

# Learning about software development with Kibana dashboards

Jesus M. Gonzalez-Barahona

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

Lightning Talks, FOSDEM  
Brussels (Belgium), January 31st 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 From the team who built MetricsGrimoire...
- 3 Your very own dashboard for your pet GitHub repo
- 4 Enjoy
- 5 Some details
- 6 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>



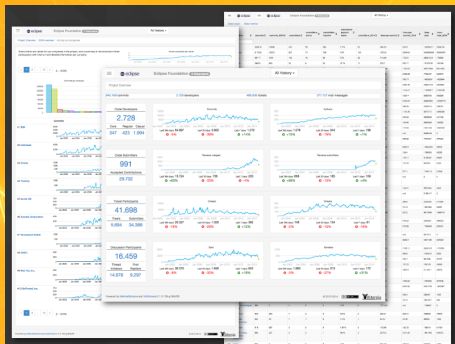
# The company



# Bitergia

## The software development analytics company

- dashboards
- reports
- consultancy
- ...

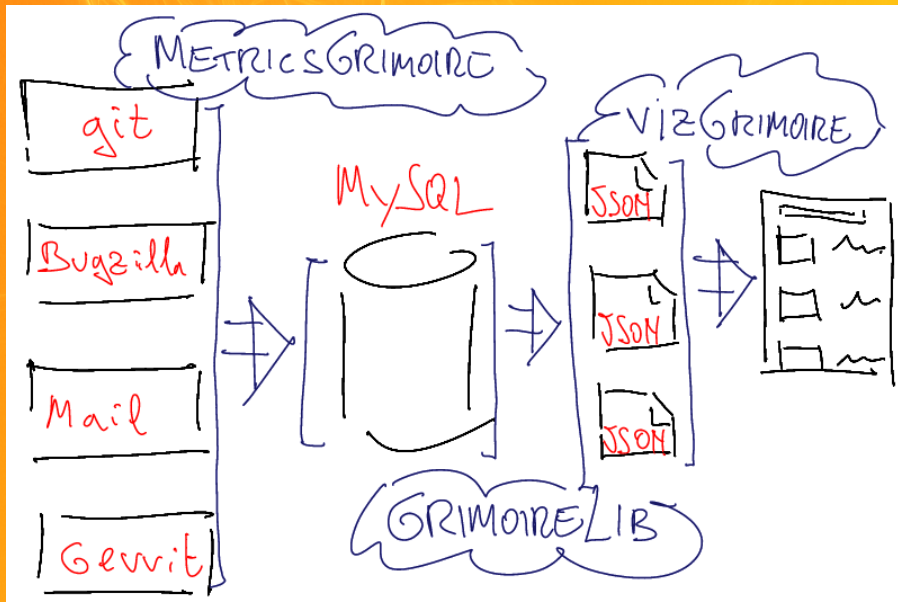


<http://bitergia.com>

# From the team who built MetricsGrimoire...



# MetricsGrimoire: the current state of affairs





# MetricsGrimoire: supported data sources



# GrimoireLab: the next big thing

- Perceval: retrieving data
- ElasticSearch: storing data
- Python / Pandas scripts: enrich, analyze, customize the data
- Kibiter: Kibana fork to interact with the data

<http://grimoirelab.github.io>  
<http://blog.bitergia.com>

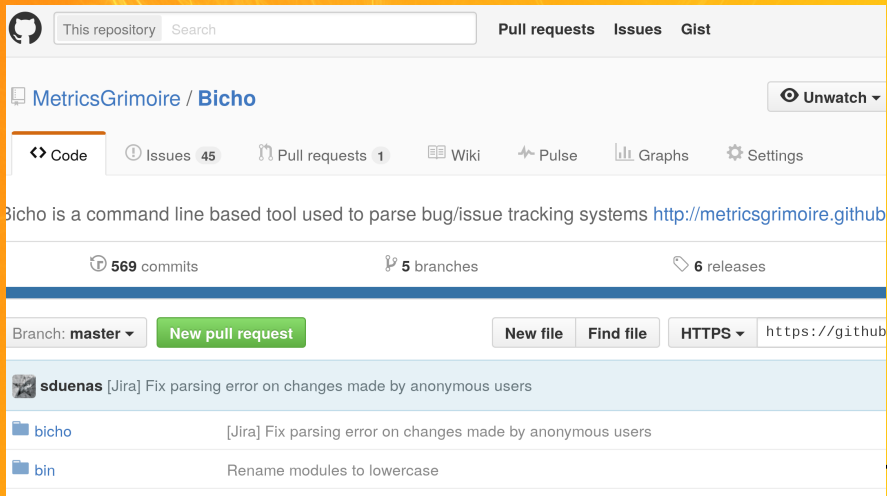


# Your very own dashboard for your pet GitHub repo



# One of my pet repos: MetricsGrimoire/Bicho

<https://github.com/MetricsGrimoire/Bicho>



The screenshot shows the GitHub repository page for MetricsGrimoire/Bicho. At the top, there's a search bar and navigation links for Pull requests, Issues, and Gist. The repository name "MetricsGrimoire / Bicho" is displayed, along with an "Unwatch" button. Below this, a navigation bar shows "Code" as the active tab, with links to Issues (45), Pull requests (1), Wiki, Pulse, Graphs, and Settings. A description of the repository states: "Bicho is a command line based tool used to parse bug/issue tracking systems <http://metricsgrimoire.github>". Statistics show 569 commits, 5 branches, and 6 releases. Action buttons include "New pull request", "New file", "Find file", and a dropdown for "HTTPS" with the URL "https://github". A commit by "sduenas" is highlighted, with the message "[Jira] Fix parsing error on changes made by anonymous users". Below the commit, a file tree shows "bicho" (linked to the commit message) and "bin" (linked to "Rename modules to lowercase").

This repository Search

Pull requests Issues Gist

MetricsGrimoire / Bicho Unwatch

Code Issues 45 Pull requests 1 Wiki Pulse Graphs Settings

Bicho is a command line based tool used to parse bug/issue tracking systems <http://metricsgrimoire.github>

569 commits 5 branches 6 releases

Branch: master New pull request New file Find file HTTPS https://github

sduenas [Jira] Fix parsing error on changes made by anonymous users

bicho [Jira] Fix parsing error on changes made by anonymous users

bin Rename modules to lowercase



# Set up the environment

Create and activate a new Python virtualenv:

```
$ virtualenv -p /usr/bin/python3 gitlab  
$ gitlab/bin/activate
```

Install dependencies

```
(gitlab) $ pip install beautifulsoup4  
(gitlab) $ pip install python-dateutil  
(gitlab) $ pip install requests  
(gitlab) $ pip install six
```



# Install GrimoireLab tools

Install Perceval from its GrimoireLab git repo:

```
(gitlab) $ git clone \  
    https://github.com/grimoirelab/perceval.git  
(gitlab) $ cd perceval  
(gitlab) $ git checkout fosdem16  
(gitlab) $ python3 setup.py install  
(gitlab) $ cd ..
```



# Install GrimoireLab tools

Clone GrimoireELK GrimoireLab git repo, checkout branch fosdem16

```
(gitlab) $ git clone \
    https://github.com/grimoirelab/GrimoireELK.git
(gitlab) $ cd GrimoireELK
(gitlab) $ git checkout fosdem16
(gitlab) $ cd ..
```





# Prepare the repository to analyze

We will use MetricsGrimoire/Bicho:

```
(gitlab) $ git clone \  
    https://github.com/MetricsGrimoire/Bicho  
(gitlab) $ cd Bicho  
(gitlab) $ git log --raw --numstat \  
    --pretty=fuller --decorate=full --parents \  
    -M -C -c --remotes=origin --all \  
> /tmp/bicho-gitlog.log
```





# Upload info to ElasticSearch

First, upload info in the git log:

```
(gitlab) $ cd ../GrimoireELK/util
(gitlab) $ python3 ./p2o.py -e http://localhost:9200 \
  --no_inc --debug git /tmp/bicho-gitlog.log
(gitlab) $ python3 ./p2o.py -e http://localhost:9200 \
  --no_inc --debug --enrich_only git /tmp/bicho-gitlog.log
...
2016-01-31 00:47:26,960 Deleted and created index
  http://localhost:9200/git_tmp_bicho-gitlog.log_enrich
```

Index name:

git\_tmp\_bicho-gitlog.log\_enrich



# Upload info to ElasticSearch

Second, upload info from GitHub issues & pull requests  
(may take a while & requires a GitHub token)

```
(gitlab) $ python3 ./p2o.py -e http://localhost:9200 \  
--no_inc --debug github --owner metricsgrimoire \  
--repository bicho --token XXXX
```

```
(gitlab) $ python3 ./p2o.py -e http://localhost:9200 \  
--no_inc --debug --enrich_only github \  
--owner metricsgrimoire --repository bicho --token XXXX
```

...

2016-01-31 00:45:37,269 Deleted and created index

[http://localhost:9200/github\\_https:\\_\\_github.com\\_metricsgrimoire\\_bicho\\_enrich](http://localhost:9200/github_https:__github.com_metricsgrimoire_bicho_enrich)

Index name:

github\_https:\_\_github.com\_metricsgrimoire\_bicho\_enrich



# Produce a Kibana dashboard

We upload template dashboards to ElasticSearch,  
and build our dashboards with them  
(notice the index names)

```
(gitlab) $ python3 ./kidash.py -e http://localhost:9200 \  
-g --import ../dashboards/git-activity.json
```

```
(gitlab) $ python3 ./kidash.py -e http://localhost:9200 \  
-g --import ../dashboards/github-pr-bubbles-geoMap.json
```

```
(gitlab) $ python3 ./e2k.py -g -e http://localhost:9200 \  
-d "Git-Activity" -i git__tmp_bicho-gitlog.log_enrich
```

```
(gitlab) $ python3 ./e2k.py -g -e http://localhost:9200 \  
-d "PRBubblesGeoMap" -i \  
github_https:__github.com_metricsgrimoire_bicho_enrich
```



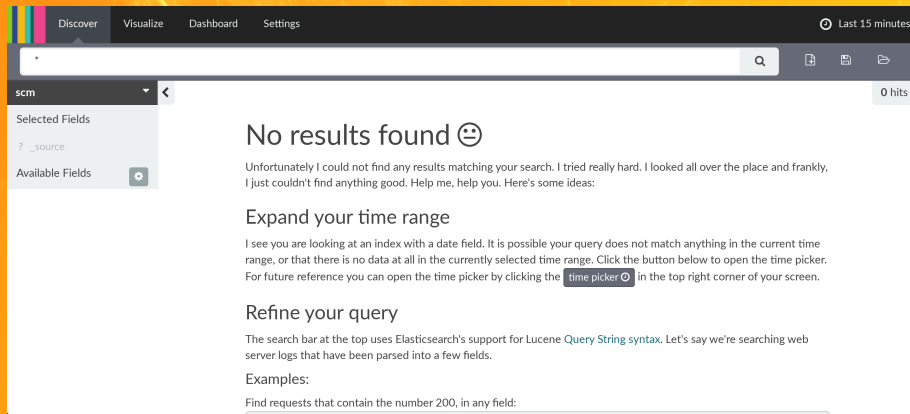


Enjoy



# See the dashboards in your browser

http://localhost:5601



The screenshot shows the Kibana web interface. At the top, there's a navigation bar with tabs: Discover, Visualize, Dashboard, and Settings. The 'Discover' tab is active. On the right of the navigation bar, it says 'Last 15 minutes'. Below the navigation bar is a search bar with a magnifying glass icon. To the left of the search bar, there's a sidebar with a dropdown menu showing 'scm' and a left arrow. Below this, there's a section for 'Selected Fields' with a list containing '? \_source'. Below that, there's a section for 'Available Fields' with a gear icon. The main content area displays 'No results found' with a sad face emoji. Below this, there's a paragraph explaining that no results were found and suggesting ideas for improvement. Then, there's a section titled 'Expand your time range' with a paragraph explaining that the search might be too narrow and suggesting to click the 'time picker' button in the top right corner. Finally, there's a section titled 'Refine your query' with a paragraph explaining that the search bar uses Elasticsearch's support for Lucene Query String syntax and suggesting to find requests containing the number 200.

Discover Visualize Dashboard Settings Last 15 minutes

scm < 0 hits

Selected Fields

? \_source

Available Fields

## No results found 😞

Unfortunately I could not find any results matching your search. I tried really hard. I looked all over the place and frankly, I just couldn't find anything good. Help me, help you. Here's some ideas:

### Expand your time range

I see you are looking at an index with a date field. It is possible your query does not match anything in the current time range, or that there is no data at all in the currently selected time range. Click the button below to open the time picker. For future reference you can open the time picker by clicking the **time picker** in the top right corner of your screen.

### Refine your query

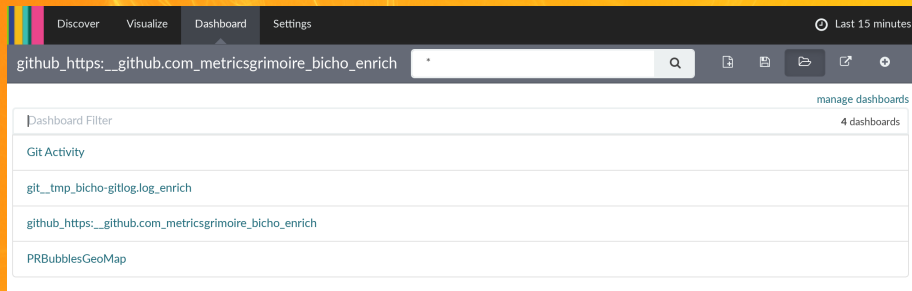
The search bar at the top uses Elasticsearch's support for Lucene [Query String syntax](#). Let's say we're searching web server logs that have been parsed into a few fields.

Examples:

Find requests that contain the number 200, in any field:



# Select the produced dashboards

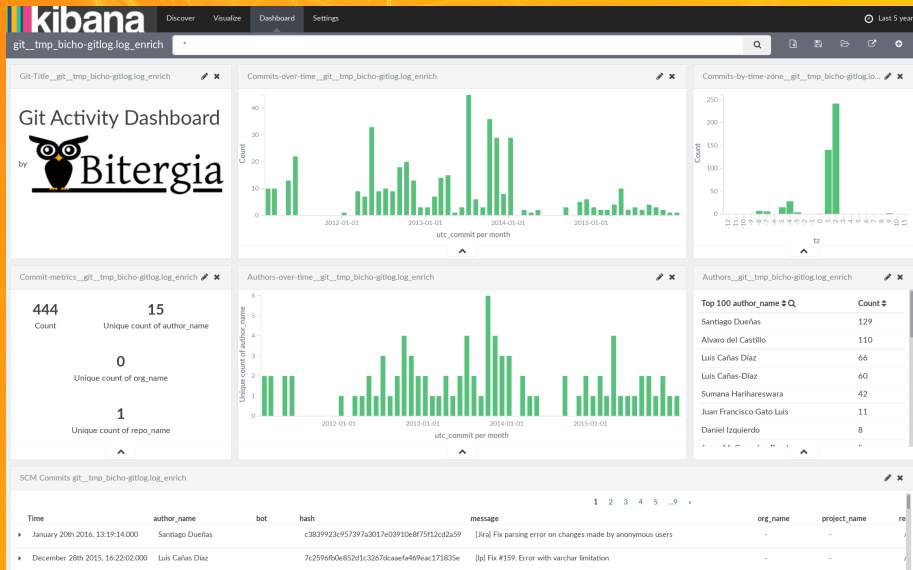


The screenshot shows the GrimoireLab dashboard interface. At the top, there is a navigation bar with tabs: Discover, Visualize, Dashboard (selected), and Settings. On the right of the navigation bar, it says "Last 15 minutes". Below the navigation bar, there is a search bar containing the text "github\_https:\_\_github.com\_metricsgrimoire\_bicho\_enrich". To the right of the search bar are icons for file operations. Below the search bar, there is a section titled "Dashboard Filter" with a sub-header "manage dashboards" and "4 dashboards". The list of dashboards includes: "Git Activity", "git\_\_tmp\_bicho-gitlog.log\_enrich", "github\_https:\_\_github.com\_metricsgrimoire\_bicho\_enrich", and "PRBubblesGeoMap".

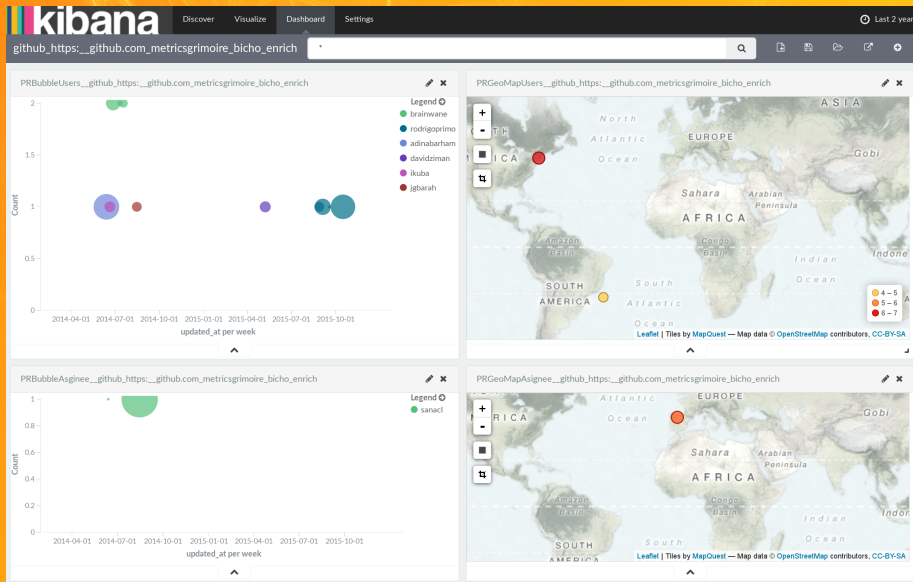
- Template dashboards: Git Activity and PRBubblesGeoMap
- Real dashboards: the other two



# Watch the results of your work!



# Watch the results of your work! Play with them!





# Some details



# Deploying Elasticsearch, Kibana

Download and uncompress Elasticsearch, from their website

```
$ cd elasticsearch-1.7.1/  
# [Configure, if needed]  
$ bin/elasticsearch
```

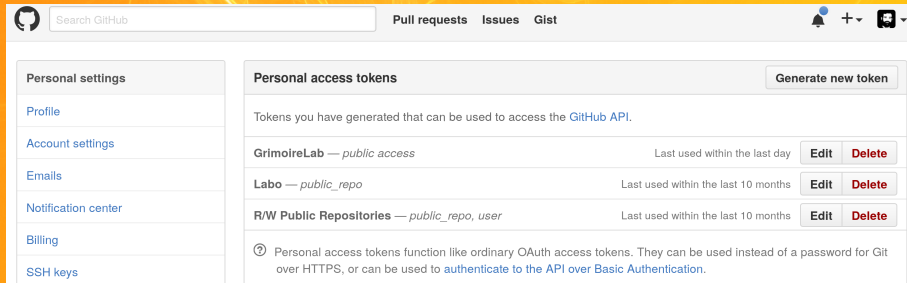
Download and uncompress Kibana, from their website

```
$ cd kibana-4.1.1-linux-x64  
# [Configure, if needed]  
$ bin/kibana
```



# Get a GitHub token

Your photo — Settings  
Personal access tokens  
Generate new token



The screenshot shows the GitHub interface for a user named 'GrimoireLab'. The top navigation bar includes 'Pull requests', 'Issues', and 'Gist'. The left sidebar lists 'Personal settings' with links to Profile, Account settings, Emails, Notification center, Billing, and SSH keys. The main content area is titled 'Personal access tokens' and includes a 'Generate new token' button. Below this, a list of tokens is shown:

Token Name	Scope	Last used	Actions
GrimoireLab	public access	Last used within the last day	Edit Delete
Labo	public_repo	Last used within the last 10 months	Edit Delete
R/W Public Repositories	public_repo, user	Last used within the last 10 months	Edit Delete

A help note at the bottom explains that personal access tokens function like ordinary OAuth access tokens and can be used instead of a password for Git over HTTPS, or to authenticate to the API over Basic Authentication.

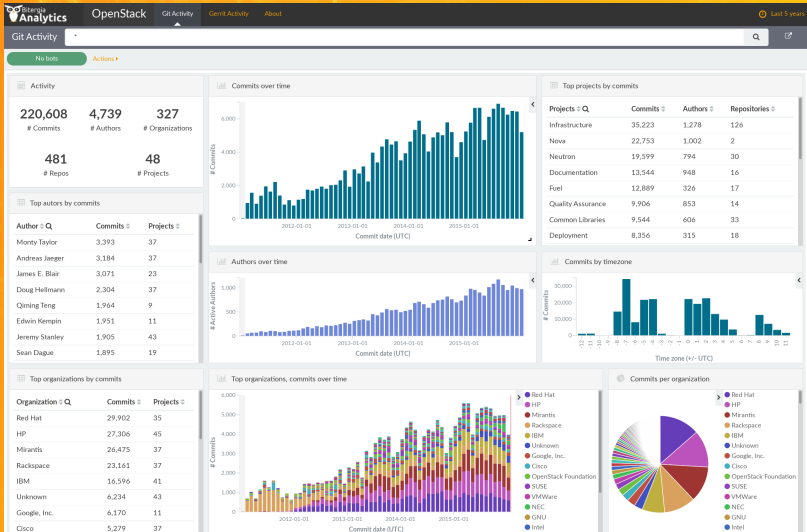
Take note of it, or you won't see it again!



# Bonus track



# Preview: our new Kibana-based dashboards



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

