

Who controls your project? Governance in the real world

Jesus M. Gonzalez-Barahona

jgb@bitergia.com @jgbarah
Bitergia / LibreSoft (URJC)

Legal and Policy Issues DevRoom, FOSDEM
Brussels (Belgium), January 30th 2016



©2016 Bitergia

Some rights reserved. This presentation is distributed under the
“Attribution-ShareAlike 3.0” license, by Creative Commons, available at
<http://creativecommons.org/licenses/by-sa/3.0/>



Structure of the presentation

- 1 A bit of context
- 2 Governance in the real world
- 3 Who is contributing?
- 4 How are changes being reviewed?
- 5 How are they fixing tickets?
- 6 Dependency
- 7 Diversity
- 8 Open development analytics
- 9 Bonus track



A bit of context



Me and my circumstances

Uni Rey Juan Carlos:

- LibreSoft research team
- Understanding free, open source software development
- Data analytics approach

Bitergia:

- From research to the real world
- Understanding free, open source software development
- Data analytics approach

<http://gsyc.es/~jgb> <http://bitergia.com>



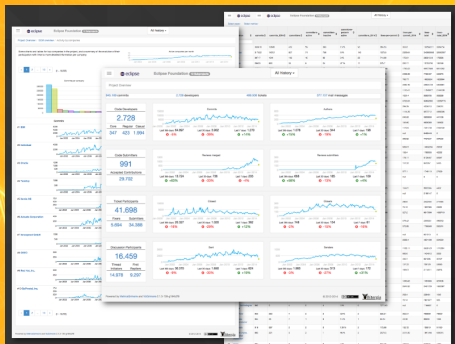
The company



Bitergia

The software development analytics company

- dashboards
- reports
- consultancy
- ...



<http://bitergia.com>

Governance in the real world



“Establishment of policies, and continuous monitoring of their proper implementation, by the members of the governing body of an organization. It includes the mechanisms required to balance the powers of the members (with the associated accountability), and their primary duty of enhancing the prosperity and viability of the organization.”

<http://businessdictionary.com>



*“Establishment of policies, and **continuous monitoring** of their proper implementation, by the members of the governing body of an organization. It includes the mechanisms required to **balance the powers** of the members (with the associated **accountability**), and their primary **duty of enhancing** the prosperity and viability of the organization.”*

<http://businessdictionary.com>



What about “non-formal” governance,
working outside governing bodies?

What about what’s happening
after specific policies
are put in place by governing bodies?



Example

I'm a company investing heavily in a project. I'm hiring developers, supporting the foundation, sponsoring activities...

- Are my developers treated according to the policies?
- Are we getting integrated in the community?
- How do we compare with other companies of similar characteristics?

Are we having reasonable metrics, according to the current stated policies and agreements?



Example

I'm an independent developer, devoting a large fraction of my time to this project.

- Are my initiatives being considered on fair terms?
- Are employees of other companies dealing with me the same way they do with their company colleagues?
- Am I considered based on my merits?

Am I having reasonable metrics,
according to the current stated policies?



Example

I'm a project foundation, interested in being neutral, inclusive, in making life easy to all contributors

- Are newcomers being treated as they should?
- Are we balancing the interests of companies and independent developers?
- Do we have subprojects which are outliers in terms of performance, inclusiveness, etc.
- Are having an impact the policies we put in place?

Do our metrics show our project is
as we intended it to be?



Who is contributing?



The influence of companies

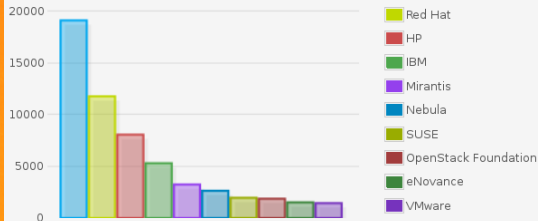
- In many projects, companies are main drivers
- They join forces to push the project...
- ...but they watch each other, look for balances
- They contribute money, resources...
- ...and direct development effort

Having an accurate, transparent picture is very important!

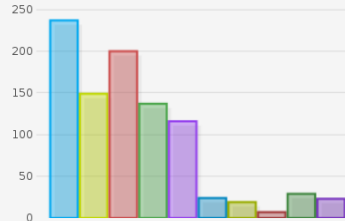


Affiliation

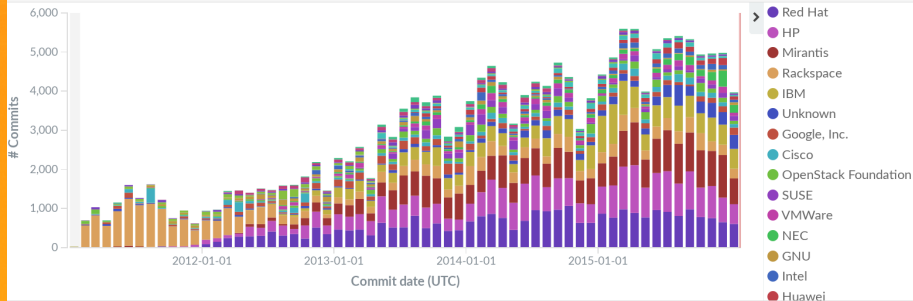
Commits per company (aggregated)



Authors per company (aggregated)



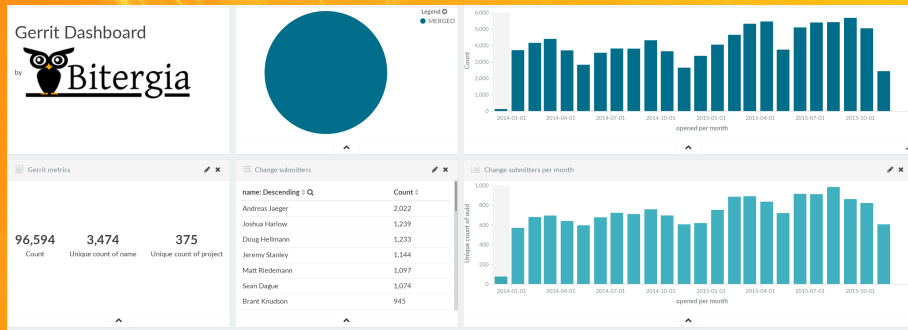
Top organizations, commits over time



are changes being reviewed?



Some reviewers are more equal than others



<http://blog.bitergia.com/2015/12/30/some-developers-are-more-equal-than-others/>



Some reviewers are more equal than others (2)



Time per review



0

50th percentile of timeopen

3

75th percentile of timeopen

17

95th percentile of timeopen

45

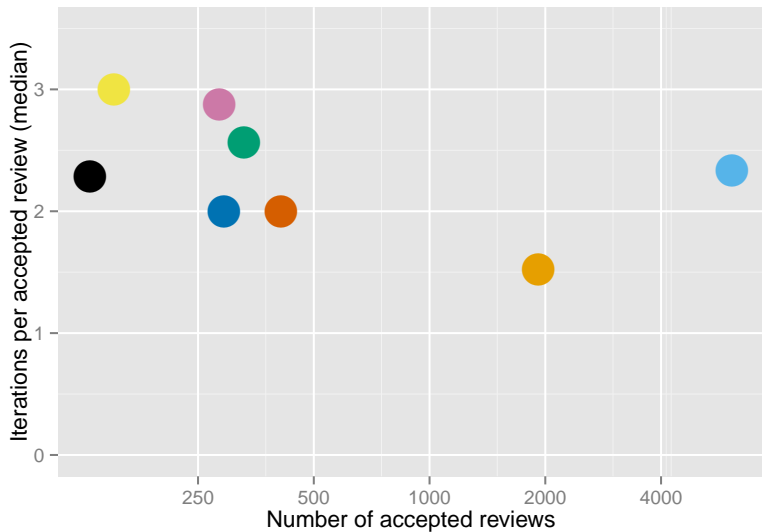
99th percentile of timeopen

80.172%

Percentile rank 4 of "timeopen"



Neutrality?



How are they fixing tickets?

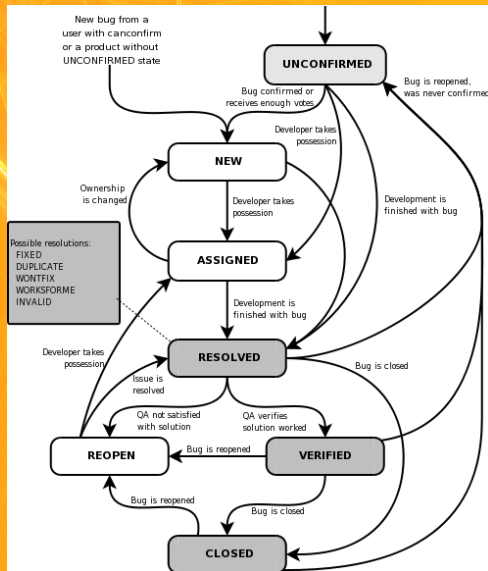


Issues may be dealt with not as intended

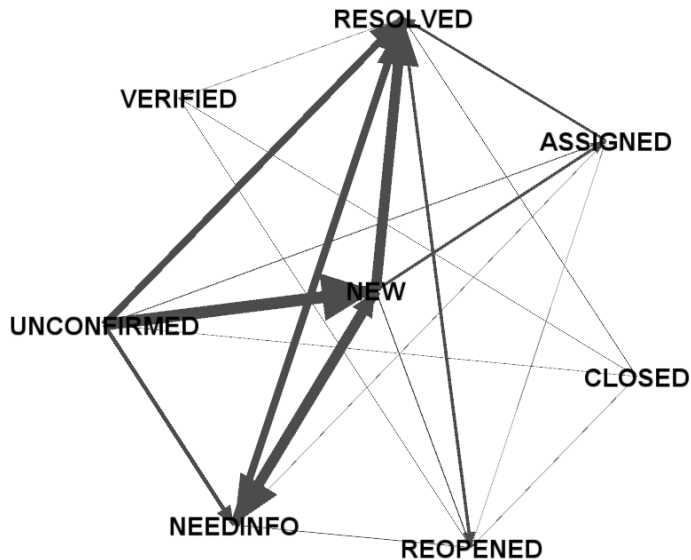
- Policy (or recommendations) may mandate transitions but are they real?
- Time to close when same company reporting / fixing?
- Time to close for external bug reports?
- Time to close depending on who reports?
- Who opens tickets that nobody cares about?



Ej: The “mandated” changes of state



The real changes of state



Dependency



Apache Pony Factor

In words of Daniel Gruno:

We [the ASF] created a term we have coined “Pony Factor” (because ASF is full of ponies, or people who think they are ponies). Pony Factor (PF) shows the diversity of a project in terms of the division of labor among committers in a project.

Pony Factor is determined as:

“The lowest number of committers whose total contribution constitutes the majority of the codebase”

<https://ke4qqq.wordpress.com/2015/02/08/pony-factor-math/>



Bitergia Elephant Factor

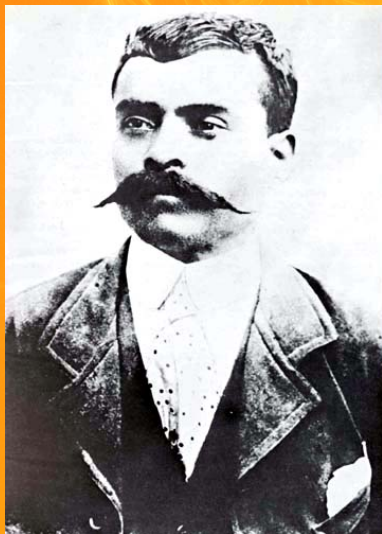
Projects can benefit from powerful collaborations from companies (elephants). The elephant factor shows the diversity of a project in terms of the division of labor among companies (by mean of developers affiliated with them).

Elephant factor is determined as:

“The lowest number of companies whose total contribution (in commits by their employees) constitutes the majority of the commits”



Code “owned”



*“The land belongs
to its workers”*

Emiliano Zapata

[Picture in the public domain]



Code “owned”

The code changes over time. The current version is “owned” by the people who produced it.

The code “belongs” to those who wrote it.

Zapata factor (work in progress):

“The lowest number of developers for whom the total number of lines of code they “own” (were last touched by them) constitutes the majority of the lines of code”



Code “owned”

The code “belongs” to companies who employ developers changing it.

United Fruit factor (work in progress):

“The lowest number of companies for whom the total number of lines of code they “own” (were last touched by their employees) constitutes the majority of the lines of code”



Pony / elephant factors for some projects

	Pony Factor	Elephant Factor	Commits (excl bots)
OpenNebula	4	1	12K
Eucalyptus	5	1	25K
CloudStack	14	1	42K
OpenStack	>100	6	126K
CloudFoundry	41	1	60K
OpenShift	10	1	15K
Docker	15	1	18K
Kubernetes	12	1	7K

[July 2015]



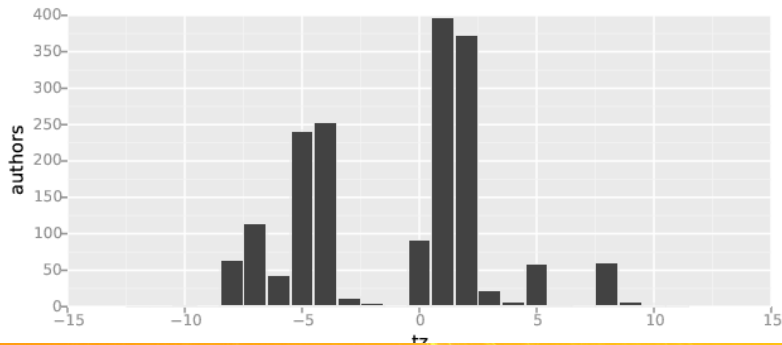


Geography

- Geographical diversity is difficult to assess
- Companies can keep detailed records, but open communities are different
- Fortunately, some tools leave traces...
- This allows for better knowledge
...and better tracking of initiatives
- Example: policies to enlarge the number of developers in XXX region



Geography: time zones in git records



Geography: GitHub profiles



Gender: Analyzing by name

Current situation of gender imbalance in OpenStack

Gender	Developers	Commits	Commits/devel
Female	750	14,647	19.5
Male	4,632	207,112	44.7

Only names with more than 80% of certainty.

[Work in progress, preliminary results]



development analytics



A new dimension of openness

When we develop in the open
we produce a great deal of data
about how we develop

“Show me the development data”
as a step beyond
“show me the code”

Let's do it!



From open development to open development analytics

Information about code, community, development
for open development projects
can be retrieved, organized, analyzed

Let's publish analytics results & data

Open Development Analytics: A new standard for transparency



Who is interested?

- Developers
- Project managers
- Community managers
- Evaluators
- ...



Open development analytics

Why?

Free software produced
with open development models
is more and more important
for IT users, producers, integrators

It is different & complex,
yet transparent,
many details are public,
and it can be improved



Some areas of interest

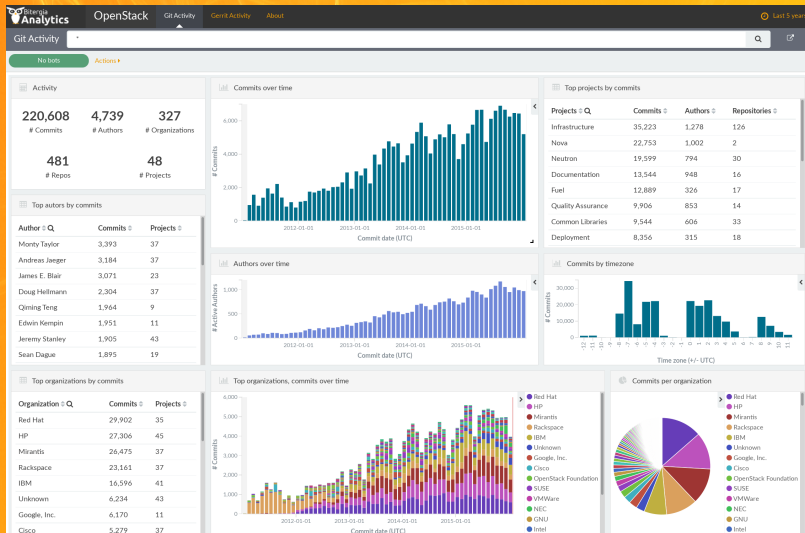
- Performance (understanding activity)
- Company participation (beyond copyright notices)
- Transparency (available information)
- Auditing (certify participation, experience, etc.)
- Profiling (key people, companies)
- Neutrality (fair treatment)



Bonus track



Preview: the new Kibana-based dashboards



<http://s.bitergia.com/db-fosdem16>