Why Do People Give Up FLOSSing?
A Study of Contributor Disengagement in Open Source

Courtney Miller, David Gray Widder
Christian Kästner, Bogdan Vasilescu
Many OSS Projects are Analogous to Critical Infrastructure
Survival Rate Among FLOSS Projects is Very Low
Having Contributors in Drivers Seat is Critical to Project Survival

41% of failed OSS projects cited developer issues
Coelho, J., & Valente, M. T. (2017, August)
Contributor On-Boarding Has Been Thoroughly Researched
The Body of Work on Established Contributors is Much Smaller
We’re Interested in Exploring Factors to Predict Disengagement from OSS
General Methodology

Survey of Disengaged Contributors

Literature Review

Operationalization of Predictive Factors

Survival Model
Used GHTorrent to Collect Survey Sample

All Public GH Contributors

Publicly Displayed Contact Information

Regular and Active Contributor

Random Sample of Population

Survey Sample
Used GHTorrent to Collect Survey Sample

All Public GH Contributors

Publicly Displayed Contact Information

Regular and Active Contributor

Random Sample of Population

Survey Sample
Identifying Regular and Active Contributors Who Disengaged

- > 100 Commits
  - 6 mo.
- > 100 Commits
  - 6 mo.
- > 100 Commits
  - 6 mo.
- < 5 Commits
  - 6 mo.

Present
Used GHTorrent to Collect Survey Sample

All Public GH Contributors

Publicly Displayed Contact Information

Regular and Active Contributor

Random Sample of Population

Survey Sample
Used GHTorrent to Collect Survey Sample

All Public GH Contributors

Regular and Active Contributor

Publicly Displayed Contact Information

Random Sample of Population

Survey Sample
Used GHTorrent to Collect Survey Sample

All Public GH Contributors

Publicly Displayed Contact Information

Regular and Active Contributor

Random Sample of Population

Survey Sample
Sent Out Simple One-Question Survey to Sample

“Could you help us understand your reasons for reducing your contributions to GitHub projects?”

21.5% Response Rate
We Used Card Sorting Analysis to Parse Raw Survey Responses
Macro Survey Results

Survey Says Out of 151 Respondents…

- **70%** Occupational Issues
- **34%** Social Issues
- **25%** Technical Issues
## Breakdown of Survey Responses

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Occupational reasons</strong></td>
<td></td>
</tr>
<tr>
<td>Got new job that doesn’t support FLOSS</td>
<td>37</td>
</tr>
<tr>
<td>Changed role/project</td>
<td>25</td>
</tr>
<tr>
<td>Left job where they contributed to FLOSS</td>
<td>16</td>
</tr>
<tr>
<td>No time: new job</td>
<td>15</td>
</tr>
<tr>
<td>No time: existing job</td>
<td>10</td>
</tr>
<tr>
<td>Left school where they contributed to FLOSS</td>
<td>12</td>
</tr>
<tr>
<td>No time: in school</td>
<td>12</td>
</tr>
<tr>
<td>FLOSS in school, now job doesn’t support FLOSS</td>
<td>7</td>
</tr>
<tr>
<td>Too much coding at work</td>
<td>4</td>
</tr>
<tr>
<td><strong>Social reasons</strong></td>
<td></td>
</tr>
<tr>
<td>Lost interest in FLOSS</td>
<td>24</td>
</tr>
<tr>
<td>No time: personal</td>
<td>23</td>
</tr>
<tr>
<td>Lack of peer support</td>
<td>16</td>
</tr>
<tr>
<td>No time: nondescript</td>
<td>15</td>
</tr>
<tr>
<td><strong>Technical reasons</strong></td>
<td></td>
</tr>
<tr>
<td>Issues w GitHub or industry</td>
<td>14</td>
</tr>
<tr>
<td>Individually moved to private repos</td>
<td>12</td>
</tr>
<tr>
<td>Changed platform</td>
<td>10</td>
</tr>
<tr>
<td>Feature complete project</td>
<td>3</td>
</tr>
</tbody>
</table>
General Methodology

Survey of Disengaged Contributors

+ Literature Review

Operationalization of Predictive Factors

Survival Model
Related Research

When contributor works …

… what contributor does.
We Were Interested In Stratifying Survey Responses

H1:

Office Hours vs. Nights & Weekends

H2:

More Support Work vs. Less Support Work
General Methodology

Survey of Disengaged Contributors

+ Literature Review

→ Operationalization of Predictive Factors

→ Survival Model
workingHours

07:00

Office Hours

Nights & Weekends

19:00

FLOSS
supportWork

More Support Work

Less Support Work

Pull Requests
Issues
Etc.

Commits

New Code!
# Stratified Survey Results

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>Count</th>
<th>Office Hrs vs Nights&amp;We</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Occupational reasons</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Got new job that doesn’t support FLOSS</td>
<td>37</td>
<td></td>
</tr>
</tbody>
</table>
## Stratified Survey Results

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>Count</th>
<th>Office Hrs vs Nights &amp; We</th>
<th>More Support vs Less</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Occupational reasons</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Got new job that doesn’t support FLOSS</td>
<td>37</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
# Stratified Survey Results

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>Count</th>
<th>Office Hrs vs Nights&amp;We</th>
<th>More Support vs Less</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Occupational reasons</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Got new job that doesn’t support FLOSS</td>
<td>37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Changed role/project</td>
<td>25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left job where they contributed to FLOSS</td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No time: new job</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No time: existing job</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left school where they contributed to FLOSS</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No time: in school</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FLOSS in school, now job doesn’t support FLOSS</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Too much coding at work</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Social reasons</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lost interest in FLOSS</td>
<td>24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No time: personal</td>
<td>23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of peer support</td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No time: nondescript</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Technical reasons</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Issues w GitHub or industry</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Individually moved to private repos</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Changed platform</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feature complete project</td>
<td>3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
TransitionFound

A Good Example: Brett Cannon

Brett Cannon
Principal software engineer at Microsoft

Vancouver, BC, Canada

brettsyk
github.com/brettcannon
http://snarky.ca

Last seen on Stack Overflow 3 days ago

TECHNOLOGIES

Preferred technologies

- python
- dart
- typescript

Non-preferred technologies

- c++
- java

Top Percentiles

- Top 10%

EXPERIENCE (10)

Principal software engineer
Microsoft

Jul 2015 → Current (3 years, 11 months)

Fellow
Python Software Foundation

Mar 2003 → Current (16 years, 3 months)

Software Engineer
Google

May 2011 → Jun 2015 (4 years, 2 months)
Identifying Transitions

**Experience**

- **Principal software engineer**
  - Microsoft
  - Jul 2015 → Current (3 years, 11 months)

- **Fellow**
  - Python Software Foundation
  - Mar 2003 → Current (16 years, 3 months)

- **Software Engineer**
  - Google
  - May 2011 → Jun 2015 (4 years, 2 months)

  - python
  - go
  - c++
  - java
  - google-app-engine
  - google-chrome-extension
  - android

- Charged with finishing the master/slave to HRD migration tool for App Engine for public release
- Designed a novel algorithm to scrape web pages for [redacted] and implemented it as an entirely client-side Chrome extension which included live image cropping
- Heavily participated in and solely maintained an email parsing pipeline for purchase receipts to collect purchase-related data (e.g., shipment tracking numbers), which included adding support for schema markup and working with ontologists to add missing structured data support to schema.org
- Took the Google Now in-store card from a UX concept to implementation and through to launch which required writing both the Android client code along with the server code that worked with disparate backends
General Methodology

Survey of Disengaged Contributors + Literature Review → Operationalization of Predictive Factors → Survival Model
But What Exactly Is A Survival Model?

Organizational Affiliation Listed
No Organizational Affiliation Listed
Survival Model Factors

Activity
WorkingHours
SupportWork
Project Popularity

Organizational Affiliation
Team Size
Transition Found
Survival Model Participants

Survey Respondents with Transition Data

+ 

Control Group
Participants with Transition 2.48 Times More Likely to Disengage

“I started employment with a company that didn’t really like open sourcing their projects.” (P25)

“I’ve moved to more of a product owner role in my position, so less time to develop.” (P50)
Working Predominantly During Office Hours Increased Risk of Disengagement by Factor of 2.20

“[It’s] quite simple. I was active because the project was sponsored by my employer, i.e., my job. I'm not active on GitHub because I've moved to another project.” (P75)
Operationalization of Predictive Factors Provided Promising Results

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>Count</th>
<th>Office Hrs vs Nights &amp; We</th>
<th>More Support vs Less</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Occupational reasons</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Got new job that doesn’t support FLOSS</td>
<td>37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Changed role/project</td>
<td>25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left job where they contributed to FLOSS</td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No time: new job</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No time: existing job</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left school where they contributed to FLOSS</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No time: in school</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FLOSS in school, now job doesn’t support FLOSS</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Too much coding at work</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Social reasons</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lost interest in FLOSS</td>
<td>24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No time: personal</td>
<td>23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of peer support</td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No time: nondescript</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Technical reasons</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Issues w GitHub or industry</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Individually moved to private repos</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Changed platform</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feature complete project</td>
<td>3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Data Driven Operationalization Of Predictive Factors

Valuable Insights About Disengaged OSS Contributors
To learn more about sustained participation of contributors in FLOSS...

Friday, May 31st at 11:20, Sophie Qiu

Going Farther Together: The Impact of Social Capital on Sustained Participation in Open Source
Key Takeaways

Occupational reasons were most commonly cited in survey, specifically occupational transitions and novel transition. Found operationalization proved successful factor in survival model.

Used operationalizations of potential disengagement factors to (1) categorize survey respondents; and (2) predict disengagement.

Established contributors who do different kinds of work and perform said work at different times tend to cite different reasons for their disengagement.

Factors such as project popularity, when contributors work, and how much they work can also be used to predict established contributor disengagement from FLOSS.

Questions?

Courtney Elta Miller

@CourtneyElta

Courtney.Miller17@ncf.edu