Meeting of the LF AI & Data Technical Advisory Council (TAC)

December 2, 2021

□LF AI & DATA

Antitrust Policy

› Linux Foundation meetings involve participation by industry competitors, and it is the intention of the Linux Foundation to conduct all of its activities in accordance with applicable antitrust and competition laws. It is therefore extremely important that attendees adhere to meeting agendas, and be aware of, and not participate in, any activities that are prohibited under applicable US state, federal or foreign antitrust and competition laws.

› Examples of types of actions that are prohibited at Linux Foundation meetings and in connection with Linux Foundation activities are described in the Linux Foundation Antitrust Policy available at http://www.linuxfoundation.org/antitrust-policy. If you have questions about these matters, please contact your company counsel, or if you are a member of the Linux Foundation, feel free to contact Andrew Updegrove of the firm of Gesmer Undergone LLP, which provides legal counsel to the Linux Foundation.

LF AI & DATA

Recording of Calls

Reminder:

TAC calls are recorded and available for viewing on the TAC Wiki

Reminder: LF AI & Data Useful Links

› Web site: lfaidata.foundation
› Wiki: wiki.lfaidata.foundation
› GitHub: github.com/lfaidata
› Landscape: https://landscape.lfaidata.foundation or https://l.lfaidata.foundation
› Mail Lists: https://lists.lfaidata.foundation
› Slack: https://slack.lfaidata.foundation
› Youtube: https://www.youtube.com/channel/UCfasaeqXJBCAJMNO9HcHfbA
› LF AI Logos: https://github.com/lfaidata/artwork/tree/master/lfaidata
› LF AI Presentation Template: https://drive.google.com/file/d/1eiDNJvXCqSZHT4Zk_-czASlz2GTBRZk2/view?usp=sharing

› Events Page on LF AI Website: https://lfaidata.foundation/events/
› Events Calendar on LF AI Wiki (subscribe available): https://wiki.lfaidata.foundation/pages/viewpage.action?pageId=12091544
› Event Wiki Pages: https://wiki.lfaidata.foundation/display/DL/LF+AI+Data+Foundation+Events

LF AI & DATA

Agenda

› Roll Call (2 mins)

› Approval of Minutes from previous meeting (2 mins)

› FLYTE graduation proposal (20minutes)

› Janusgraph update (20 minutes)

› Updates from the outreach committee (10 minutes)

› LF AI General Updates (2 min)

› Open Discussion (2 min)

TAC Voting Members - Please note

Please ensure that you do the following to facilitate smooth procedural quorum and voting processes:

- Change your Zoom display name to include your First/Last Name, Company/Project Represented
    - example: Nancy Rausch, SAS
- State your First/Last Name and Company/Project when submitting a motion
    - example: First motion, Nancy Rausch/SAS

LF AI & DATA

Challenge with TAC Quorum

›   19 voting members requiring 10 voting members to achieve quorum

›   Proposing updating charter to reflect the following changes:

    ›   A TAC voting member who misses 2 TAC meetings in a row will lose their voting seat until they attend twice in a row.

›   Process: Socialize with GB and TAC. Propose amendment to the Charter and have the GB vote on it.

⊓LF AI & DATA

TAC Voting Members

* = still need backup specified on [wiki](#)

Board Member	Contact Person	Email
AT&T	Anwar Atfab*	anwar@research.att.com
Baidu	Ti Zhou	zhouti@baidu.com
Ericsson	Rani Yadav-Ranjan*	rani.yadav-ranjan@ericsson.com
Huawei	Huang Zhipeng	huangzhipeng@huawei.com
IBM	Susan Malaika	malaika@us.ibm.com
Nokia	Jonne Soininen	jonne.soininen@nokia.com
OPPO	Jimin Jia*	jiajimin@oppo.com
SAS	Nancy Rausch	nancy.rausch@sas.com
Tech Mahindra	Amit Kumar	Kumar_Amit@techmahindra.com
Tencent	Bruce Tao	brucetao@tencent.com
Zilliz	Jun Gu	jun.gu@zilliz.com
ZTE	Wei Meng	meng.wei2@zte.com.cn

Graduate Project	Contact Person	Email
Acumos	Nat Subramanian	natarajan.subramanian@techmahindra.com
Angel	Bruce Tao	brucetao@tencent.com
Egeria	Mandy Chessell	mandy_chessell@uk.ibm.com
Horovod	Travis Addair*	taddair@uber.com
Milvus	Xiaofan Luan	xiaofan.luan@zilliz.com
ONNX	Jim Spohrer (Chair of TAC)	spohrer@us.ibm.com
Pyro	Fritz Obermeyer*	fritz.obermeyer@gmail.com

LF AI & DATA

Minutes approval

Approval of November 18th, 2021 Minutes

Draft minutes from the November 18th TAC call were previously distributed to the TAC members via the mailing list

Proposed Resolution:

› That the minutes of the November 18th meeting of the Technical Advisory Council of the LF AI & Data Foundation are hereby approved.

LF AI & DATA

Flyte LF AI & Data 2021 Update & Request for Graduation

Ketan Umare   @kumare3

Agenda

- **Flyte Overview**
  - Adoption
  - Flyte Community
  - Collaboration
- **Graduation Requirements**
  - Top Contributors
  - Commit History
  - Best Practices Badge
  - Governance & Procedures
- **Roadmap**
- **Summary**

Flyte

Flyte Overview

Flyte

Adoption

- **GitHub**
  - 2K+ stars (all repos) (Nov 2021)
  - 175 forks
  - 27 repos
- **Docker**
  - Hub Downloads?
- **Google Analytics**
  - 9,100+ users
  - 120+ countries
  - 1,900+ cities

Star history

https://star-history.t9t.io/#flyteorg/flyte
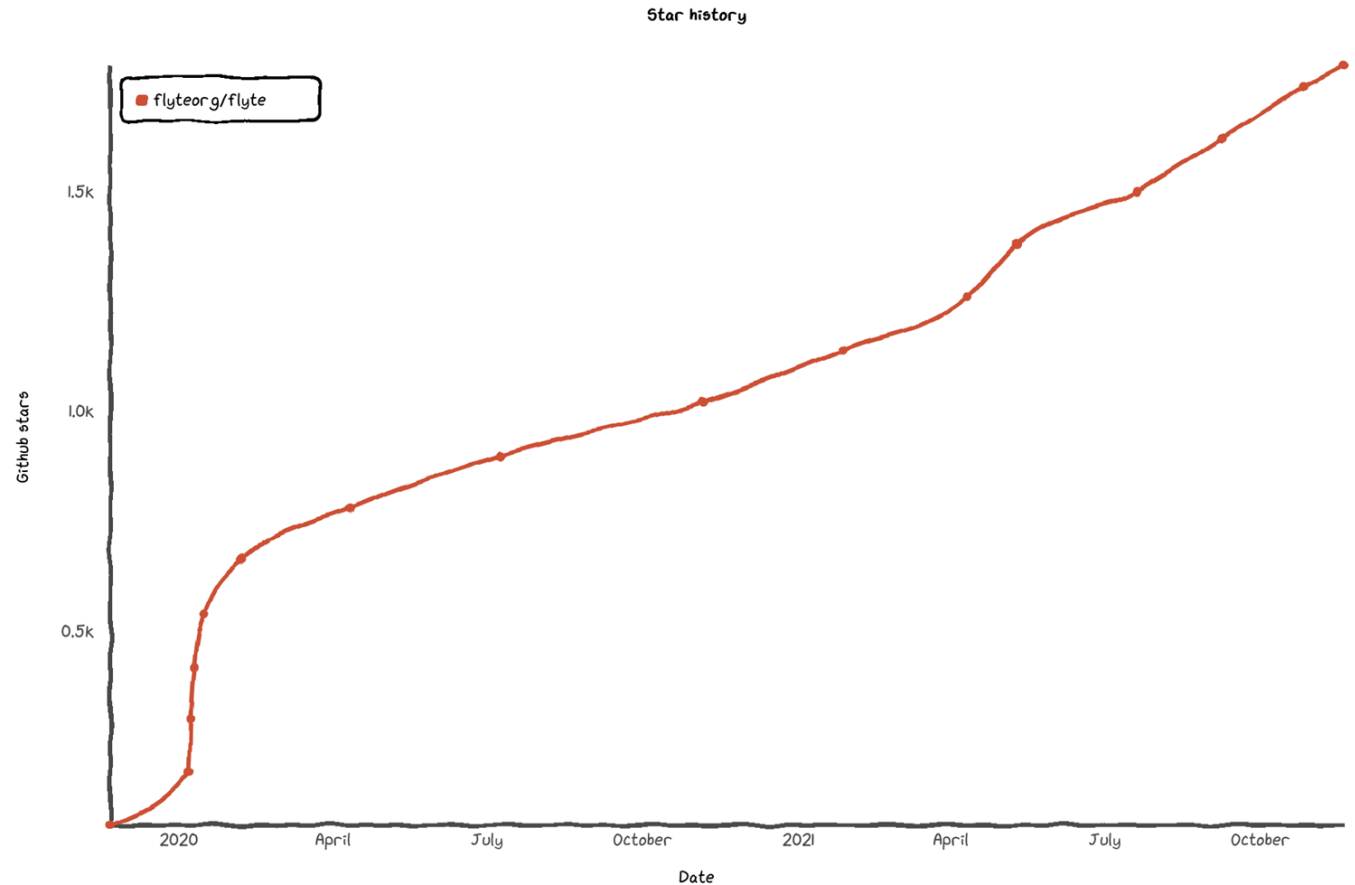
Flyte Community

- Communication
  - Flyte Slack (650+)
    https://slack.flyte.org

  - Flyte GitHub Discussions
    https://github.com/flyteorg/flyte/discussions

- Announcements
  - Slack
  - Twitter (220+)
    https://twitter.com/flyteorg

  - Google Group (140+)
    https://groups.google.com/u/0/a/flyte.org/g/users

- Publications
  - Flyte Blog
    https://blog.flyte.org/

  - Flyte Monthly
    https://www.getrevue.co/profile/flyte

- Sync Up
  - Biweekly OSS Sync Up
    https://www.addevent.com/calendar/kE355955

  - YouTube
    https://bit.ly/3jL6xOX

Collaborators, Contributors & Integrations

- Lyft
- Spotify
- Striveworks
- USU
- Freenome
- Level 5
- Intel
- GoJek
- Union.ai
- Blackshark
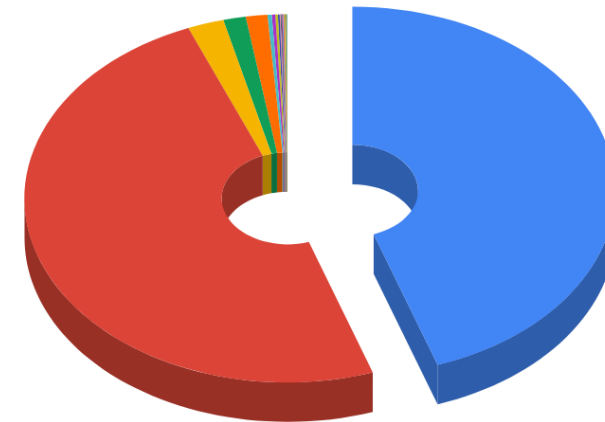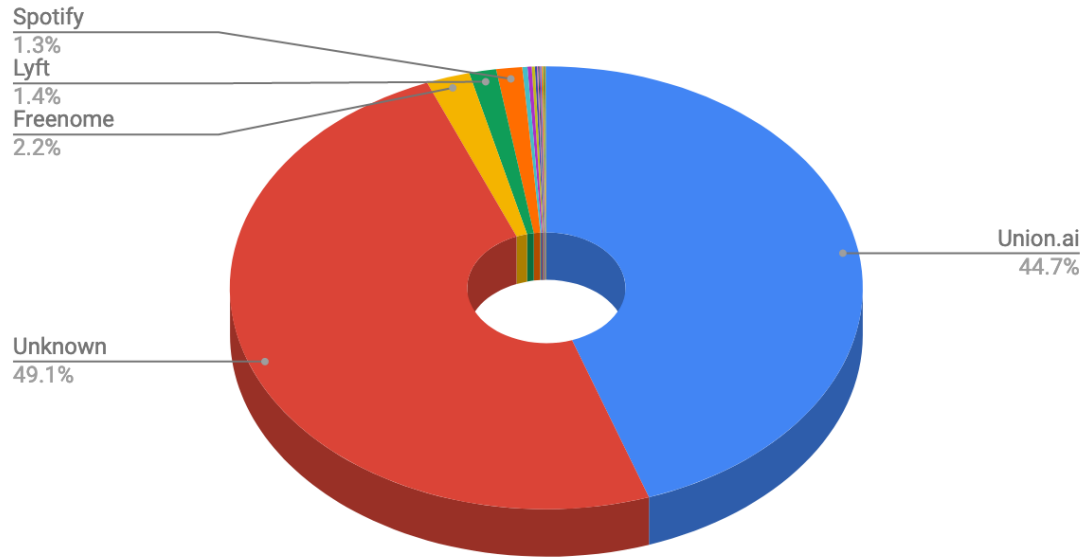- Wolt
- RunX
- MethaneSAT, ...

Graduation Requirements
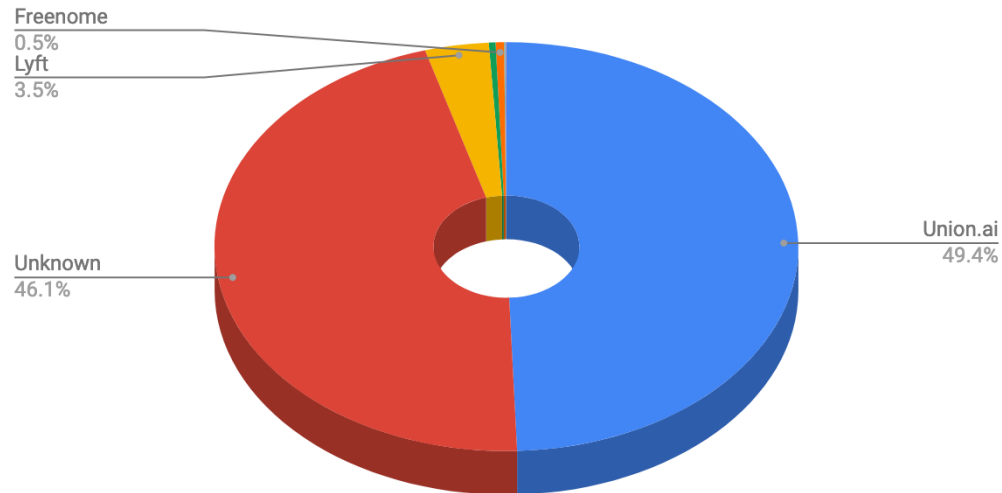
Flyte

Top Contributors in 2021

1.88K commits

Spotify
1.3%
Lyft
1.4%
Freenome
2.2%
Union.ai
44.7%
Unknown
49.1%

Union.ai · Unknown · Freenome · Lyft · Spotify · Snyk · Wolt · Scarf
Run-x · Toss · Intel Corporation · Onfido · GoJek · Merantix · Dropbox
Red Hat · Clarisights · GitHub

Repos: flyte, flytadmin, flytepropeller, flytekit, flyteconsole, flytestdlib, flytesnacks, flyteorg.github.io, flytectl, flytekit-java, flyteidl, flyteplugins, boilerplate, datacatalog, flytecopilot, flytekit-python-template, flytelab

Top Contributors in 2021

Number of lines added:
901926

Freenome
0.5%
Lyft
3.5%

Union.ai
49.4%

Unknown
46.1%

Union.ai   Unknown   Lyft   Spotify   Freenome   GoJek   Run-x
Scarf   Merantix   Dropbox   Intel Corporation   Red Hat   Wolt   GitHub
Synk   Clarisights   Toss   Onfido

Repos: flyte, flytadmin, flytepropeller, flytekit, flyteconsole, flytestdlib, flytesnacks, flyteorg.github.io, flytectl, flytekit-java, flyteidl, flyteplugins, boilerplate, datacatalog, flytecopilot, flytekit-python-template, flytelab
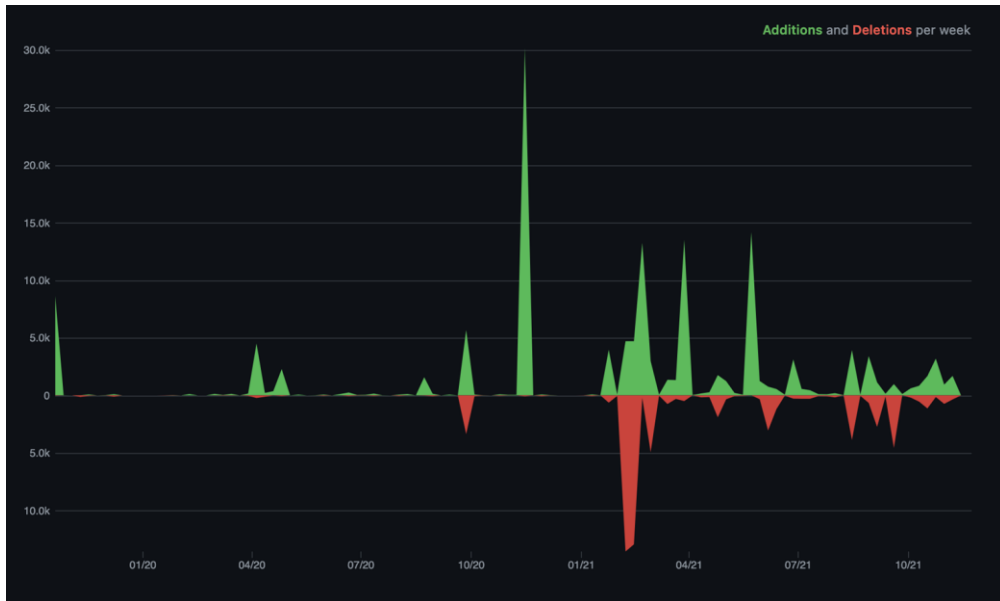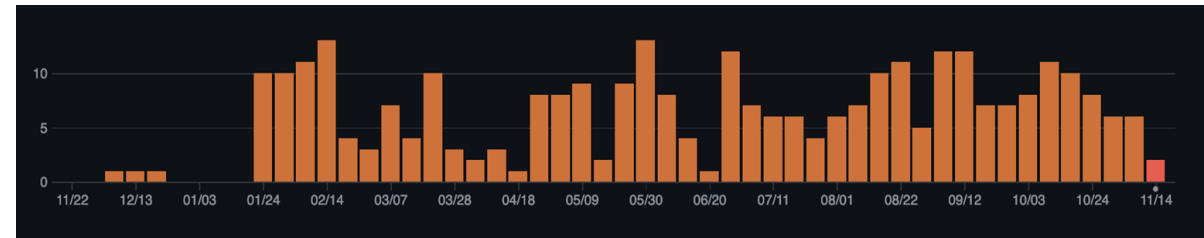
Commit History (2/27 repos)

flyteorg/flyte repo
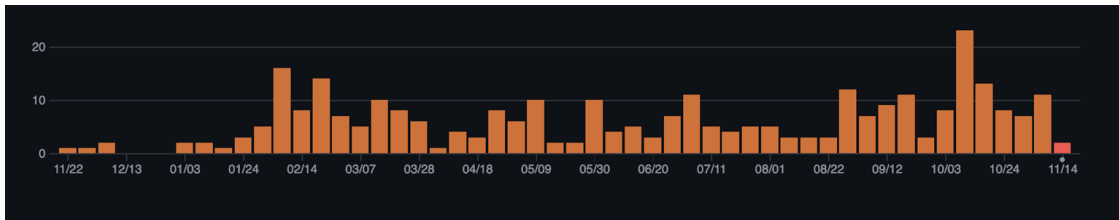
flyteorg/flytekit repo

Commits growth

Commits Growth
The growth in terms of the aggregated count of total number of unique commits during the selected time period.
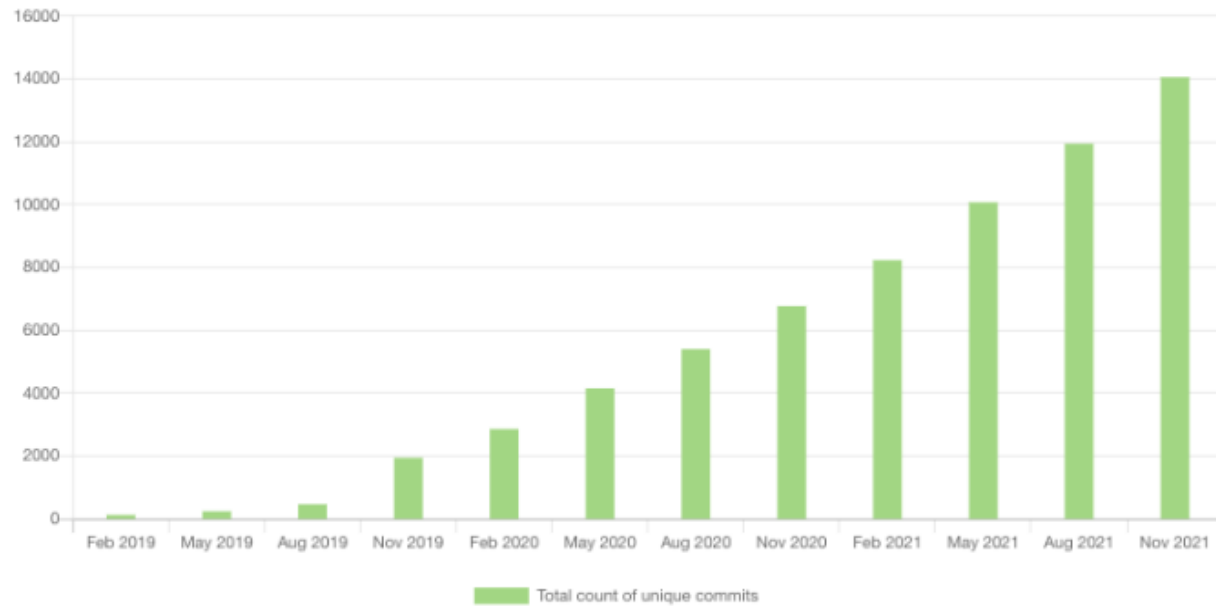
Key Insights
• An average of **5.52K** commits were pushed by active code contributors during the last 3 Years.

Contributor strength – avg 134

Contributor Strength

The growth in the aggregated count of unique contributors analyzed during the selected time period. A contributor is anyone who is associated to the project by means of a commit or any activity on Pull Requests / Changesets.
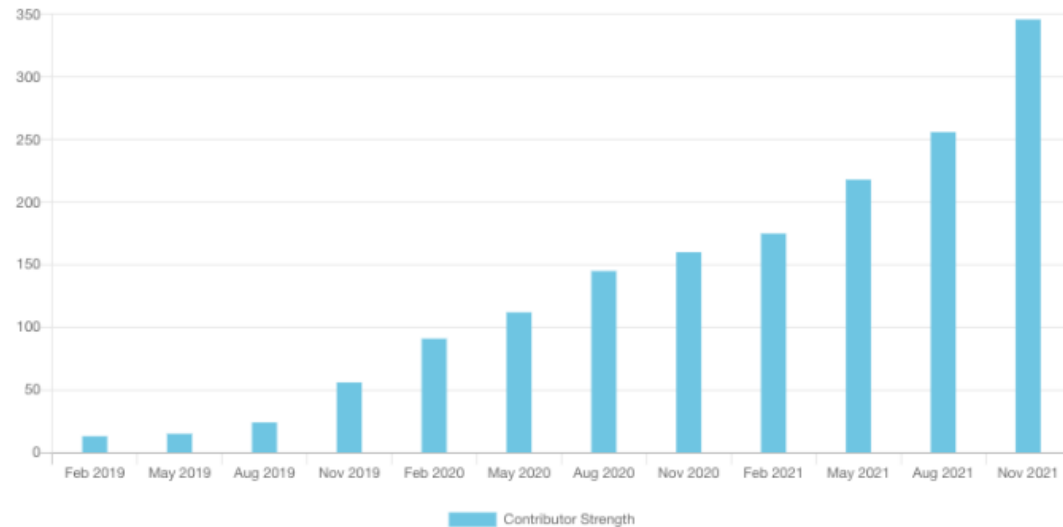
Key Insights

- The contributor strength **increased** by **2.56K%** during the last 3 Years.
- The contributor strength grew the most between the time period **Aug 2021 and Nov 2021** during the last 3 Years.
- The average contributor strength was **134** during the last 3 Years.

Best Practices Badge

https://bestpractices.coreinfrastructure.org/en/projects/4670

Governance & Procedures

- Contributor's Guide
  - https://github.com/flyteorg/flyte/blob/master/CONTRIBUTING.md
- Code of Conduct
  - https://github.com/flyteorg/flyte/blob/master/CODE_OF_CONDUCT.md
- Governance Information
  - Details

Roadmap

Flyte

2021 Roadmap – Features and Improvements

- Core Features
  - Getting started and local execution: Launch a Flyte cluster locally in one docker command
  - Flytekit python fully native pythonic interface (First of its kind - leading the way) - 20+ UI Features, 40+ Usability features
  - Flytekit-java/scala - write workflows in java & scala
  - Native cross cloud Interval scheduler for Flyte
- Performance and Reliability
  - Full recoverability from system failures
  - Sharded Flyte execution engine and 10+ performance improvements
  - Improved documentation, examples and getting started
  - Published nightly test reports for the core platform on various cloud providers
- Deployment features
  - One line deployment to AWS and GCP
  - Detailed deployment guide for deploying to AWS, GCP, other clouds or onPrem
  - Helm and Kustomize charts

2021 Roadmap – Collaboration

- **Collaborate with LF AI & Data projects**
  - Feast
  - Horovod
- **Collaborate with other projects**
  - Great Expectations
  - Pandera
  - Weights & Biases
  - Dolt
  - Modin
  - Kubeflow operators
- **Cloud services**
  - AWS: Sagemaker, Athena
  - GCP: BigQuery
  - Snowflake

2022 Roadmap – Features and improvements

- UX and Usability
  - Improvements in UI - diving deeper into runs, search by inputs and outputs, inline docs for users, better graphs
  - Flytekit-java production ready. Potentially explore flytekit-R (and other SDK layers)
  - Ability to use Bazel to build projects
  - Single Python script deploy
- Core features
  - Automatic model checkpointing (for long running models)
  - Ability to run micro-workflows (runtime in seconds)
- Reliability
  - Nightly test reports for Flyte plugin reliability on various cloud platforms
- Deployment features
  - One line deployment to AWS and GCP
  - Detailed deployment guide for deploying to AWS, GCP, other clouds or onPrem
  - Helm charts and Kustomize manifests

2022 Roadmap – Collaboration

- **Collaborate with LF AI & Data projects**
  - ONNX
  - Ludwig
  - kServe
  - Amundsen / Open Lineage / Marquez
  - Milvus (explore)
- **Collaborate with other open-source projects**
  - FastAPI (Union)
  - KNative (Union)
  - CloudEvents (Union)
  - Ray (Applied AI)
  - Flink (Spotify)
  - Hydra (Merantix)
  - MLFlow (Wolt / Doordash)

- Cloud services
  - GCP Services - Dataproc, Vertex AI
  - AWS Services - EMR
  - Databricks

Long Term roadmap

- Set the gold standard in UX and Usability for ML and Data workflows
- Provide a singular platform to execute serverless workflows in production @ low cost, high scale and low overhead
- Provide a singular platform for outstanding open source DS/Data projects to collaborate
- Make a standardized, structure approach to Data Science pipelines a reality - MLOps for everyone

Summary

Flyte

Summary

- Flyte has been incubating for the past 10 months
- During this period adoption has grown more than 5x
- The community trusts Flyte as a reliable software and has deployed it in production at multiple organizations
- All of the graduation criteria has been met
- Graduating the project would give many organizations confidence to adopt it and deploy for their critical workloads
- We look forward to your positive vote to graduate the project

Thank You!

Flyte

TAC Vote on Flyte Project Proposal

Proposed Resolution:

The TAC approves the Flyte proposal as a graduation project of the LF AI & Data Foundation

JanusGraph

Brief Description:

JanusGraph is a scalable graph database optimized for storing and querying graphs containing hundreds of billions of vertices and edges distributed across a multi-machine cluster.

Contributed by:
Initially launched at the Linux Foundation by Expero, Google, GRAKN.AI, Hortonworks, IBM and others in January 2017, and added as an Incubation Project in December 2020.

Key Links:

Github: https://github.com/janusgraph

Website: https://janusgraph.org

Artwork:

https://artwork.lfaidata.foundation/projects/janusgraph/

Mailing lists:
› janusgraph-announce
› janusgraph-users
› janusgraph-dev
› janusgraph-tsc

Project overview

Fork of TitanDB (being developed since 2012).
The initial release of JanusGraph was made in April 2017.

Current repositories:
- **janusgraph** – Main database repository. Includes Java Gremlin Language Variant (GLV)
- **janusgraph-dotnet** - JanusGraph .NET Gremlin Language Variant (GLV)
- **janusgraph-python** - JanusGraph Python Gremlin Language Variant (GLV) (initial stage,
- not officially released but can be compiled from `initial-pr` branch)
- **janusgraph-docker** - JanusGraph Docker images
- **janusgraph-ambari** - JanusGraph Ambari plugin
- **janusgraph-cassandra** - Cassandra Thrift and Embedded Backend Adapter for JanusGraph
- **janusgraph-foundationdb** - FoundationDB storage adapter for JanusGraph
- **janusgraph.org** – Main web page repository
- **docs.janusgraph.org** – Documentation repository
- **logos** – JanusGraph logos
- **blog.janusgraph.org** – was intended to be a blog but the development didn't start
- **legal** – Repository to support CLA paper signers (deprecated as it was replaced by EasyCLA)

□LF AI & DATA

Architectural Overview
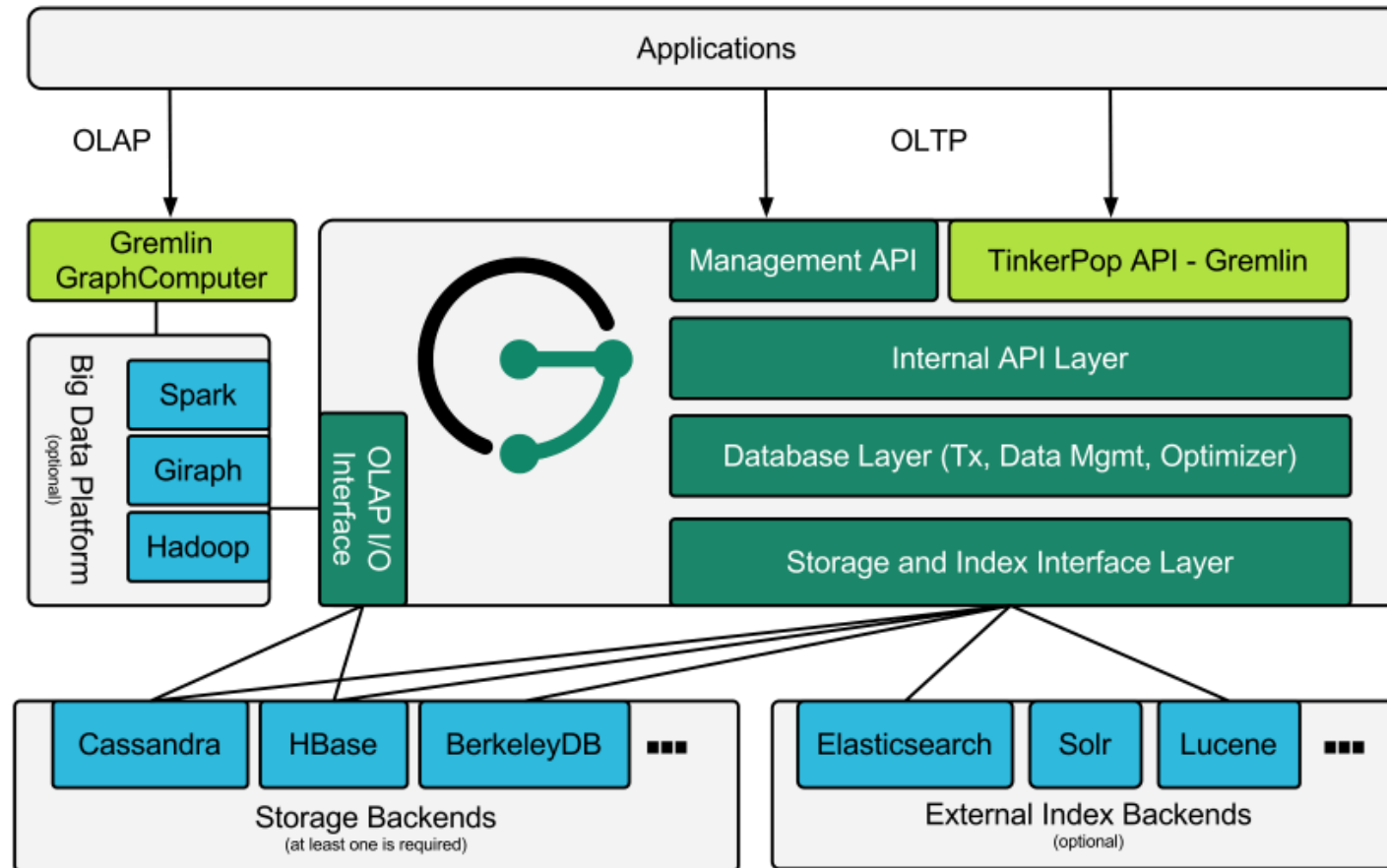
Supported storage backends: Cassandra, HBase, Bigtable, BerkeleyDB, ScyllaDB
Supported storage backends via adapters: Aerospike, DynamoDB, FoundationDB
Supported index backends: ElasticSearch, Solr, Lucene

Organizations contributing

Datadog
eBay
Expero
FireEye
G DATA
Goldman Sachs
Google
IBM
Mapped

and others

Contributions for the past year

936.40K
Lines Of Code Changed

590
Commits

32
Contributors

1
No Of Sub Projects

7
Repositories

Top 10 Contributors By Commits

View All

NAME	LINES OF CODE	COMMITS	%
Oleksandr Porun...	83.91K	197	8.96%
Boxuan Li	57.87K	172	6.18%
Jan Jansen	772.72K	146	82.52%
Clement de Groc	1.23K	15	0.13%
Florian Hockmann	1.05K	9	0.11%
Ted Wilmes	390	7	0.04%
Pavel Ershov	2.52K	7	0.27%
Florian Grieskamp	5.10K	4	0.54%
Robert Yokota	284	3	0.03%

Top 10 Organizations By Commits

View All

590
Commits

- Mapped Inc.
- Goldman Sachs & Co. LLC
- G DATA CyberDefense AG
- Datadog, Inc
- Individual - No Account
- G DATA Software
- Expero
- Google LLC
- FireEye
- International Business Machines Corporation
- Others

LF AI & DATA
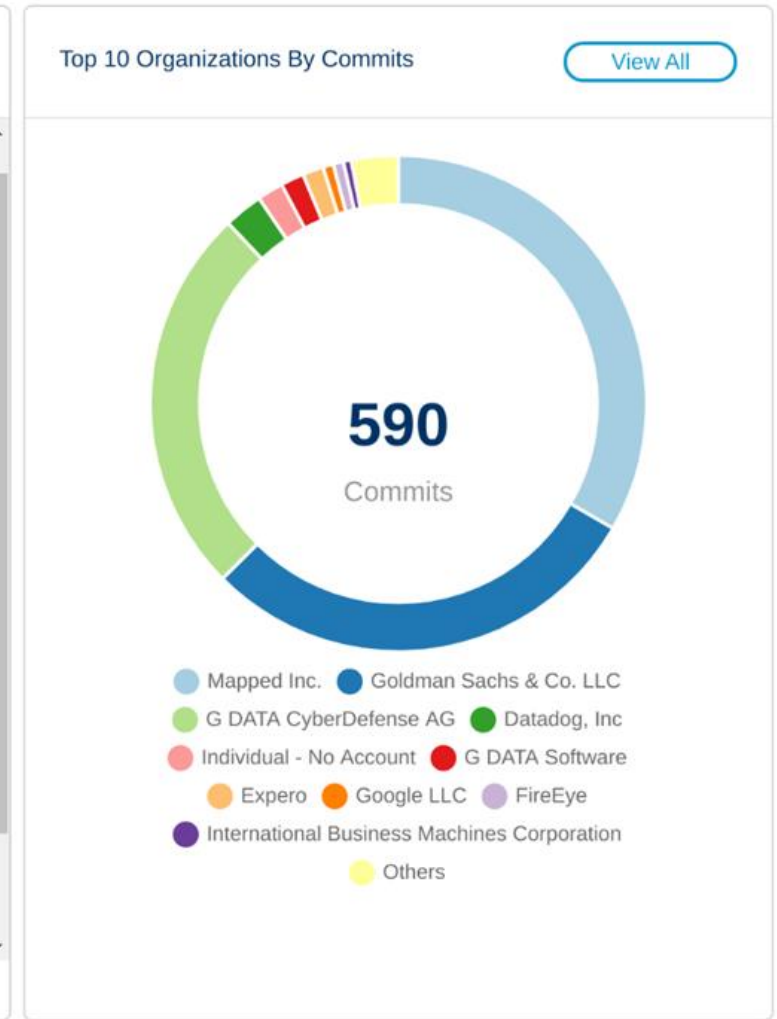
Key Achievements in the past year

0.5.3 and 0.6.0 releases

Added Donation option

1 new TSC member (Boxuan Li)

1 new JanusGraph Online Meetup held

2 new production use cases approved (eBay and Crédit Agricole CIB)

560+ likes in GitHub for the past year

LF AI & DATA

Areas the project could use help on

- Core code contributions
- Documentation improvement
- Social networks management

Collaboration opportunities

- Egeria (graduate project) has a number of open JanusGraph connectors. A direct collaboration could be done to improve these connectors.

- Amundsen (incubating project) has "Amundsen Metadata Service" project which can use either Neo4j or Apache Atlas as a persistent layer. Apache Atlas uses JanusGraph internally which creates indirect collaboration between Amundsen and JanusGraph but we can go even farther and add direct JanusGraph support to the project.

Feedback on working with LF AI & Data

We are grateful for being able to work with LF AI & Data.

Currently LF AI & Data provided JanusGraph with the:

- New mailing lists
- Helped to setup LFX Crowdfunding page

Questions

Outreach Committee

December 2, 2021

Jamil Chawki

Announcements

https://lfaidata.foundation/news/

Recent Announcements

1. Egeria, What Next After You Have Built a Catalog part 1, Nov 30
2. New Realse Adlik Deer Release (v0.4.0), Nov 30
3. International Conference AI for People: Towards Sustainable AI (CAIP'21), Nov 10

Upcoming Announcements

1. KubeCon, CloudNativeCon, Open Source Summit China - Shanghai, China, Dec 9-10
2. LF AI & Data Booth at OSS Open Source Summit Japan, Tokyo, Dec 14-15
3. LF AI & Data AI/ML/DL Track at Open Source Summit Japan, Dec 14-15
4. What Next After You Have Built a Data Catalog, Dec 6
5. LF AI & Data New Member Welcome 2021 Q4
6. Announcement of new LF AI & Data projects OpenBytes & KServe

THE **LINUX** FOUNDATION

LF AI & DATA

Promoting Upcoming Project Releases

We promote project releases via a blog post and on LF AI & Data Twitter and/or LinkedIn social channels

For links to details on upcoming releases for LF AI & Data hosted projects visit the Technical Project Releases wiki

If you are an LF AI & Data hosted project and would like LF AI & Data to promote your release, reach out to pr@lfaidata.foundation to coordinate in advance (min 2 wks) of your expected release date.
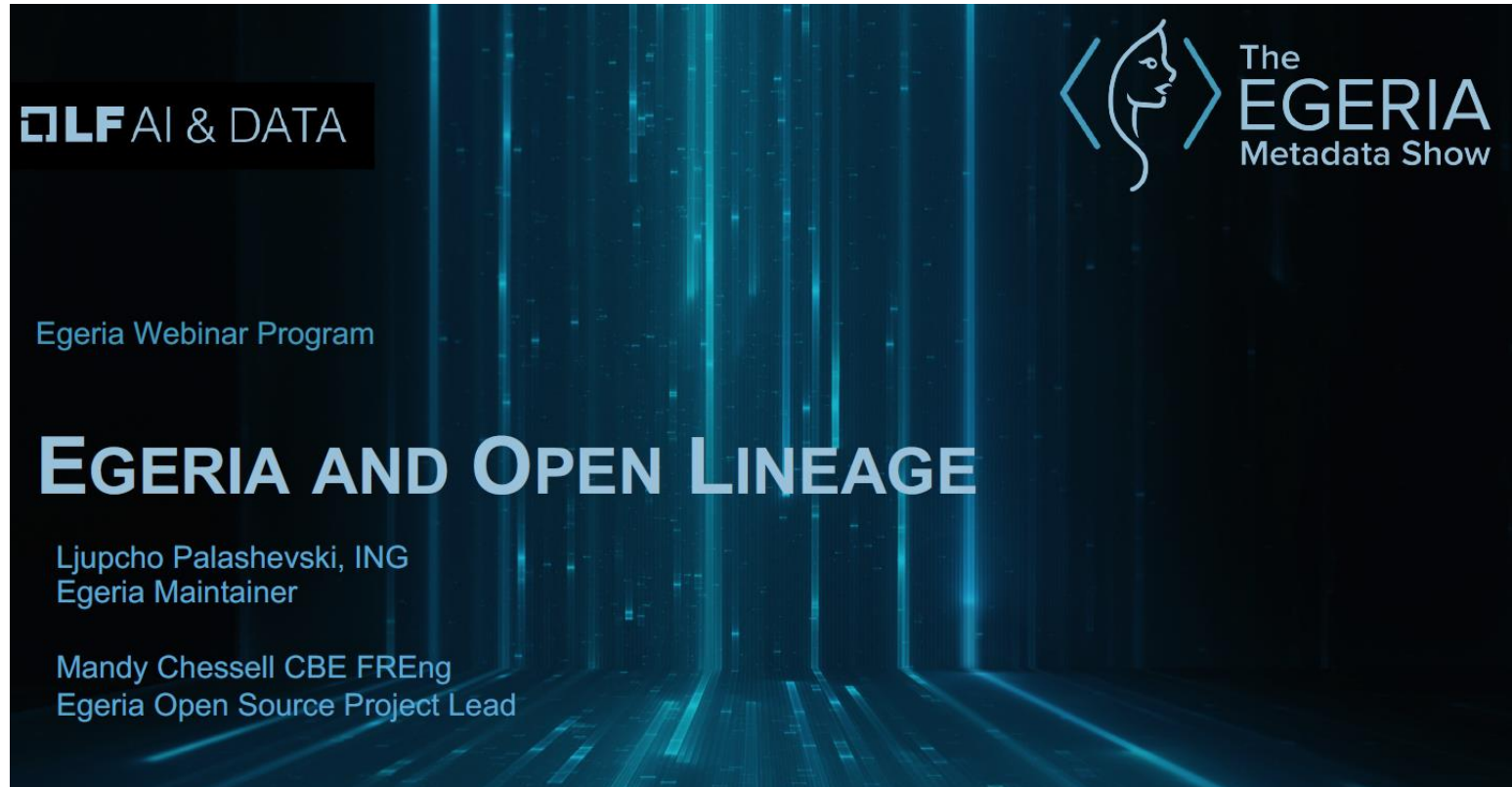
LF AI & Data PR/Comms

› **View all recent announcement on the LF AI & Data Blog**

› Please follow LF AI & Data on Twitter (@LFAIDataFdn) & LinkedIn and help amplify news via your social networks - Please retweet and share!
  › Also watch for news updates via the tac-general mail list

› Open call to publish project/committee updates or other relevant content on the LF AI & Data Blog

› To discuss more details on participation or upcoming announcements, please email pr@lfai.foundation

THE **LINUX** FOUNDATION   **LF** AI & DATA

Recent Events

Egeria, Open lineage, Nov 8

› Presentation: https://odpi.github.io/egeria-docs/education/previous-webinars/november-2021/

Upcoming Events

KubeCon, CloudNativeCon, Open Source Summit, China, Dec 9-10

OSS Event site

https://www.lfasiallc.com/kubecon-cloudnativecon-open-source-summit-china/

LF AI & Data Booth @ OSS 2021, Tokyo, Dec 14-15

OSS Event site [Open Source Summit Japan | Linux Foundation Events](#)

All Events

› Upcoming Events
  › Visit the LF AI & Data Events Calendar or the LF AI & Data 2021 Events wiki for a list of all events
  › To participate visit the LF AI & Data 2021 Events wiki page or email info@lfaidata.foundation

› To discuss participation in an event or to host an event, please email events@lfaidata.foundation

Upcoming TAC Meetings

Upcoming TAC Meetings (Tentative)

› December 16, 2021: DataPractices.org, others TBD

› December 30, 2021: Canceled for the holiday

› January 13, 2021: Meetings resume, ART graduation proposal, AI Fairness 360, AI Explainability 360, Adversarial Robustness Toolbox annual project review

Please send agenda topic requests to tac-general@lists.lfaidata.foundation

Open Discussion

□LF AI & DATA

TAC Meeting Details

› To subscribe to the TAC Group Calendar, visit the wiki: https://wiki.lfaidata.foundation/x/cQB2 _____

› Join from PC, Mac, Linux, iOS or Android: https://zoom.us/j/430697670

› Or iPhone one-tap:

  › US: +16465588656,,430697670# or +16699006833,,430697670#

› Or Telephone:

  › Dial(for higher quality, dial a number based on your current location):

  › US: +1 646 558 8656 or +1 669 900 6833 or +1 855 880 1246 (Toll Free) or +1 877 369 0926 (Toll Free)

› Meeting ID: 430 697 670

› International numbers available: https://zoom.us/u/achYtcw7uN

LF AI & DATA

Legal Notice

› The Linux Foundation, The Linux Foundation logos, and other marks that may be used herein are owned by The Linux Foundation or its affiliated entities, and are subject to The Linux Foundation's Trademark Usage Policy at https://www.linuxfoundation.org/trademark-usage, as may be modified from time to time.

› Linux is a registered trademark of Linus Torvalds. Please see the Linux Mark Institute's trademark usage page at https://lmi.linuxfoundation.org for details regarding use of this trademark.

› Some marks that may be used herein are owned by projects operating as separately incorporated entities managed by The Linux Foundation, and have their own trademarks, policies and usage guidelines.

› TWITTER, TWEET, RETWEET and the Twitter logo are trademarks of Twitter, Inc. or its affiliates.

› Facebook and the "f" logo are trademarks of Facebook or its affiliates.

› LinkedIn, the LinkedIn logo, the IN logo and InMail are registered trademarks or trademarks of LinkedIn Corporation and its affiliates in the United States and/or other countries.

› YouTube and the YouTube icon are trademarks of YouTube or its affiliates.

› All other trademarks are the property of their respective owners. Use of such marks herein does not represent affiliation with or authorization, sponsorship or approval by such owners unless otherwise expressly specified.

› The Linux Foundation is subject to other policies, including without limitation its Privacy Policy at https://www.linuxfoundation.org/privacy and its Antitrust Policy at https://www.linuxfoundation.org/antitrust-policy. each as may be modified from time to time. More information about The Linux Foundation's policies is available at https://www.linuxfoundation.org.

› Please email legal@linuxfoundation.org with any questions about The Linux Foundation's policies or the notices set forth on this slide.

LF AI & DATA