

Joining the Scurry of Squirrels:

# Contributing to Apache Flink®



Tzu-Li (Gordon) Tai

@tzulitai

tzulitai@apache.org

**dataArtisans**

# Who?

---

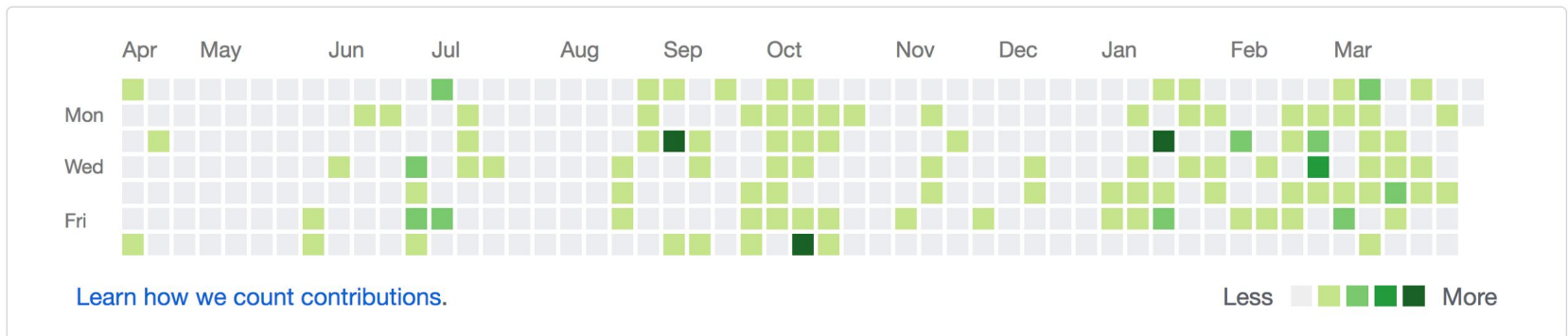


- Apache Flink Committer
  - Streaming connectors (Kafka, Kinesis, Elasticsearch)
- Software Engineer at data Artisans

# Who?



- Definitely not a veteran contributor of Apache Flink ;-)



***... my life of Apache Flink contribution ;)***



*The usual first question ...*

***How can I contribute?***

# It never is just about the code ...



**Yehuda Katz** ✓

@wycats

Following



As a tech community, we must treat documentation, marketing, logistics, infrastructure, art, etc. work with as much respect as engineering

RETWEETS

1,004

LIKES

1,456



10:11 AM - 12 Dec 2015

← 33

↻ 1.0K

♥ 1.5K

# Different forms of Contribution

---



## ■ **As an user ...**

- file bug reports
- join discussions / propose new features on the mailing list
- testing release candidates
- talk about Flink

## ■ **As a developer ...**

- submit & review patches for features and fixes
- help answer user questions

*Know your resources!*

**Mailing Lists, JIRA, and Wiki**



# Know your resources!

---



- The resources *is* the community, and where everything happens
  - Mailing lists
  - JIRA
  - Wiki

# Mailing lists

---



- Developer mailing list
  - [dev@flink.apache.org](mailto:dev@flink.apache.org)
  - Ask questions about Flink development
  - Discuss new Flink features
  
- User mailing list
  - [user@flink.apache.org](mailto:user@flink.apache.org)
  - General questions about using Flink

# JIRA board

---



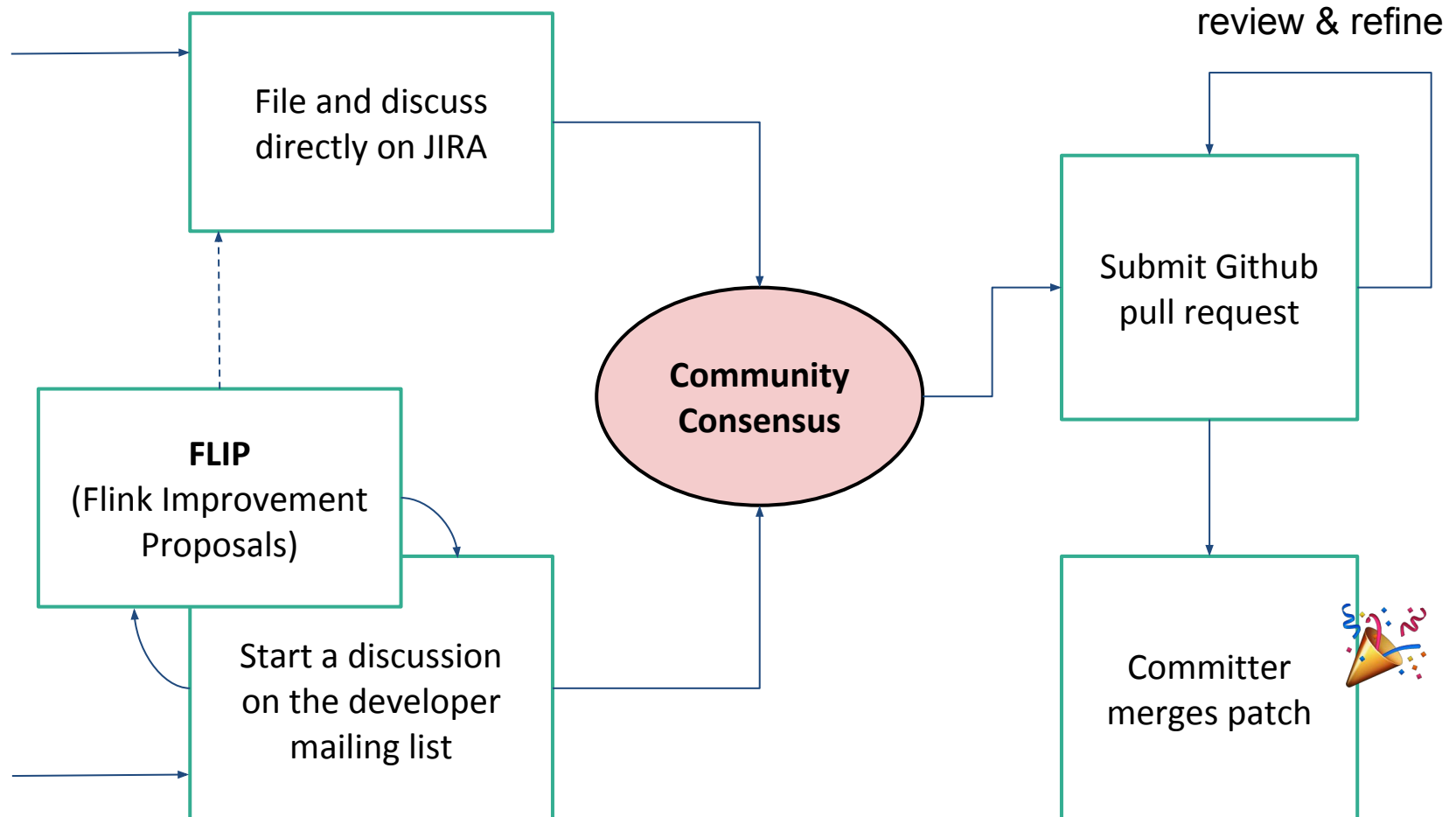
- Issue tracking
  - <https://issues.apache.org/jira/browse/FLINK>
  - Where bugs and new features / improvements are filed
  - All code contributions must always have a corresponding JIRA issue ticket opened
  - Do NOT ask questions here (--> mailing lists)



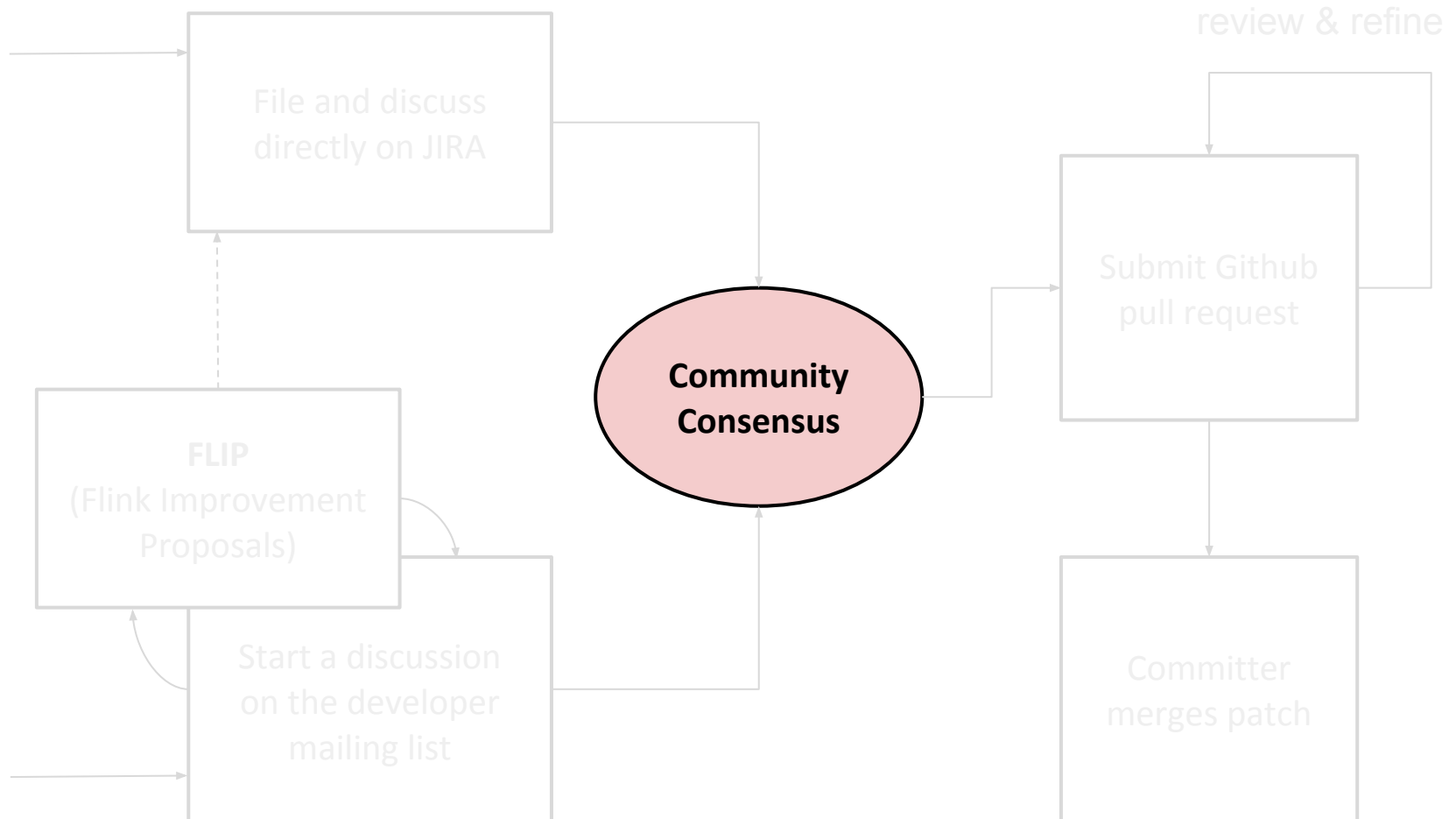
- General information about the project
  - <https://cwiki.apache.org/confluence/display/FLINK>
  - Mostly relevant for FLIPs (Flink Improvement Proposals)
  - Other info somewhat outdated

# Apache Flink Patch Submission

# Patch Submission Process



# Patch Submission Process



# The Apache Way: Community Over Code

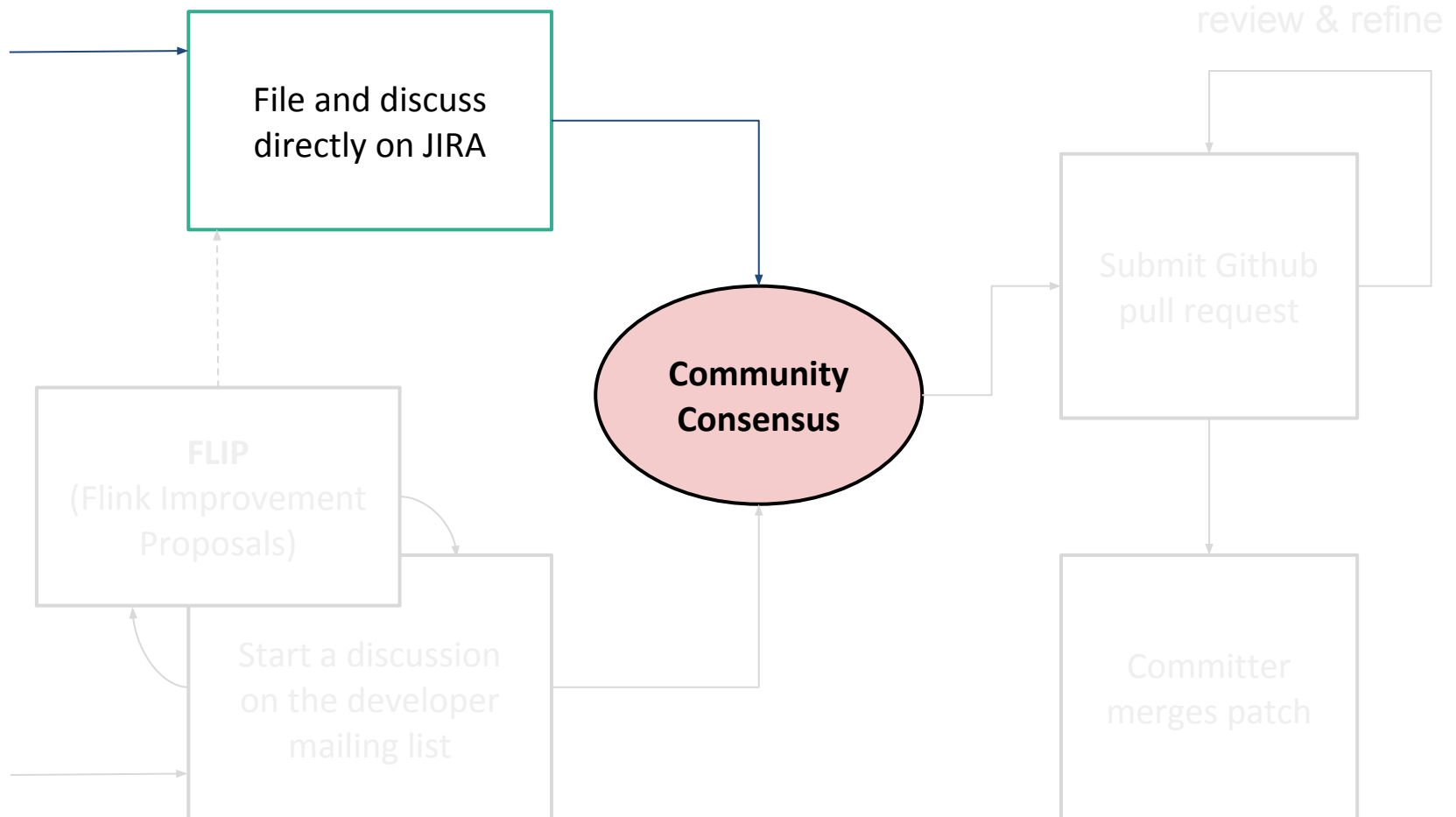
---



- Changes, especially new features, should be agreed upon by the community
  - Does not apply just to the developers
  - Beneficial to users as a whole
  
- Always seek to reach some level of consensus throughout the code contribution process
  - Avoid needing to go back to feature discussion after coding



# Patch Submission Process



# Filing a JIRA

---



- Always have a good description of what the issue is
  - *Bugs* - the cause and the proposed solution
  - *Improvements and features* - bootstrap potential discussion with the implementation you have in mind
- Remember to appropriately set “Components”

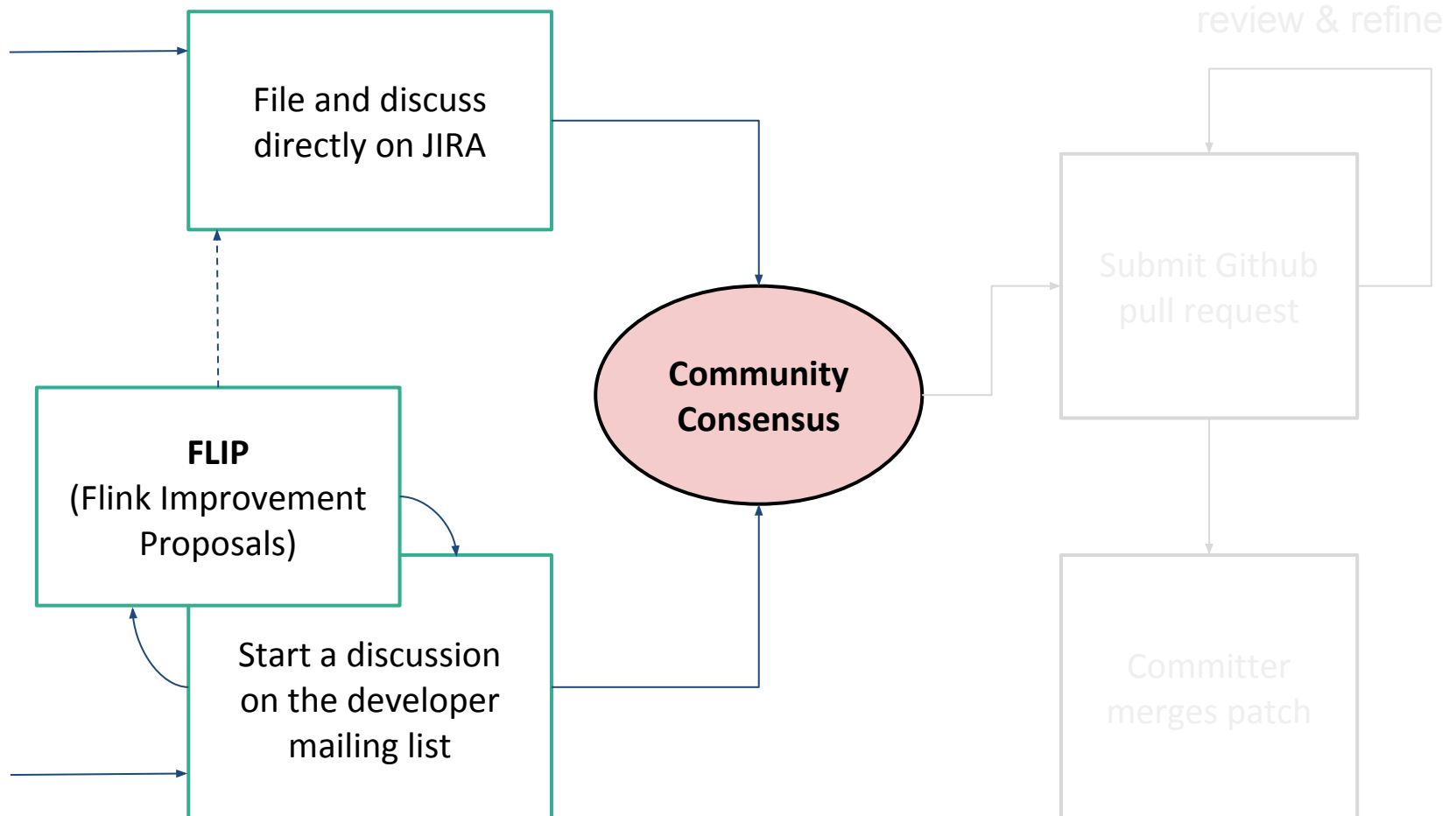
# Assign yourself to a JIRA

---



- Set the “Assignee” field to yourself
  - Notifies others that this issue is already in good hands
- Simply request JIRA permission on the developer mailing list

# Patch Submission Process



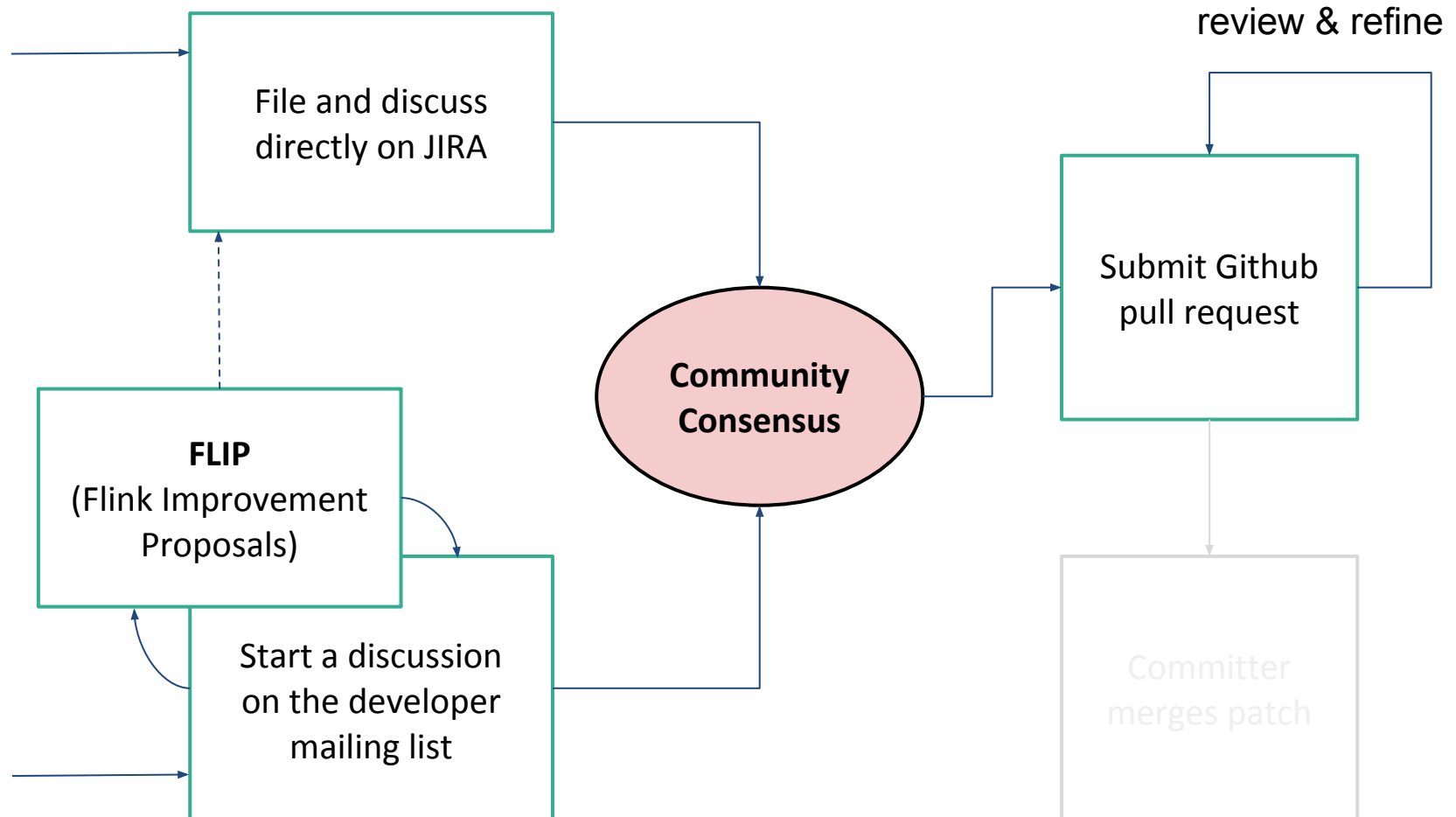
# FLIP (Flink Improvement Proposals)

---



- Adopted from the Apache Kafka community
- Official design documentation for major features
  - Serves as the basis of discussion on the dev mailing list
  - Allows the community to decide on the best future-proof design

# Patch Submission Process



# Patch Submission Best Practices

---



- Rebase onto latest master prior to opening pull request
  - `git pull --rebase origin master`
- Squash your commits to appropriate set
- Commit message example:
  - `[FLINK-6025] [cep] Implement skip till next match strategy`

# Patch Submission Best Practices

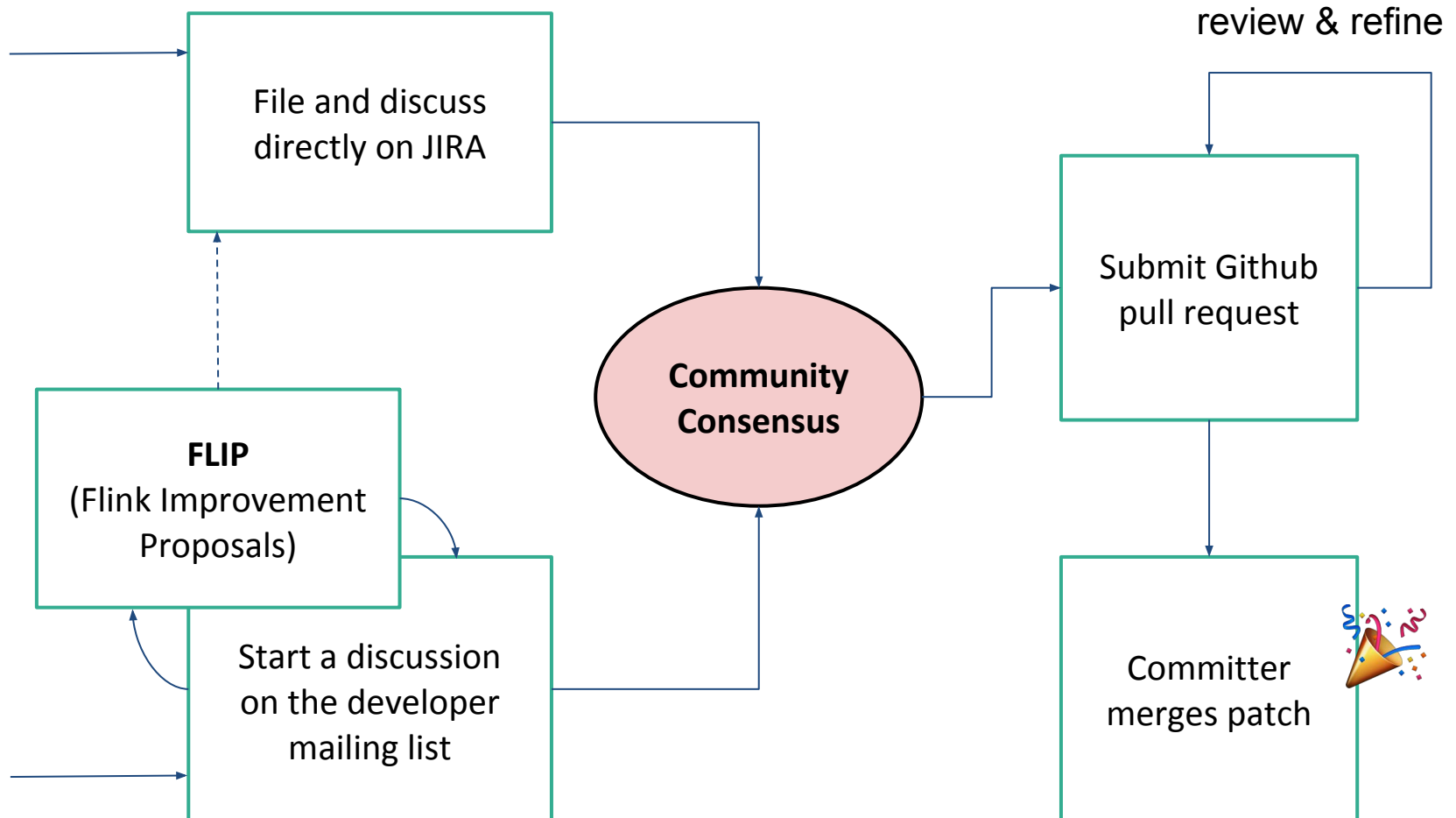
---



- Usually the pull request naming follows the primary commit's message:
  - `[FLINK-6025] [cep] Implement skip till next match strategy`
- Address all discussions happening on the pull request
- Push follow-up commits
  - Retain history of the patch review
  - Also have meaningful msgs for follow-up commits



# Patch Submission Process



**Closing**

# Final Takeaways

---



- Development of Apache Flink is heavily centered around its community
  - The Apache Way - community over code
  - Contributing is never just about writing code
- Know the resources and do not hesitate to approach.
- Extra pair of helping hands is always welcome ;-) !

# Shoutouts

---



- Flink CEP
  - Complex event processing
  - Pattern detection for streams
  - → If interested, contact  
Kostas Kloudas ([kk1oudas@apache.org](mailto:kk1oudas@apache.org))
  
- Flink ML
  - Online learning
  - Incremental learning
  - Model serving
  - → If interested, contact  
Theodore Vasiloudis ([tvas@apache.org](mailto:tvas@apache.org))

# Thank you!

@tzulitai  
@ApacheFlink  
@dataArtisans