



Introduction to GrimoireLab and more ...

Venu Vardhan Reddy Tekula

CHA OSS Beijing Meetup
08 July, 2021



chaoss.community

Links and slides

vchrombie.github.io/presentations



Agenda

- What is GrimoireLab
- A Brief History
- How does it work
- Demo of GrimoireLab
- How to Get Started
- Current Work
- How to Get Involved



About myself

Hyderabad, India

Graduate from Amrita Vishwa
Vidyapeetham, Member [@amfoss](#)

GrimoireLab Community Manager

Previously Backend Developer at
Bitergia

Google Summer of Code w/ CHAOSS
and GrimoireLab



Community Projects Health



What are the tools?

What is the data?

How do we collect the data?

Where do we store it?

How do we want to show it?

What are the goals?

What about contributors information?

GrimoireLab



FREE, LIBRE OPEN SOURCE TOOLS FOR
SOFTWARE DEVELOPMENT ANALYTICS

CHA^{CO}SS

grimoirelab.github.io



A Brief History

2001-2015



Counting Potatoes: the Size of Debian 2.2

Jesús M. González-Barahona, Miguel A. Ortuño Pérez, Pedro de las Heras Quirós, José Centeno González and Vicente Matellán Olivera

Debian is the largest Free Software distribution, with well over 2,800 source packages in the latest stable release (Debian 2.2) and more than 4,000 source packages in the release currently in preparation. But, how large is “the largest”? In this paper, we use David Wheeler’s sloccount system to determine the number of physical source lines of code (SLOC) of Debian 2.2 (aka Potato). We show that Debian 2.2 includes over 56,000,000 physical SLOC (almost twice than Red Hat 7.1, released about 8 months later), showing that the Debian development model (based on the work of a large group of voluntary developers spread around the world) is at least as capable as other development methods (like the more centralized one, based on the work of employees, used by Red Hat or Microsoft) to manage distributions of this size.

Keywords: Debian, Free Software, Libre Software, SLOC, Lines of Code, Linux

1 Introduction

On August 14th of 2000 the Debian Project announced Debian GNU/Linux 2.2, the “Joel ‘Espy’ Klecker” release [Debian22Ann] [Debian22Rel]. Code named “Potato”, it is the latest (to date) release of the Debian GNU/Linux Operating System. In this paper, we have counted this distribution, show-

2 Some Background about Debian


The Debian 2.2 GNU/Linux distribution is put together and maintained by the Debian project. In this section, we offer some background data about Debian as a project, and about the Debian 2.2 release.


Jesús M. González Barahona is a lecturer at the *Universidad Rey Juan Carlos*, and collaborator of *BarraPunto.Com*. He began working on the promotion of Free Software in 1991, in the PD-SOFT group (later the Software group). Since then he has been


Research group at
URJC, 2001

FLOSSMetrics project

Not secure | flossmetrics.org


Project funded by the European Commission
under contract number FP6-033982


is part of the FLOSSQuality initiative



Free/Libre and Open Source Software Metrics

Highlights

- Final Report About FLOSSMetrics Melquiades (data about >2,800 FLOSS projects) FLOSS Guide for SMEs:
 - Latest version (PDF)
 - FLOSSMetrics deliverable Exploiting FLOSS Research results for Industry Booklet

The FLOSSMetrics project

FLOSSMetrics stands for Free/Libre Open Source Software Metrics.

The main objective of FLOSSMETRICS is to construct, publish and analyse a large scale database with information and metrics about libre software development coming from several thousands of software projects, using existing methodologies, and tools already developed.

FLOSSMetrics is providing access to dumps of this database (along with charts, tables and other quantitative information about FLOSS development projects) in the [Melquiades website](#)

FLOSSMetrics is also providing the [FLOSS Guide for SMEs \(small and medium enterprises\)](#), which explains the benefits of FLOSS form SMEs, and how to take advantage of it.

More information about the project, including results and main achievements, can be found in [the final report document](#).

Project News

- New SCM aggregated database available (fm3_aggregated_scm-2009-08-19)

FLOSSMetrics main targets

- Identify and evaluate sources of data and develop a comprehensive database structure, built upon the results of CALIBRE
- Integrate already available tools to extract and process such data into a complete platform
- Build and maintain an updated empirical database applying extraction tools to thousands of open source projects
- Develop visualisation methods and analytical studies, especially relating to benchmarking, identification of best practices, measuring and predicting success



FLOSSMetrics,
2008-2010

Metrics Grimoire

metricsgrimoire.github.io

MetricsGrimoire

Tools for software & development community analytics

@ GitHub


Update: Check the new GrimoireLab platform

MetricsGrimoire

[Update]: Currently, main development related with *MetricsGrimoire* has stopped. Main contributors are working on a new platform: **GrimoireLab**

MetricsGrimoire (pronounced /metrɪksgrɪmʊər/) is a toolset to obtain data from repositories related to software development: source code management (aka *version control*), issue tracking (aka *bug reporting*) systems, mailing lists, etc. Data and metadata about the software development processes is retrieved from those repositories (information about commits, ticket management, communication in mailing lists, etc.), and then organized and stored into SQL databases that can later be mined for specific patterns or summaries of activity.


MetricsGrimoire tools support many kinds of repositories, including those provided by GitHub (git & GitHub issue tracking).



MetricsGrimoire has been already used to analyze many different projects and together with visualization tools like **VizGrimoire** it is possible to get dashboards or reports like **the ones**



MetricsGrimoire,
2007-2011

 vizGrimoire

Not secure | vizgrimoire.bitergia.org

vizGrimoire

Analyzing and visualizing software development

[GitHub Profile](#)

The vizGrimoire project

vizGrimoire (usually pronounced / vizgrimwar/) is a toolset and framework to analyze and visualize data about software development. Currently, it is focused on data produced by the [MetricsGrimoire tools](#) (CVSAnalY, Bicho and MailingListStats).

vizGrimoire is promoted by [Bitergia](#), the company providing software development analytics services, but it is a project open to the community: you are welcome to participate.

The following is a list of the current set of tools with a brief description. For more information, visit each of tool repositories, and/or read their install and readme files. You can also have a glimpse of the intended [vizGrimoire architecture](#).

vizGrimoireR

Analyze data in databases produced by MetricsGrimoire tools from software development repositories (source code management systems, issue tracking systems, mailing lists, etc.).

[Repository at GitHub](#)

vizGrimoireJS

JavaScript framework to create dashboards and interactive reports and visualizations from data produced by vizGrimoireJS

[Repository at GitHub](#)





bitergia.com
2011

Extracting Knowledge From CHAOSS

CHAOSS will develop metrics, including:

- › Project lifecycle
- › Diversity and inclusion
- › Risk and provenance
- › Ecosystem impact

In addition...

- › Develop a FLOSS reference implementation of defined metrics
- › Integrate GrimoireLab, Prospector and Cregit into a collaborative framework
- › Develop a better understanding of how contributions happen

Backed by:

- › Red Hat, Bitergia, Eclipse Foundation, Linaro, Mozilla, OpenStack, Polytechnique Montreal, Sauce Labs, Software Sustainability Institute, Symphony Software Foundation, University of Missouri, University of Mons, University of Nebraska at Omaha, and University of Victoria, Laval University, and Jono Bacon Consulting

CHAOSS

 THE **LINUX** FOUNDATION



 **OPEN SOURCE SUMMIT**
NORTH AMERICA

THE LINUX FOUNDATION





EVENTS HOW IT WORKS COMPONENTS BUILT WITH ABOUT DOCS



Join us!

CHA^{SS}

GrimoireLab is one of CHAOSS Software founding projects



FREE, LIBRE, OPEN SOURCE TOOLS FOR
SOFTWARE DEVELOPMENT ANALYTICS

The background of the slide is a complex, abstract network diagram. It consists of numerous nodes, represented by circles of various sizes and colors (including purple, blue, green, red, and grey), interconnected by thin, light grey lines. Some nodes are highlighted with larger, concentric circles in blue, green, or red. The overall structure is dense and interconnected, suggesting a complex system or network.

How does it work



CHA^{SS}



Where is the community? Where is the data?



Git

GitHub

GitLab

BitBucket

Jira

Gerrit

...

Where is the community? Where is the data?



Git

GitHub

GitLab

BitBucket

Jira

Gerrit

...

Community doesn't
mean just code

Community means
people,
communications,
and more

Where is the community? Where is the data?



Git

GitHub

GitLab

BitBucket

Jira

Gerrit

...

Wiki

Discourse

Mailing List

Slack

Meetup.com

StackOverflow

...

Raw data:

- Get data from the data sources

Enriched data:

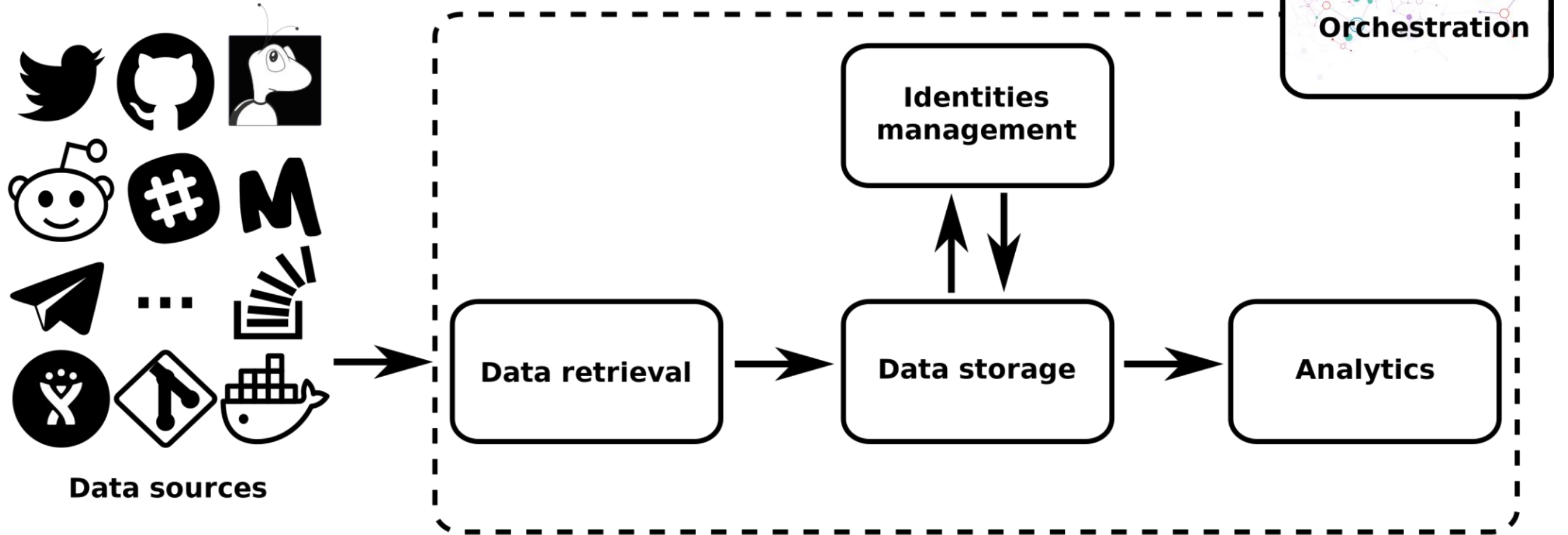
- Unify data, extract what is required
- Manage identities
- Calculate metrics

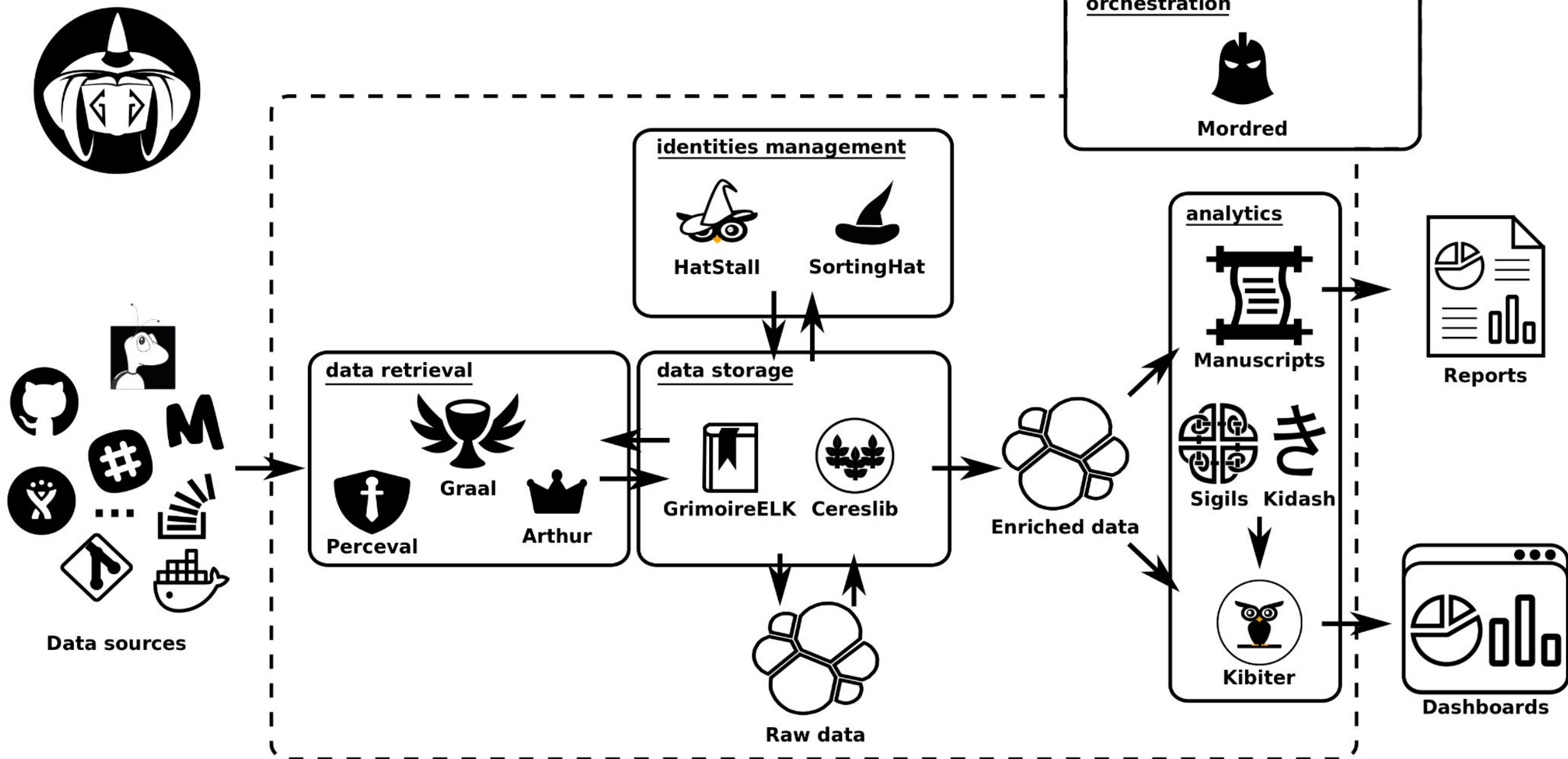
Useful data:

- Visualize, report



GrimoireLab





Demo of GrimoireLab

sneak-peek



GrimoireLab powers Metrics aaS



There are some software that offer community metrics as a service, with GrimoireLab working under the hood.

- cauldron.io
- lfx.linuxfoundation.org/tools/insights

Demo of GrimoireLab



There are some instances already deployed, you can visit any of the links

- chaoss.biterg.io
- dashboard.mautic.org

How to Get Started



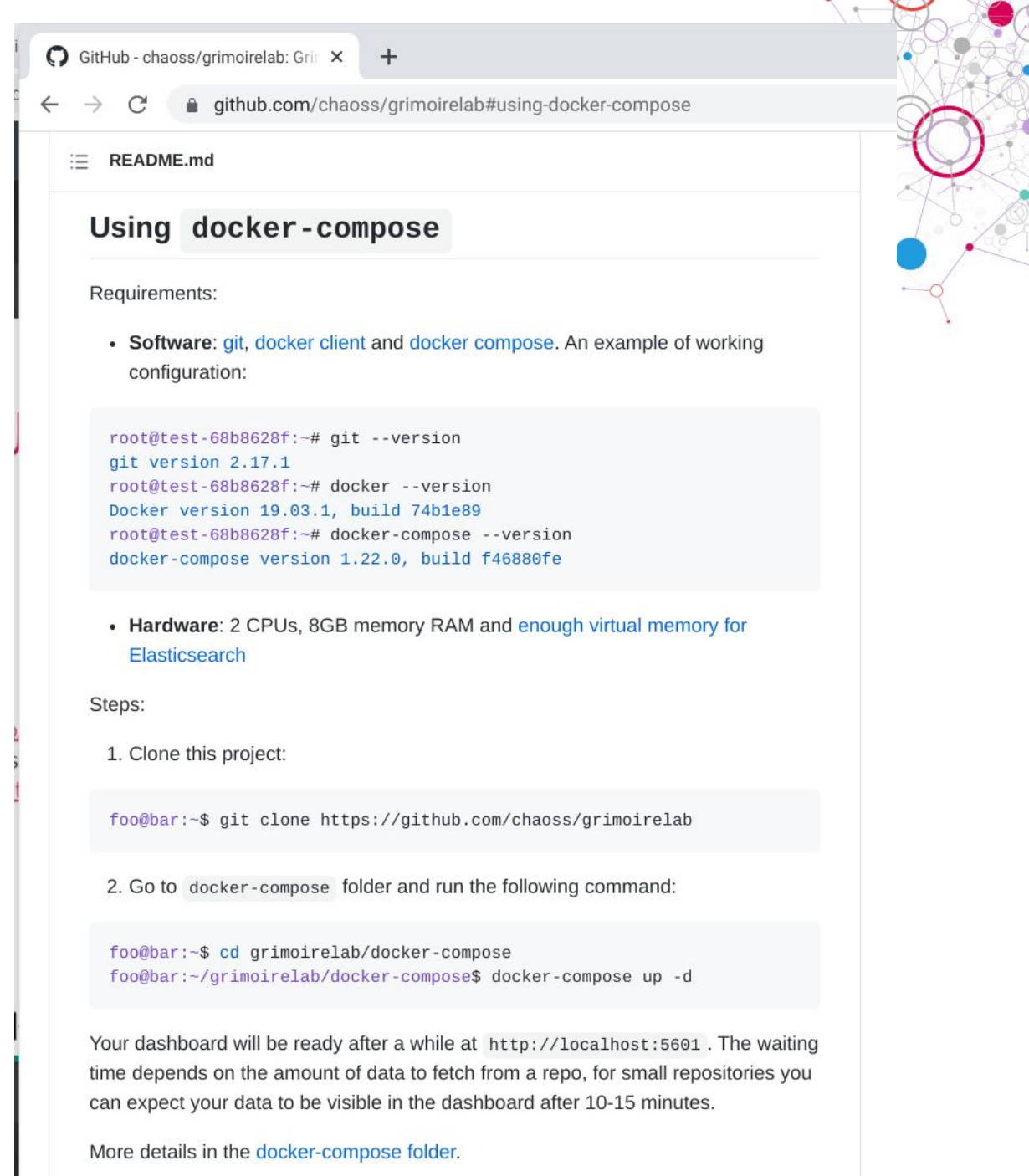
How to Get Started

1. docker-compose (easiest of all)
2. docker run
3. pip packages
4. source code (mainly, for developers)



docker-compose method

1. Install git, docker & docker-compose
2. Clone chaoss/grimoirelab repository
3. Go to docker-compose folder
4. Run docker-compose up -d
5. Open your browser to see the dashboard at localhost:5601



GitHub - chaoss/grimoirelab: Gri

github.com/chaoss/grimoirelab#using-docker-compose

Using docker-compose

Requirements:

- **Software:** [git](#), [docker client](#) and [docker compose](#). An example of working configuration:

```
root@test-68b8628f:~# git --version
git version 2.17.1
root@test-68b8628f:~# docker --version
Docker version 19.03.1, build 74b1e89
root@test-68b8628f:~# docker-compose --version
docker-compose version 1.22.0, build f46880fe
```

- **Hardware:** 2 CPUs, 8GB memory RAM and [enough virtual memory for Elasticsearch](#)

Steps:

1. Clone this project:

```
foo@bar:~$ git clone https://github.com/chaoss/grimoirelab
```

2. Go to `docker-compose` folder and run the following command:

```
foo@bar:~$ cd grimoirelab/docker-compose
foo@bar:~/grimoirelab/docker-compose$ docker-compose up -d
```

Your dashboard will be ready after a while at `http://localhost:5601`. The waiting time depends on the amount of data to fetch from a repo, for small repositories you can expect your data to be visible in the dashboard after 10-15 minutes.

More details in the [docker-compose folder](#).

[Blog Post](#)

GrimoireLab Install

By CHAOSS Community | October 8, 2019

Simplifying how to install GrimoireLab

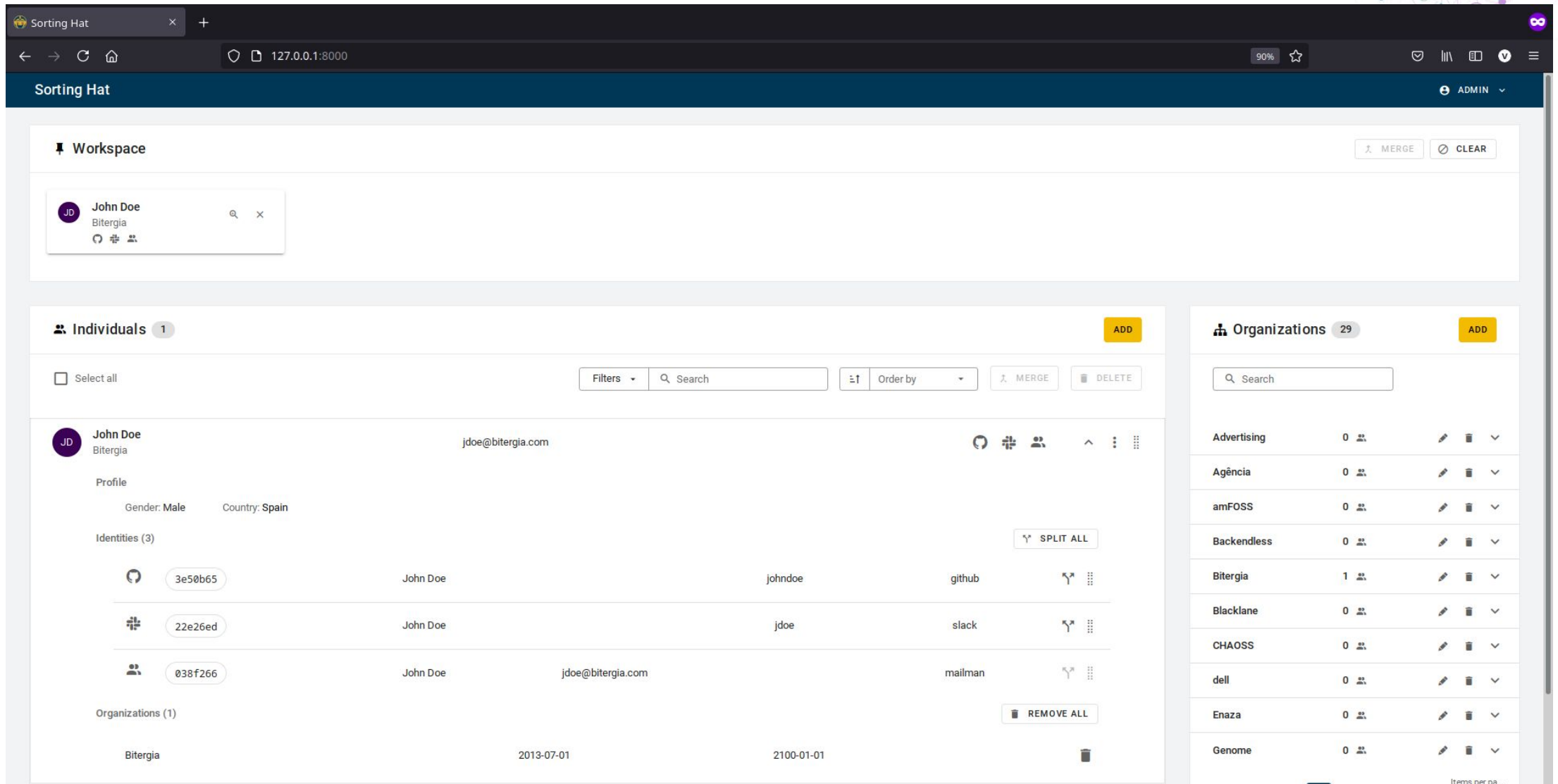
By [Georg Link](#)

When I talk to people who tried to install GrimoireLab, I get one consistent answer. It is difficult and our [GrimoireLab tutorial](#) is too complicated. I believe this status quo is hurting the adoption of GrimoireLab software and CHAOSS metrics. This blog post is about how to make it easier for anyone to start using GrimoireLab.

Current Work



SortingHat (muggle branch)



The screenshot displays the SortingHat web interface. At the top, a browser window shows the URL 127.0.0.1:8000. The interface has a dark blue header with the "Sorting Hat" logo and an "ADMIN" dropdown. Below the header, a "Workspace" section contains a card for "John Doe" from "Bitergia". The main area is divided into two panels: "Individuals" and "Organizations".

Individuals Panel: Shows a list of individuals. The first individual is "John Doe" from "Bitergia" with email "jdoe@bitergia.com". Below the name, there is a "Profile" section with "Gender: Male" and "Country: Spain". Under "Identities (3)", there are three entries:

Identity	Name	Username	Service
3e50b65	John Doe	johndoe	github
22e26ed	John Doe	jdoe	slack
038f266	John Doe	jdoe@bitergia.com	mailman

Below the identities, there is a section for "Organizations (1)" showing "Bitergia" with dates "2013-07-01" and "2100-01-01".

Organizations Panel: Shows a list of organizations. The first organization is "Advertising" with 0 members. Other organizations include "Agência", "amFOSS", "Backendless", "Bitergia" (1 member), "Blacklane", "CHAOSS", "dell", "Enaza", and "Genome".

Bestiary (unicorn branch)



```
default-grimoirelab-settings — vim setup.cfg — 83x73

[git]
raw_index = git_demo_raw
enriched_index = git_demo_en
latest-items = true
studies = [enrich_demogr


[github]
api-token = <YOUR_API_TOKEN>
#enterprise-url = <ENTERPRISE_URL>
raw_index = github
sleep-for-rate = 1
sleep-time = "300"
enriched_index =
#gitlab-issue

default-grimoirelab-
sh — 83x73

MacBook-A1320:~$ cd default-grimoirelab-settings/
MacBook-A1320:default-grimoirelab-settings$ cat projects.json
{
  "grimoire": {
    "meta": {
      "title": "GrimoireLab Settings"
    },
    "git": [
      {
        "url": "https://github.com"
      }
    ],
    "github": [
      {
        "url": "https://github.com/chaoss/grimoirelab"
      }
    ]
  }
}
```

Google Summer of Code 2021





Sorting Hat - Extend data model and user interface to capture better information about contributors

STUDENT

Rashmi K A

PROJECT NAME

Sorting Hat


ORGANIZATION

CHA OSS Project

MENTORS

Venu Vardhan Reddy


SHARE LINK



Grimoirelab is an open-source toolset for software development analytics. Grimoirelab provides a set of tools to collect, analyze and visualize software development metrics from a variety of sources like Git, Jira, Confluence, Slack, etc. In order to manage the identities of people across these different sources, Grimoirelab developed Sorting Hat. Sorting Hat manages the identities of people and related metadata.

As part of the metadata collected around identities, Sorting Hat stores organizational information such as the name and domains related to the organization. This project aims to add to this information by extending the existing Organization model to capture the internal structure of organizations such as departments, sub-organizations, and teams. This will help in annotating the identity information more meaningfully.

Summer 2021 Open Source Promotion Plan

 开源软件供应链点亮计划—暑期2021

ABOUT [PROJECT LIST](#) MASTER CLASS HOW IT WORKS RESULT DATA GET STARTED HELP Summer2020 中文 | [ENG](#)

Expanding and restyling the GrimoireLab tutorial

Project ID: 210190824

Selected Students: Veerasamy Sevagen

Degree of Difficulty: Medium

Language: English

Numbers of Applicants: 1

GrimoireLab is a powerful open-source platform that provides support for monitoring and in-depth analysis of software projects. It produces a rich set of metrics with data extracted from more than 30 tools related to contributing to Open Source development such as version control systems, issue trackers and forums. These metrics are shown and exploited on Web dynamic dashboards, which can be easily inspected by decision-makers to help them understand the evolution and health of their projects. The main entry point to learn about GrimoireLab is the tutorial, which provides a walkthrough of the platform and its components.

Mentor of Project: Venu Vardhan

Contact Information: [<venuvardhanreddytekula8@gmail.com>](mailto:venuvardhanreddytekula8@gmail.com)

[Project Details >](#)

How to Get Involved



Check out the CONTRIBUTING Guidelines.

Feel free to contact the maintainers.



github.com/chaoss/grimoirelab/blob/master/CONTRIBUTING.md

A screenshot of a web browser displaying the GitHub repository page for 'chaoss/grimoirelab'. The browser's address bar shows the URL 'github.com/chaoss/grimoirelab/blob/master/CONTRIBUTING.md'. The repository page header includes the repository name 'chaoss / grimoirelab' with statistics: 19 watchers, 256 stars, and 103 forks. Below the header, there are tabs for 'Code', 'Issues' (87), 'Pull requests' (7), 'Actions', 'Projects' (1), and 'Wiki'. The main content area shows the 'master' branch selected, with the file 'grimoirelab / CONTRIBUTING.md' highlighted. A commit by 'vchrombie' is shown, dated 'Latest commit d4cebbf on Mar 30'. Below the commit, it says '5 contributors' with their avatars. The file content is displayed, showing the title 'Contributing to GrimoireLab' and the text: 'These are some general guidelines and information related to how we contribute to GrimoireLab. GrimoireLab is a part of the Software Technical Committee of the CHA OSS Collaborative Project.' followed by a section 'Communication means' and a list of communication channels: 'Mailing list', 'IRC channel', and 'Issues / pull requests'. A note at the bottom states: 'Each of them is intended for an specific purpose. Please understand that you may be redirected to some other mean if the communication you intend to have is considered to fit better elsewhere.'

Join the community



github.com/chaoss/grimoirelab



[#grimoirelab](#) on CHAOSS Slack



[@grimoirelab](#)

CHAOSS

chaoss.github.io/grimoirelab

You can also help us with the different
working groups

chaoss.community/participate



Questions?

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

Mail to: venu@chaoss.community

Thanks to Georg Link, Yehui Wang
for the support and providing the
resources.



chaoss.community

xiè xie nǐ

