# License Changes and Forking Data

## OpenUK Research Show & Tell Sept 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!

CHAOSS

@geekygirldawn

# Whoami



- Geek, traveler, reader
- 20+ yr tech career focused on 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

# Assessing Organizational

# Participation

# Company Controlled OSS

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

CHAOSS

@geekygirldawn

# Organizational Participation

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

Key metrics:

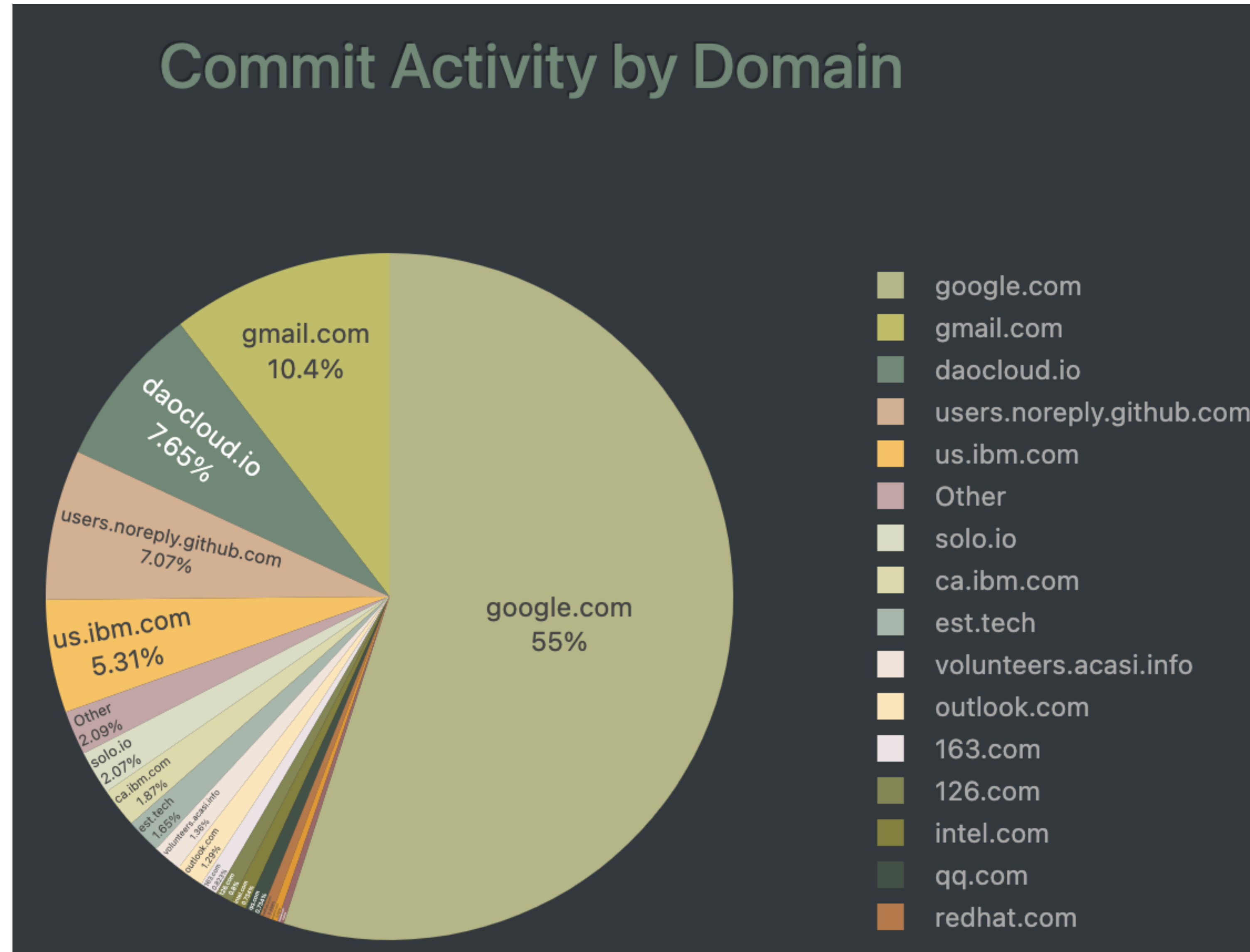- Elephant Factor
- Organizational Influence
- Organizational Diversity

CHAOSS

https://chaoss.community/practitioner-guide-organizational-participation/
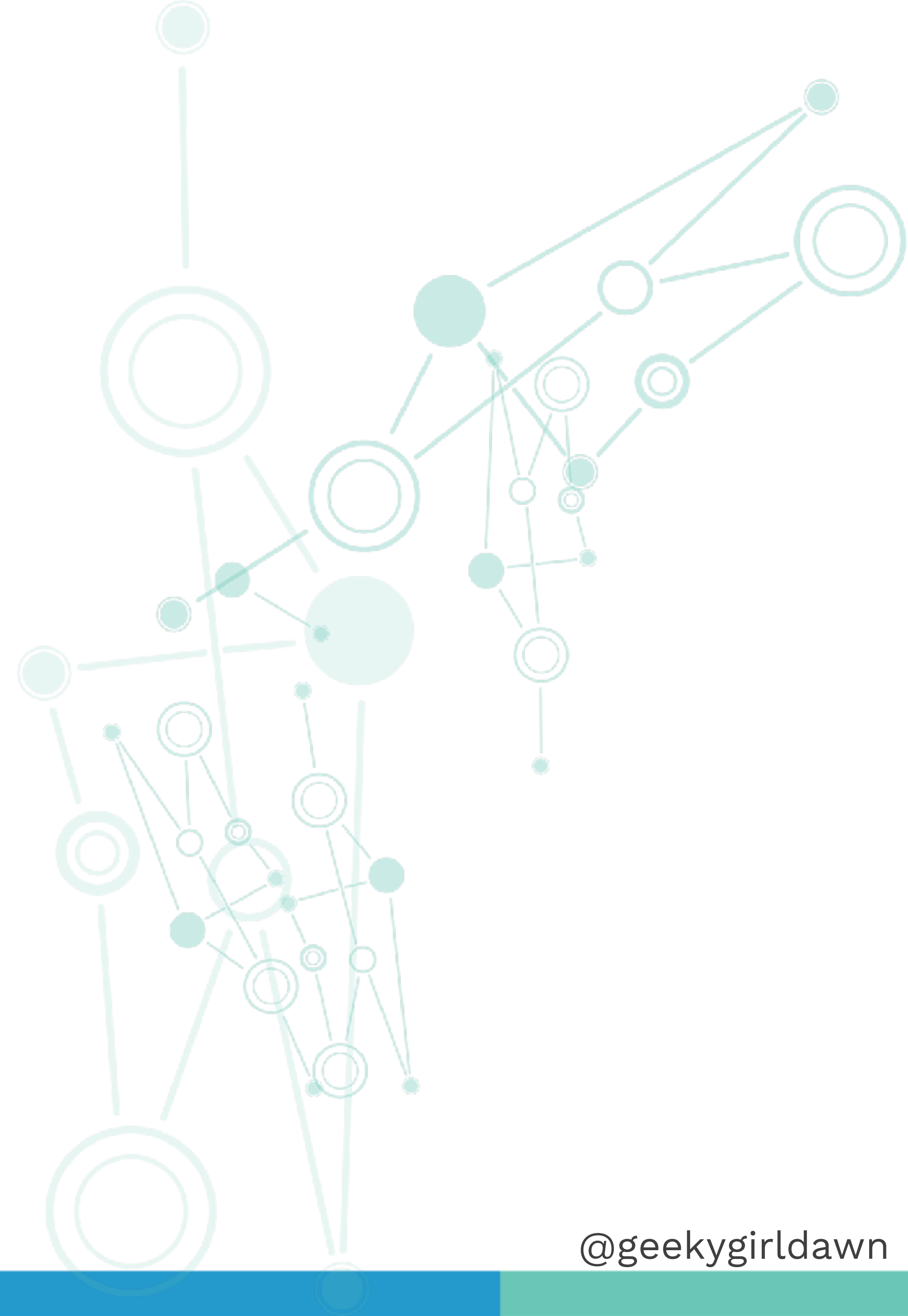
@geekygirldawn

# 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



Pie chart: Commit Activity by Domain
- 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%

Legend:
- 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

# Impacts and Data

# About the Data
## (Contributions Welcome!)

### Caveat:

- **All of this is WIP data that hasn't been validated by the projects.**

### Current Focus:

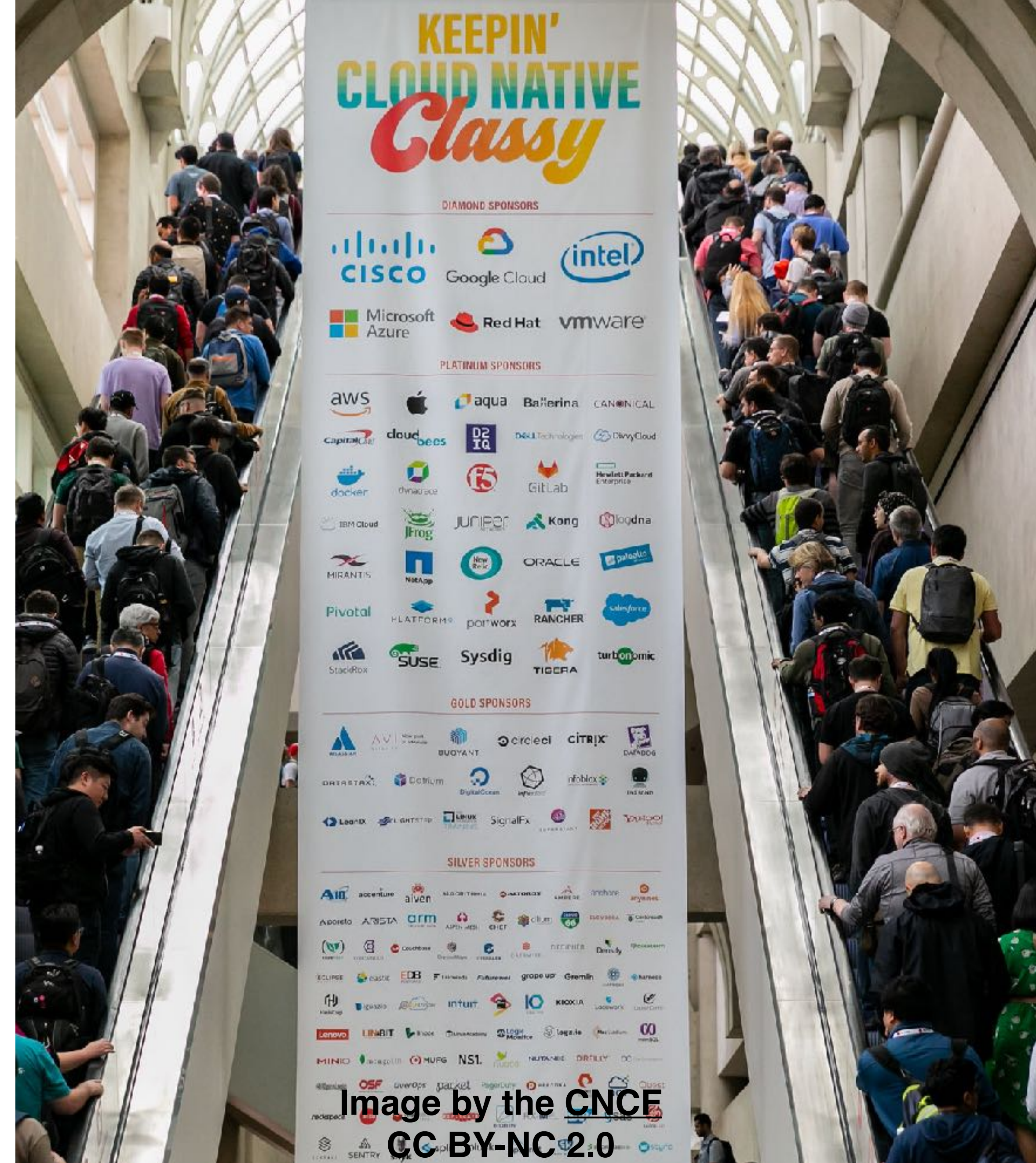- **Organizations / Employers (challenging to gather)**

### Commits:

- **Used for thresholds**

### Additions / Deletions:

- **Lines of code for magnitude**

### Coming Soon:

- **Additional community health measurements**



Image by the CNCF
CC BY-NC 2.0

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

@geekygirldawn

# Relicensing: Users Drove Fork

Relicense had bigger impact on users than contributors.



**Key Dates:**

- **2021-02-03: Elastic relicenses to SSPL & Elastic**
- **2021-04-12: OpenSearch fork owned by AWS**
- **2024-08-29: Elastic adds AGPL**
- **2024-09-16: OpenSearch moves under LF**



CHAOSS

@geekygirldawn

# Elasticsearch repo (https://github.com/elastic/elasticsearch)

## 1 year before orig relicense (2020-02-03 to 2021-02-03)

- Elastic employees with >= 10 commits (67 people):

  - 96% of total additions

  - 97% of total deletions

- Other people with >= 10 commits (3 people):

  - 0.3% of total additions

  - 0.1% of total deletions

## 1 year before AGPL relicense (2023-08-29 to 2024-08-29)

- Elastic employees with >= 10 commits (99 people):

  - 95% of total additions

  - 97% of total deletions

- Other people with >= 10 commits (1 person):

  - 0.01% of total additions

  - 0.03% of total deletions

CHAOSS

@geekygirldawn

# OpenSearch Repo (https://github.com/opensearch-project/OpenSearch)

## First Year of the Fork (2021-04-12 to 2022-04-12)

- Amazon / AWS Employees with >= 10 commits (7 people):

  - 80% of total additions

  - 91% of total deletions

- Other people with >= 10 commits (2 people):

  - 7% of total additions

  - 4% of total deletions

## Final Year under AWS before LF (2023-09-16 to 2024-09-16)

- Amazon / AWS Employees with >= 10 commits (40 people):

  - 63% of total additions

  - 65% of total deletions

- Other people with >= 10 commits (6 people):

  - 11% of total additions

  - 13% of total deletions

https://github.com/chaoss/wg-data-science/blob/main/dataset/license-changes/fork-case-study/notebooks/OpenSearch.ipynb

@geekygirldawn

# Relicensing: Contributors Drove Fork

The external contributors making significant contributions to Redis have moved to Valkey.

**Key Dates:**
- **2024-03-20: Redis relicenses to SSPL & RSAL**
- **2024-03-28: Valkey fork launched as an LF project**

# Redis Repo (https://github.com/redis/redis)

## 1 year before relicense (2023-03-20 - 2024-03-20)

- Redis Employees with >= 5 commits (6 people):
  - 80% of total additions (189,656)
  - 74 % of total deletions (83,122)
- Other people with >= 5 commits (12 people):
  - 12% of total additions
  - 15% of total deletions

## After relicense (2024-03-20 - 2024-08-21)

- Redis Employees with >= 5 commits (7 people):
  - 75% of total additions (38,270)
  - 72% of total deletions (10,464)
- Other people with >= 5 commits (0):

@geekygirldawn

CHAOSS

# Valkey Repo (https://github.com/valkey-io/valkey )

**After 2024-03-28 Fork (until 2024-08-20):**

- **Companies with >= 5 commits (27 people at 9 companies):**

- **79% of total additions (70,708)**

- **76% of total deletions (50,158)**

| People | Company | Commits | Additions | Addition% | Deletions | Deletion% |
|---|---|---|---|---|---|---|
| 2 | Google | 27 | 40191 | 44.82% | 38406 | 58.35% |
| 12 | Amazon | 122 | 15072 | 16.81% | 5772 | 8.77% |
| 2 | Alibaba | 6 | 303 | 0.34% | 51 | 0.08% |
| 2 | ByteDance | 5 | 3913 | 4.36% | 562 | 0.85% |
| 4 | Huawei | 66 | 3205 | 3.57% | 2764 | 4.20% |
| 1 | Tencent Cloud | 46 | 1690 | 1.88% | 703 | 1.07% |
| 1 | Intel | 10 | 559 | 0.62% | 455 | 0.69% |
| 2 | Ericsson | 37 | 5671 | 6.32% | 1372 | 2.08% |
| 1 | @gnet-io | 8 | 104 | 0.12% | 73 | 0.11% |

CHAOSS

# Relicensing: Users Drove Fork

OpenTofu contributors were not Terraform contributors, so there was an impact to users, but not contributors.

**Key Dates:**
- **2023-08-10: Terraform relicenses to BSL**
- **2023-09-05: OpenTofu launched as fork under the LF**

CHAOSS

@geekygirldawn

# Terraform Repo (https://github.com/hashicorp/terraform)

## 1 year before relicense (2022-08-10 - 2023-08-10

- HashiCorp Employees with >= 5 commits (21 people):

  - 93% of total additions

  - 95% of total deletions

- Other - people with >= 5 commits (2):

  - 0.04% of total additions

  - 0.04% of total deletions

## After relicense (2023-08-10 to 2024-08-10)

- Hashicorp employees with >= 5 commits (24 people):

  - 90% of total additions

  - 93% of total deletions

- Other people with >= 5 commits (2 people - 1 might be employee):

  - 0.05% of total additions

  - 0.01% of total deletions

@geekygirldawn

# OpenTofu Repo (https://github.com/opentofu/opentofu)

## After 2023-09-05 fork until 2024-09-05:

- **Companies with >= 5 commits (27 people at 11 companies):**
  - **80% of total additions**
  - **82% of total deletions**

| People | Company | Commits | Additions | Addition% | Deletions | Deletion% |
|---|---|---|---|---|---|---|
| 1 | Chainguard | 6 | 266 | 0.17% | 93 | 0.10% |
| 3 | Harness | 17 | 2948 | 1.85% | 366 | 0.40% |
| 6 | Env0 | 99 | 26507 | 16.61% | 12248 | 13.23% |
| 8 | Spacelift | 305 | 81271 | 50.91% | 59111 | 63.87% |
| 1 | @claranet | 6 | 118 | 0.07% | 20 | 0.02% |
| 1 | @Cooby-inc | 5 | 72 | 0.05% | 69 | 0.07% |
| 1 | Hangzhou Dianzi University | 6 | 891 | 0.56% | 242 | 0.26% |
| 3 | Red Hat | 15 | 1605 | 1.01% | 159 | 0.17% |
| 1 | @OpenTofu | 44 | 12350 | 7.74% | 3373 | 3.64% |
| 1 | @lessops | 6 | 2017 | 1.26% | 226 | 0.24% |
| 1 | @nvdnc | 5 | 68 | 0.04% | 11 | 0.01% |

@geekygirldawn

# Resources

chaoss.community

github.com/chaoss/wg-data-science

github.com/chaoss/wg-data-science/tree/main/dataset/
license-changes/fork-case-study

chaoss.community/about-chaoss-practitioner-guides

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.
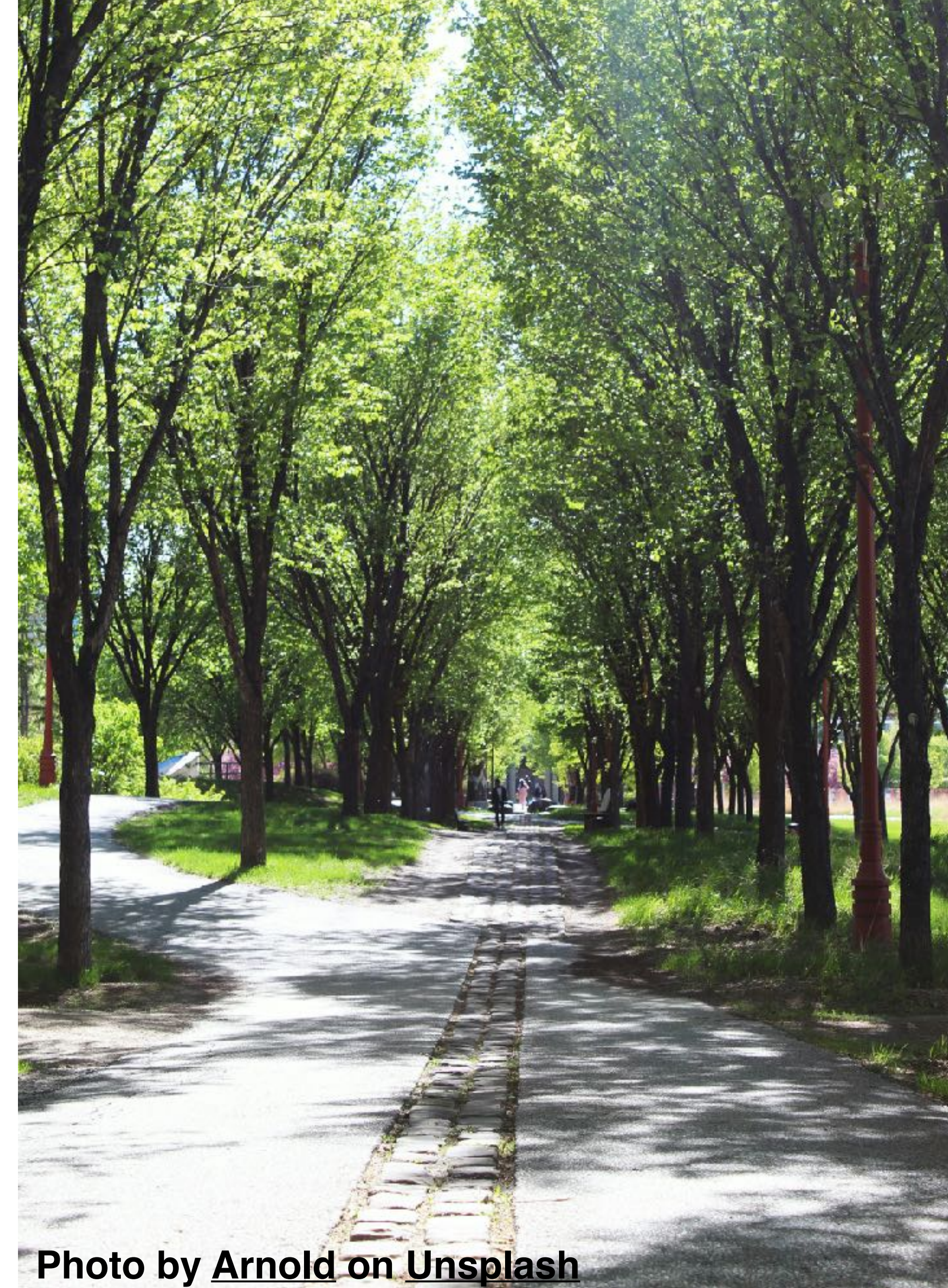
Photo by <u>Arnold</u> on <u>Unsplash</u>

CHAOSS

@geekygirldawn

# THANK YOU!
## Any Questions?

https://chaoss.community/          https://fastwonderblog.com/

https://github.com/chaoss          https://github.com/geekygirldawn

@chaoss@fosstodon.org          @geekygirldawn@hachyderm.io

CHAOSS