

Divining Purpose: Finding Intelligent Design in Software Evolution

Abram Hindle, Michael W. Godfrey, Richard C. Holt

Software Architecture Group

David R. Cheriton School of Computer Science

University of Waterloo

Canada

<http://swag.uwaterloo.ca/>

{ahindle,migod,holt}@cs.uwaterloo.ca

Purpose

- Our goal is to semi-automatically
 - label commits by aspects of their purpose
 - label developer topics by non functional requirements
 - recover and infer the underlying software development process

What is this commit about?

A yellow rectangular sticky note with a folded top-left corner and a black border. In the center of the note is a large, semi-transparent orange warning triangle. The text is centered on the note.

**Added a test for bug
#1326 on OSX**

What is this commit about?



**Added a test for bug
#1326 on OSX**

What is this commit about?

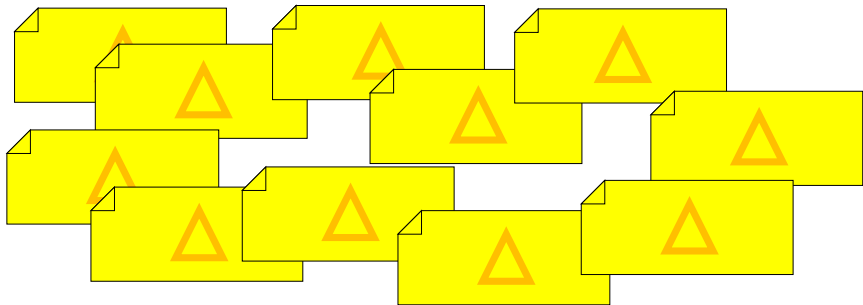
**Added a test for bug
#1326 on OSX**

Corrective

Testing

Portability

But we have many commits..



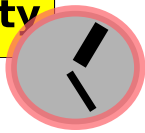
Corrective



Testing



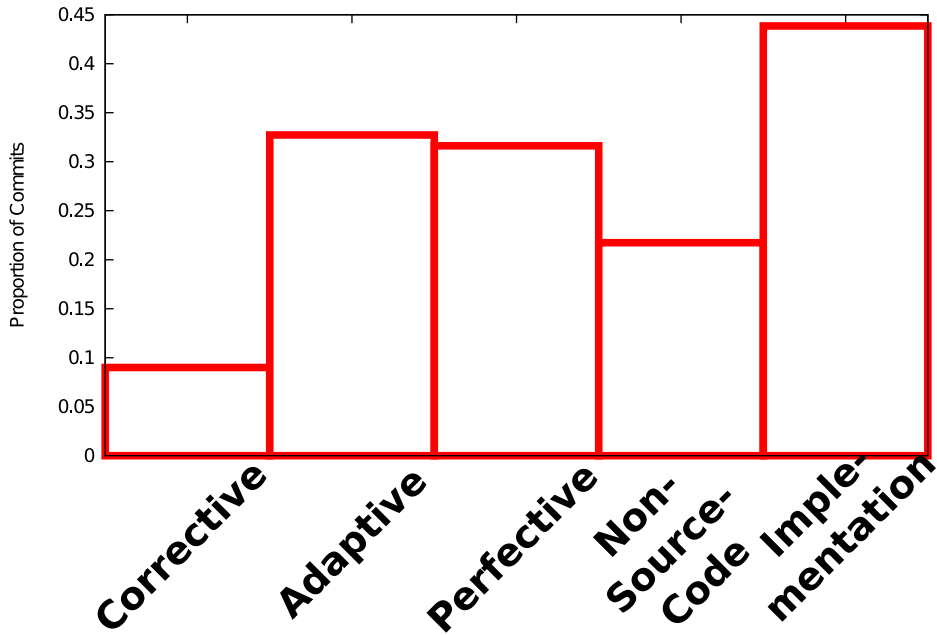
Portability



Purpose

- The immediate issue
- Change in quality
- Kind of maintenance
- Often there's more than one purpose

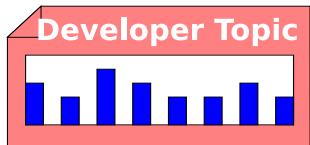
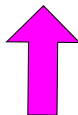
We could label commits by their Swanson Maintenance Classification



Automatic Maintenance Category Classification

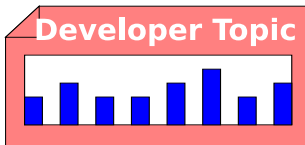
- B to A level results (AUC/ROC of $\sim .7$ to $.8$)
- Suffers from lack of multilabels
- Changes sometimes are broad with multiple purposes

Developer Topics

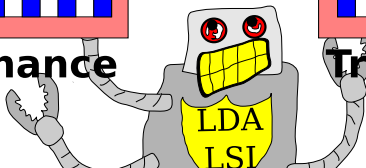


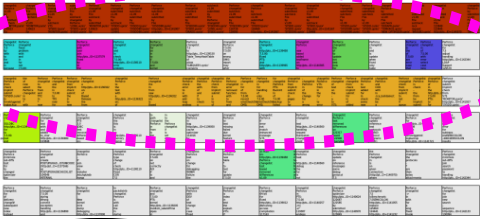
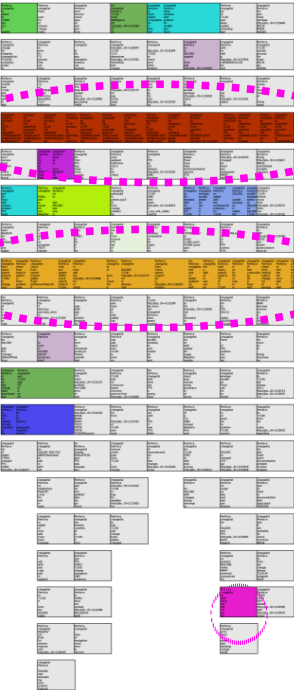
Maintenance

purpose?

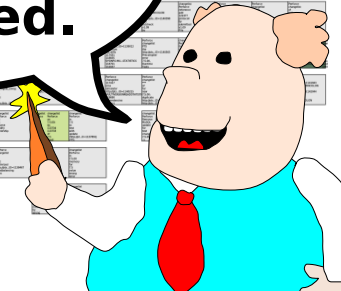


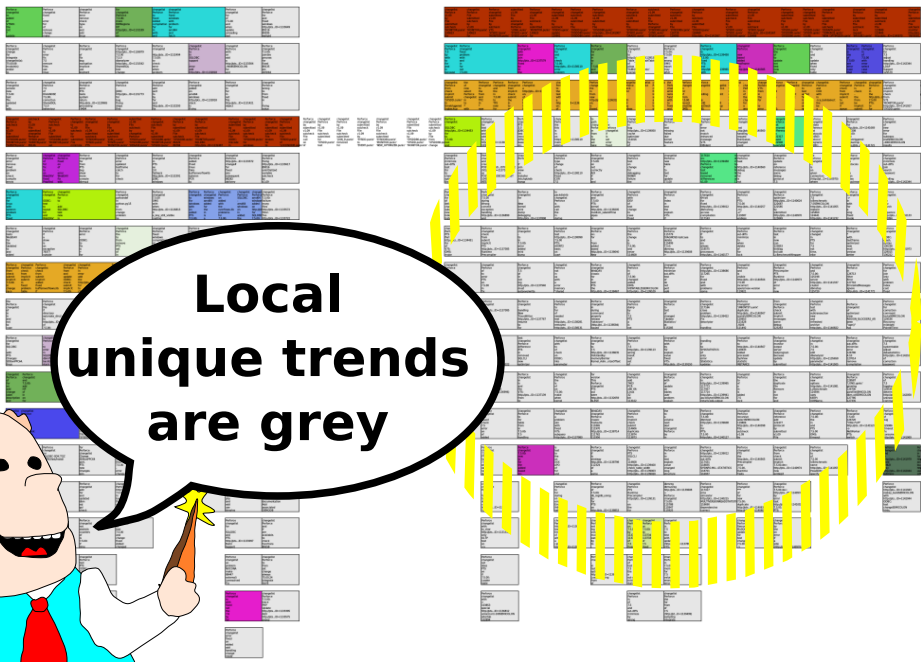
Transactions



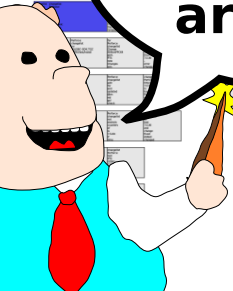


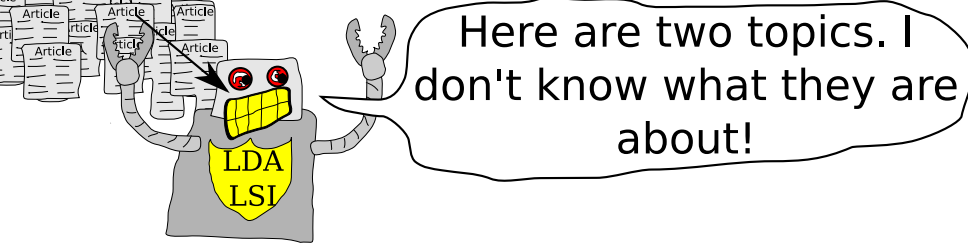
Repeating trends are colored.





**Local
unique trends
are grey**



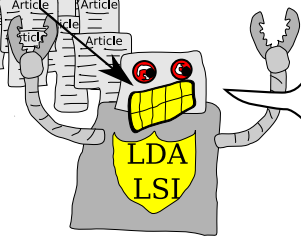


Topic 1

- * win32
- * backport
- * fix
- * bug
- * ...

Topic 2

- * version
- * bump
- * number
- * up
- * ...



Here are two topics. I don't know what they are about!

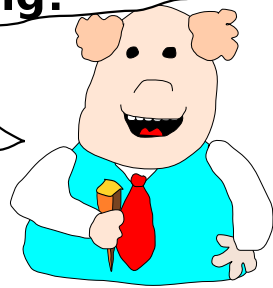
These word lists might be: **Windows backport** and **versioning!**

Topic 1

- * win32
- * backport
- * fix
- * bug
- * ...

Topic 2

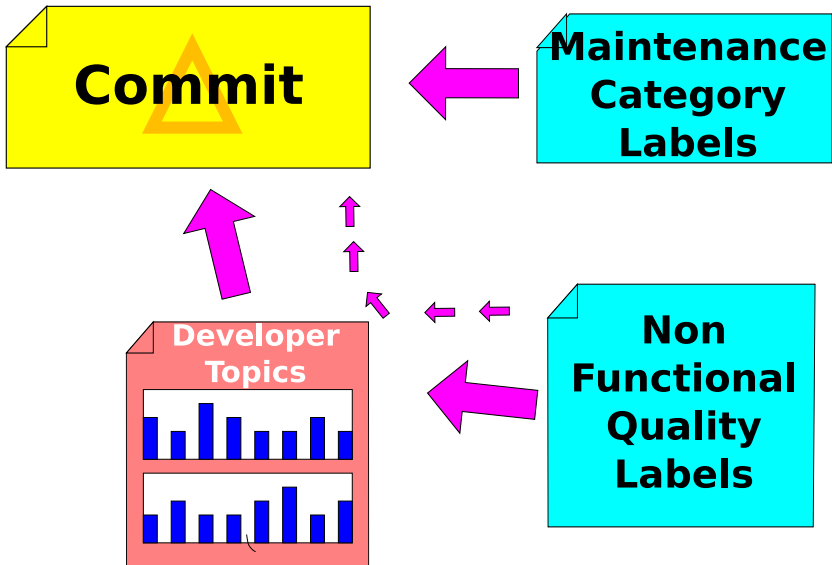
- * version
- * bump
- * number
- * up
- * ...



Label topics by quality

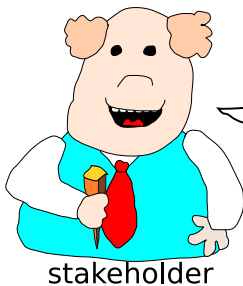
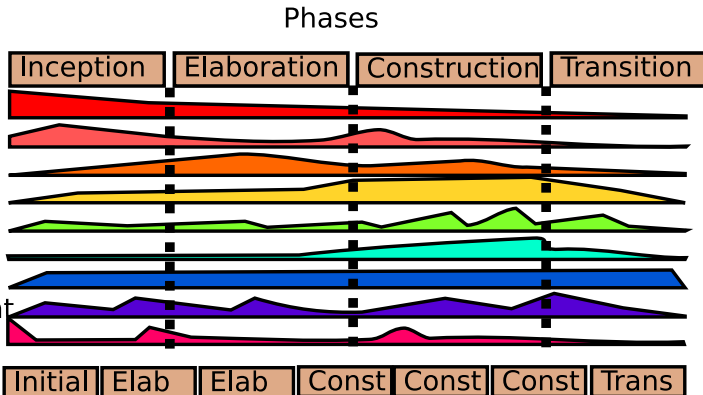
- Non functional requirements
 - Maintainability
 - Functionality
 - Portability
 - Efficiency
 - Usability
 - Reliability

Automatic Labelling



Disciplines

Business Modeling
Requirements
Analysis & Design
Implementation
Test
Deployment
CM and SCS
Project Mangement
Environment



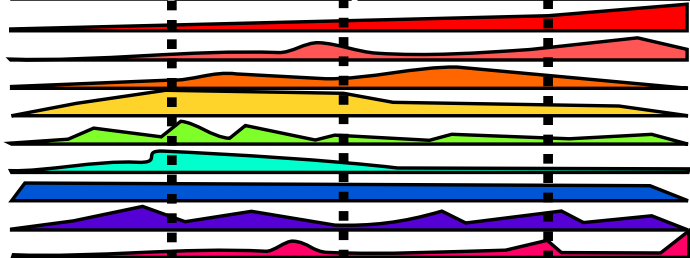
This **Rational Unified Process** diagram shows different disciplines are used at different times.

Labels

Maintainability
Functionality
Efficiency
Usability
Reliability
Portability
Corrective
Adaptive
Prefrective

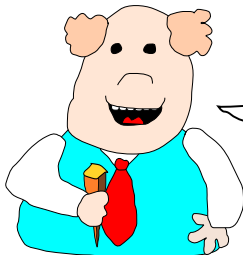
Phases

Inception Elaboration Construction Transition



Initial | Elab | Elab | Const | Const | Const | Trans

Iterations



stakeholder

We could also visualize non-functional requirements or maintenance categories

**Business
Modelling**

Requirements

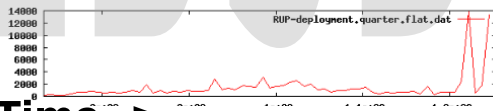
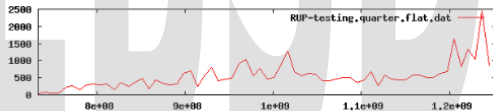
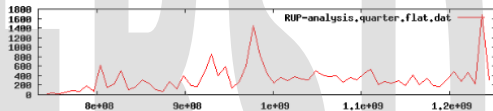
Analysis/Design

Implementation

Testing

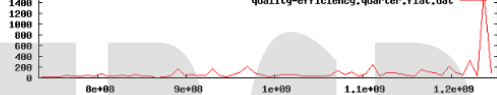
Deployment

Events ->

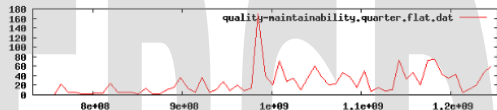


Time ->

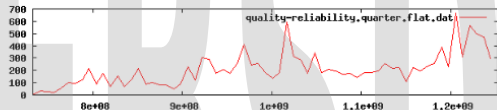
Efficiency



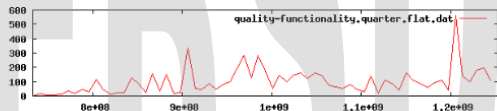
Maintainability



Reliability



Functionality

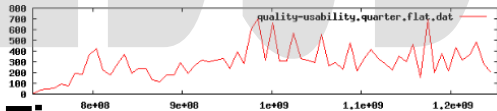


Portability

Events ->

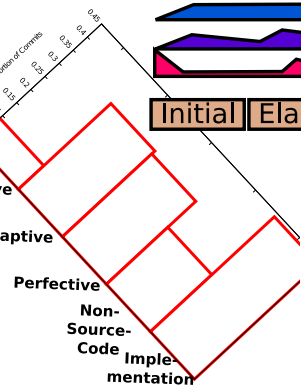
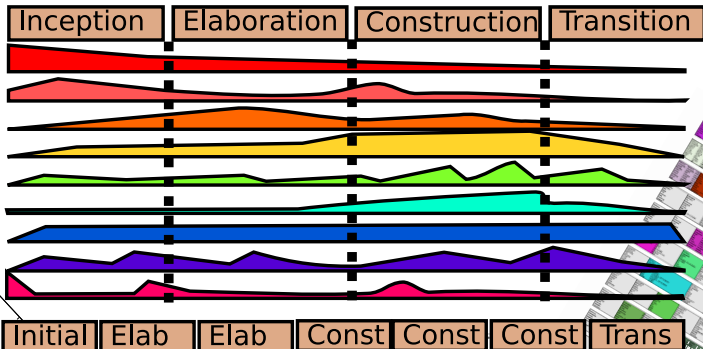


Usability



Time ->

Conclusion



Non-Source-Code Implementation

RUP-Requirements.quarter.flat.dat

RUP-analysis.quarter.flat.dat

RUP-implementation.quarter.flat.dat

RUP-testing.quarter.flat.dat

RUP-deployment.quarter.flat.dat

RUP-Requirements.quarter.flat.dat

RUP-analysis.quarter.flat.dat

RUP-implementation.quarter.flat.dat

RUP-testing.quarter.flat.dat

RUP-deployment.quarter.flat.dat

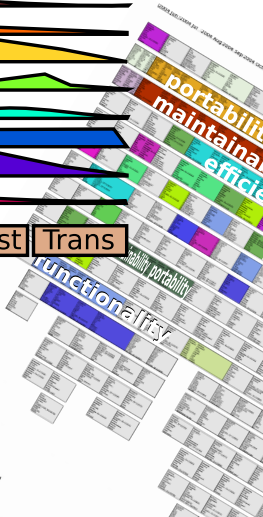
RUP-Requirements.quarter.flat.dat

RUP-analysis.quarter.flat.dat

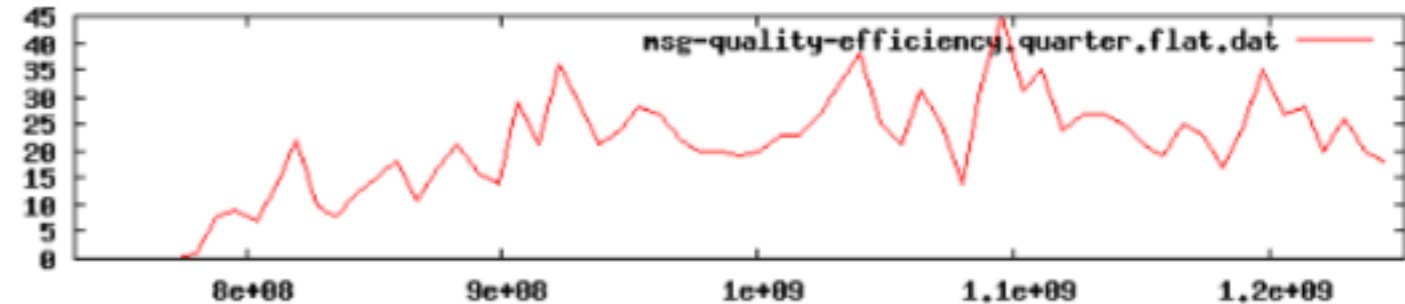
RUP-implementation.quarter.flat.dat

RUP-testing.quarter.flat.dat

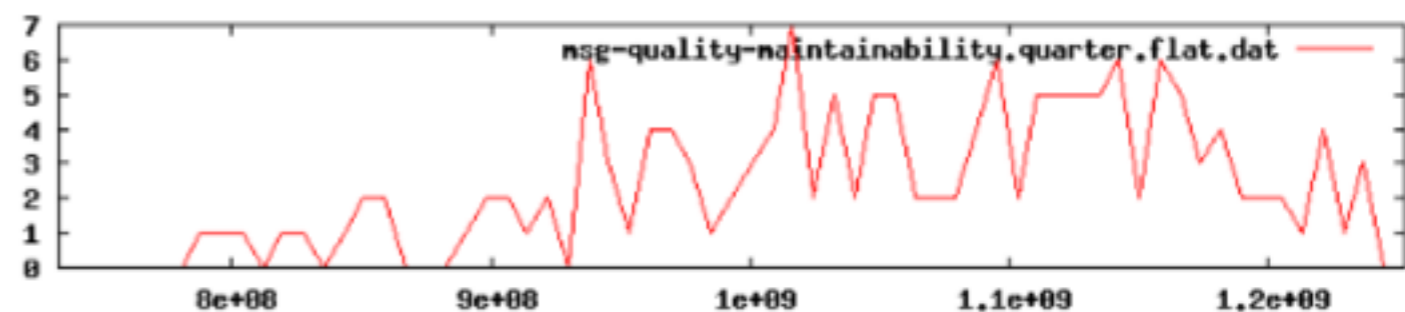
RUP-deployment.quarter.flat.dat



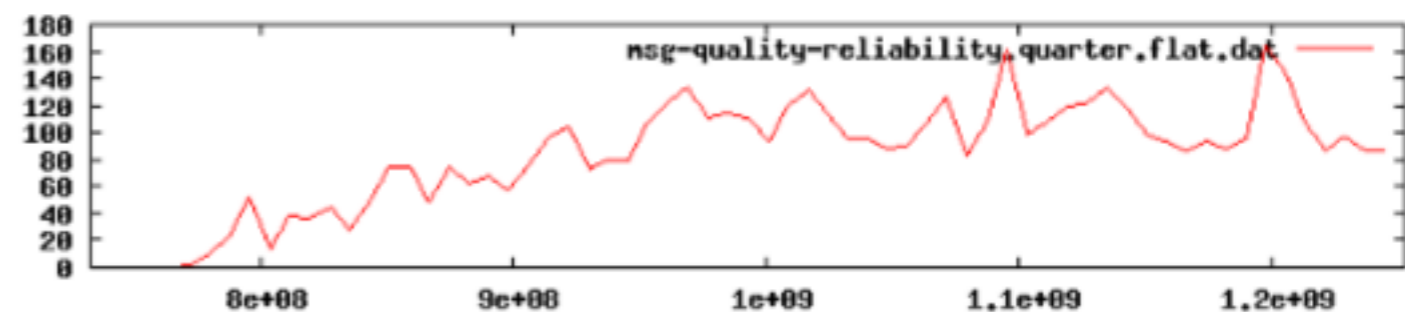
msg-quality-efficiency.quarter.flat.dat



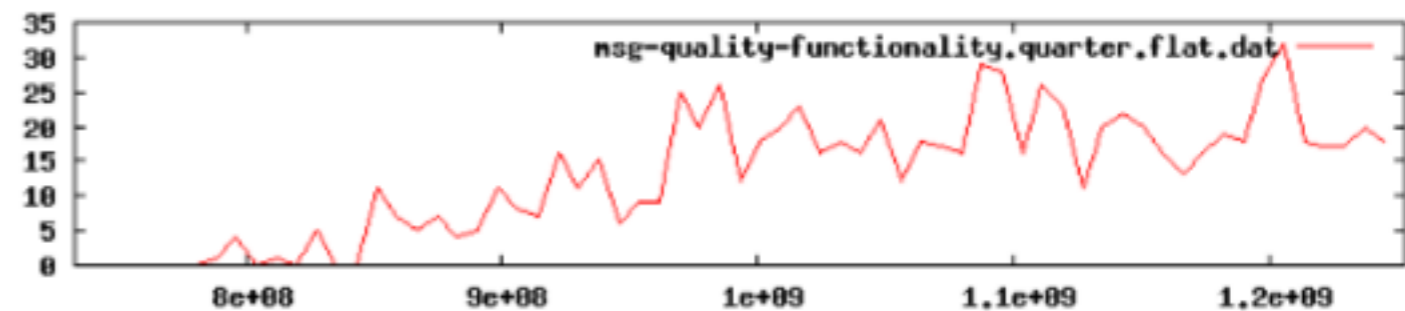
msg-quality-maintainability.quarter.flat.dat



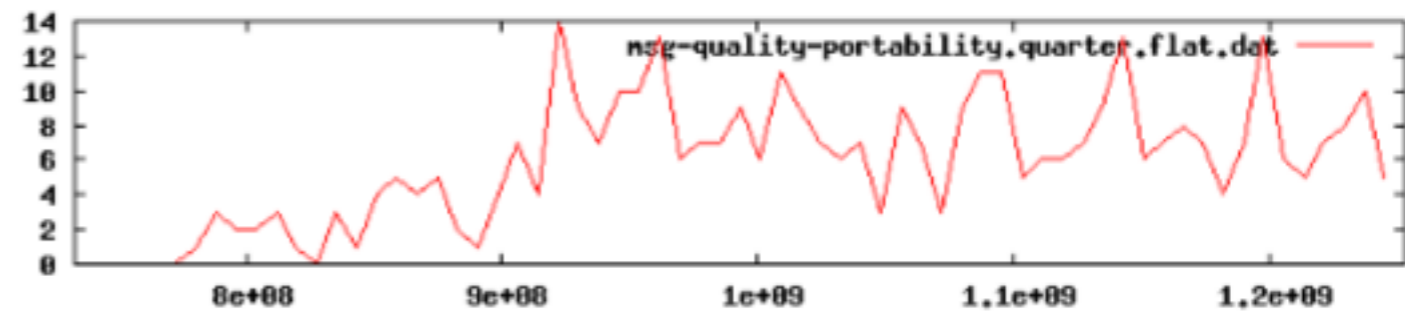
msg-quality-reliability.quarter.flat.dat



msg-quality-functionality.quarter.flat.dat



msg-quality-portability.quarter.flat.dat



msg-quality-usability.quarter.flat.dat

