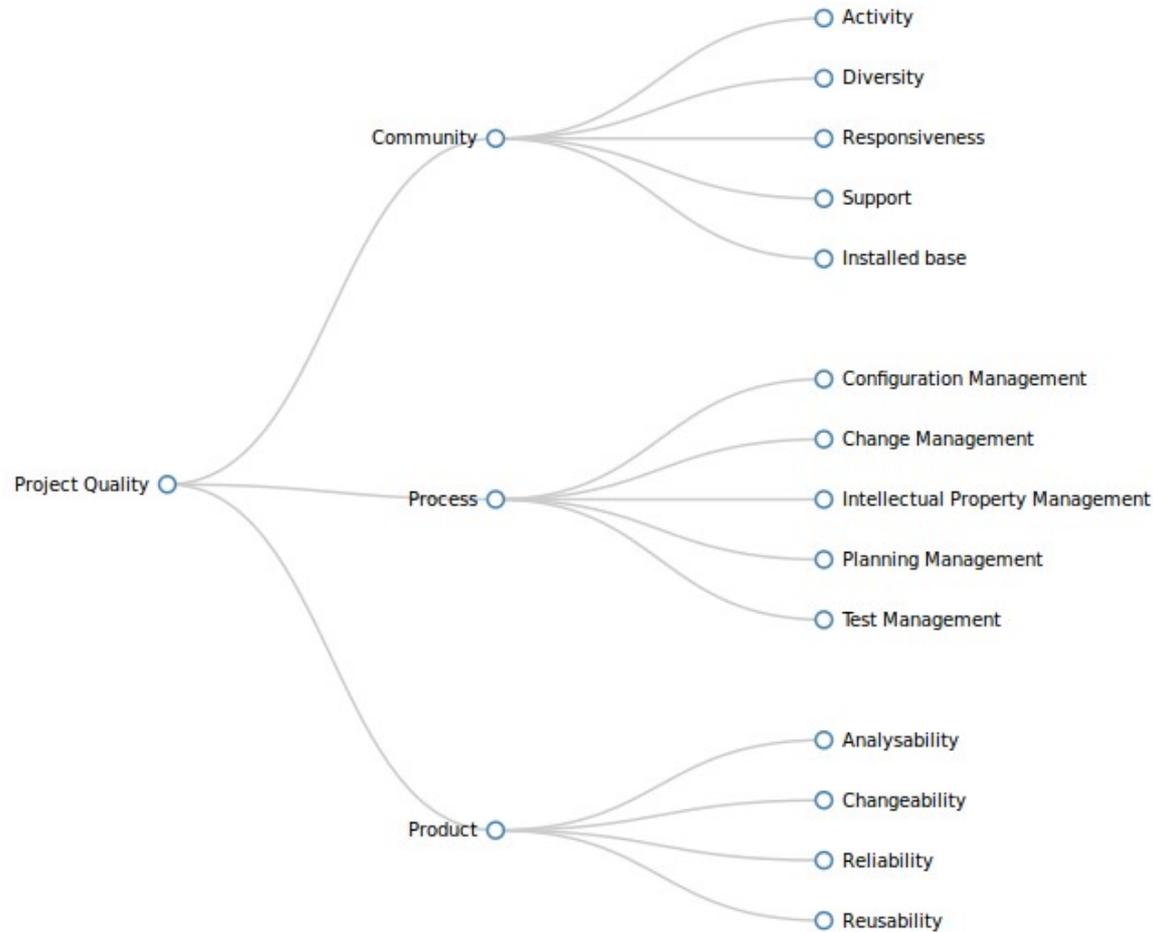
PolarSys is an Eclipse Industry Working Group that provides an open-source tool chain for embedded systems.

A maturity assessment task force was initiated to ensure the overall consistency of projects regarding quality:

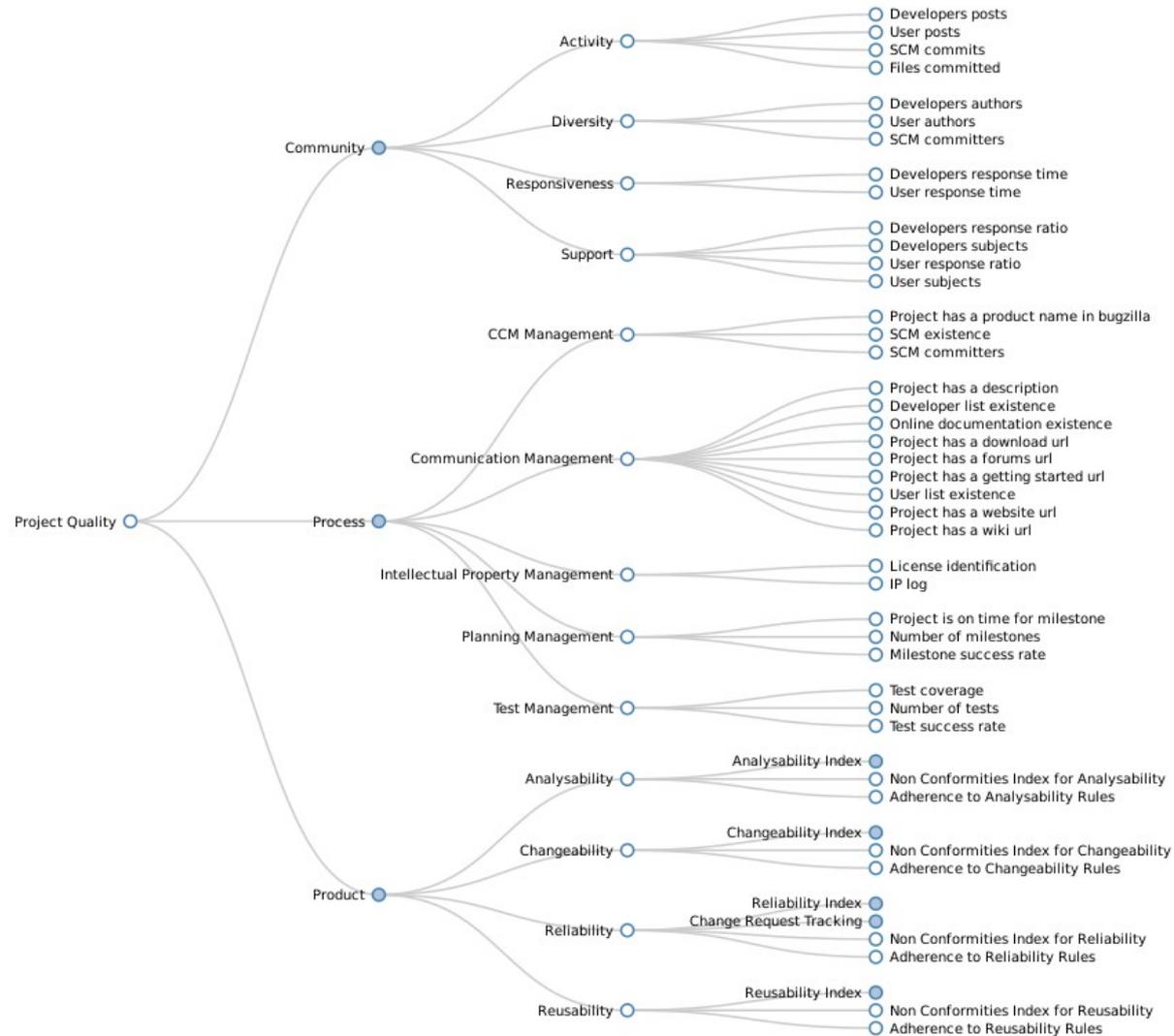
- Define a quality model with specific quality concerns
- Define means to measure quality attributes
- Define presentation of data (dashboard)



Use Cases: The PolarSys dashboard



Use Cases: The PolarSys dashboard



Use Cases: The PolarSys dashboard

PolarSys Dashboard — v1.4 β

tools.cdt **Summary**

Summary QM Attributes Questions Metrics Practices Manual data Plugins Download Errors

Description Completeness: **45 / 45**

Rating for main quality attributes

Completeness of metrics 45 / 45

Ecosystem → (13 / 13 metrics) 3.7 / 5

Process → (12 / 12 metrics) 3.8 / 5

Product → (20 / 20 metrics) 2.7 / 5

PMI Summary

- Eclipse Dashboard Not defined
- Web <http://www.eclipse.org/cdt>
- Wiki <http://wiki.eclipse.org/index.php/CDT>
- Downloads <http://www.eclipse.org/cdt/downloads.php>
- Documentation <http://wiki.eclipse.org/index.php/CDT>
- Mailing lists [[Developer ML](#)] []

Comments

Add a comment? (Authentication required) +

Manual data

Ecosystem

Activity → (4 / 4 metrics) 4.2 / 5

Diversity → (3 / 3 metrics) 4.0 / 5

Responsiveness → (2 / 2 metrics) 2.5 / 5

Support → (4 / 4 metrics) 4.0 / 5

Process

Configuration Management → (3 / 3 metrics) 4.5 / 5

Change Management → (3 / 3 metrics) 4.5 / 5

Planning Management → (2 / 2 metrics) 3.0 / 5

Test Management → (4 / 4 metrics) 3.2 / 5

Product

Analysability → (5 / 5 metrics) 2.4 / 5

Changeability → (4 / 4 metrics) 2.9 / 5

Reliability → (7 / 7 metrics) 2.6 / 5

Reusability → (4 / 4 metrics) 2.9 / 5

Home About Documentation Quality Model Attributes Questions Metrics Scales Practices References Projects CDT Capella EASE EMF Compare ESF Gendoc Kitalpha Sirius TraceCompass Admin panel Contact

Use Cases: The PolarSys dashboard

PolarSys Dashboard — v1.4 β
👤

- Home
- About
- Documentation
 - Quality Model
 - Attributes
 - Questions
 - Metrics
 - Practices
 - References
- Projects
 - CDT
 - Capella
 - EASE
 - EMF Compare
 - ESF
 - Gendoc
 - Kitalpha
 - Sirius
 - TraceCompass
- Admin panel
- Contact

tools.cdt Quality model

Summary
QM
Attributes
Questions
Metrics
Practices
Manual data
Plugins
Download
Errors

The Quality model

The quality model shows the complete hierarchy tree, from quality attributes to measurement concepts (questions) and metrics. Following Basil's Goal-Question-Metric approach [Basil1994], the quality attributes (the 3 first columns) are goals for our measurement, concepts (4th col) are mapped to questions, and metrics (right col) are the base measures.

Activity (4.0)	Developer ML activity (5.0)	Developer ML posts (462 / 5)
	User ML activity (4.0)	User ML posts (23 / 4)
	SCM activity (3.0)	SCM Commits (12 / 3)
		Files committed (68 / 3)
Diversity (4.0)	Developer ML diversity (5.0)	Developer ML authors (38 / 5)
	User ML diversity (3.0)	User ML authors (5 / 3)
	SCM diversity (4.0)	SCM committers (4 / 4)
Responsiveness (2.5)	Developer ML responsiveness (4.0)	Developer ML response time (0.1 / 4)
	User ML responsiveness (1.0)	User ML response time (1278.7 / 1)
Support (4.0)	Developer ML support (4.5)	Developer ML response ratio (2.8 / 4)
	User ML support (3.5)	Developer ML subjects (115 / 5)
		User ML subjects (15 / 5)
		User ML subjects (1 / 2)
Visibility	Visibility in research	Number of publications
Usage	Downloads	Number of downloads on the web site
		Number of downloads on update site
User feedback	Installations	Installation survey
	Marketplace feedback	Number of successful installs on the Marketplace
Configuration Management (4.2)	SCM Information access (5.0)	Number of favourites on the Marketplace
	SCM usage (3.5)	SCM information (5 / 5)
Change Management (4.5)	ITS Information access (5.0)	SCM Commits (12 / 3)
	ITS usage (4.0)	SCM committers (4 / 4)
		ITS information (5 / 5)
		ITS updates (461 / 4)
		ITS authors (56 / 4)
		License identification
		IP Log
		IP log code coverage
		Plan on time
		Project is on time for next milestone
		Milestones (3.0)
		Number of milestones (1)
		Reviews success rate (80 / 5)
		Test coverage (3.0)
		Percentage of lines of code covered by tests (39.9 / 3)
		Percentage of branches covered by tests (38.1 / 4)
		Number of tests relative to the code size (0.8 / 1)
		Test success density (5.0)
		Test success density (100 / 5)
		Code size (2.0)
		Source Lines Of Code (898719 / 2)
		Access points
		Average number of attributes
		Average number of methods
		Average Cyclomatic Complexity (3.1 / 3)
		Depth of Inheritance Tree
		Comment rate (16.3 / 3)
		Average number of non-conformities for analysability (6.8 / 1)
		Adherence to analysability rules (16.7 / 2)
		Public access points (2.0)
		Public API (40282 / 2)
		Average Cyclomatic Complexity (3.1 / 3)
		Depth of Inheritance Tree
		Cloning density (8.1 / 5)
		Average number of non-conformities for changeability (4.7 / 1)

Metric: User ML authors (MLS_USR_AUTH_3M) [More info](#)

Active: True **Value:** 5

Data source: polarsys_grimoire **Computed indicator:** 3

Description:

Number of distinct senders for messages dated during the last three months, in user mailing list archives.

User mailing list is the list or lists considered as 'for users' in the project documentation. The date used is the mailing list server date, as stamped in the message. Time range is measured as three calendar months period starting the day before the data retrieval (example: if retrieval is on Feb 3rd, period is from Jan 3rd to Feb 2nd, both included). Distinct senders are those with distinct email addresses. Email addresses used are the strings found in 'From:' fields in messages.

Note that developer communications are measured through the mailing list, and user communications are measured through the forums.

Castalia Solutions

16

Use Cases: The Eclipse forge

Use the Eclipse Foundation's infrastructure to retrieve information from the various data sources and tools of the Eclipse forge:

- The PMI for project details (description, web site, downloads, etc.)
- Issue tracking (Bugzilla)
- Configuration Management (Git)
- Communication: Mailing lists and forums
- Hudson CI, PMD..



Also use external sources:

- Stack Overflow questions about the project

Use Cases: The Eclipse forge

Alambic dashboard for Eclipse - v2.0 (beta)

technology.egit **Summary**

Summary QM Attributes Questions Metrics Manual data Plugins Download Errors

Description Completeness: **13 / 13**

Highlights

13 / 13 Completeness	81 Metrics	
4.2 / 5 Ecosystem →	4.5 / 5 Process →	
44 Commits ↗ last month	40 Issues closed ↗ last month	46 Mails sent ↗ last month

PMI Summary

- Eclipse Dashboard [Not defined](#)
- Web <http://www.eclipse.org/egit/>
- Wiki <http://wiki.eclipse.org/EGit>
- Downloads <http://eclipse.org/egit/download>
- Documentation <http://www.eclipse.org/egit/documentation>
- Mailing lists [[Developer ML](#)] [[User Forums](#)]

Manual Data

Add a comment? (Authentication required) +

Visualisation plugins

- [pmi_checks](#)
- [eclipse_mls](#)
- [eclipse_its](#)
- [eclipse_scm](#)
- [stack_overflow](#)
- [hudson](#)

Configuration management

Overall:
4668 commits made by 127 authors and 127 committers, representing 3621 files.

Last month:
44 commits by 8 authors.

Last year:
480 commits by 34 authors.

Issue Tracking System

Overall:
2492 issues opened by 688 people
1867 issues closed by 111 people.

Last month:
263 % of issues were closed by 4 people.

Last year:
11 % of issues were closed by 18 people.

Mailing lists

Overall:
6196 messages sent by 189 authors in 1807 threads.

Last month:
46 messages sent by 4 authors.

Last year:
478 messages sent by 25 authors.

Stack Overflow

Home About Documentation Projects Andmore CDT CDT TCF EASE EMF Egit GEF GMF Hudson JDT OCL PDT Papyrus Scout Sirius Target Management Titan Trace Compass Admin panel Contact

Use Cases: The Eclipse forge

Alambic dashboard for Eclipse - v2.0 (beta)

Home About Documentation Projects Andmore CDT CDT TCF EASE EMF Egit GEF GMF Hudson JDT OCL PDT Papyrus Scout Sirius Target Management Titan Trace Compass

technology.egit Plugin stack_overflow

Summary QM Attributes Questions Metrics Manual data Plugins Download Errors

StackOverflow Questions

This section shows information about SO posts concerning the tag `egit` during the last 5 years (i.e. from 2011-05-21T22:01:38).

Quick Summary

- There is a total of 1013 questions and 2438 posts related to `egit`.
- Posts analysed span from 2011-05-21 to 2016-05-19.
- Last activity registered was on 2016-05-19 15:31.
- 712 of them (70%) have an accepted answered.
- 175 actually have answers but none of them is accepted.
- 126 have no answer at all.
- There is an average of 1.41 answers to questions.

The wordcloud on the right details frequent words used in the question titles. The size and colour of the words depend on their number of appearances in questions. It shows what are the concerns and issues of people when dealing with the project.

Recent questions waiting for an answer

The following recent questions (less than 1 month old) have been asked more than 2 days ago and still have no answer. Would you like to check them?

Question	Created on	Score	Views
Eclipse cdt Egit commit run configuration	2016-05-05 10:12:57	0	10
Export deployable plug-ins and fragments for projects stored in Git repository from Eclipse Kepler	2016-05-09 20:40:53	0	25
Eclipse egit cannot not detecting changed files or add new file	2016-05-12 16:39:01	0	14
Eclipse Git - Making changes at the same time	2016-05-14 13:33:27	1	21
Commit and Push keeps making new branch (Eclipse)	2016-05-19 01:51:15	0	16



Alambic 3.0

Alambic 3.0 adds many features:

- Architecture / Use Cases
 - More data-centric, positioned as a central data management platform.
 - Address new use cases: DevOps, system engineering.
 - Address new forges: GitHub, Apache..
- Plugins
 - New types of plugins: post and global triggers for reports
 - Wizards to create new projects: Eclipse PMI, GitHub, Apache
 - Tools: easily use external tools (e.g. R for analysis and reporting)

Conclusion & Perspective

Alambic project:

- Home: <http://alambic.castalia.camp>
- Source, issues, documentation: <http://bitbucket.org/BorisBaldassari/alambic>

Check the live instances:

- Polarsys dashboard: <http://dashboard.polarsys.org>
- Eclipse forge dashbaord: <http://eclipse.castalia.camp>

Check the full article! →

