

How to improve the quality of your application

A practical guide for developers

Ioannis Kolaxis – Software Engineer / Senior Expert

Friday 1st February 2019, Java Meetup Thessaloniki / Greece

Our application: CMP

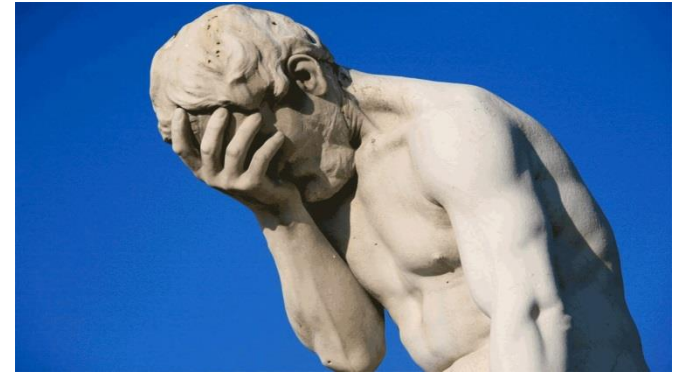
- Automates the configuration & provisioning of our switches, achieving significant time savings for our service.

The screenshot displays the UNIFY Common Management Platform (CMP) interface. The top navigation bar includes tabs for Configuration, Maintenance, User Management (active), Fault Management, Performance Management, and Accounting. The right side of the header shows the user as administrator@system and various status indicators (4 red, 138 orange, 17 yellow). The left sidebar lists categories: Administration (Users & Resources, Localities, User Templates, Extension Ranges) and Provisioning (CSV Import, Users from HPUM, LDAP, Users from UC Accounts, Users from OSV Subscribers). The main content area is titled 'Users & Resources' and contains a search bar with filters for Last Name, All Localities, and Any Resource. Below the search bar is a table of users with columns for selection, last name, first name, bus phone 1, resources, user template, and status. The table lists six users: Desarti, Fotopoulos, Karakatselos, Kolaxis, Nanouris, and Pegiou.

	Last name	First name	Bus. Phone 1	Resources	User Template	Status
<input type="checkbox"/>	Desarti	Athina	+30 (210) 8189-613		osvUserTemplate	✓
<input type="checkbox"/>	Fotopoulos	Dimitrios	+30 (210) 8189-140			●
<input type="checkbox"/>	Karakatselos	Konstantinos	+30 (210) 8189-847		osvUserTemplate	✓
<input type="checkbox"/>	Kolaxis	Ioannis	+302108189858		osvUserTemplate	✓
<input type="checkbox"/>	Nanouris	Ioannis	+498970071230			●
<input type="checkbox"/>	Pegiou	Vasiliki	+30 (210) 8189-793		osvUserTemplate	✓

Software quality issues

- Are you working for a software product, where ...?
 - *New features* take **too much time** to be implemented
 - *Customers* keep complaining about **bugs**



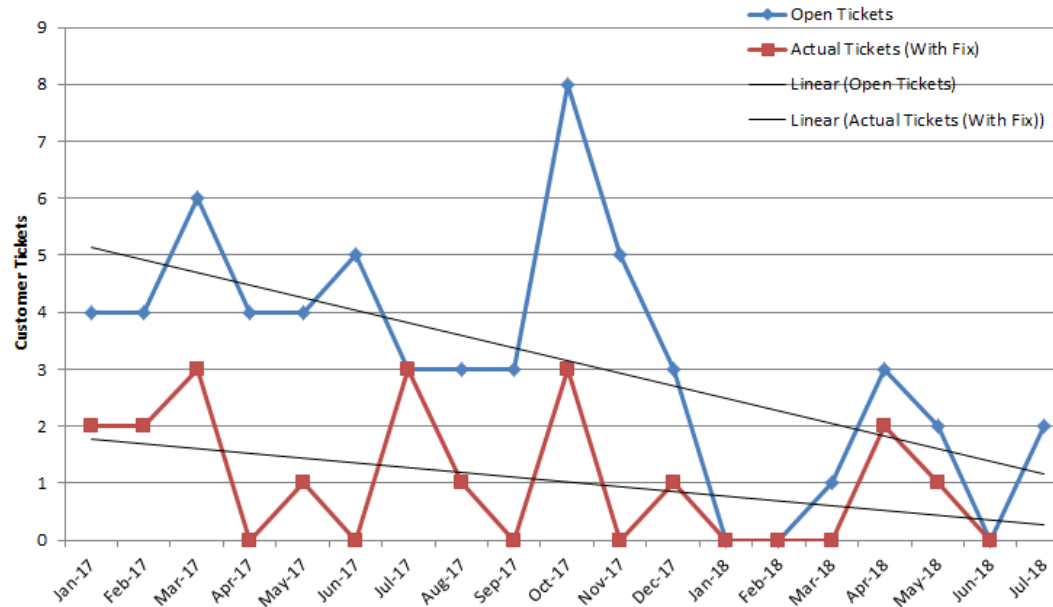
What can you do?

- Can you improve the quality of your software?
- How?



Customer tickets ✓

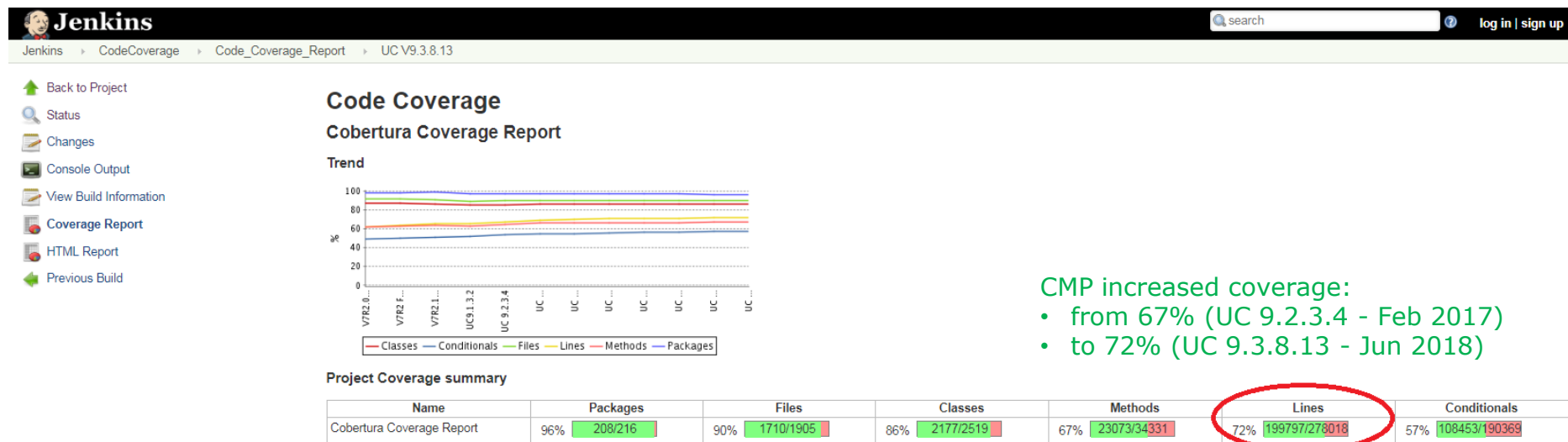
- We usually measure *quality* indirectly, via customer tickets:



CMP Customer tickets

Code coverage ✓

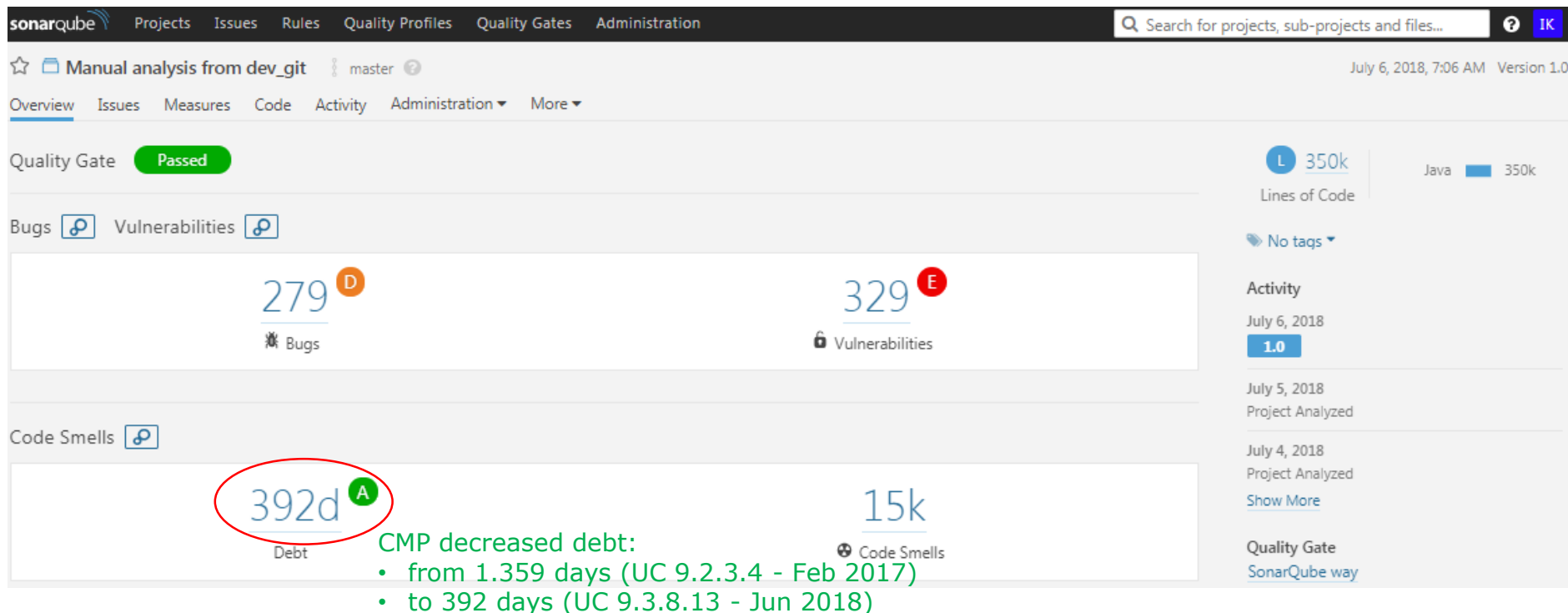
- When we refer to *quality*, we usually think of code coverage!



CMP increased coverage:

- from 67% (UC 9.2.3.4 - Feb 2017)
- to 72% (UC 9.3.8.13 - Jun 2018)

Can/Should you pay off your debt?



Older code is more reliable!

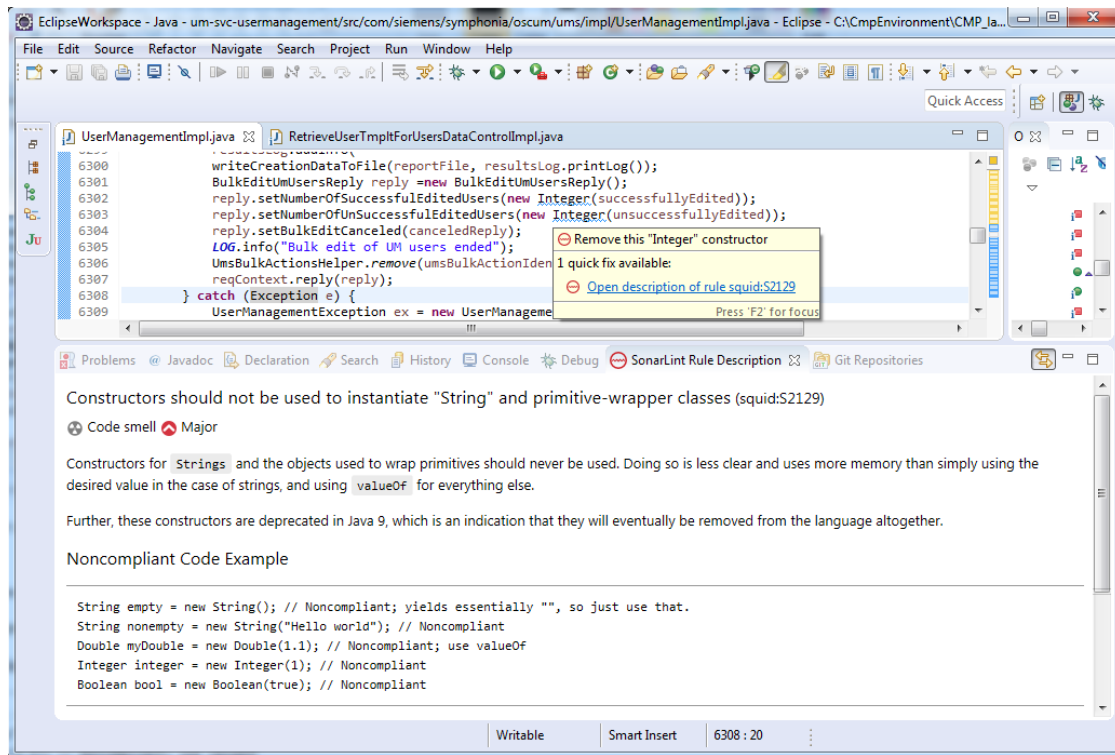
- Do not touch older code, in order to fix Sonar* findings; You are likely to introduce new defects!

"If a module is, on the average, a year older than an otherwise similar module, the older module will have roughly a third fewer faults."

T. L. Graves, A. F. Karr, J. S. Marron and H. Siy, "Predicting fault incidence using software change history" in *IEEE Transactions on Software Engineering*, vol. 26, no. 7, pp. 653-661, Jul 2000.

Stop creating new debt

- Install SonarLint plugin in your IDE.
- It helps you detect, and fix quality issues as you write code.
- Download at:
www.sonarlint.org



Stop creating new debt

- Setup Quality Gates in SonarQube

The screenshot shows the SonarQube web interface. The top navigation bar includes 'sonarqube', 'Projects', 'Issues', 'Rules', 'Quality Profiles', 'Quality Gates' (selected), and 'Administration'. A search bar on the right says 'Search for projects, sub-projects and files...'. On the left sidebar, 'Quality Gates' is selected with a 'Create' button, and 'SonarQube way' is highlighted. The main content area is titled 'SonarQube way' and has buttons for 'Rename', 'Copy', 'Set as Default', and 'Delete'. Below this is the 'Conditions' section, which states: 'Only project measures are checked against thresholds. Sub-projects, directories and files are ignored. [More](#)'. A table lists five conditions:

Metric	Over Leak Period	Operator	Warning	Error		
New Blocker Issues	Always	is greater than ▼	<input type="text"/>	0	<input type="button" value="Update"/>	<input type="button" value="Delete"/>
New Bugs	Always	is greater than ▼	<input type="text"/>	0	<input type="button" value="Update"/>	<input type="button" value="Delete"/>
New Critical Issues	Always	is greater than ▼	<input type="text"/>	0	<input type="button" value="Update"/>	<input type="button" value="Delete"/>
New Major Issues	Always	is greater than ▼	<input type="text"/>	0	<input type="button" value="Update"/>	<input type="button" value="Delete"/>
New Vulnerabilities	Always	is greater than ▼	<input type="text"/>	0	<input type="button" value="Update"/>	<input type="button" value="Delete"/>

Just lie back and think ...

- As a *developer*, where do you spend most of your time?
 - A. Reading existing code,
 - B. Writing new code,
 - C. Waiting for a full build to complete,
 - D. Other activities
- In which parts of your codebase are you spending most of your time?

Just lie back and think ...

- As a *developer*, where do you spend most of your time?
 - A. Reading existing code,
 - B. Writing new code,
 - C. Waiting for a full build to complete,
 - D. Other activities
- In which parts of your codebase are you spending most of your time?

Data never lies

- Use **git** to find out where you spend most of your development efforts:

```
git log --format=format: --name-only | egrep -v '^$' | sort | uniq -c | sort -r >  
files_change_frequency.txt
```

Commits
per file



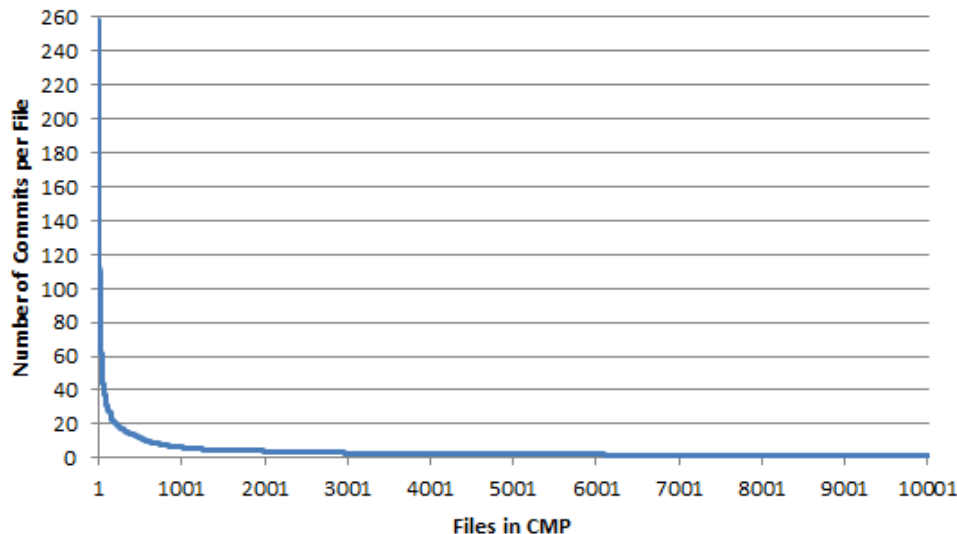
258	usermanagementportlet/.../UserManagement_de.properties
250	usermanagementportlet/.../UserManagement_en.properties
227	usermanagement/.../RetrieveUserTpltForUsersDataControlImpl.java
205	usermanagement/.../UserManagementImpl.java
154	usermanagement/.../EditUserResourceTemplateRulesBean.java
135	usermanagementportlet/.../AddEditUserBean.java
109	usermanagementportlet/.../ConfigureNewUserResourceBean.java
103	usermanagementportlet/.../addEditUser.jsp

Focus on files that change more often

- Only a few files change frequently.
- This is where you spend most of your time!
- Refactor those files, and you will become *more productive* by spending **less time** to:
 - Read existing code,
 - Extend functionality.

CMP consists of 10.007 files:

- 11 files → more than 101 commits
- 91 files → $31 < \text{commits} < 100$
- 455 files → $10 < \text{commits} < 30$
- 9.450 files → less than 10 commits



Changing files predict system failures

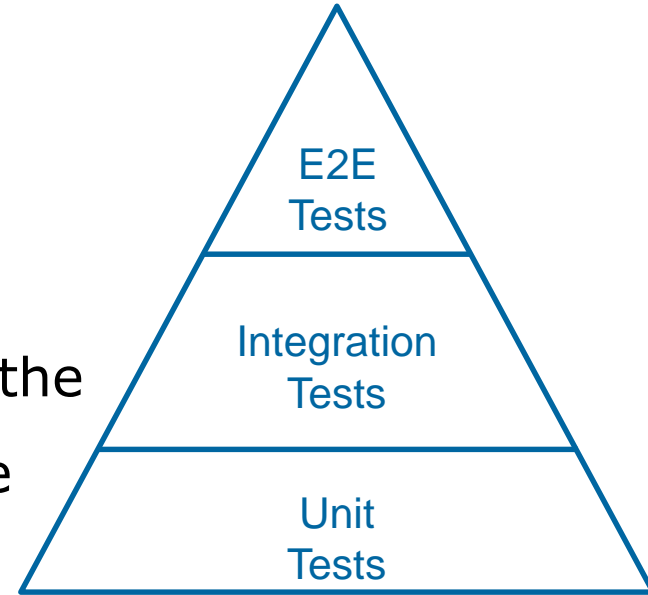
- *"Churn measures based on counts of lines added, deleted, and modified are very effective for fault prediction."*
- *Files involved in a lot of bug fixing activities are most likely to be defective*

R. M. Bell, T. J. Ostrand, and E.J. Weyuker, "Does Measuring Code Change Improve Fault Prediction?", ACM Press, 2011.

R. Moser, W. Pedrycz, and G. Succi, "A Comparative Analysis of the Efficiency of Change Metrics and Static Code Attributes for Defect Prediction", Proceedings of the 30th International Conference on Software Engineering, 181-190, 2008.

Focus your Quality Assurance efforts

- Do not waste your time testing *mature* functionality (=components that do not change).
- Instead, focus all your testing efforts on the parts that are changing frequently; those parts are the ones most likely to fail!
- What is the coverage of your new/changing code?



Identify stable components

- Files not changed in the past years → *stable components* → ***mature features***
- Are those *mature features* still being used by our customers?
 - If they are **not** used, then **delete** this code!
 - Else, **extract** stable features in separate libraries (e.g. JAR).
 - Save time from your builds (Local ones & Jenkins/CI).
 - New developers will *not* have to familiarize with old/stable code.
→ Achieve better onboarding, by focusing only on actively developed code.

Measure code complexity

- Gain more insight about frequently changing files, by measuring *code complexity* for each of those files.
- How many times did you provide a bug fix, by adding a nested conditional in your code?

```
if (...) {  
  for (...) {  
→ if (customerSpecificSetup) {  
    // Do some magic, so that our application  
    // will work for this customer setup!  
  }  
}  
}
```

Measure tabs for code complexity

- Number of tabs = a *language-neutral* metric of code complexity

```
python git_complexity_trend.py --start 729da6c --end 3132260 --file  
managementportal/.../UserManagementImpl.java
```

```
rev,n,total,mean,sd
```

```
95eb233,2859,8378,2.93,1.8
```

```
6a9c070,2938,8551,2.91,1.79
```

```
1b84371,3491,10699,3.06,1.91
```

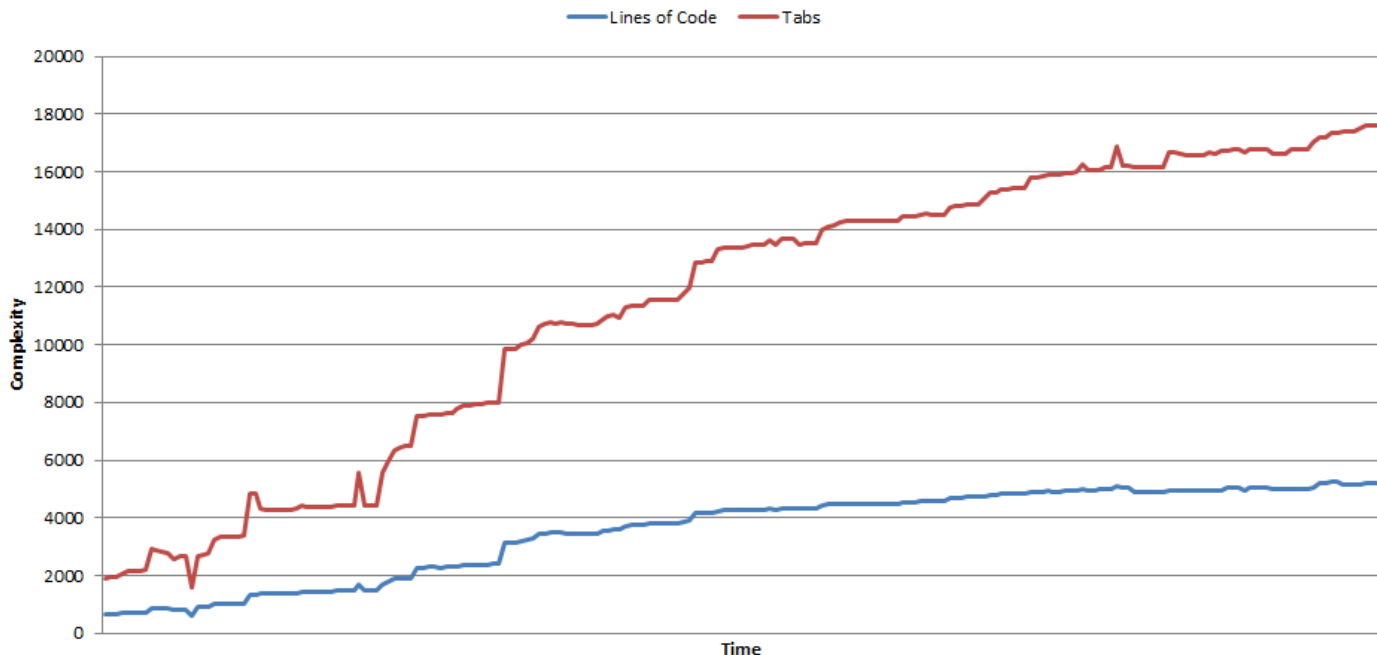
```
1613426,3829,11783,3.08,1.93
```

- The scripts that calculate the complexity are available here:

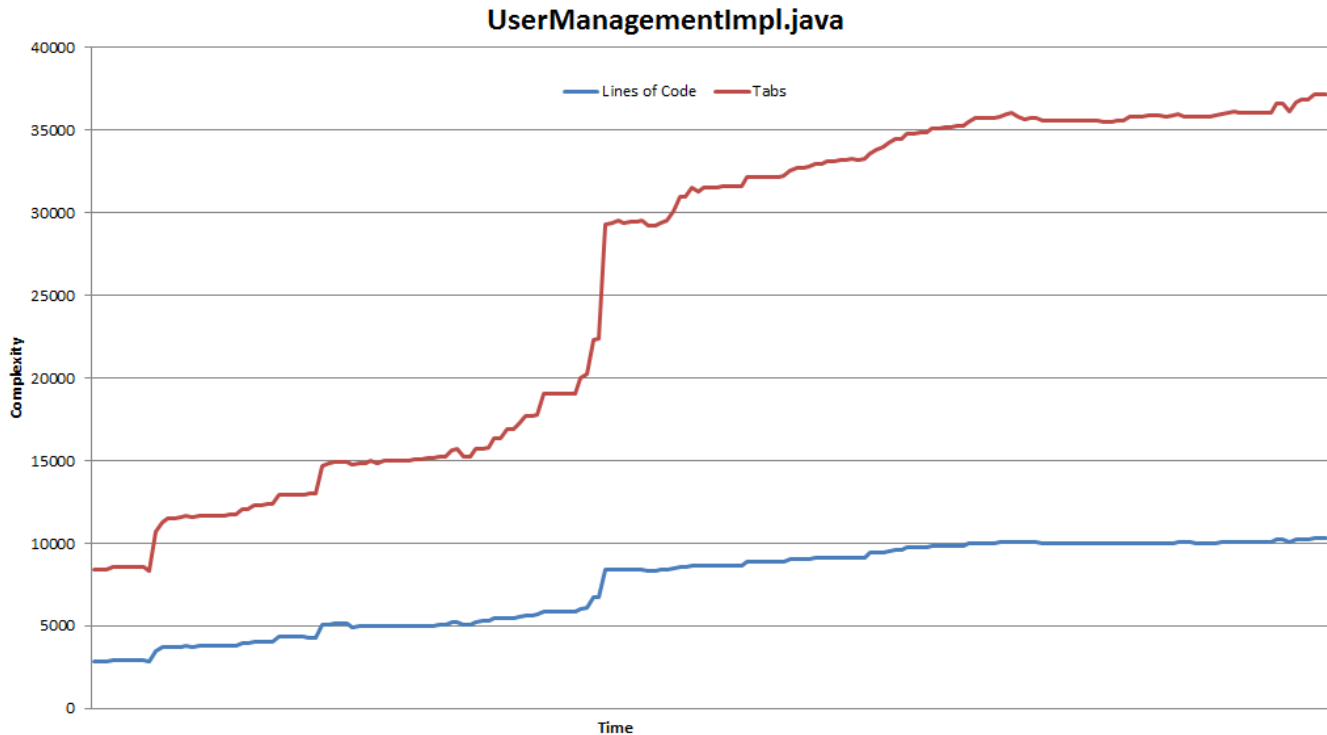
<https://github.com/adamtornhill/maat-scripts/tree/master/miner>

Rising complexity calls for refactoring

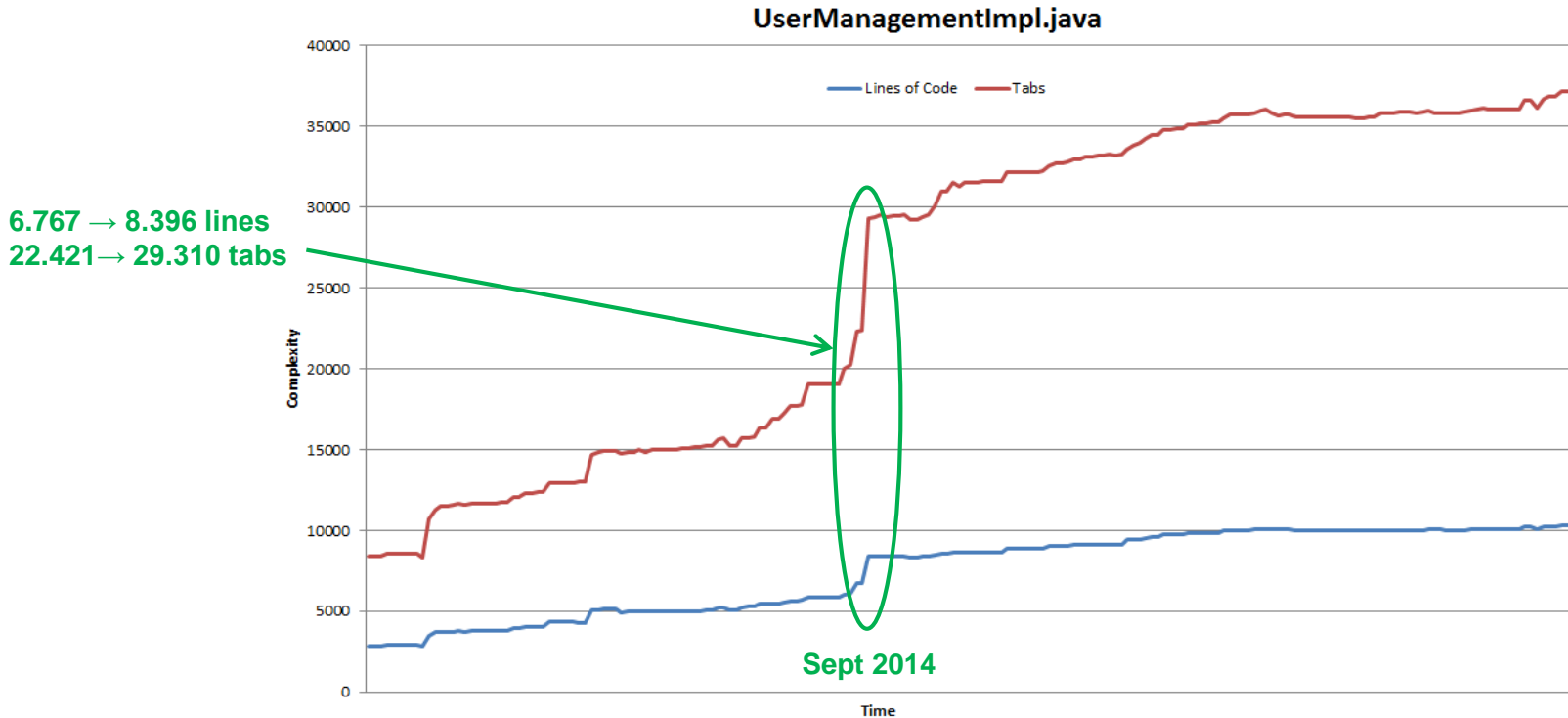
RetrieveUserTpltForUsersDataControlImpl.java



Our #1 priority for refactoring



Our #1 priority for refactoring



Refactor frequently changing files

- The identified files are being changed by many developers in parallel.
- Is it feasible to perform refactoring on a private branch?
- Can we afford to stop development, while someone works for a *long time* on refactoring the identified files?

Break large file by responsibilities

Original file has too many responsibilities

UserManagementImpl.java

```
findUser()
addUser()
editUser()
deleteUser()
```

```
findExtensionRange()
addExtensionRange()
editExtensionRange()
deleteExtensionRange()
```

```
getAssignedPhones()
getUnassignedPhones()
```

....
....

Refactor



Delegate old method calls to new classes

UserManagementImpl.java



UserMgmt.java

```
findUser()
addUser()
editUser()
deleteUser()
```

ExtensionRangeMgmt.java

```
findExtensionRange()
addExtensionRange()
editExtensionRange()
deleteExtensionRange()
```

PhonesMgmt.java

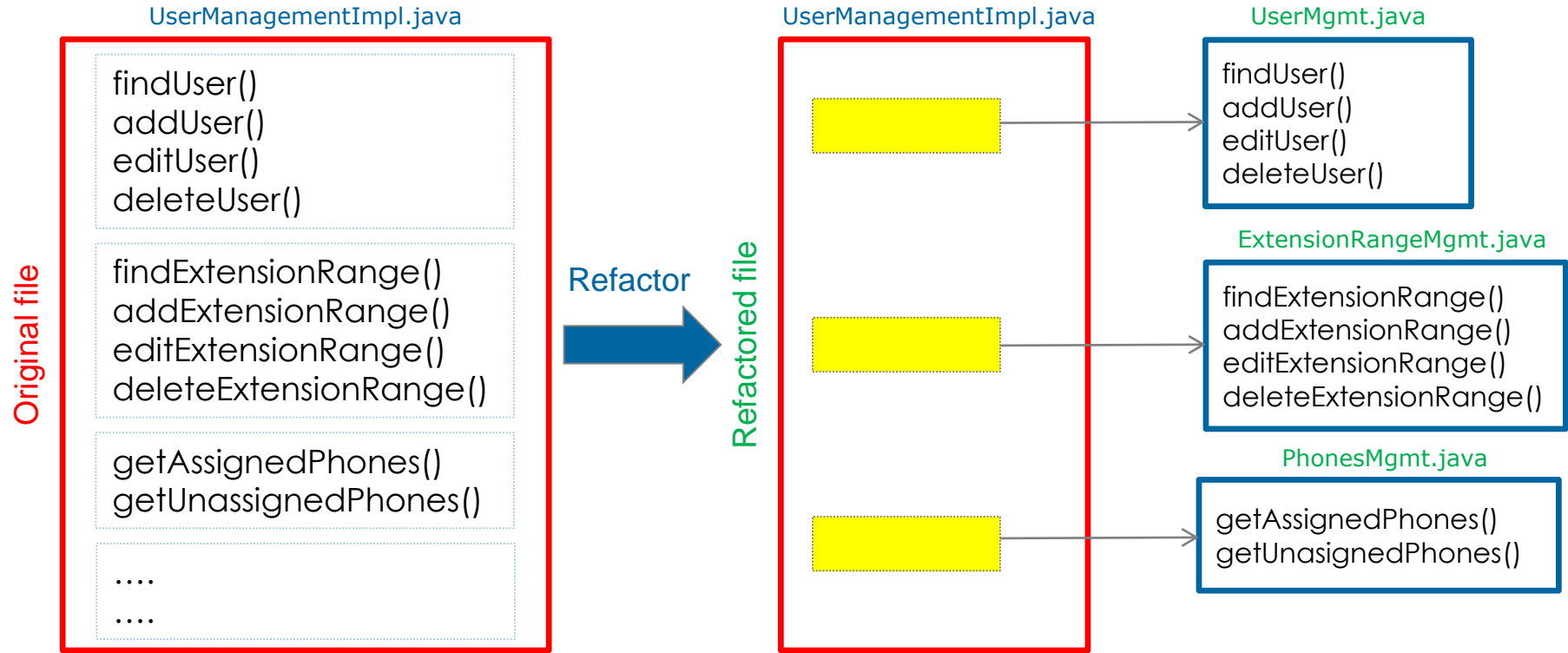
```
getAssignedPhones()
getUnassignedPhones()
```


Divide and conquer



- When you refactor,
always try to stabilize new/changing code!

Stabilizing code by refactoring



Do you remember Windows Vista?

- Released on 8th November 2006.
 - > 50 million lines of code.
 - ~ 2.000 developers.



Organizational structure vs Quality

- Microsoft measured several *organizational metrics*, and studied their correlation with the *defects* of Windows Vista.

Organizational metric	Assertion
Number of Engineers	The more people who touch the code, the lower the quality.
Number of Ex-Engineers	A large loss of team members affects the knowledge retention, and thus quality.
Organization Intersection Factor	The more diffused the different organizations contributing code, the lower is the quality.

- Can the ***structure of your organization*** affect the ***quality*** of your software application?

[N. Nagappan, B. Murphy, and V.R. Basili, "The Influence of Organizational Structure on Software Quality: An Empirical Case Study", ACM, 2008.](#)

Organizational structure impacts Quality

- **Organizational metrics** are better predictors of **failure-proneness** than the traditional metrics used so far, such as *code coverage*, *code complexity*, etc.

Model	Precision
Organizational structure	86,2%
Code coverage	83,8%
Code complexity	79,3%
Code churn	78,6%
Dependencies	74,4%
Pre-release bugs	73,8%

[N. Nagappan, B. Murphy, and V.R. Basili, "The Influence of Organizational Structure on Software Quality: An Empirical Case Study", ACM, 2008.](#)

More organizational metrics

- In another research, focused on Windows 7, Microsoft distinguished between the following kinds of developers, depending on their commits for a given component:
 - **Owner:** has the most commits to that component.
 - **Major contributor:** has *more than 5%* of total commits.
 - **Minor contributor:** has *less than 5%* of total commits.



[C.Bird, N. Nagappan, B. Murphy, H. Gall, and P. Devanbu, "Don't Touch My Code! Examining the Effects of Code Ownership on Software Quality", ACM, 2011.](#)

Effects of minor contributors

- The researchers concluded that:
 - *"The number of **minor contributors** has a strong positive relationship with both pre- and post-release failures ..."*
 - *"Higher levels of **ownership** for the top contributor to a component results in fewer failures when controlling for the same metrics, but the effect is smaller than the number of minor contributors"*

[C. Bird, N. Nagappan, B. Murphy, H. Gall, and P. Devanbu, "Don't Touch My Code! Examining the Effects of Code Ownership on Software Quality", ACM, 2011.](#)

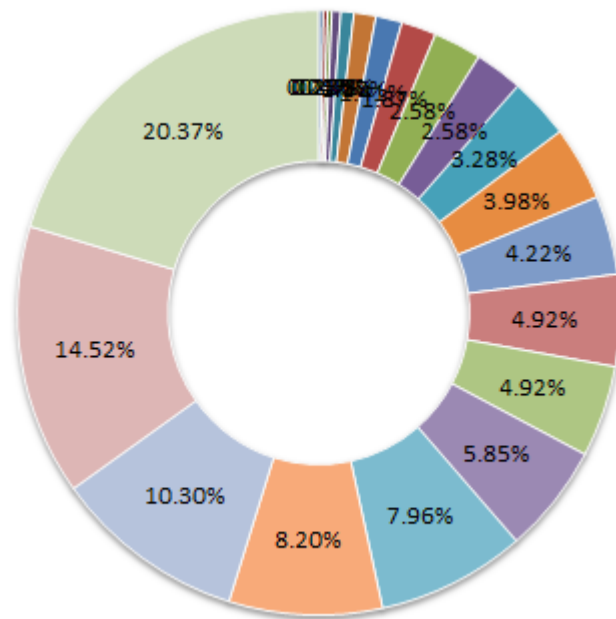
Gain insight into your components

- In one of our software components, we had a total of 427 commits:

- The top contributing developer made 87 commits:
 $87/427 = \mathbf{20,37\%}$ ownership

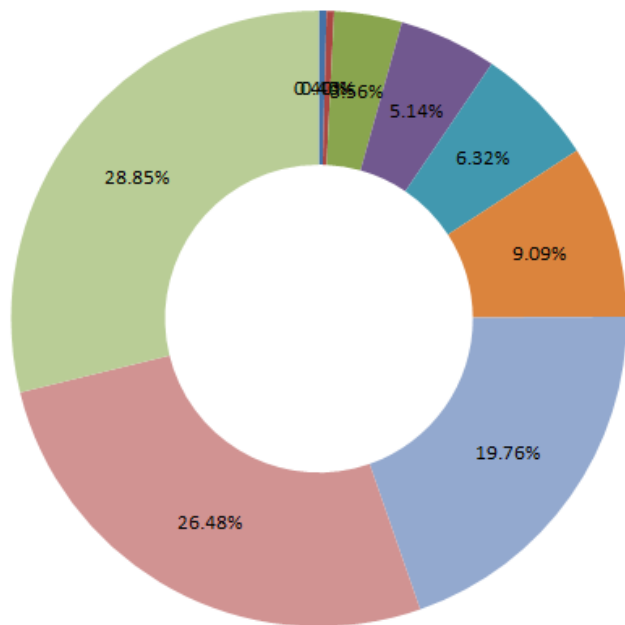
Metric	Value
Minor contributors	15
Major contributors	6
Total contributors	21
Ownership	20,37%

Commits per developer



Gain insight into your components

Commits per developer



- In another software component, we had a total of 253 commits for the same period:

Metric	Value
Minor contributors	3
Major contributors	6
Total contributors	9
Ownership	28,85%

- The top contributing developer made 73 commits:
 $73/253 = \mathbf{28,85\%}$ ownership

Know where you are standing ...

Metric	Component A	Component B
Minor contributors	15	3
Major contributors	6	6
Total contributors	21	9
Ownership	20,37%	28,85%

- Which component will probably have more defects?
- Where would you focus your testing efforts?

Beware of minor contributors!

Metric	Component A	Component B
Minor contributors	15	3
Major contributors	6	6
Total contributors	21	9
Ownership	20,37%	28,85%

- More **minor contributors**
→ More defects
- Bigger **ownership**
→ Less defects

Using data to build better software

- How can those metrics help us achieve higher quality?
 - Minor contributors must be **consulting** a major contributor of a component before making any changes to it.
 - Pay more attention when **reviewing** code submitted by minor contributors.
 - More **extensive testing** should be performed for components with ***low ownership***.

Planning new features

- A customer asks for a **new feature** to be implemented, but the **major contributors** of that component are not available. What will you do?
- Ask from **minor contributors**, to start implementing this new feature right away, or
- Delay the implementation of the feature, until one or more **major contributors** are available?

Learn your contributors

- Use **git** to find out all the contributors for a component (e.g. located under folder *ucaas-usermanagement*):

```
git shortlog -s -- ucaas-usermanagement > contributors.txt
```

Commits per developer	Folder
17	Ioannis Kolaxis
18	...
34	...

- Or, to limit the results to contributors *after a given date* (e.g. due to an organizational restructuring)

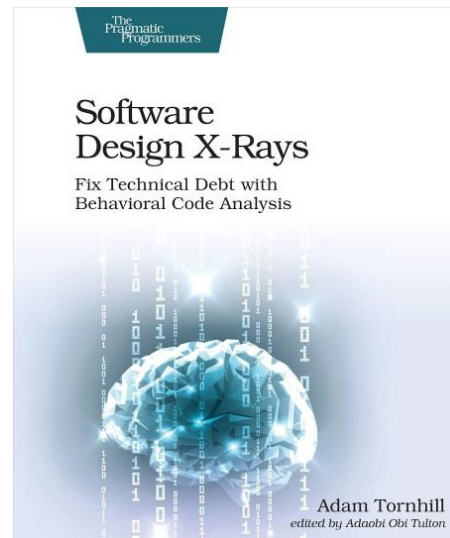
```
git shortlog -s --after=2018-05-01 -- ucaas-usermanagement > contributors.txt
```

Summary of proposed actions

- Stop creating new debt
- Do not touch older code in order to fix Sonar* findings.
- Find the files in your project/product that change more often.
Don't forget to measure their complexity!
- Refactor your most complex, frequently changing files.
- Focus your QA Efforts on the files that change more frequently.
- Find out who are the major contributors in each component.
Pay special attention when a minor contributor changes code.

Further reading

- A. Tornhill, "Software Design X-Rays – Fix Technical Debt with Behavioral Code Analysis", The Pragmatic Programmers, 2017.
- [N. Nagappan, B. Murphy, and V.R. Basili, "The Influence of Organizational Structure on Software Quality: An Empirical Case Study", ACM, 2008.](#)
- [C.Bird, N. Nagappan, B. Murphy, H. Gall, and P. Devanbu, "Don't Touch My Code! Examining the Effects of Code Ownership on Software Quality", ACM, 2011.](#)



Questions?

Email : ioannis.kolaxis@atos.net

Twitter : [@loanniskolaxis](https://twitter.com/loanniskolaxis)

Atos, the Atos logo, Atos Codex, Atos Consulting, Atos Worldgrid, Worldline, BlueKiwi, Bull, Canopy the Open Cloud Company, Unify, Yunano, Zero Email, Zero Email Certified and The Zero Email Company are registered trademarks of the Atos group. June 2016. © 2016 Atos. Confidential information owned by Atos, to be used by the recipient only. This document, or any part of it, may not be reproduced, copied, circulated and/or distributed nor quoted without prior written approval from Atos.

