Your Company Cares About Open Source Sustainability. But Are You Measuring and Encouraging Upstream Contributions?

CHAOSScon
@IndeedEng | opensource.indeed.com
Hi, I’m Dani.

Dani Gellis, Software Engineer
Indeed’s Open Source Program Office

@danisyellis | dgellis@indeed.com
Overview

- Program Overview
- What We Decided To Measure And Why
- Tools We Looked At, Implemented, And Built
- What We Learned So Far
- What We’ll Measure and Build Next
We help people get jobs.
Your Company’s Open Source “Legacy Code”

Boxcar
Launched: 2010
Deprecated: 2013

Proctor
Launched: 2013
2018: 1 Contributor
Total: 11 Contributors

Imhotep
Launched: 2014
1 known external user
Mistakes to Avoid when Open Sourcing Proprietary Tech


(or search for Open Source Summit North America 2018 Jack Humphrey)
indeed
open source
established:
November 2017
TO BUILD A CULTURE OF WIDESPREAD, ACTIVE AND RECURRING OPEN SOURCE PARTICIPATION
Why contribute back to projects?
IF YOU CAN’T MEASURE IT, YOU CAN’T IMPROVE IT - PETER DRUCKER
YOU GET WHAT YOU MEASURE
WHAT COUNTS AS A CONTRIBUTION?
WHAT NOT TO MEASURE AND WHY

Lines of Code

- Encourages Code Quantity With No Eye Toward Quality

Contribution Counts

- Competition Discourages New Contributors
- Projects Have To Manage Low Quality Contributions
ACTIVE RECURRING PARTICIPANTS

- **Active** - triggered an event that we care about
- **Recurring** - two or more days of activity
- **Participation** - a broad category of events
Contributions can be More than just Code
Tell us more about your open source contribution: *

Your answer

Are there any relevant links for the work we should be aware of? *

Your answer

Did you make a contribution this month? Please explain:

Your answer
Don’t Judge Candidates by Their GitHub Profiles: How open source participation plays into hiring

DUANE O'BRIEN
EVENTS WE’RE MEASURING, AND WHY

Events we are measuring:

- Comments
- Pull Requests
- Issues
- Code Reviews

Events we’re not:

- Push Events
Tools We Looked At, Implemented, And Built
THE CHALLENGE (AND OPPORTUNITY)

We’re measuring something different
THE CHALLENGE (AND OPPORTUNITY)

We’re measuring something different

What Existing Tools Measure: Open Source Project Health

- Org/Repo Focused

What We Want: Indeed Employee Contributions

- User Focused
How standards proliferate:

(See: A/C chargers, character encodings, instant messaging, etc.)

Situation:

There are 14 competing standards.

14?! Ridiculous!

We need to develop one universal standard that covers everyone's use cases. Yeah!

Soon:

Situation:

There are 15 competing standards.
The TODO Group is a group of companies who want to collaborate on practices, tools, and other ways to run successful and effective open source projects and programs.

https://todogroup.org
GitHub Metrics and Dashboards

- **oss-dashboard** - A dashboard for viewing many GitHub organizations, and/or users, at once.
- **osstracker** - OSS Tracker is an application that collects information about a GitHub organization and aggregates the data across all projects within that organization into a single user interface to be used by various roles within the owning organization.
- **ghcrawler** - GHcrawler is a GitHub API crawler that crawls a GitHub-hosted project and automatically tracks, retrieves, and stores its contents. GHcrawler is primarily intended for people trying to track sets of organizations and data repositories.
- **devstats** - A toolset to visualize GitHub archives using Grafana dashboards used by the Cloud Native Computing Foundation and Kubernetes
- **MeasureOSS** - A contributor relationship management system
- **GrimoireLab** - Software development analytics platform supporting more than 30 different data sources, part of CHA OSS Software project from The Linux Foundation

https://github.com/todogroup/awesome-oss-mgmt#github-metrics-and-dashboards
MICROSOFT’S GITHUB CRAWLER

Gets information about events on GitHub

- But, it's Org/Repo focused
- Yes, we CAN adapt it
- Pros and Cons
THE INTERIM DASHBOARD

This came first to get us some numbers

Nothing to look at, but very useful!
const eventsWeCareAbout = {
  CommitCommentEvent: 'comment',
  IssueCommentEvent: 'comment',
  IssuesEvent: 'issue',
  PullRequestEvent: 'pull_request',
  PullRequestReviewEvent: 'pull_request',
  PullRequestReviewCommentEvent: 'comment'
}
Dashboards
DASHBOARD SOLUTIONS

Tried MeasureOSS

- It was also Org/Repo focused
- Adding user views would have been confusing
DASHBOARD SOLUTIONS

Use Imhotep and IQL (Imhotep Query Language)

- 2 Different Dashboards
  - Contributor Dashboard
  - Aggregate Dashboard
Hmmm, I guess this initiative worked!

Up and to the right!

Source: The metrics I gathered
Employee Contributions Over Time
FROM ossContributions 2018-10-01 today GROUP BY time(1day) /* time ...

COUNT() by time(1day)

Graph
FROM ossContributions 2018-10-01 today GROUP BY repo SELECT count() as all, event_type="CommitCommentEvent" as commit_comments, event_type="IssueCommentEvent" as issue_comments, event_type="IssueEvent" as issues, event_type="PullRequestEvent" as pull_requests, event_type="PullRequestReviewCommentEvent" as pr_review_comments
Show 50 entries

<table>
<thead>
<tr>
<th>repo</th>
<th>all</th>
<th>commit_comments</th>
<th>issue_comments</th>
<th>issues</th>
<th>pull_requests</th>
<th>pr_review_comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>5kg/libswift-cv</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>9ballsyndrome/WebGL_Compute_shader</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ACollectionOfAtoms/disco-bot</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
CONTRIBUTOR DASHBOARD

Used by Individuals

Provides Contribution Insights

Useful For Quarterly Reviews
**Breakdown by GitHub Repository and Event Type**

<table>
<thead>
<tr>
<th>repo</th>
<th>all</th>
<th>commit_comments</th>
<th>issue_comments</th>
<th>issues</th>
<th>pull_requests</th>
<th>pr_review_comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>DuaneOBrien/ContributionVelocity</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>MeasureOSS/Measure</td>
<td>6</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Microsoft/ghcrawler</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>daniyellis/GithubAPIPaginator</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>daniyellis/ghcrawler</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>indeedeng/starfish</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>rhc2104/sqliteaching</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>usahidi/opendesign</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
What’s Next?
What Have We Learned So Far?
HIGHLY ACTIVE USERS : OUTLIERS

How Active? 10x? 100x!!!

Competition Discourages New Contributors

Don't Build For Edge Cases

Should Even Out Over Time
Employee Contributions - By Month

FROM ossContributions 2018-10-03 today GROUP BY time(1month) /* time ...
FOSS Contributor Fund

Learn more about the fund from these blog posts:

https://engineering.indeedblog.com/blog/2019/07/foss-fund-six-months-in/
https://blog.opencollective.com/indeeds-open-source-sustainability-strategy/
Run your own FOSS Contributor Fund - use starfish to find eligible voters

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>.env.template</td>
<td>Revert &quot;js&quot;</td>
<td>2 months ago</td>
</tr>
<tr>
<td>.gitignore</td>
<td>WIP Initial commit of working code that needs a good refactoring/desi...</td>
<td>3 months ago</td>
</tr>
<tr>
<td>CODE_OF_CONDUCT.md</td>
<td>bring repo up to OSS standards. Also, move timezone offset to env var...</td>
<td>2 months ago</td>
</tr>
<tr>
<td>LICENSE</td>
<td>bring repo up to OSS standards. Also, move timezone offset to env var...</td>
<td>2 months ago</td>
</tr>
<tr>
<td>OSSMETADATA</td>
<td>bring repo up to OSS standards. Also, move timezone offset to env var...</td>
<td>2 months ago</td>
</tr>
<tr>
<td>README.md</td>
<td>adds newline</td>
<td>2 months ago</td>
</tr>
<tr>
<td>index.js</td>
<td>make warning message clearer #2</td>
<td>2 months ago</td>
</tr>
<tr>
<td>package-lock.json</td>
<td>fix package vulnerabilities, including lodash</td>
<td>28 days ago</td>
</tr>
</tbody>
</table>
Free Soda when you Buy a Burger and Fries!
Virtual Community Building
Results to Date

#OSS members

#OSS-help members

Unique Views
Wiki Pages | Internal Blog

Participants in Hacktoberfest
And 24-Pull Requests
Q4 2018
Internal Events
Results to Date

Open Source Summit 2019

Attendees
Net Promoter Score

Fireside chats
In-Person Attendees
Virtual Attendees

/// OSS Contribution Workshop

Attendees
Thank You!

Let me know if you’d like to talk more about
- The FOSS Contributor Fund
- How you can measure upstream contributions
  - Other ways we can work together to support open source sustainability

Dani Gellis  |  @danisyellis  |  dgellis@indeed.com