Introduction to GrimoireLab and more ...

Venu Vardhan Reddy Tekula

CHAOSS Beijing Meetup
08 July, 2021

Links and slides

vchrombie.github.io/presentations

CHAOSS

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?

CHAOSS

GrimoireLab

FREE, LIBRE OPEN SOURCE TOOLS FOR

SOFTWARE DEVELOPMENT ANALYTICS

CHAOSS

grimoirelab.github.io

A Brief History

2001-2015

CHAOSS

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

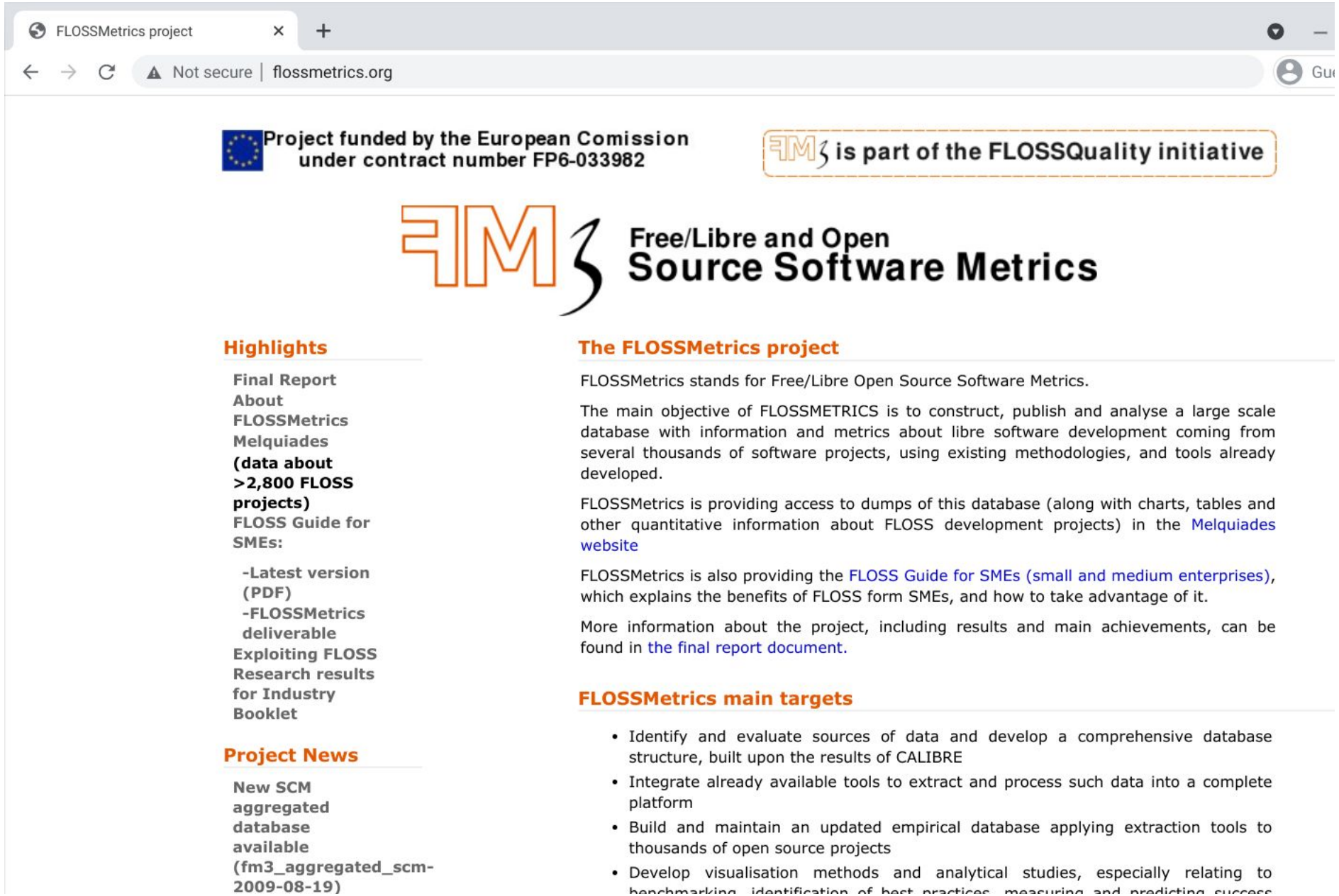
The Debian 2.2 GNU/Linux distribution is put together and maintained by the Debian project. In this section, we offer some background data about Debian as a project, and about the Debian 2.2 release.

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-

Research group at URJC, 2001

CHAOSS

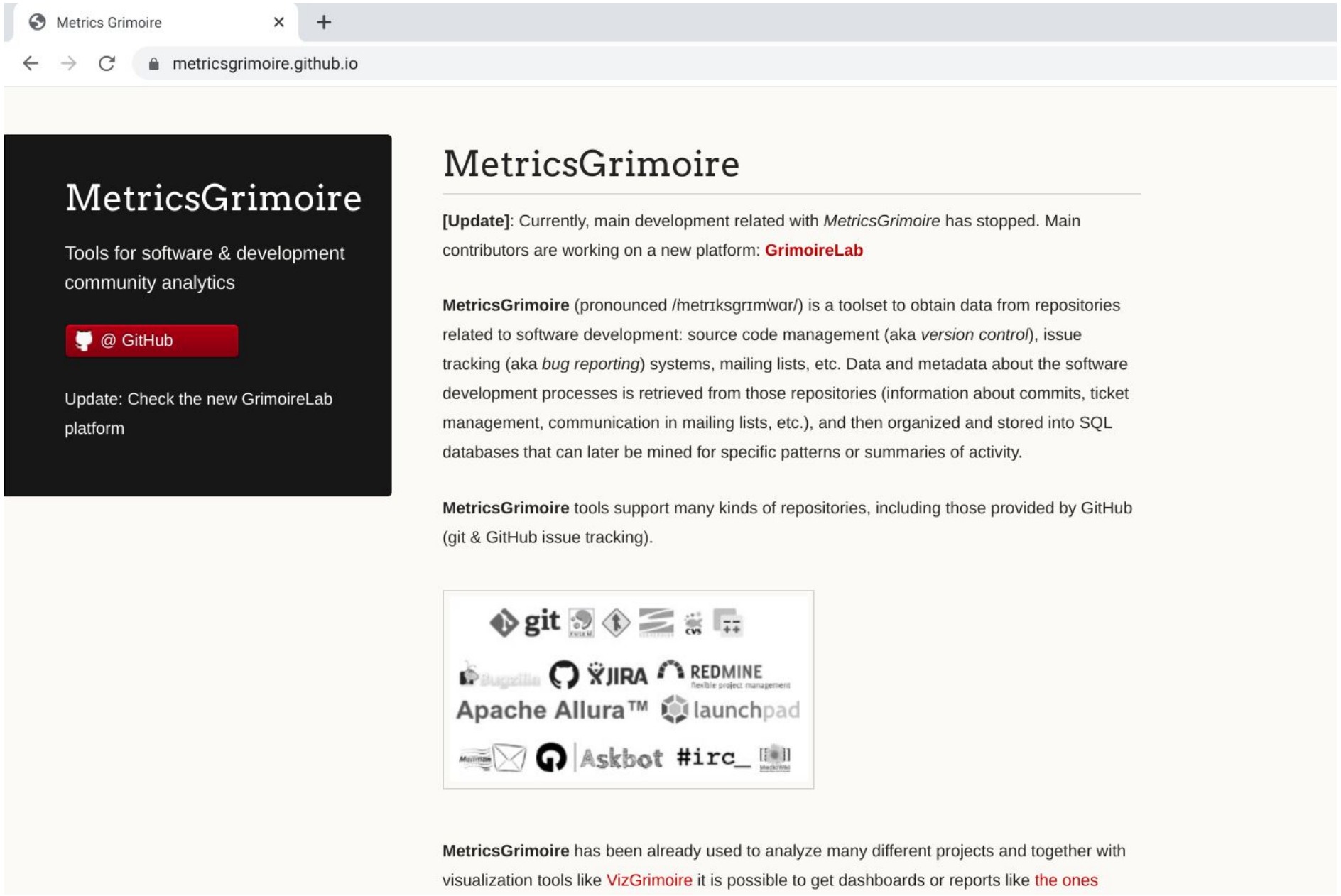
robotica.unileon.es/vmo/pubs/upgrade.pdf



FLOSSMetrics, 2008-2010

CHAOSS

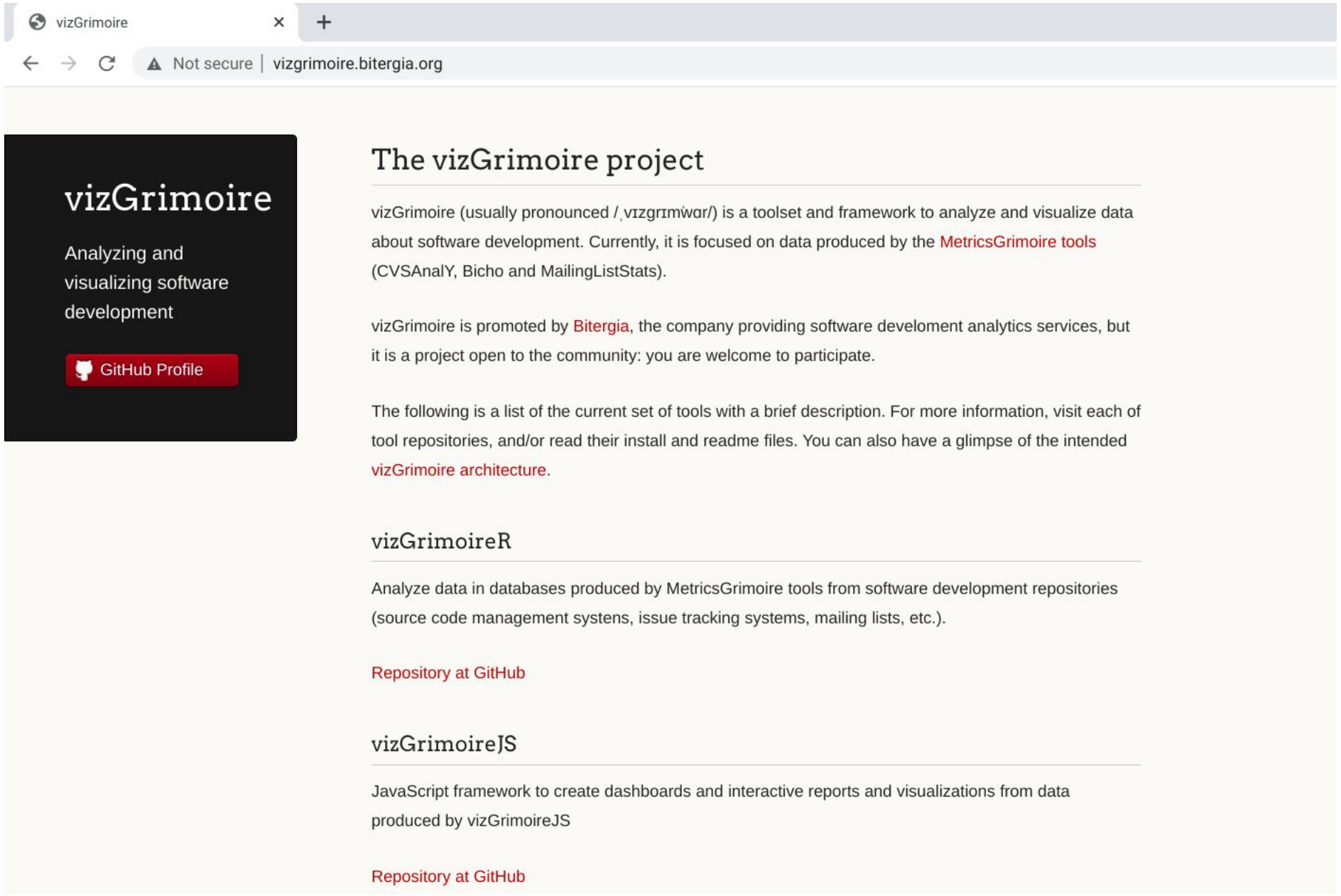
flossmetrics.org



MetricsGrimoire, 2007-2011

CHAOSS

metricsgrimoire.github.io



The vizGrimoire Project, 2011-2014

CHAOSS

vizgrimoire.github.io

bitergia.com

2011

CHAOSS

Extracting Knowledge From CHAOSS

CHAOSS will develop metrics, including:

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

CHAOSS

In addition…

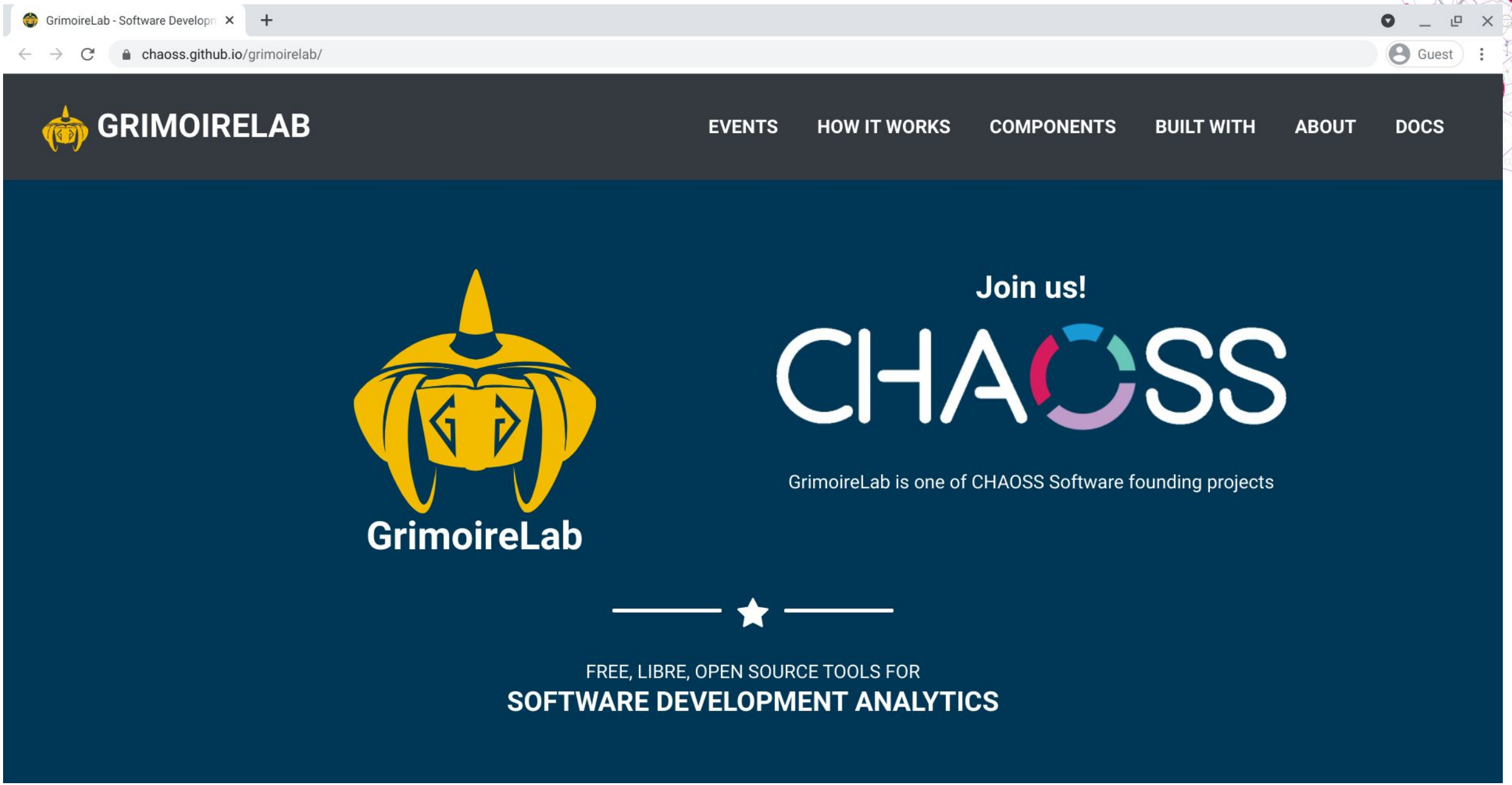
- Develop a FLOSS reference implementation of defined metrics
- Integrate GrimoireLab, Prospector and Cregit into a collaborative framework
- Develop a better understanding of how contributions happen

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

THE LINUX FOUNDATION

THE LINUX FOUNDATION
OPEN SOURCE SUMMIT
NORTH AMERICA


GrimoireLab - Software Developm
chaoss.github.io/grimoirelab/
Guest
GRIMOIRELAB
EVENTS
HOW IT WORKS
COMPONENTS
BUILT WITH
ABOUT
DOCS
GrimoireLab
Join us!
CHAOSS
GrimoireLab is one of CHAOSS Software founding projects
FREE, LIBRE, OPEN SOURCE TOOLS FOR
SOFTWARE DEVELOPMENT ANALYTICS

How does it work

CHAOSS

Discourse
git
slack
GitLab
JIRA
Mailman
GitHub
Confluence
stackoverflow
Bitbucket
CHAOSS
NORTH AMERICA
ASIA
AFRICA
SOUTH AMERICA
OCEANIA
ANTARCTICA
bundles
app
plugins
CoreBundle
63.013

Where is the community? Where is the data?

Git

GitHub

GitLab

BitBucket

Jira

Gerrit

...

CHAOSS

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

CHAOSS

Where is the community? Where is the data?

Git	Wiki
GitHub	Discourse
GitLab	Mailing List
BitBucket	Slack
Jira	Meetup.com
Gerrit	StackOverflow
...	...

CHAOSS

Raw data:

- Get data from the data sources

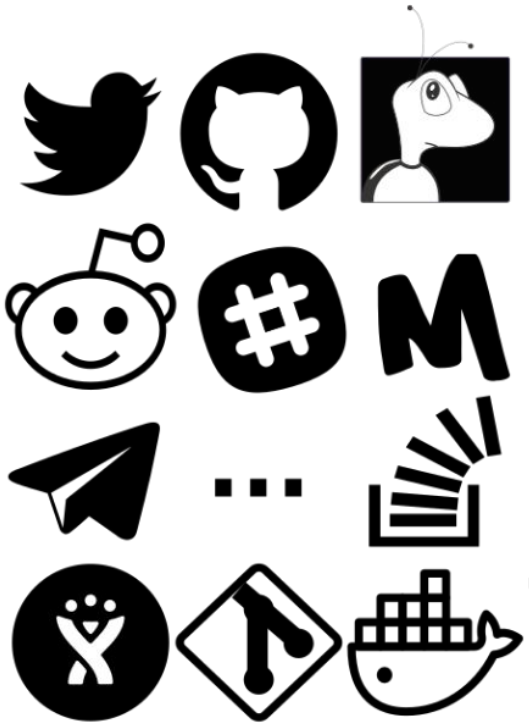
Enriched data:

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

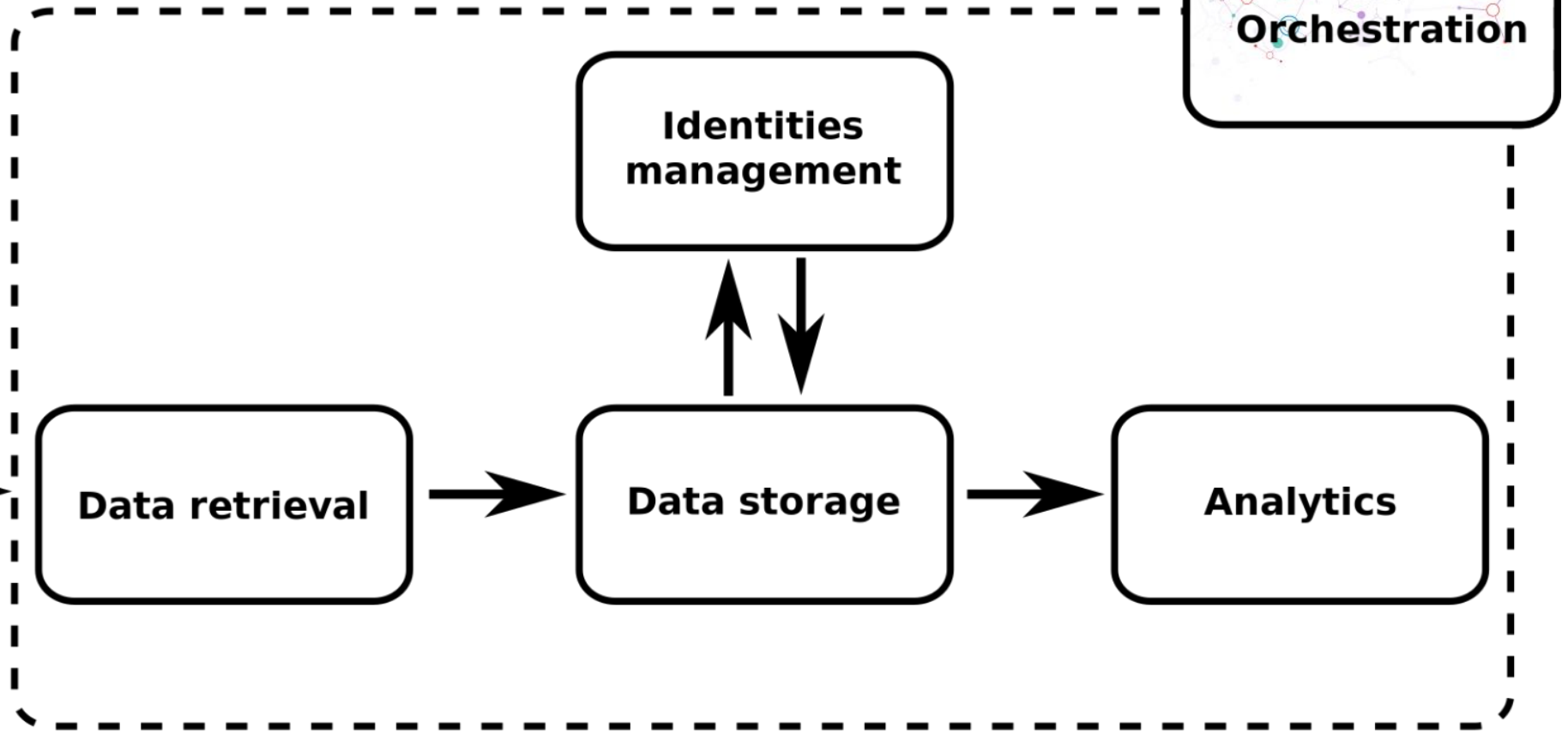
Useful data:

- Visualize, report

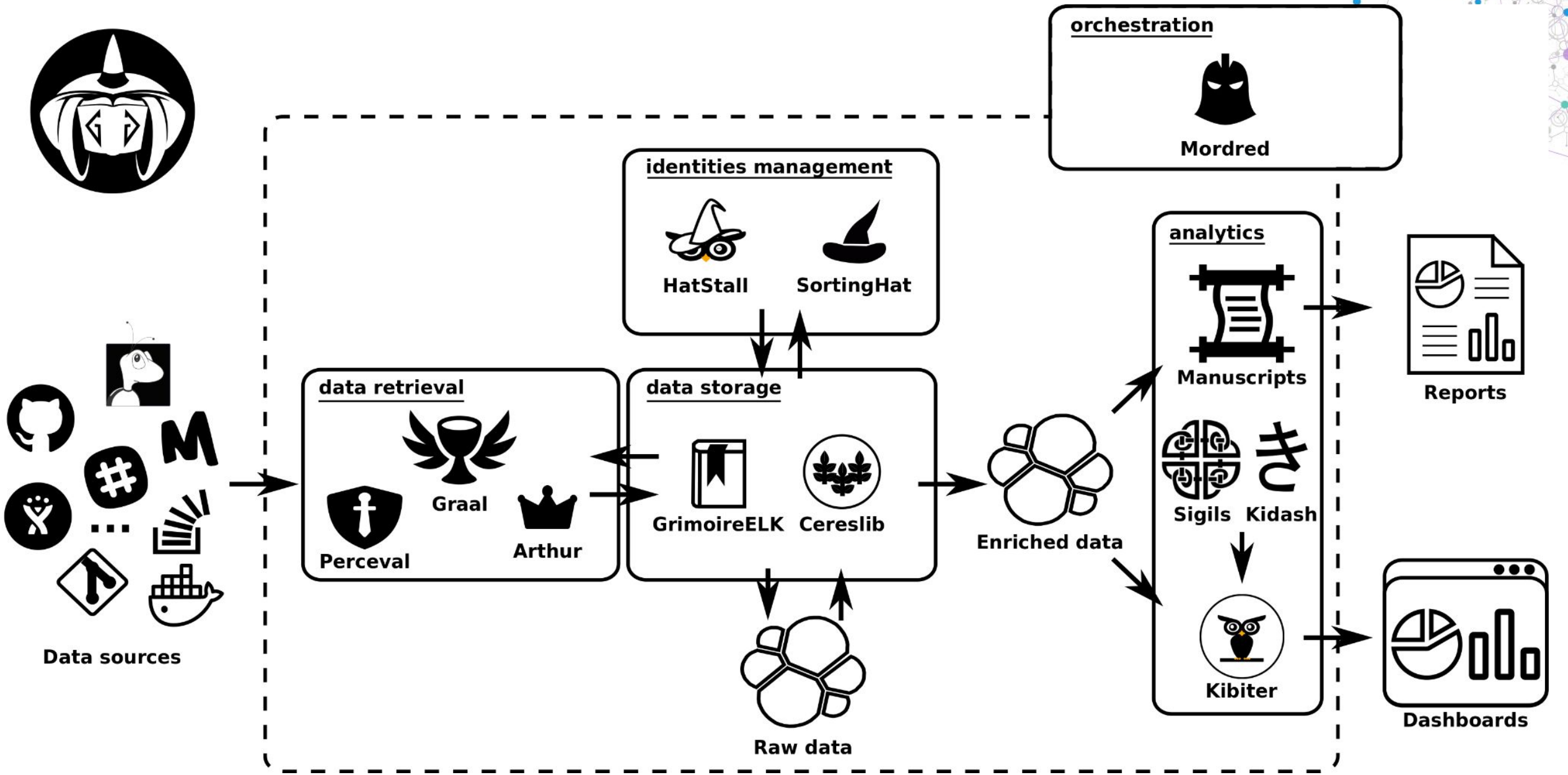
GrimoireLab



Data sources



CHAOSS



CHAOSS

github.com/chaoss/grimoirelab#grimoirelab-components

Demo of GrimoireLab

sneak-peek

CHAOSS

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

CHAOSS

Overview - CHAOSS Bitergia Ana

chaoss.biterg.io/app/kibana#/dashboard/Overview?_g=h@44136fa&_a=h@9b19fb9

Guest

CHAOSS

Overview | Data Sources | Community | Activity | Performance | Data Status | About

Dashboard / Overview

Full screen | Share | Edit | Documentation | Auto-refresh | Last 5 years

>_ *

Options | Update

NOT Empty Commits | NOT Bots | Add a filter +

Actions

Git Top Authors

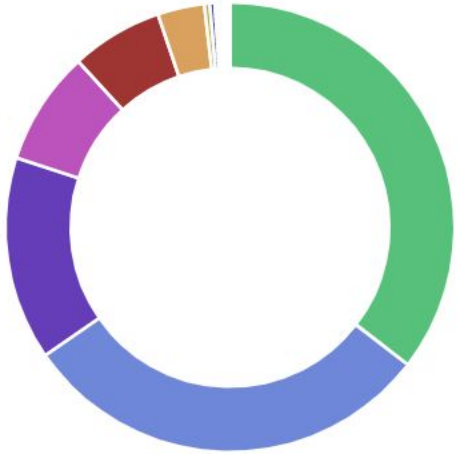
Search

Author	Commits	Projects	Added Lines	Removed Lines	Avg. Files
Valerio Cosentino	2,276	3	226818	151945	2.416
Alvaro del Castillo	1,978	1	904297	193310	1.941
Sean Patrick Goggins	1,476	7	1636574	782028	4.978
Santiago Dueñas	1,240	2	103864	345982	3.317
Georg Link	1,036	10	163499	27671	1.619

Export: Raw Formatted

1 2 3 4 »

Organizations



- Bitergia
- Unknown
- University of Missouri
- Elasticsearch
- University of Nebrask...
- CHAOSS Mentees
- University of Victoria
- Pivotal
- CHAOSS
- EZserver
- GitLab
- GSoC
- Red Hat
- Acquia
- O2
- The Scale Factory
- IBM
- Independent
- Linux Foundation
- University of Mons

Git Top Projects

Search

Project	Commits	Authors	Repositories
GrimoireLab	16,996	353	19
Augur	4,529	72	2
Coordination	1,608	119	3
Evolution WG	570	27	1
D&I WG	374	28	1
Metrics TC	181	28	1
Risk WG	162	11	1
	24,886	**675**	**32**

Export: Raw Formatted

1 2 »

Git

GitHub Pull Requests

Mailing Lists

How to Get Started

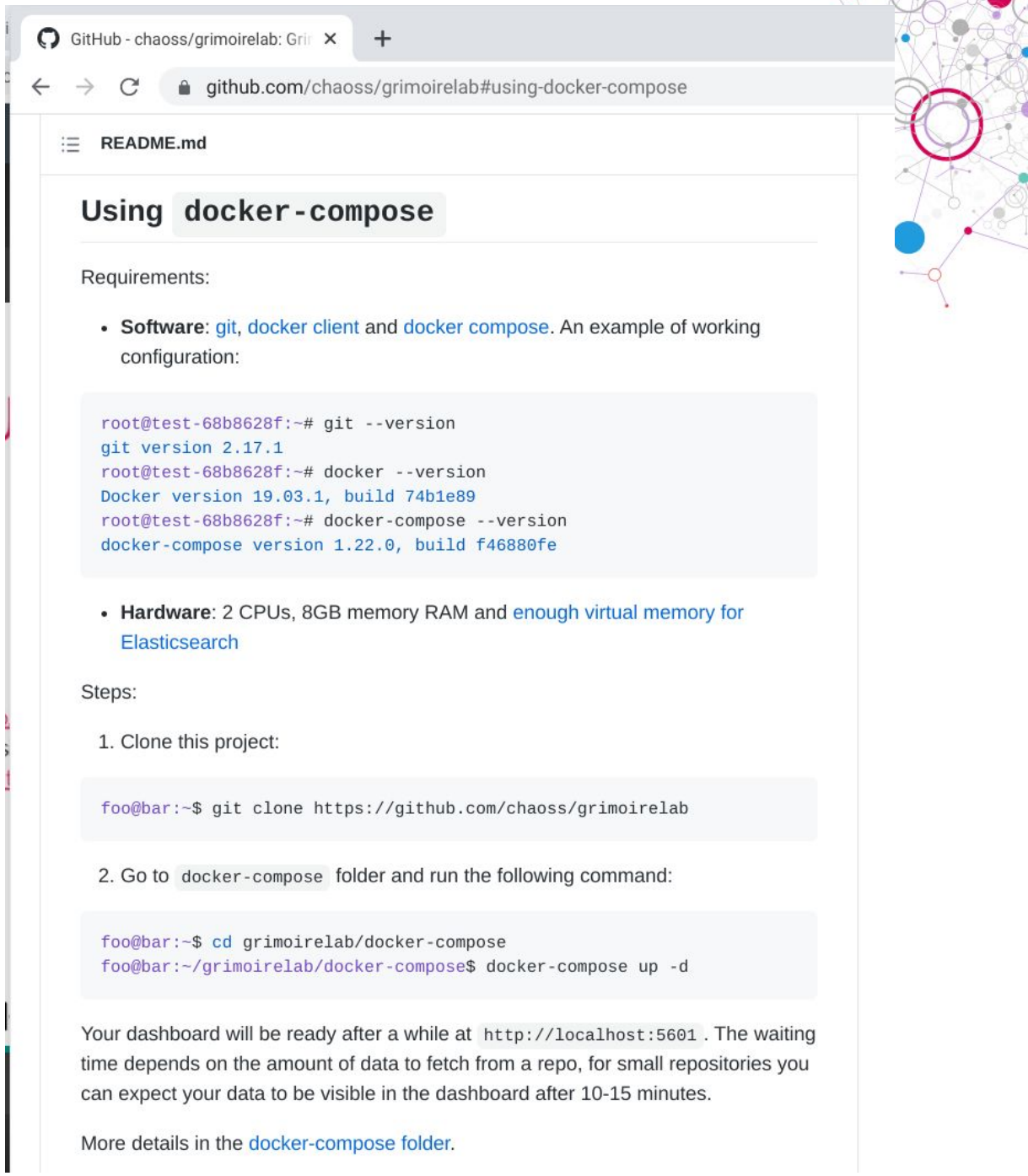
CHAOSS

How to Get Started

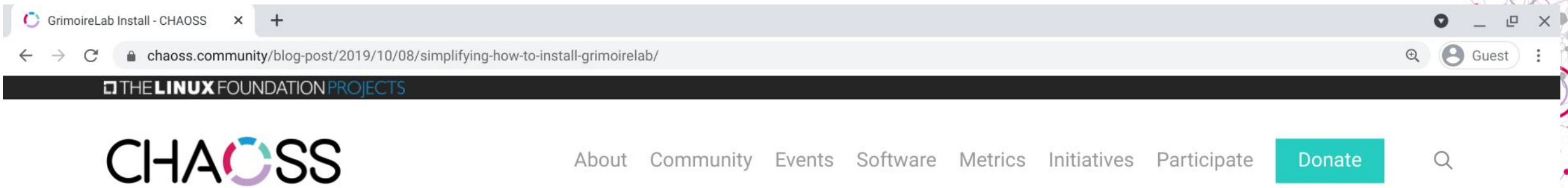
1. docker-compose (easiest of all)
2. docker run
3. pip packages
4. source code (mainly, for developers)

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



CHAOSS



Blog Post

GrimoireLab Install

By CHAOSS Community | October 8, 2019

Simplifying how to install GrimoireLab

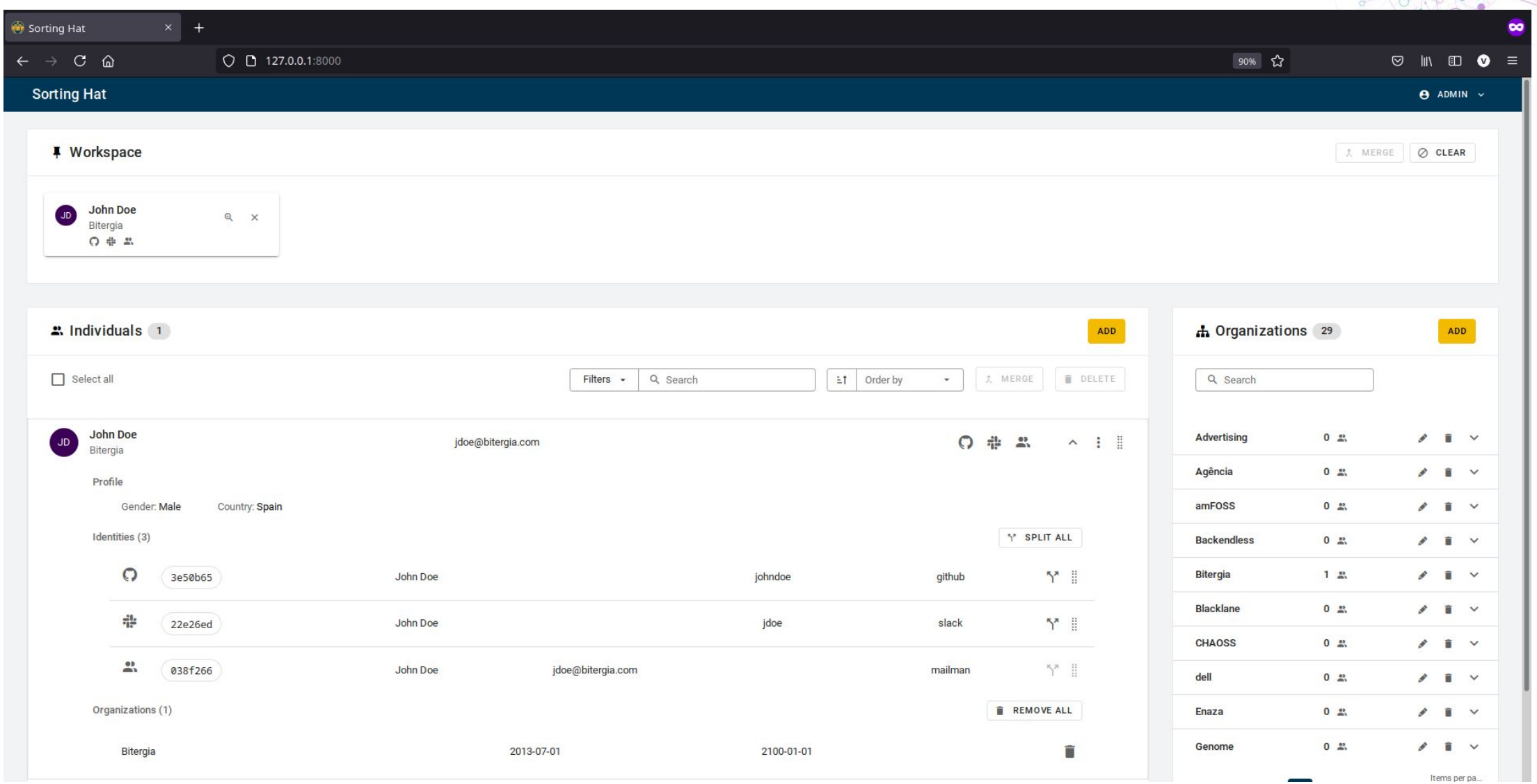
By Georg Link

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.

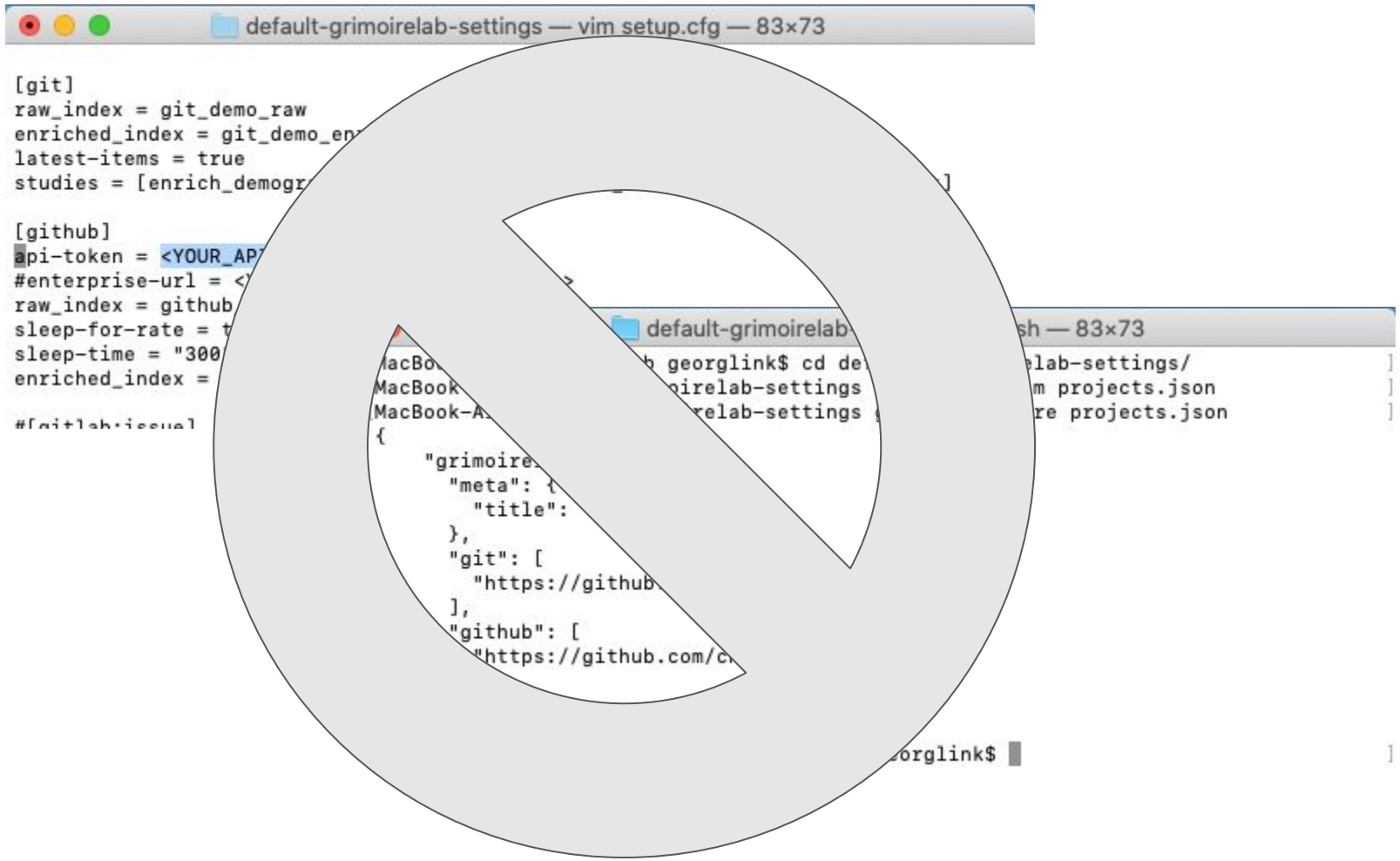
Current Work

CHAOSS

SortingHat (muggle branch)



Bestiary (unicorn branch)



CHAOSS

Google Summer of Code 2021

STUDENT

Rashmi K A

PROJECT NAME

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

ORGANIZATION

CHAOSS Project

MENTORS

Venu Vardhan Reddy

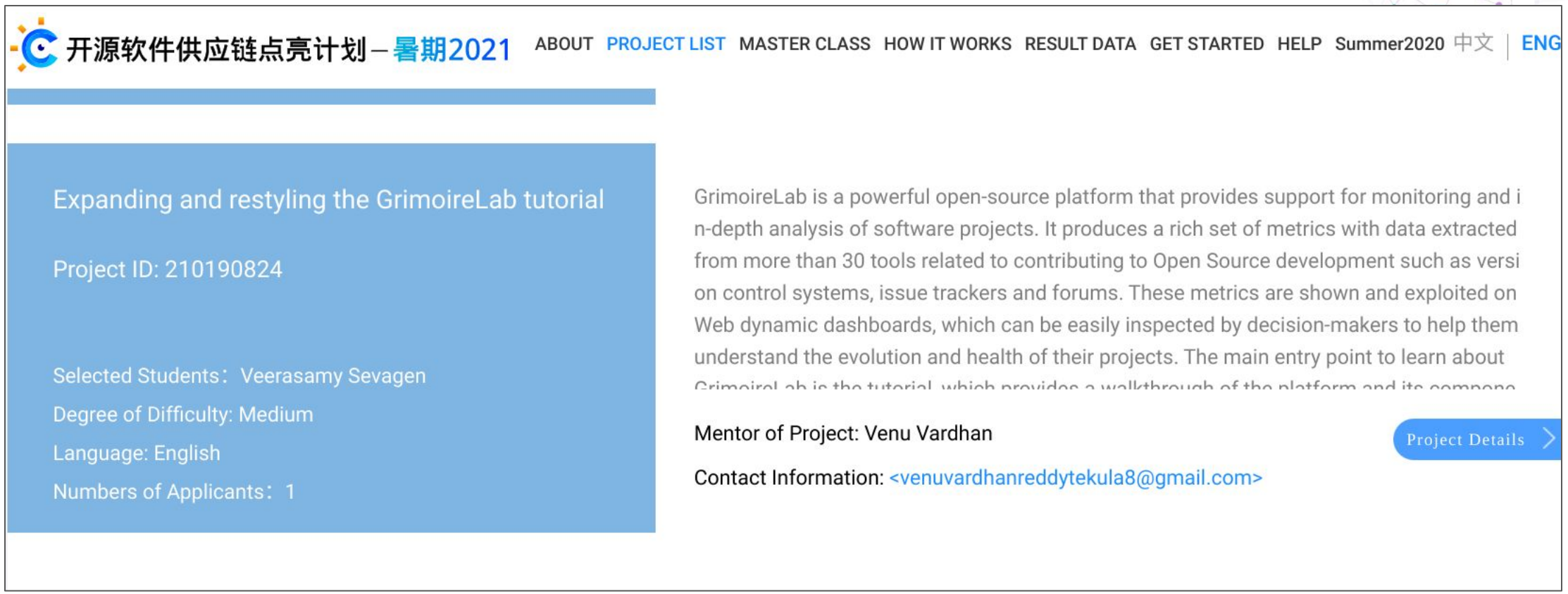
SHARE LINK

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.

CHAOSS

Summer 2021 Open Source Promotion Plan



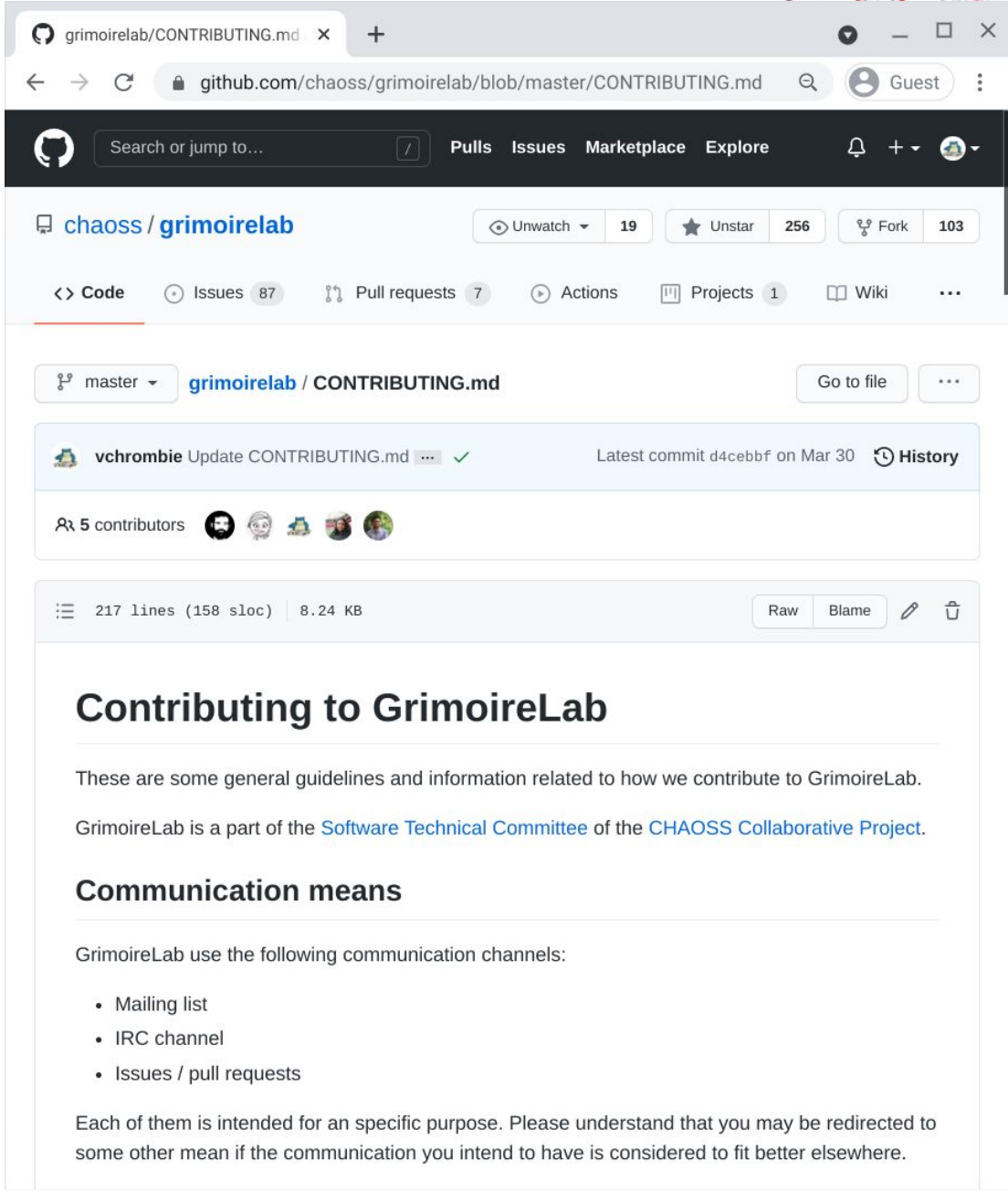
CHAOSS

How to Get Involved

CHAOSS

Check out the CONTRIBUTING Guidelines.

Feel free to contact the maintainers.



CHAOSS

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

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

CHAOSS

Questions?

xiè xie nǐ

Twitter: @vchrombie

Mail to: venu@chaoss.community

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