License Changes, Forks, and Project Health

OpenUK Online Meetup July 2024
Dr. Dawn M. Foster
Director of Data Science for CHAOSS

https://chaoss.community/
https://github.com/chaoss
@chaoss@fosstodon.org

https://fastwonderblog.com/
https://github.com/geekygirldawn
@geekygirldawn@hachyderm.io

CHAOSS

Thank You!

Whoami

- Geek, traveler, reader
- 20+ yr tech career focused on community & open source (VMware, Intel, Puppet, ...)
- OpenUK Board
- CHAOSS Board and Maintainer
- CNCF TAG Contrib Strategy co-chair
- PhD on Linux kernel collaboration

Photos by Mom, Josh Bancroft, Don Park

Agenda

- **Company Dynamics:**
  Increased potential for license change / fork / other disruptions

- **Assessing Organizational Participation Data**

- **Impact on Community and Project Health**

CHAOSS

@geekygirldawn

Ownership:

Foundation vs. Company

Company Controlled

A single company in control has a higher risk of forks and other disruptions.

CHAOSS

@geekygirldawn

Neutral Foundations

Forks are less common in open source projects under neutral foundations.

@geekygirldawn

Assessing Organizational

Participation

CHAOSS

CHAOSS: Community Health Analytics for Open Source Software

We are:

- An Open Source, Linux Foundation Project

- Globally distributed with 2000+ members in Slack

- A lovely place to spend your time

We are NOT:

- Not focused on medical health

- Not for-profit with any specific mandate

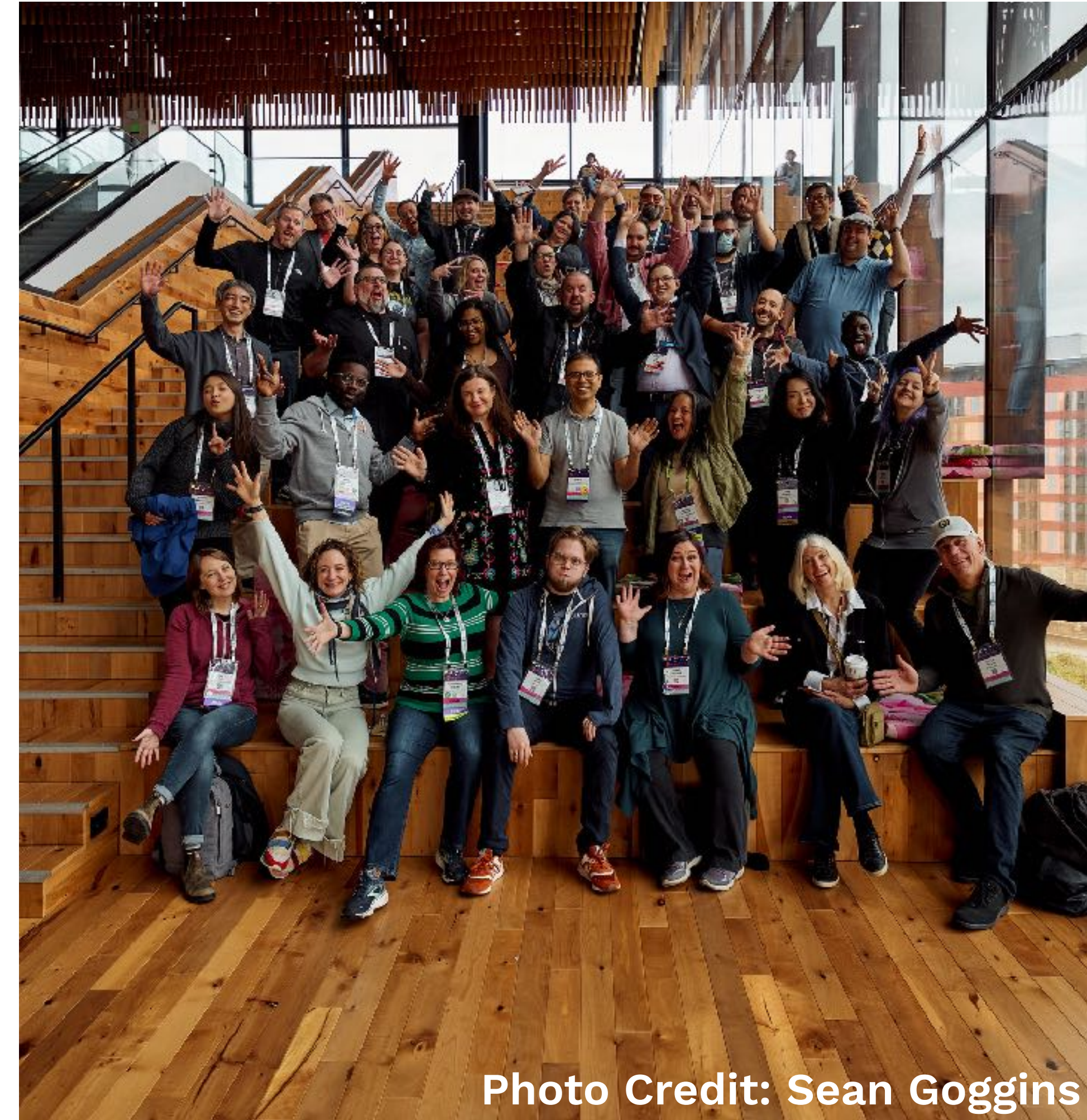- Not only focused on code contributions

Photo Credit: Sean Goggins

https://chaoss.community/kb-getting-started/

@geekygirldawn

New CHAOSS Practitioner Guides

Practitioner Guides:

- Introduction

- Responsiveness

- Contributor Sustainability

- **Organizational Participation**

- More guides coming soon

Photo by Martin Wilner on Unsplash

CHAOSS

https://chaoss.community/about-chaoss-practitioner-guides/

@geekygirldawn

Organizational Participation

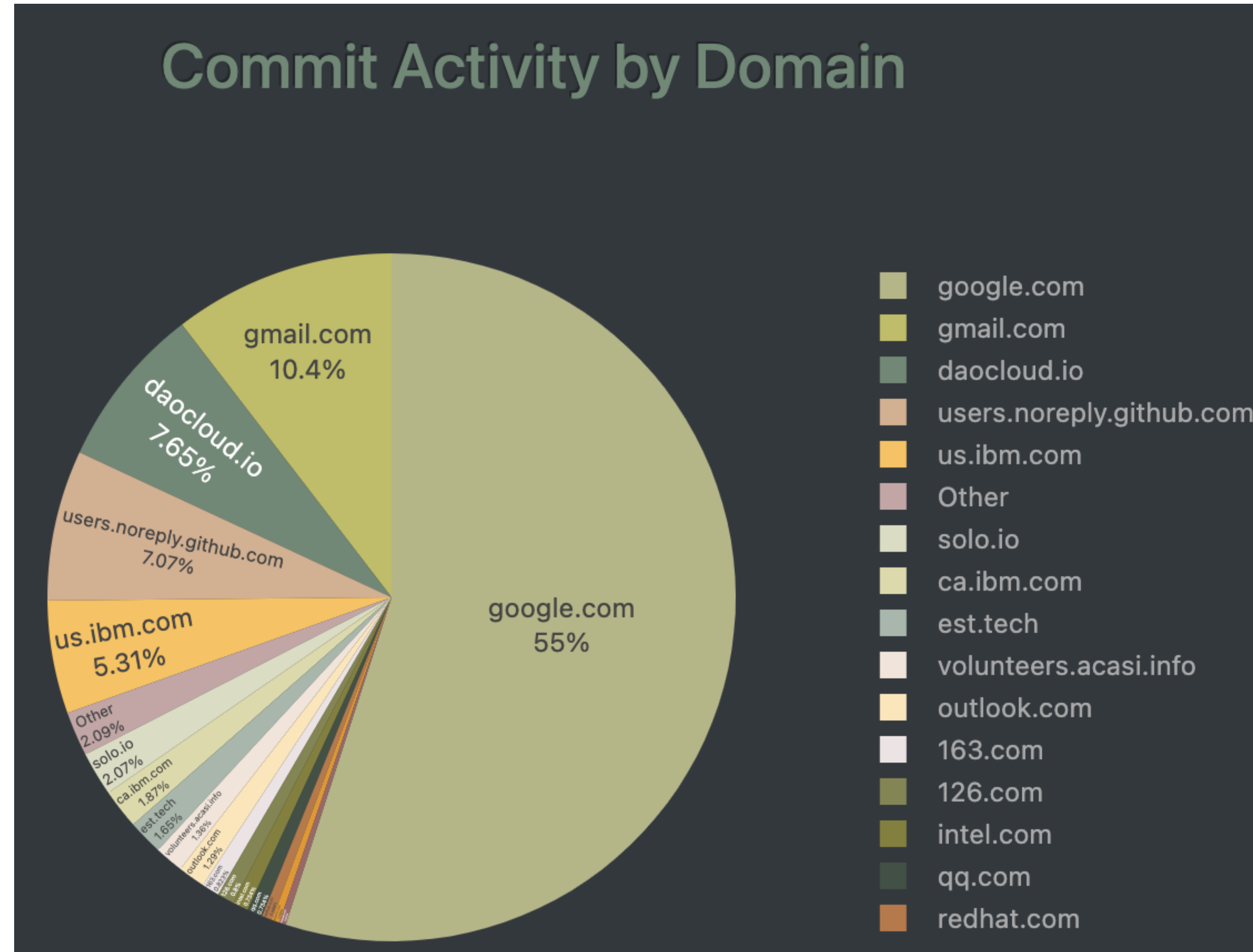Projects dominated by a single organization might be less sustainable and more likely to fork.

Key metrics:

- Elephant Factor
- Organizational Influence
- Organizational Diversity

CHAOSS https://chaoss.community/practitioner-guide-organizational-participation/

@geekygirldawn

Organizational Participation

Getting clean organizational affiliation data can be a challenge. It just needs to be good enough to understand the dominant players.

Commit Activity by Domain

- google.com
- gmail.com
- daocloud.io
- users.noreply.github.com
- us.ibm.com
- Other
- solo.io
- ca.ibm.com
- est.tech
- volunteers.acasi.info
- outlook.com
- 163.com
- 126.com
- intel.com
- qq.com
- redhat.com

google.com 55%
gmail.com 10.4%
daocloud.io 7.65%
users.noreply.github.com 7.07%
us.ibm.com 5.31%
Other 2.09%
solo.io 2.07%
ca.ibm.com 1.87%
est.tech 1.65%
volunteers.acasi.info 1.39%
outlook.com 1.29%

Impact on Community and Project Health

Acquisition: Contributors Drove Fork

Both are successful.
MariaDB has a larger
contributor community,
but MySQL continues
to be widely used.

Acquisition: Contributors Drove Fork

LibreOffice became more successful.

Both contributors and users embraced the fork.

Relicensing: Users Drove Fork

Relicense had bigger impact on users than contributors. OpenSearch is gaining traction with open source folks.

Strategy / Relicense: CNCF -> Company

Foundation projects with a single company dominating the contributions are still at risk of a fork.

CHAOSS License Change Dataset

We welcome contributions!

project	relicense_date	orig_license	new_license
Airbyte	2021-09-27	MIT	Elastic License v2.0
Akka	2022-09-07	Apache-2.0	Business Source License
ArangoDB	2024-02-23	Apache-2.0	Business Source License
Aseprite	2016-08-29	GPL-2.0	EULA that permits personal use but forbids redistribution
CockroachDB	2019-06-04	Apache-2.0	Business Source License
Confluent	2018-12-14	Apache-2.0	Confluent Community License Agreement
Consul	2023-08-10	MPL-2.0	Business Source License
Elasticsearch	2021-02-03	Apache-2.0	Elastic License and Server Side Public License
Elasticsearch	2021-02-03	Apache-2.0	Elastic License and Server Side Public License
LiveCode	2021-08-31	GPL-3.0-only	project archived
MongoDB	2018-10-16	AGPL-3.0-only	Server Side Public License
OTRS	2021-01-27	GPL-3.0-or-later	project archived
Redis	2024-03-20	BSD-3-Clause	dual: custom license and Server Side Public License
Sentry	2019-11-06	BSD-3-Clause	Business Source License
Sourcegraph	2018-10-27	Apache-2.0	Sourcegraph Enterprise license
Terraform	2023-08-10	MPL-2.0	Business Source License
Timescale	2018-12-29	Apache-2.0	Timescale License
Vagrant	2023-08-10	MIT	Business Source License
Vault	2023-08-10	MPL-2.0	Business Source License

https://github.com/chaoss/wg-data-science/tree/main/dataset/license-changes

@geekygirldawn

Resources

chaoss.community

chaoss.community/about-chaoss-practitioner-guides

chaoss.community/chaoss-guide-oss-viability

github.com/chaoss/wg-data-science/tree/main/dataset

CHAOSS

@geekygirldawn

Final Thoughts

As companies face increasing
pressure to generate revenue
from open source efforts,
it's likely that we'll continue
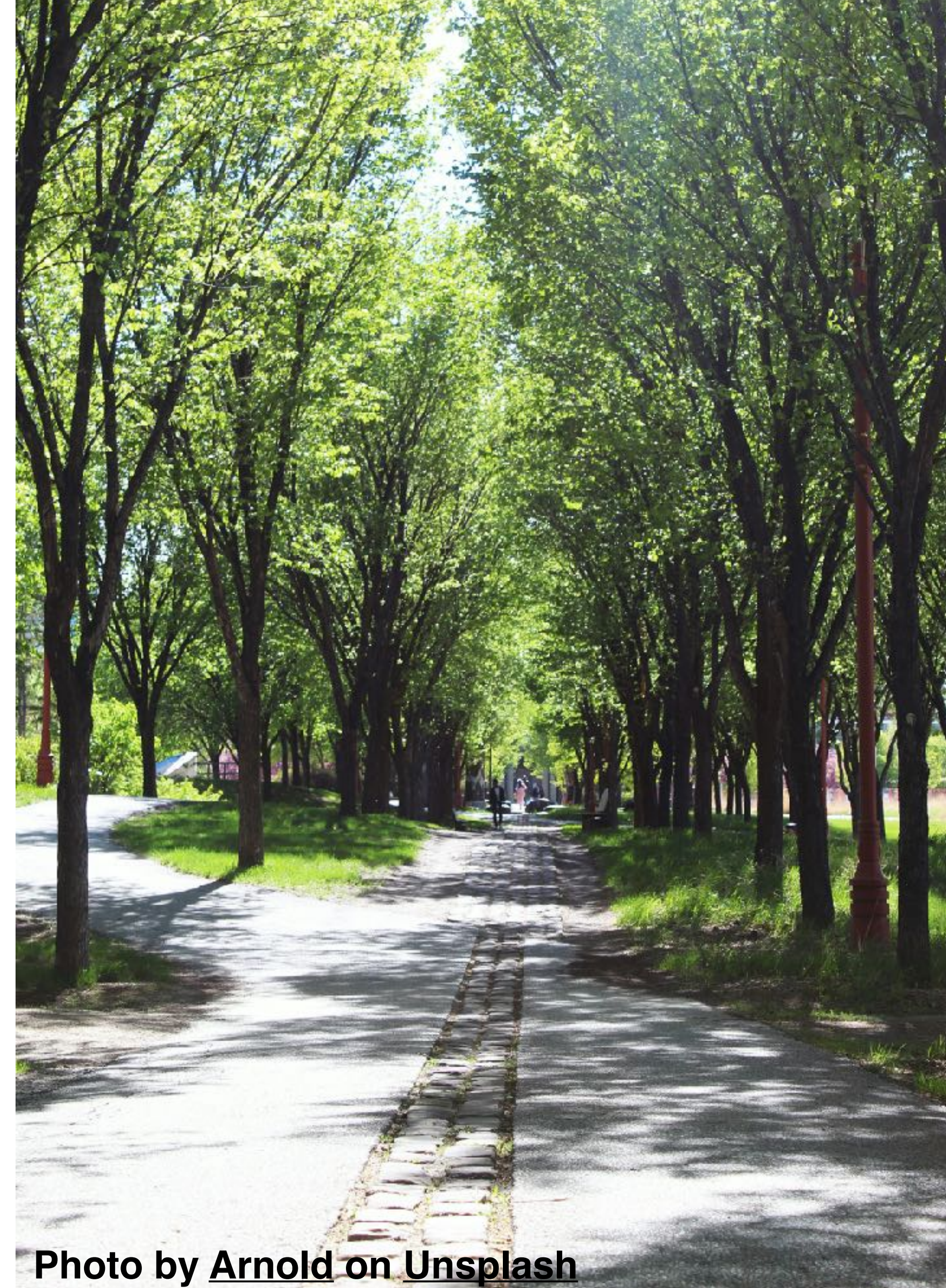to see an increase in
license changes and forks.

CHAOSS

https://chaoss.community/about-chaoss-practitioner-guides/

@geekygirldawn

THANK YOU!
Any Questions?

🌐 https://chaoss.community/
🐙 https://github.com/chaoss
🐘 @chaoss@fosstodon.org

🌐 https://fastwonderblog.com/
🐙 https://github.com/geekygirldawn
🐘 @geekygirldawn@hachyderm.io

CHAOSS