Making Your First Commit To Open Source

A Casual Conversation

AI Alliance Office Hours
20 Feb 2025

Agenda

- About Me / My Patron
- Choosing your ~~mark~~ project
- Open Source Politika
- What project leads care about (What I care about when I review a PR)
- Roadmap to getting involved with OSS projects
- Reviewing A PR

About me and the patron of my artistic endeavors

Trevor Grant

- Member of the Apache Software Foundation
- PMC (Project Mentorship Committee) Apache Mahout
  - Formerly of Apache Streams (project retired), and Apache Community Development (resigned), and mentor for Apache SDAP-incubating(graduated)
- Project Lead of `gofannon` (An AI-Alliance Project)
- Author of [Kubeflow for Machine Learning: From Lab to Production](#)
- From Chicago, best. city. ever.

About me **and the patron** of my artistic endeavors

I am a line item in the IBM Research budget

They believe that I work for something called The AI Alliance

The AI Alliance is an international community of researchers, developers and organizational leaders committed to support and enhance open innovation across the AI technology landscape to accelerate progress, improve safety, security and trust in AI, and maximize benefits to people and society everywhere. Members of the AI Alliance believe that open innovation is essential to develop and achieve safe and responsible AI that benefit society rather than benefit a select few big players.
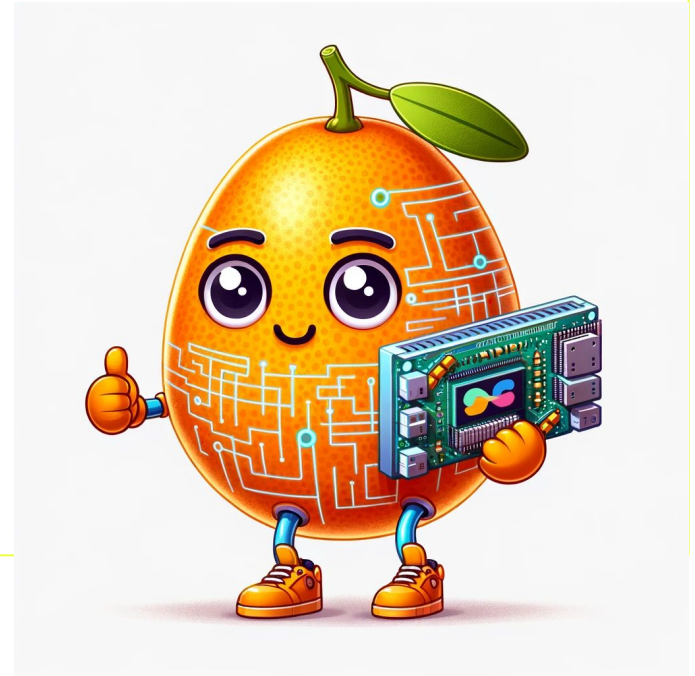
AI Alliance

About me and the patron of **my artistic endeavors**

Gofannon

Apache Mahout (qumat)

Choosing A Project

- Project Size / Popularity
  - Large / Popular / Established / commercial projects
    - Apache Spark | KServe | Data Prep Kit
  - Small / Less popular / Start up or maintenance mode / hobby projects
    - Gofannon | Apache Mahout | Apache Streams (retired)
- Foundation Owned?
  - The ASF | CNCF | GNU - Comes with Governance (usually more community oriented)
  - MongoDB | Data Prep Kit - bashing on DPK, I'm too close to it, and it makes a good example
  - *The Curious Case of Kubeflow*
- [My first open source commit](#)
  - Apache Flink was a start up community -> they needed others to contrib
  - I made the PR not (really) knowing Scala (this is pre chat gpt, so it was pretty gnarly)
  - 37 back and forths + a prior thread - finally, not even merged, another maintainer forked my branch, cleaned it up, then merged that

The Politics of Dancing and FOSS

- **Apache Spark** - Nominally Foundation Run Open Source - Actually a commercial product of Databricks (who run an optimized fork).
- **MongoDB** - Not Foundation Owned - You're just doing free work for them. Their 'community' is an outgrowth of their sales/marketing arm.
- **Data Prep Kit** - Not Foundation Owned - IBMers working in the open (but trying to change)
- **Kubeflow -** Not foundation owned…until it's time passed, then donated to CNCF - I used to work for a company that was trying to be the Databricks of Kubeflow, but they missed their Series B
- **Apache Mahout - qumat** - We ARE a open hobby proj, but you have to understand a bit about quantum computing- which is a different sort of barrier to entry.
- **PUNCHLINE:** *There are more reasons than not you and your PR won't be welcomed into a community that have* NOTHING *to do with code quality.*

What project leads (ie me) care about

- On your first commit I don't know who you are- so I assume you're a newb just getting started, and probably doing a **drive by**
- I don't want some huge whiz-bang feature as a drive by from a newb (because then it falls to me to maintain it)
- I **DO** want small patches / bug fixes. No, but I like them as drive bys (because it't nothing new to maintain, it is maintenance of existing things)
- I *never* want a big feature dump. You get to do those by yourself once you're a comitter project lead, and even then you get buy in- different communities have different rules. Usually you need to build up some karma (with small patches and bug fixes)
- OR Small enhancements so long as they are
  - Well documented
  - Have good test coverage

Getting Involved

- FOSS Projects are like parties
  - **Mandatory Fun:** A party your boss makes you go to
  - **P-Diddy White Parties:** A party that 'everyone wants to be at' but the real stuff is happening behind closed doors so you'll never see it and even if you do it's pretty unseemly
  - **Loyal Order of the Water Buffalo:** A party you go to on your own time because you like the people there and it's more of a community thing, software is secondary.
  - **The Club:** (aka Benevolent Despots) Someone (or a few people) own the project- no foundational oversight, if you don't play by their rules you get booted.
  - **House Party**: People already knew each other, just getting together to have fun, best to be invited, hard (but not impossible) to crash a house party. (Catalina Wine Mixer)
  - The list goes on. Any conceivable party type has an open source project analog.

Getting Involved - But Why?

- Lot's of reasons - your reasons will direct you to different projects
- Your boss said to go contribute to open source
    - Find an open project- do a drive by - call it a day.
- You want the prestige of having commits on a certain project or type of project
    - This is the White Party - you can find some big name project - follow my advice on making a very small manageable PR. You won't be on the inside. After months (or years) or making small PRs, you *might* get invited to the inner circle, but you might not like what you find.
- Building your 'portfolio' for when you apply somewhere else
    - You want to know how I code and communicate with others? Here's links to me contributing code / dealing with feed back / etc
- Industry networking / making real friends
    - This could be The Club or it could be the Loyal Order of the Water Buffalo
    - Club / Benevolent Despot example: Bike Lane Uprising
    - Loyal Order of the Water Buffalo: Apache Mahout and Apache Streams(retired)
- You and some friends have a fun cool idea
    - Cataline Wine Mixer
    - Gofannon started as this, is growing - it's like a college kegger. It is a house party, but we've marketed it to freshman the first week school is in session. Some day we hope to grow into a LOoWB not a Club

Getting Involved - But How?

- Depends what kind of community/party it is
  - Different parties have different things that work
  - Some things never work, like asserting you're a big deal and everyone should bow down because you're there.
- How does the community communicate?
  - Mailing List
  - Slack
    - Secret Slack Kabals
  - Discord
  - Beer on Friday Afternoons
- Making a small bug fix / patch is a good way to introduce yourself
  - https://github.com/apache/mahout/pull/492
- Sometimes there is an entry path laid out (like in gofannon)

Let's pick apart a PR together

https://github.com/The-AI-Alliance/gofannon/pull/39

Appendix

Follow up qs: trevor.grant@thealliance.ai

For more gossip - find me at a conference / buy me a beer.

Gofannon - https://github.com/The-AI-Alliance/gofannon

Apache Mahout - https://github.com/apache/mahout/pull