

TIL...: 10+ Findings from 2021 Global FLOSS Research in Five Minutes

Kaylea Champion – kaylea@uw.edu

<http://communitydata.science>

University of Washington

SeaGL—November 5-6, 2021



```
#!/usr/global/bin/research  
import FLOSS
```

Caveats:

1. Most of this is not my work!

So Much FLOSS in Research!

Caveats:

1. Most of this is not my work!
2. But it is still awesome!

So Much FLOSS in Research!

Caveats:

1. Most of this is not my work!
2. But it is still awesome!
3. Time is short!

So Much FLOSS in Research!

Caveats:

1. Most of this is not my work!
2. But it is still awesome!
3. Time is short!
4. Read these papers for full details!

So Much FLOSS in Research!

Caveats:

1. Most of this is not my work!
2. But it is still awesome!
3. Time is short!
4. Read these papers for full details!
5. Free box of exclamation points!

[Literally crumbling infrastructure. J.C. Burns (jcburns) via flickr, CC BY-NC-ND 2.0]

1 - Underproduction is a problem in FLOSS

- Who: Kaylea Champion & Benjamin Mako Hill (me! and my advisor)—Univ. Washington

K. Champion and B. Hill, “Underproduction: An Approach for Measuring Risk in Open Source Software,” in 2021 IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER), Honolulu, HI, USA, 2021 pp. 388-399. doi: 10.1109/SANER50967.2021.00043

1 - Underproduction is a problem in FLOSS

- Who: Kaylea Champion & Benjamin Mako Hill (me! and my advisor)—Univ. Washington
- Why: Looking for software infrastructure risk

K. Champion and B. Hill, “Underproduction: An Approach for Measuring Risk in Open Source Software,” in 2021 IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER), Honolulu, HI, USA, 2021 pp. 388-399. doi: 10.1109/SANER50967.2021.00043

1 - Underproduction is a problem in FLOSS

- Who: Kaylea Champion & Benjamin Mako Hill (me! and my advisor)—Univ. Washington
- Why: Looking for software infrastructure risk
- How: Studying packages and bugs in Debian

K. Champion and B. Hill, “Underproduction: An Approach for Measuring Risk in Open Source Software,” in 2021 IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER), Honolulu, HI, USA, 2021 pp. 388-399. doi: 10.1109/SANER50967.2021.00043

1 - Underproduction is a problem in FLOSS

- Who: Kaylea Champion & Benjamin Mako Hill (me! and my advisor)—Univ. Washington
- Why: Looking for software infrastructure risk
- How: Studying packages and bugs in Debian
- Key Finding: Substantial underproduction – highly important, low quality packages – is **widespread** in Debian

K. Champion and B. Hill, “Underproduction: An Approach for Measuring Risk in Open Source Software,” in 2021 IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER), Honolulu, HI, USA, 2021 pp. 388-399. doi: 10.1109/SANER50967.2021.00043

2 - Persisting FLOSS Communities

- Who: Wm Salt Hale—Univ. Washington

Hale, W. S. (2021). Resilience in Free/Libre/Open source software: Do founder decisions impact development activity after crisis events? μ

2 - Persisting FLOSS Communities

- Who: Wm Salt Hale—Univ. Washington
- Why: Does licensing help FLOSS communities develop?

Hale, W. S. (2021). Resilience in Free/Libre/Open source software: Do founder decisions impact development activity after crisis events? μ

2 - Persisting FLOSS Communities

- Who: Wm Salt Hale—Univ. Washington
- Why: Does licensing help FLOSS communities develop?
- How: Studying licensing of packages in Debian

Hale, W. S. (2021). Resilience in Free/Libre/Open source software: Do founder decisions impact development activity after crisis events? μ

2 - Persisting FLOSS Communities

- Who: Wm Salt Hale—Univ. Washington
- Why: Does licensing help FLOSS communities develop?
- How: Studying licensing of packages in Debian
- Key Finding: Protective licenses (like the GPL) are associated with **higher rates of development**.

Hale, W. S. (2021). Resilience in Free/Libre/Open source software: Do founder decisions impact development activity after crisis events? μ

3 - FLOSS Foundations

- Who: Tadeusz Chetkowski, Dariusz Jemielniak, Kacper Macikowski—Poland

Chetkowski T, Jemielniak D, Macikowski K (2021) Free and Open Source Software organizations: A large-scale analysis of code, comments, and commits frequency. PLoS ONE 16(9): e0257192. <https://doi.org/10.1371/journal.pone.0257192>

3 - FLOSS Foundations

- Who: Tadeusz Chełkowski, Dariusz Jemielniak, Kacper Macikowski—Poland
- Why: How are FLOSS foundations changing over time?

Chełkowski T, Jemielniak D, Macikowski K (2021) Free and Open Source Software organizations: A large-scale analysis of code, comments, and commits frequency. PLoS ONE 16(9): e0257192. <https://doi.org/10.1371/journal.pone.0257192>

3 - FLOSS Foundations

- Who: Tadeusz Chełkowski, Dariusz Jemielniak, Kacper Macikowski—Poland
- Why: How are FLOSS foundations changing over time?
- How: Code-oriented activity metrics from repositories for Apache, Debian, Eclipse, Gentoo, Gnome, JBoss, KDE, Mozilla, NASA, OpenStack, OSGeo, OW2, OWASP, TDF, Wikimedia

Chełkowski T, Jemielniak D, Macikowski K (2021) Free and Open Source Software organizations: A large-scale analysis of code, comments, and commits frequency. PLoS ONE 16(9): e0257192. <https://doi.org/10.1371/journal.pone.0257192>

3 - FLOSS Foundations

- Who: Tadeusz Chełkowski, Dariusz Jemielniak, Kacper Macikowski—Poland
- Why: How are FLOSS foundations changing over time?
- How: Code-oriented activity metrics from repositories for Apache, Debian, Eclipse, Gentoo, Gnome, JBoss, KDE, Mozilla, NASA, OpenStack, OSGeo, OW2, OWASP, TDF, Wikimedia
- Key Finding: Activity levels are **flat or declining** at these long-standing organizations.

Chełkowski T, Jemielniak D, Macikowski K (2021) Free and Open Source Software organizations: A large-scale analysis of code, comments, and commits frequency. PLoS ONE 16(9): e0257192. <https://doi.org/10.1371/journal.pone.0257192>

4 - Onboarding Contributors

- Who: Hyuga Horiguchi, Itsuki Omori, and Masao Ohira—Japan

H. Horiguchi, I. Omori and M. Ohira, “Onboarding to Open Source Projects with Good First Issues: A Preliminary Analysis,” 2021 IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER), 2021, pp. 501-505, doi: 10.1109/SANER50967.2021.00054.

4 - Onboarding Contributors

- Who: Hyuga Horiguchi, Itsuki Omori, and Masao Ohira—Japan
- Why: FLOSS communities need new members

H. Horiguchi, I. Omori and M. Ohira, “Onboarding to Open Source Projects with Good First Issues: A Preliminary Analysis,” 2021 IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER), 2021, pp. 501-505, doi: 10.1109/SANER50967.2021.00054.

4 - Onboarding Contributors

- Who: Hyuga Horiguchi, Itsuki Omori, and Masao Ohira—Japan
- Why: FLOSS communities need new members
- How: 11 GitHub repos that made heavy use of the 'Good First Issue' tag.

H. Horiguchi, I. Omori and M. Ohira, "Onboarding to Open Source Projects with Good First Issues: A Preliminary Analysis," 2021 IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER), 2021, pp. 501-505, doi: 10.1109/SANER50967.2021.00054.

4 - Onboarding Contributors

- Who: Hyuga Horiguchi, Itsuki Omori, and Masao Ohira—Japan
- Why: FLOSS communities need new members
- How: 11 GitHub repos that made heavy use of the 'Good First Issue' tag.
- Key Finding: Tagging bugs in the bug tracker as e.g. 'Good First Issue' helps people get started—but doesn't bring in documentation help.

H. Horiguchi, I. Omori and M. Ohira, "Onboarding to Open Source Projects with Good First Issues: A Preliminary Analysis," 2021 IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER), 2021, pp. 501-505, doi: 10.1109/SANER50967.2021.00054.

5 - Modern Code Review Antipatterns

- Who: Moataz Chouchen, Ali Ouni, Raula Gaikovina Kula, Dong Wang, Patanamon Thongtanunam, Mohamed Wiem Mkaouer and Kenichi Matsumoto—Tunisia, Japan, Australia, USA

M. Chouchen et al., “Anti-patterns in Modern Code Review: Symptoms and Prevalence,” 2021 IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER), 2021, pp. 531-535, doi: 10.1109/SANER50967.2021.00060.

5 - Modern Code Review Antipatterns

- Who: Moataz Chouchen, Ali Ouni, Raula Gaikovina Kula, Dong Wang, Patanamon Thongtanunam, Mohamed Wiem Mkaouer and Kenichi Matsumoto—Tunisia, Japan, Australia, USA
- Why: Key practice for improving quality.

M. Chouchen et al., “Anti-patterns in Modern Code Review: Symptoms and Prevalence,” 2021 IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER), 2021, pp. 531-535, doi: 10.1109/SANER50967.2021.00060.

5 - Modern Code Review Antipatterns

- Who: Moataz Chouchen, Ali Ouni, Raula Gaikovina Kula, Dong Wang, Patanamon Thongtanunam, Mohamed Wiem Mkaouer and Kenichi Matsumoto—Tunisia, Japan, Australia, USA
- Why: Key practice for improving quality.
- How: Existing lit + sample from Open Stack

M. Chouchen et al., “Anti-patterns in Modern Code Review: Symptoms and Prevalence,” 2021 IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER), 2021, pp. 531-535, doi: 10.1109/SANER50967.2021.00060.

5 - Modern Code Review Antipatterns

- Who: Moataz Chouchen, Ali Ouni, Raula Gaikovina Kula, Dong Wang, Patanamon Thongtanunam, Mohamed Wiem Mkaouer and Kenichi Matsumoto—Tunisia, Japan, Australia, USA
- Why: Key practice for improving quality.
- How: Existing lit + sample from Open Stack
- Key Finding: there are **recognizable patterns** in bad code reviews (e.g.: divergent review, low review participation, shallow review, toxic review)

M. Chouchen et al., “Anti-patterns in Modern Code Review: Symptoms and Prevalence,” 2021 IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER), 2021, pp. 531-535, doi: 10.1109/SANER50967.2021.00060.

6 - Consequences of Toxicity

- Who: Isabella Ferreira, Jinghui Cheng, Bram Adams—Canada

Isabella Ferreira, Jinghui Cheng, and Bram Adams. 2021. The “Shut the f**k up” Phenomenon: Characterizing Incivility in Open Source Code Review Discussions. Proc. ACM Hum.-Comput. Interact. 5, CSCW2, Article 353 (October 2021), 35 pages.
<https://doi.org/10.1145/3479497>

6 - Consequences of Toxicity

- Who: Isabella Ferreira, Jinghui Cheng, Bram Adams—Canada
- Why: What happens in a volunteer project when civility breaks down?

Isabella Ferreira, Jinghui Cheng, and Bram Adams. 2021. The “Shut the f**k up” Phenomenon: Characterizing Incivility in Open Source Code Review Discussions. Proc. ACM Hum.-Comput. Interact. 5, CSCW2, Article 353 (October 2021), 35 pages.
<https://doi.org/10.1145/3479497>

6 - Consequences of Toxicity

- Who: Isabella Ferreira, Jinghui Cheng, Bram Adams—Canada
- Why: What happens in a volunteer project when civility breaks down?
- How: Analysis of Linux kernel mailing list patch rejections

Isabella Ferreira, Jinghui Cheng, and Bram Adams. 2021. The “Shut the f**k up” Phenomenon: Characterizing Incivility in Open Source Code Review Discussions. Proc. ACM Hum.-Comput. Interact. 5, CSCW2, Article 353 (October 2021), 35 pages.
<https://doi.org/10.1145/3479497>

6 - Consequences of Toxicity

- Who: Isabella Ferreira, Jinghui Cheng, Bram Adams—Canada
- Why: What happens in a volunteer project when civility breaks down?
- How: Analysis of Linux kernel mailing list patch rejections
- Key Finding: Toxic rejections often **drive people away**.

Isabella Ferreira, Jinghui Cheng, and Bram Adams. 2021. The “Shut the f**k up” Phenomenon: Characterizing Incivility in Open Source Code Review Discussions. Proc. ACM Hum.-Comput. Interact. 5, CSCW2, Article 353 (October 2021), 35 pages.
<https://doi.org/10.1145/3479497>

7 - Python Security

- Who: Mahmoud Alfadel, Diego Elias Costa and Emad Shihab—Canada

M. Alfadel, D. E. Costa and E. Shihab, “Empirical Analysis of Security Vulnerabilities in Python Packages,” 2021 IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER), 2021, pp. 446-457, doi: 10.1109/SANER50967.2021.00048.

7 - Python Security

- Who: Mahmoud Alfadel, Diego Elias Costa and Emad Shihab—Canada
- Why: Python is a key technology

M. Alfadel, D. E. Costa and E. Shihab, “Empirical Analysis of Security Vulnerabilities in Python Packages,” 2021 IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER), 2021, pp. 446-457, doi: 10.1109/SANER50967.2021.00048.

7 - Python Security

- Who: Mahmoud Alfadel, Diego Elias Costa and Emad Shihab—Canada
- Why: Python is a key technology
- How: 550 vulnerabilities in 252 packages on PyPi

M. Alfadel, D. E. Costa and E. Shihab, “Empirical Analysis of Security Vulnerabilities in Python Packages,” 2021 IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER), 2021, pp. 446-457, doi: 10.1109/SANER50967.2021.00048.

7 - Python Security

- Who: Mahmoud Alfadel, Diego Elias Costa and Emad Shihab—Canada
- Why: Python is a key technology
- How: 550 vulnerabilities in 252 packages on PyPi
- Key Finding: [security fixes are slow!](#) Vulnerabilities on average take more than 3 years to be discovered and most are only fixed after being announced.

M. Alfadel, D. E. Costa and E. Shihab, “Empirical Analysis of Security Vulnerabilities in Python Packages,” 2021 IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER), 2021, pp. 446-457, doi: 10.1109/SANER50967.2021.00048.

8 - Impact of COVID on FLOSS

- Who: Linux Foundation; Google

8 - Impact of COVID on FLOSS

- Who: Linux Foundation; Google
- Why: Massive shift to remote work amid stressful conditions

8 - Impact of COVID on FLOSS

- Who: Linux Foundation; Google
- Why: Massive shift to remote work amid stressful conditions
- How: Repository activity metrics

8 - Impact of COVID on FLOSS

- Who: Linux Foundation; Google
- Why: Massive shift to remote work amid stressful conditions
- How: Repository activity metrics
- Key Findings: Linux Kernel – largely unchanged. Google-affiliated – increased activity.

9 - Great Maintainers

- Who: Edson Dias, Paulo Meirelles, Fernando Castor, Igor Steinmacher, Igor Wiese, Gustavo Pinto
—USA and Brazil

E. Dias, P. Meirelles, F. Castor, I. Steinmacher, I. Wiese and G. Pinto, "What Makes a Great Maintainer of Open Source Projects?," 2021 IEEE/ACM 43rd International Conference on Software Engineering (ICSE), 2021, pp. 982-994, doi: 10.1109/ICSE43902.2021.00093.

9 - Great Maintainers

- Who: Edson Dias, Paulo Meirelles, Fernando Castor, Igor Steinmacher, Igor Wiese, Gustavo Pinto
—USA and Brazil
- Why: What makes great maintainers great?

E. Dias, P. Meirelles, F. Castor, I. Steinmacher, I. Wiese and G. Pinto, "What Makes a Great Maintainer of Open Source Projects?," 2021 IEEE/ACM 43rd International Conference on Software Engineering (ICSE), 2021, pp. 982-994, doi: 10.1109/ICSE43902.2021.00093.

9 - Great Maintainers

- Who: Edson Dias, Paulo Meirelles, Fernando Castor, Igor Steinmacher, Igor Wiese, Gustavo Pinto
—USA and Brazil
- Why: What makes great maintainers great?
- How: Interviews with maintainers of top projects and survey of contributors

E. Dias, P. Meirelles, F. Castor, I. Steinmacher, I. Wiese and G. Pinto, "What Makes a Great Maintainer of Open Source Projects?," 2021 IEEE/ACM 43rd International Conference on Software Engineering (ICSE), 2021, pp. 982-994, doi: 10.1109/ICSE43902.2021.00093.

9 - Great Maintainers

- Who: Edson Dias, Paulo Meirelles, Fernando Castor, Igor Steinmacher, Igor Wiese, Gustavo Pinto
—USA and Brazil
- Why: What makes great maintainers great?
- How: Interviews with maintainers of top projects and survey of contributors
- Key Finding: **Communication** is the most important and essential trait—but the role also requires technical skill, vision for the project, and attention to detail.

E. Dias, P. Meirelles, F. Castor, I. Steinmacher, I. Wiese and G. Pinto, "What Makes a Great Maintainer of Open Source Projects?," 2021 IEEE/ACM 43rd International Conference on Software Engineering (ICSE), 2021, pp. 982-994, doi: 10.1109/ICSE43902.2021.00093.

10 - Contributor Motivations

- Who: Marco Gerosa, Igor Wiese, Bianca Trinkenreich, Georg Link, Gregorio Robles, Christoph Treude, Igor Steinmacher, Anita Sarma—USA, Brazil, Spain, Australia

M. Gerosa et al., “The Shifting Sands of Motivation: Revisiting What Drives Contributors in Open Source,” 2021 IEEE/ACM 43rd International Conference on Software Engineering (ICSE), 2021, pp. 1046-1058, doi: 10.1109/ICSE43902.2021.00098.

10 - Contributor Motivations

- Who: Marco Gerosa, Igor Wiese, Bianca Trinkenreich, Georg Link, Gregorio Robles, Christoph Treude, Igor Steinmacher, Anita Sarma—USA, Brazil, Spain, Australia
- Why: World has changed since the bulk of this research was done 20 years ago.

M. Gerosa et al., “The Shifting Sands of Motivation: Revisiting What Drives Contributors in Open Source,” 2021 IEEE/ACM 43rd International Conference on Software Engineering (ICSE), 2021, pp. 1046-1058, doi: 10.1109/ICSE43902.2021.00098.

10 - Contributor Motivations

- Who: Marco Gerosa, Igor Wiese, Bianca Trinkenreich, Georg Link, Gregorio Robles, Christoph Treude, Igor Steinmacher, Anita Sarma—USA, Brazil, Spain, Australia
- Why: World has changed since the bulk of this research was done 20 years ago.
- How: Survey of 242 contributors

M. Gerosa et al., “The Shifting Sands of Motivation: Revisiting What Drives Contributors in Open Source,” 2021 IEEE/ACM 43rd International Conference on Software Engineering (ICSE), 2021, pp. 1046-1058, doi: 10.1109/ICSE43902.2021.00098.

10 - Contributor Motivations

- Who: Marco Gerosa, Igor Wiese, Bianca Trinkenreich, Georg Link, Gregorio Robles, Christoph Treude, Igor Steinmacher, Anita Sarma—USA, Brazil, Spain, Australia
- Why: World has changed since the bulk of this research was done 20 years ago.
- How: Survey of 242 contributors
- Key Finding: Experts shift toward altruism over time, newer folks are drawn toward social rewards and having fun

M. Gerosa et al., “The Shifting Sands of Motivation: Revisiting What Drives Contributors in Open Source,” 2021 IEEE/ACM 43rd International Conference on Software Engineering (ICSE), 2021, pp. 1046-1058, doi: 10.1109/ICSE43902.2021.00098.

In Other News....2021 gave us:

- FLOSS in Space and on Mars

In Other News....2021 gave us:

- FLOSS in Space and on Mars
- OSPOs (Open Source Program Offices) Are Everywhere

In Other News....2021 gave us:

- FLOSS in Space and on Mars
- OSPOs (Open Source Program Offices) Are Everywhere
- Mozilla's Continuing Success with Outreachy Internships

In Other News....2021 gave us:

- FLOSS in Space and on Mars
- OSPOs (Open Source Program Offices) Are Everywhere
- Mozilla's Continuing Success with Outreachy Internships
- Cyberattacks—Software Supply Chain Security—Software Bill of Materials



Neglected packages need help!

Protective licenses are powerful!

Foundation activity is flattening!

Good First Issue brings in newcomers!

Code review anti-patterns persist!

Toxic environments lose contributors!

Python security work needed!

We prevailed despite COVID!

Great maintainers communicate precisely!
We are altruistic experts + joyful newcomers!

Space-OSPOs-Outreachy-Supply Chain!

Thank You!

`kaylea@uw.edu`—`@kayleachampion`

<https://communitydata.science>

`@comdatasci`