

Meeting of the Technical Advisory Council (TAC)

January 14, 2021

 **DLF** AI & DATA

Anti-Trust 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 Updegrave of the firm of Gesmer Undergone LLP, which provides legal counsel to the Linux Foundation.

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>
- ›
- › 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>

Agenda

- › Roll Call (2 mins)
- › Approval of Minutes (2 mins)
- › Welcome new Associate Member (2 minutes)
 - › University of Washington Tacoma
- › Project Incubation Proposal + Q&A + TAC Vote (35 minutes)
 - › Data Lifecycle Framework (Yiannis Gkoufas <YIANNISG@ie.ibm.com>)
- › Invited Presentation (15 minutes)
 - › Project Stages (Ibrahim Haddad)
- › LF AI General Updates (2 minutes)
- › Open Discussion (2 minutes)

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 |
| SAS | Nancy Rausch | nancy.rausch@sas.com |
| Tech Mahindra | Nikunj Nirmal | nn006444@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 |
| ONNX | Jim Spohrer (Chair of TAC) | spohrer@us.ibm.com |

Approval of December 17th, 2020 Minutes

Draft minutes from the December 17th TAC call were previously distributed to the TAC members via the mailing list
















Proposed Resolution:

- › That the minutes of the December 17th meeting of the Technical Advisory Council of the LF AI & Data Foundation are hereby approved.

Welcome new associate member!

University of Washington Tacoma

LF AI & Data Member Company - Associate (15)

| | | | | | |
|---|---|---|---|---|---|
|  <p>AI for people AI For People</p> |  <p>aivancity PARIS-CACHAN</p> |  <p>AMBIANIC</p> |  <p>EIL</p> |  <p>The International Society of Service Innovation Professionals</p> |  <p>MAIEI</p> |
| <p>AI for People AI For People</p> | <p>aivancity School for Technology, Business & Society aivancity school for Technology, Business & Society</p> | <p>Ambianic Ambianic</p> | <p>Institute of Ethical AI and Machine Learning The Institute for Ethical AI & Machine Learning</p> | <p>International Society of Service Innovation Professionals International Society of Service Innovation Professionals</p> | <p>Montreal AI Ethics Institute Montreal AI Ethics Institute</p> |
|  <p>NYU</p> |  <p>Open 启智 新一代人工智能开源开放平台</p> |  <p>Peng Cheng Laboratory</p> |  <p>Penn State Great Valley</p> |  <p>PSIT Kanpur</p> |  <p>shopen 上海开源信息技术协会</p> |
| <p>New York University New York University Funding: \$5M</p> | <p>Open Open</p> | <p>Peng Cheng Laboratory Peng Cheng Laboratory</p> | <p>Penn State University Pennsylvania State University Funding: \$13.4M</p> | <p>Pranveer Singh Institute Of Technology Pranveer Singh Institute Of Technology</p> | <p>Shanghai Open Source Information Technology Association Shanghai OpenSource Information Technology Association</p> |
|  <p>UNIVERSITY of WASHINGTON TACOMA</p> |  <p>ULT UNIVERSITÉ LIBRE DE TUNIS</p> |  <p>XPRIZE</p> | | | |
| <p>University of Washington - Tacoma University of Washington Funding: \$87.53M</p> | <p>Université Libre de Tunis Université Libre de Tunis</p> | <p>XPRIZE Foundation XPRIZE Funding: \$1.75M</p> | | | |

Project Hosting Proposal - Data Lifecycle Framework

Project Incubation Proposal Review/Discussion/Vote: Data Lifecycle Framework

Description:

DLF is enabling and accelerating data access for Kubernetes/OpenShift workloads in a transparent and declarative way. An open source project since September of 2019 with growing multi-vendor community support and use-cases related to data access in AI projects, the projects benefits include:

Data scientists/engineers: Focus on workload/experiments development and not on configuring/tuning data access

Storage Providers: Increase adoption since the framework is extensible without hindering the User Experience

Data-oriented Frameworks: Can build capabilities (caching, scheduling) on top of DLF using a declarative way to access/manage data sources

Currently it supports connection to S3,NFS,HostPath-based data sources and will expand to support more in the future.

Presenter: Yiannis Gkoufas <YIANNISG@ie.ibm.com>

Resources:

Github: <https://github.com/IBM/dataset-lifecycle-framework>

LF AI & Data website: <https://github.com/lfaidata/proposing-projects/pull/32/commits/88da22cf4984d0d942cadba6bb26a464e528a484>

Project Level Request to TAC: Incubation

Dataset Lifecycle Framework

Yiannis Gkoufas

+ 3 Committers from **IBM Research – Europe**

github.com/IBM/dataset-lifecycle-framework

ibm.github.io/dataset-lifecycle-framework

*A meta-framework to
transparently **enable** and
accelerate Data access for
Kubernetes/OpenShift workloads*

Background: Kubernetes and Data

Kubernetes < 1.6

Mostly used for stateless applications. Users could only leverage local storage

Kubernetes 1.6

Introduction of Dynamic Storage Provisioning and Storage Classes

Kubernetes 1.13

Container Storage Interface becomes GA

Motivation for our work

- **Non-power kubernetes users** face complexities when accessing remote data
- Becomes increasingly common to store/retrieve data from remote stores in **Machine Learning and AI workloads**
- **Kubernetes administrators** have tools to limit/track resource usage (CPUs, RAM) but not for data access

Dataset Lifecycle Framework: Overview

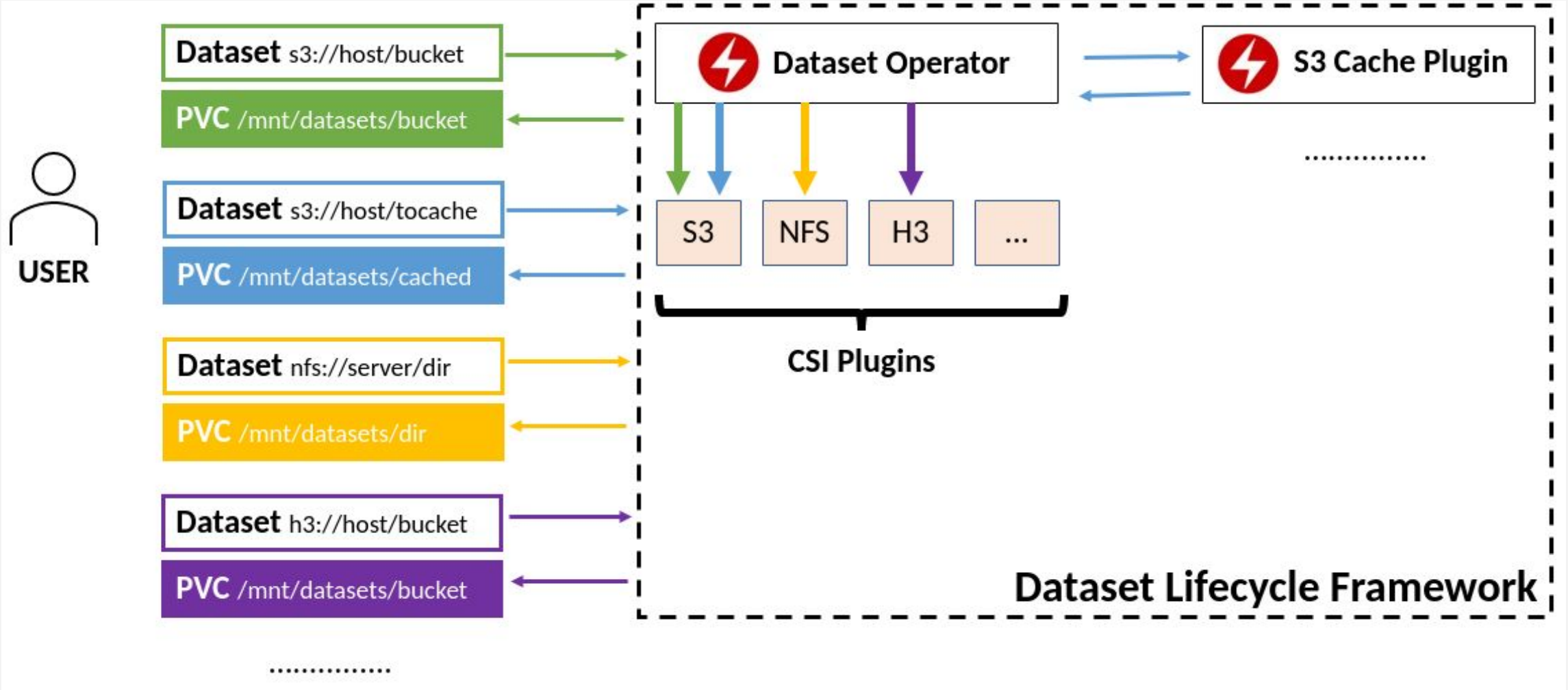
- A Kubernetes meta-Framework to provide easy access to S3 and NFS **Datasets** within pods.
- Orchestrates the provisioning of **Persistent Volume Claims** and **ConfigMaps** needed for each **Dataset**.

Goals

Improve **User Experience** and **Performance** of workloads

Be **highly extensible** to support all major types of Data sources

High-level Functionality



Dataset Definition

```
apiVersion: com.ie.ibm.hpsys/v1alpha1
kind: Dataset
metadata:
  name: test
spec:
  local:
    type: "COS"
    accessKeyId: "testKeyId"
    secretAccessKey: "testKey"
    endpoint: "https://s3.eu.cloud-object-storage.appdomain.cloud"
    bucket: "test-yiannis"
    region: "" #it can be empty
```

```
apiVersion: v1
kind: Pod
metadata:
  name: nginx
  labels:
    dataset.0.id: "test"
    dataset.0.useas: "mount"
spec:
  containers:
    - name: nginx
      image: nginx
```



MINIO

Benefits

- **Data scientists/engineers:** Focus on workload/experiments development and not on configuring/tuning data access
- **Storage Providers:** Increase adoption since the framework is extensible without hindering the User Experience
- **Data-oriented Frameworks:** Can build capabilities (caching, scheduling) on top of DLF using a declarative way to access/manage data sources

Opensource Community Engagements/Integrations

Presentations in Community Calls: Kubernetes SIG Storage, CNCF Sig Storage, Kubeflow, OpenDataHub, Noobaa



European Bioinformatics Institute

<https://www.ebi.ac.uk/>

Leveraged DLF within their Kubeflow Pipelines to access 1000genomes Data stored in S3 Buckets

<https://github.com/IBM/dataset-lifecycle-framework/issues/38>
<https://github.com/IBM/dataset-lifecycle-framework/issues/36>
<https://github.com/IBM/dataset-lifecycle-framework/issues/30>



KFData Proposal (Pachyderm)

<https://github.com/pachyderm/kfdata>

Proposal for Kubeflow to standardize data access in Kubeflow Pipelines

Forked DLF in their repo and presented demo in Kubeflow community using DLF <https://ibm.biz/BdqkEq>

Opensource Community Engagements/Integrations



OpenDataHub (Red Hat)

<https://opendatahub.io/>

Expressed interest in integrating DLF for data access

<https://github.com/IBM/dataset-lifecycle-framework/issues/40>



Noobaa (Red Hat)

<https://www.noobaa.io/>

Creation of Caching Plugin, presentation to their community call

Github Metrics

Released: September 2019

63 Stars, 20 Forks (as of Jan 2020)

Top 6.5% in IBM Github Organization

36 Closed Pull Requests, 16 Closed Issues

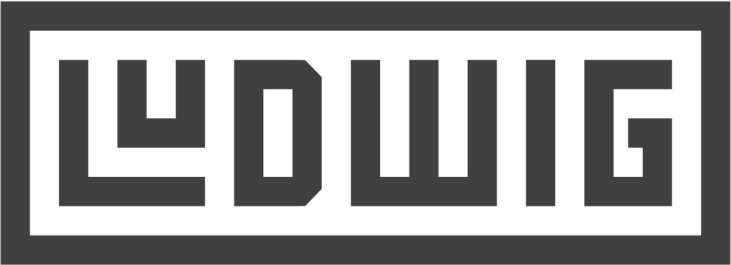
4 IBM Maintainers

Contributions from at least 4 external organizations

Roadmap

- **Establish Caching Interface Architecture and APIs**
Implementation of Noobaa-, Spectrum Scale-based Caching Plugins
Sample implementation for developers
- **Security Hardening**
Exploration for Vault-based keys management
- **Scheduling Hints**
Assist the default Kubernetes scheduler to increase data-locality for workloads
- **Improvements on integration tests**

Potential Adopters/Users within LF AI & Data



THANK YOU!

Github: <https://github.com/IBM/dataset-lifecycle-framework>

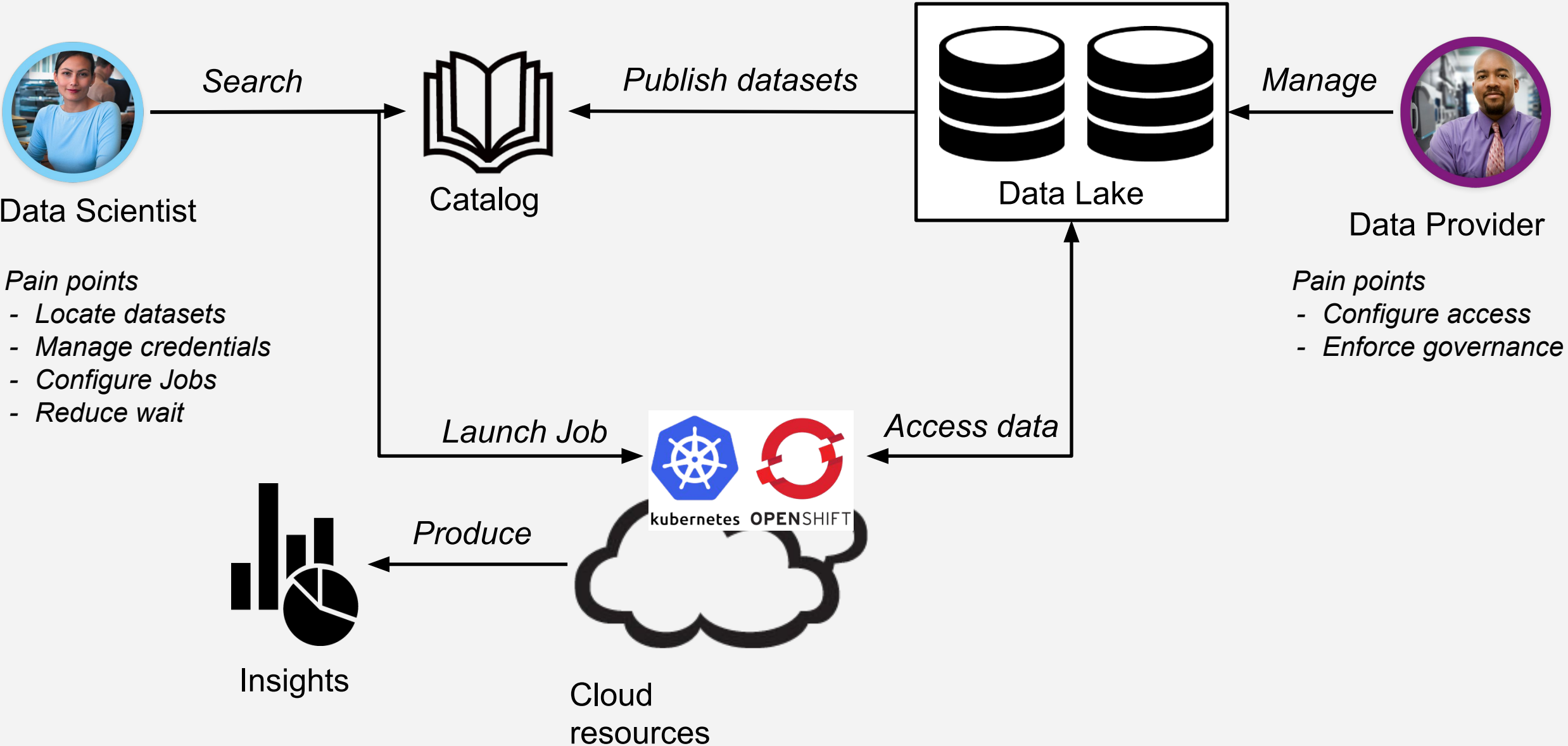
Website: <https://ibm.github.io/dataset-lifecycle-framework/>

Demo: <https://asciinema.org/a/276331>

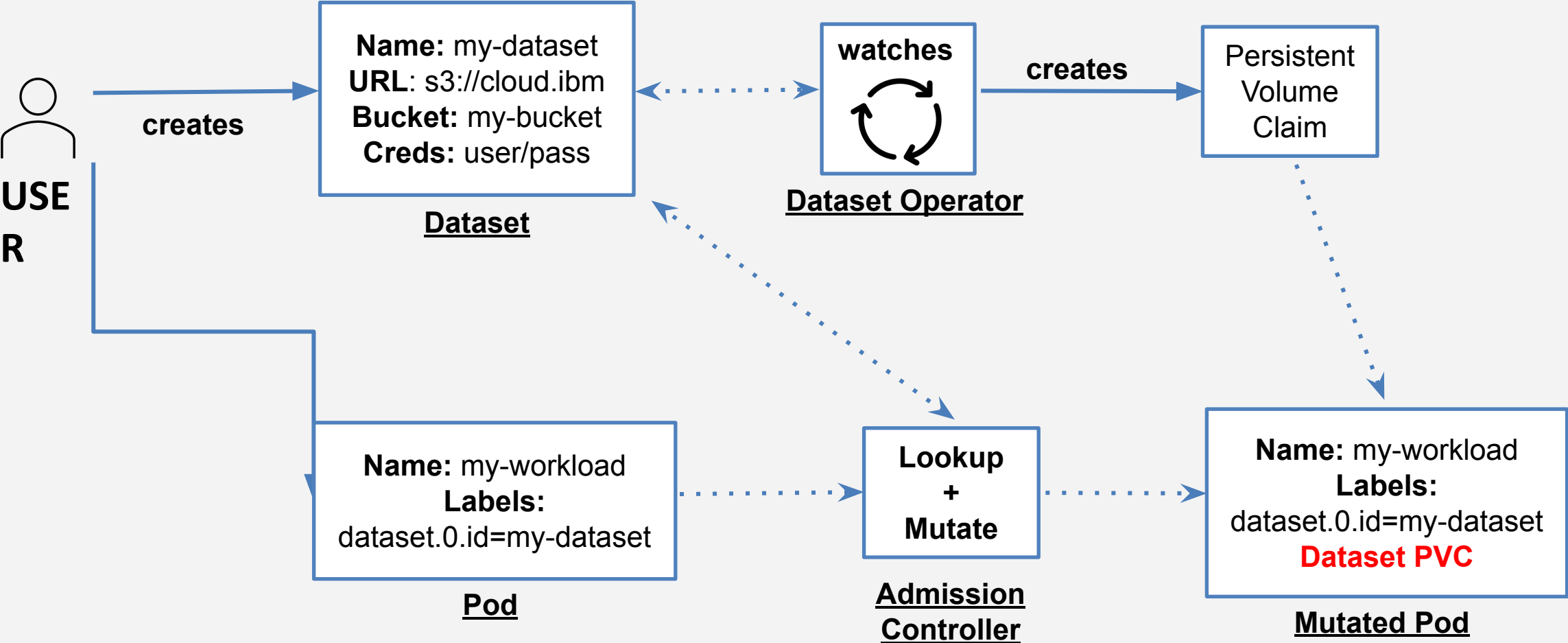
Contact: yiannisg@ie.ibm.com

APPENDIX

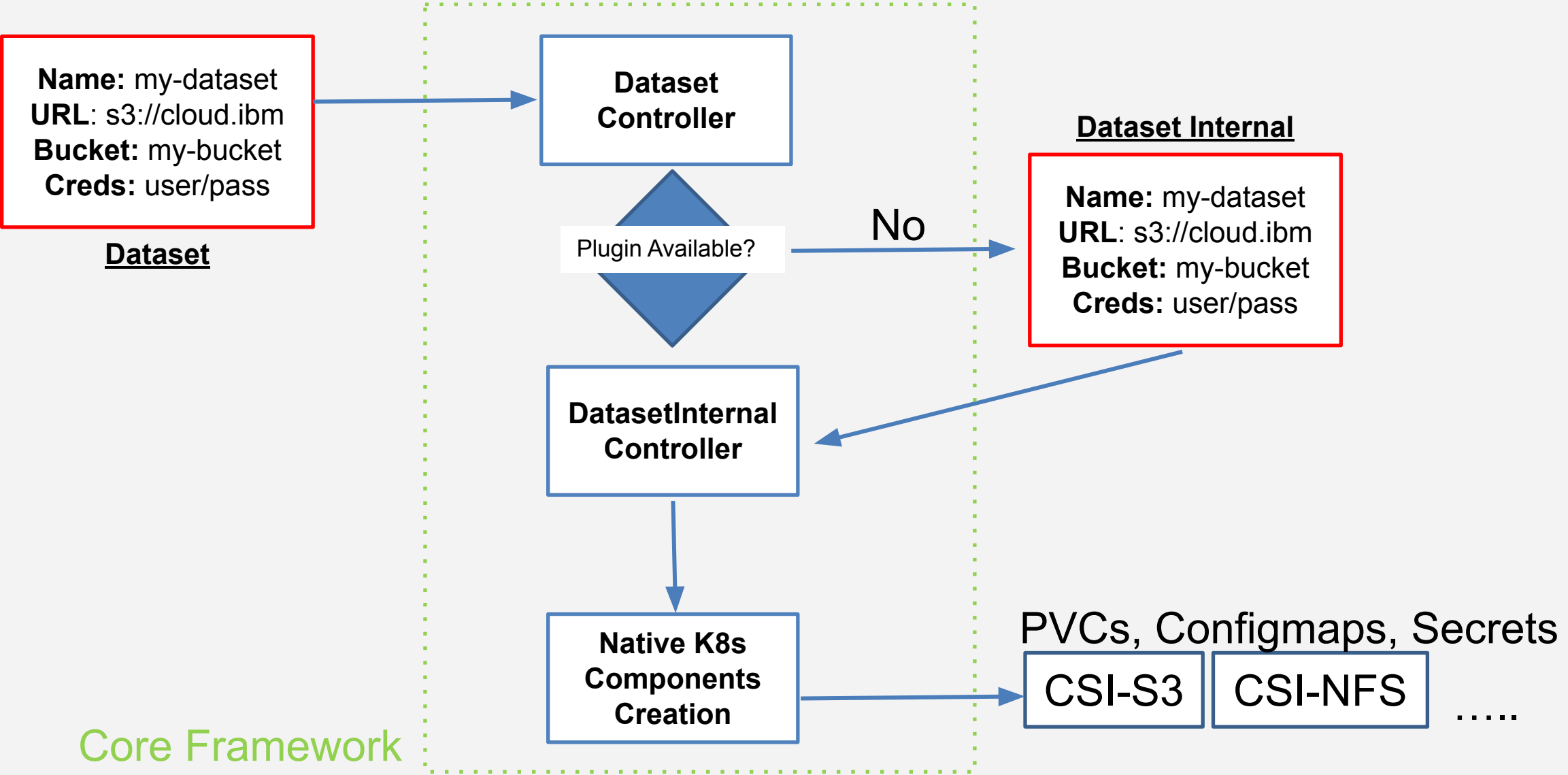
Scenario



Datasets Lifecycle Framework - Components

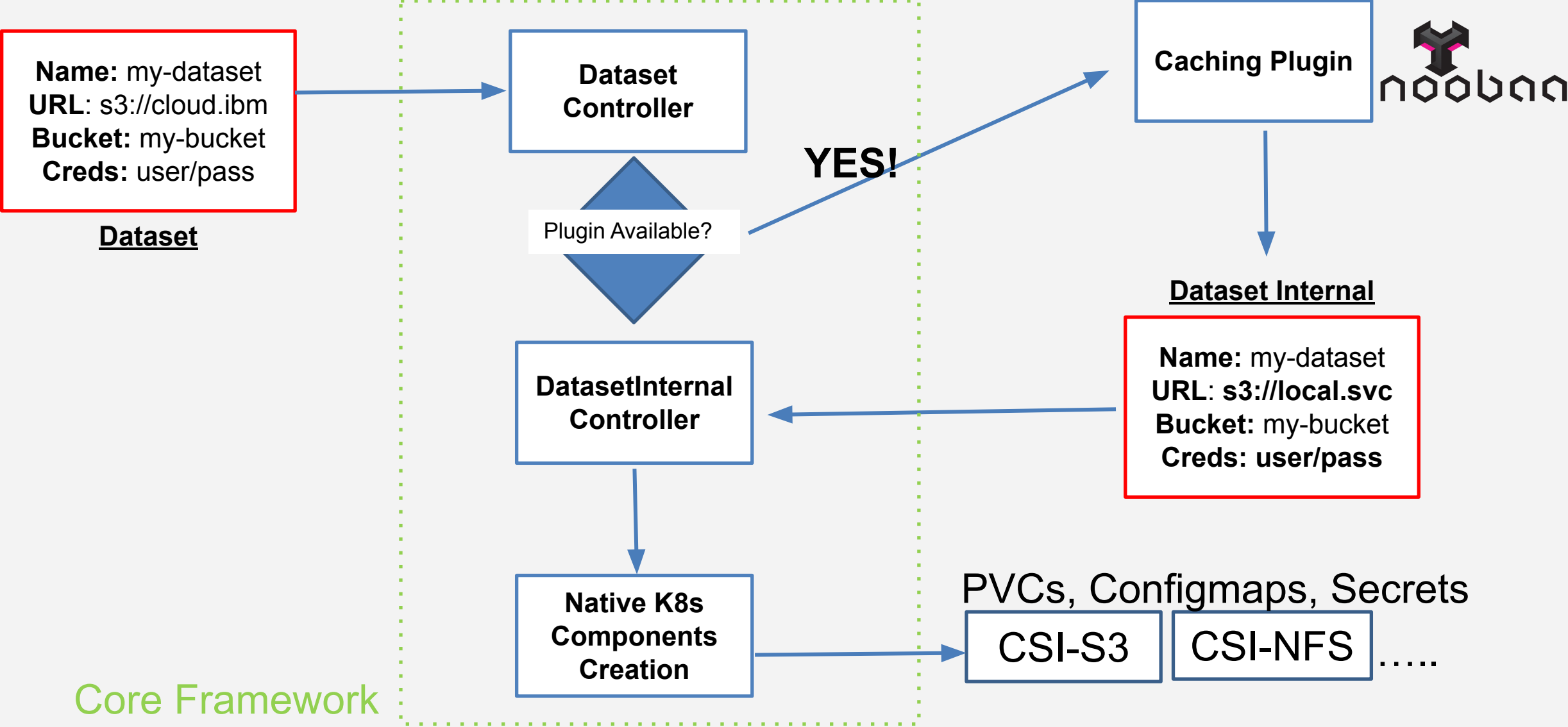


Datasets Lifecycle Framework – Transparent Caching



Core Framework

Datasets Lifecycle Framework – Transparent Caching



TAC Vote on Project Incubation Proposal: Data Lifecycle Framework

Proposed Resolution:

The TAC approves the Data Lifecycle Framework as an Incubation project of the LF AI & Data Foundation

Next Steps

LF AI & Data staff will work with Data Lifecycle Framework to onboard the project leading to the announcement of the project joining LF AI & Data

Explore potential integrations between the project and other LF AI & Data projects

Integrate the project with LF AI & Data operations

Invited Presentation - Project Stages

Presenter

- Ibrahim Haddad

LF AI & Data Project Lifecycle Document

Ibrahim Haddad, Ph.D.
Executive Director, LF AI & Data
Ibrahim@LinuxFoundation.org

 LF AI & DATA

Background

The LF AI & Data Project Lifecycle Document defines the project levels, requirements to be accepted in each level, process and various associated details.

It is approved by the Technical Advisory Council (TAC) and then the Governing Board (GB).

Current version dates May 2018.

Revisiting the Document

Over 2 years since the document was last revised. A lot of progress has been made in terms of new projects joining.

A lot of experience gained in onboarding projects and insights on improvements to be made including higher the bar to join the foundation and to graduate.

We've received numerous feedback and examined how more mature umbrella foundation operate and structure their projects' stages and lifecycle.

Key Updates Introduced to the Document

1. Introducing Sandbox stage
2. Improving requirements to incubate projects
3. Improving requirements to graduate projects
4. Adding specific language to clarify the benefits for projects hosted in every stage
5. Elaborating on the Archive Stage projects to eliminate ambiguities
6. Adding information on the Annual Review of projects
7. General edits for the purpose of clarity

1. Introducing Sandbox Stage

This stage is specific to projects that meet one of the following requirements:

- Any project that intends to join LF AI & Data Incubation in the future and wishes to lay the foundations for that.
- New projects that are designed to extend one or more LF AI & Data projects with functionality or interoperability libraries.
- Independent projects that fit the LF AI & Data mission and provide the potential for a novel approach to existing functional areas (or are an attempt to meet an unfulfilled need).



2. Improving requirements to incubate projects

To be accepted into the Incubation stage, a project must meet all the requirements of the Sandbox stage plus:

- Have [at least two organizations](#) actively contributing to the project.
- Have a defined Technical Steering Committee (TSC) with a chairperson identified, with open and transparent communication.
- Have [a sponsor who is an existing LF AI & Data member. Alternatively, a new organization would join LF AI & Data and sponsor the project's incubation application.](#)
- Have [at least 300 stars on GitHub](#); this is an existing requirement for a project to be listed on the LF AI & Data landscape.
- Have achieved and maintained a Core Infrastructure Initiative [Best Practices Silver Badge](#).
- In addition to the affirmative vote of the TAC, incubation stage projects also require the affirmative vote of the Governing Board.

3. Improving requirements to graduate projects

To be accepted into the Graduation stage, a project must meet the Incubation stage requirements plus:

- Have a healthy number of code contributions coming from [at least five organizations](#).
- Have reached a [minimum of 1000 stars on GitHub](#).
- Have achieved and maintained a Core Infrastructure Initiative [Best Practices Gold Badge](#).
- Have demonstrated a substantial ongoing flow of commits and merged contributions for the past 12 months*.
- Receive the affirmative vote of two-thirds of the TAC and the affirmative vote of the Governing Board.
- Have completed at least one collaboration with another LF AI & Data hosted project
- Have a technical lead appointed for representation of the project on the LF AI & Data Technical Advisory Council.

4. Adding specific language to clarify the benefits for projects hosted in every stage

5. Elaborating on the Archive Stage projects to eliminate ambiguities

To archive a project:

- A proposal must be put forth to the TAC
- The proposal must remain open for at least 2 weeks of discussion
- A vote must be finalized with 2/3 approval from the TAC

What does archiving for an LF AI & Data project mean?

- LF AI & Data will no longer provide support for the project
- LF AI & Data will list archived projects online
- Trademarks and domain names of archived projects are still hosted by the LF AI & Data and the Linux Foundation
- LF AI & Data can provide services such as documentation updates to help transition users.
- Other LF AI & Data marketing activities will no longer be provided for the project

Reactivating an Archived Project

Any project can be reactivated into LF AI & Data by following the normal project proposal.

6. Adding information on the Annual Review of projects

- The TAC will undertake an annual review of all LF AI & Data projects.
- This annual review will include an assessment as to whether projects in Sandbox and Incubation are making adequate progress towards the Graduation stage; and that projects in the Graduation stage are maintaining positive growth and adoption
- Any project may be moved to Archive stage by affirmative vote of the TAC, provided, that in the case of any Graduation stage project both of the following conditions must be met: (a) the affirmative vote of the TAC must be of at least two-thirds of the TAC and (b) the transition to Archive stage must also be approved by the affirmative vote of the Governing Board.

Existing Projects

No impact.

New projects coming into the Foundation after the Governing Board has approved the new Document will need to follow the updated process, requirements.

Request for Feedback

Call for comment is open until Friday Jan 22nd.

The finalized document will be put for a TAC vote on Jan 28th then passed to the GB for approval.

After GB approval, the document will be posted online and goes in effect.

Joining LF AI & Data is Easy!



General Inquiries:

info@lfaidata.foundation

 LF AI & DATA

LF AI & Data - General Updates

 LF AI & DATA

| | | | | | | | | | |
|------------------|-----------|----------|---------|-----------|----------|---------|------|------------------------|-------------|
| Machine Learning | Framework | Platform | Library | Framework | Platform | Library | Tool | Reinforcement Learning | Programming |
| | | | | | | | | | |




| | | | | | | | | | | |
|----------------------|------------|----------------|------------|-------------------|------------|---------------------|---------------|---------------------|-------------------------|------------|
| Notebook Environment | Versioning | Store & Format | Operations | Stream Processing | SQL Engine | Feature Engineering | Visualization | Pipeline Management | Labeling and Annotation | Governance |
| | | | | | | | | | | |







| | | | | | | | | | | | |
|-------|--------------|----------|-----------|--------------------|-------------|----------|-----------|------|----------------|-------------|-----------------|
| Model | Benchmarking | Training | Parameter | Format & Interface | Marketplace | Workflow | Inference | Tool | Explainability | Adversarial | Bias & Fairness |
| | | | | | | | | | | | |








| | | | | | |
|-----------------------|------------------------|-----------|--------------------|-----------------------------|-----------|
| Distributed Computing | Computing & Management | Interface | Security & Privacy | Natural Language Processing | Education |
| | | | | | |

The LF AI & Data landscape explores open source projects in Artificial Intelligence and Data and their respective domains.

l.fai.foundation

| | | | | | | | | | |
|------------------|-----------|--|--|-----------|----------|---------|------|------------------------|--|
| Machine Learning | Framework | Platform | Library | Framework | Platform | Library | Tool | Reinforcement Learning | Programming |
| | |  LF AI & Data |  LF AI & Data | | | | | |  LF AI & Data |

| | | | | | | | | | | | |
|----------------------|----------------------|------------|---|--|---|------------|---------------------|---------------|---------------------|-------------------------|---|
| Notebook Environment | Notebook Environment | Versioning | Store & Format | Operations | Stream Processing | SQL Engine | Feature Engineering | Visualization | Pipeline Management | Labeling and Annotation | Governance |
| | | |  LF AI & Data |  LF AI & Data  LF AI & Data  LF AI & Data <small>Incubating</small> |  LF AI & Data | | | | | |  LF AI & Data |

| | | | | | | | | | | | |
|-------|--------------|---|---|---|---|----------|---|------|---|---|---|
| Model | Benchmarking | Training | Parameter | Format & Interface | Marketplace | Workflow | Inference | Tool | Explainability | Adversarial | Bias & Fairness |
| | |  LF AI & Data |  LF AI & Data |  LF AI & Data |  LF AI & Data | |  LF AI & Data | |  LF AI & Data |  LF AI & Data |  LF AI & Data |

| | | | | | | | | | |
|-----------------------|---|---|--|---|---|--|--------------------|---|---|
| Distributed Computing | Computing & Management | Interface |  The LF AI & Data landscape explores open source projects in Artificial Intelligence and Data and their respective sub-domains. lfaidata.foundation | | | | Security & Privacy | Natural Language Processing | Education |
| |  LF AI & Data |  LF AI & Data |  LF AI & Data |  LF AI & DATA Landscape |  LF AI & DATA | | |  LF AI & Data |  LF AI & Data  LF AI & Data <small>Incubating</small> |

Suggested Additions

Project Key

Yellow = not in [Landscape](#), maybe should be added

Programming

[Numpy](#)
[Numba](#)
[SciPy](#)
[Dask](#)
[Julia](#) (*)
[Python](#)
[Rstudio](#)

Notebooks

[Flyra](#)
[I-python](#)
[Jupyter Notebooks](#)
[PixieDust](#)
[Rmarkdown](#)

Security & Privacy

[HE-Lib](#) (*)
[TensorFlow Privacy](#)
[TF-Encrypted](#)

Distributed Computing

Management
[OpenShift](#)
[Kubernetes](#)
[Mesos](#)
[Ranger](#)
[Storm](#)

Interface
[Sparklyr](#)
[Toree](#)
[Livy](#)
[Spark-NLP](#)

Data

Versioning
[Pachyderm](#) (*)

Store & Format
[Alluxio](#)
[Arrow](#)
[Avro](#)
[Delta Lake](#) (*)

[Druid](#)
[JanusGraph](#)
[Parquet](#)
[Ceph](#)

Stream Processing

[Flink](#)
[Kafka](#)
[Logstash](#) (*)
[FluentD](#) (*)

Relational DB

[Postgres](#)
[MySQL](#)
[CouchDB](#)

SQL Engine
[Presto](#) (*)

Visualization

[Bokeh](#)
[D3](#)
[Plotly](#)
[Facets](#)
[Grafana](#)
[Seaborn](#)
[Superset](#) (*)
[TensorBoard](#)
[Prometheus](#)

Data

Governance
[Egeria](#)
[CLDA](#)

Feature Engineering
[Tsfresh](#)

Operations
[FEAST](#) (*)
[Amundsen](#) (*)
[Hive](#) (*)
[Snorkel](#) (*)

Pipeline Management
[Beam](#)

Labeling & Annotation
[Vott](#) (*)

Exploration
[Hue](#)
[Kibana](#)

Machine Learning

Framework
[LightGBM](#)
[Mahout](#)
[Ray](#) (*)

Platform
[Kubeflow](#)
[H2O](#)
[SystemML](#)
[Mlflow](#) (*)
[Seldon](#) (*)
[Marvin-AI](#) (*)

Library
[Scikit-learn](#)
[XGBoost](#)
[cat-boost](#)
[SparkML](#)

Deep Learning

Framework
[TensorFlow](#)
[PyTorch](#)
[MX-Net](#)

Library
[Keras](#)

Reinforcement Learning

[DeepMind Lab](#) (*)
[OpenAI Gym](#) (*)

Model

Inference
[TensorRT](#)
[TensorRT Inference](#)

Benchmarking
[MLPerf](#)

Training
[Horovod](#) (*)

Parameter
[HyperOpt](#)
[Katib](#)

Format & Interface
[ONNX](#)

Marketplace
[MAX](#) (*)

Workflow
[Kubeflow Pipelines](#)
[Tekton](#)

[Airflow](#) (*)
[Nifi](#) (*)
[Argp](#) (*)
[Mleap](#) (*)
[Volcano](#) (*)

Tool
[KFServing](#)
[ONNX Runtime](#)
[TorchServe](#) (*)
[Clipper](#) (*)
[MMS](#) (*)

Trusted AI

Explainability
[AI Explainability 360](#)
[Alibi](#) (*)
[LIME](#)
[SHAP](#)

Bias & Fairness
[AI Fairness 360](#)

Adversarial Attacks
[Adversarial Robustness Toolbox](#)

Natural Language Processing

[UIMA](#)
[BERT](#)
[Core NLP](#)
[Lucene](#)
[PyText](#)
[Spacy](#)
[Transformers](#) (*)

Education
[OpenDS4All](#)

2020 TAC Meetings Summary

| | | | |
|-------------------|---|---|---|
| Jan Feb Mar | 16: Milvus (Zilliz)* | 13: <i>MLOps Work (LF CD)</i> 27: <i>Collective Knowledge (Coral Reef)</i> | 12: NNStreamer (Samsung)* 26: ForestFlow (?)* |
| Apr May Jun | 9: <i>Trusted AI & ML Workflow (LF)</i> 23: <i>Open Data Hub (Red Hat)</i> | 7: Ludwig (Uber)* 21: <i>SnapML (IBM)</i> | 4: <i>Trusted AI (AI for Good, Ambianic.ai, MAIEI)</i> 18: Fairness, Explainability, Robustness (IBM)* |
| Jul Aug Sep | 16: <i>Mindspore (Huawei)</i> 30: Amundsen (Lyft)* | 16: <i>Delta (Didi)</i> 16: Horovod (Uber/LF)** 30: <i>ModelDB (?)</i> 30: <i>Egeria, OpenDS4All, BI&AI (LF ODPI)</i> | 10: SOAJS (HeronTech)* 10: Delta (Didi)* 24: FEAST (Gojek)* 24: Egeria, (LF ODPI)** 24: OpenDS4All (ODPI)* 24: BI&AI Committee (ODPI) |
| Oct Nov Dec | 8: <i>Fairness, Explainability, Robustness (LF)</i> 22: <i>OpenLineage (DataKins)</i> 22: <i>IDA (IBM/Salesforce)</i> | 5: DataPractices.Org (WorldData/LF)* 5: <i>Kubeflow-On-Prem (Google, Arrikto/Intel)</i> 19: <i>OpenDS4All, DataPractices.Org, edX Ethical AI (LF)</i> | 3: TBD - JanusGraph (LF)* 3: <i>TBD - RosaeGL (?)</i> 17: TBD – Seldon Core (Seldon)* 17: TBD – Pyro (Uber/LF)** |

(Entity)* = incubating vote

**** bold = graduate vote**

Italics = invited project presentation

2021 TAC Meetings Pipeline Summary

| | | | |
|-------------------|--|---|---|
| Jan Feb Mar | 14: Data Lifecycle Framework (IBM)* 28: Tentative: Verse (Seldon) | 11: MARS (Aliabab) 25: Flyte (Lyft) | 11: Streams (IBM) 25: Tentative: Substra Framework |
| Apr May Jun | 8: Adlik (ZTE)** 22: Kubeflow-On-Prem (Google, Arrikto, Intel) | ?: Ray (Anyscale.io) ?: Pachyderm (Pachyderm) ?: DataHub (LinkedIn) | ?: Common Knowledge (Code Reef) ?: Couler (Ant Financial) |
| Jul Aug Sep | ?: KubeflowServing (Google, Arrikto, Seldon) | ?: Kubeflow Pipeline (Google, Bloomberg) | ?: Open Data Hub (Red Hat) |
| Oct Nov Dec | ?: Vespa (Verizon Media) | ?: Snorkle (Snorkle) ?: Plotly (DASH) ?: Mellody (Substra) ?: mloperator (Polyaxen) ?: SnapML (IBM) | ?: PMML/PFA (DMG.org) ?: Mindspore, Volcano (Huawei) ?: TransmorgrifAI (Salesforce) ?: AIMET (Qualcomm) ?: Elyra-AI (IBM) |

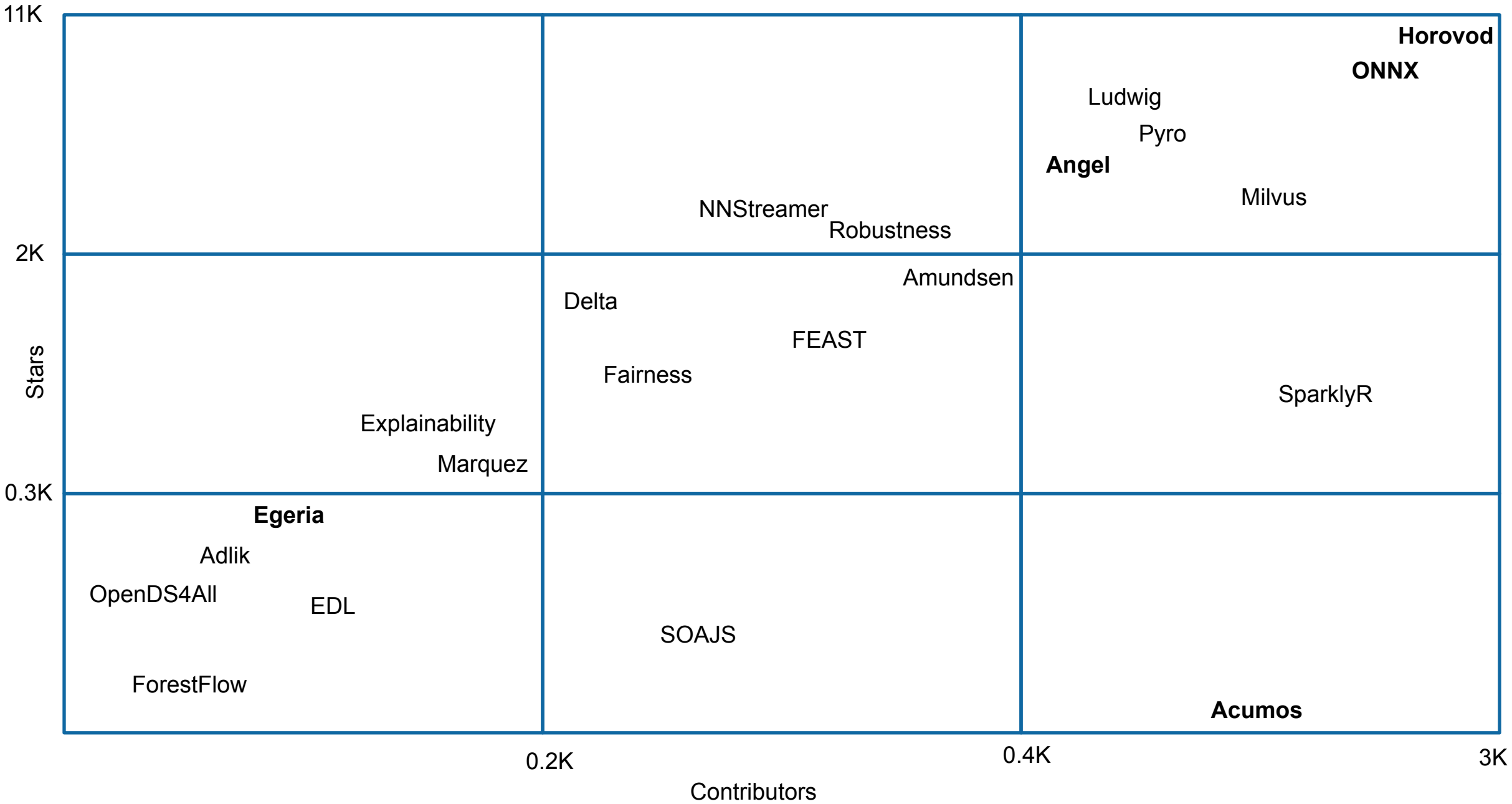
(Entity)* = incubating vote

** **bold** = graduate vote

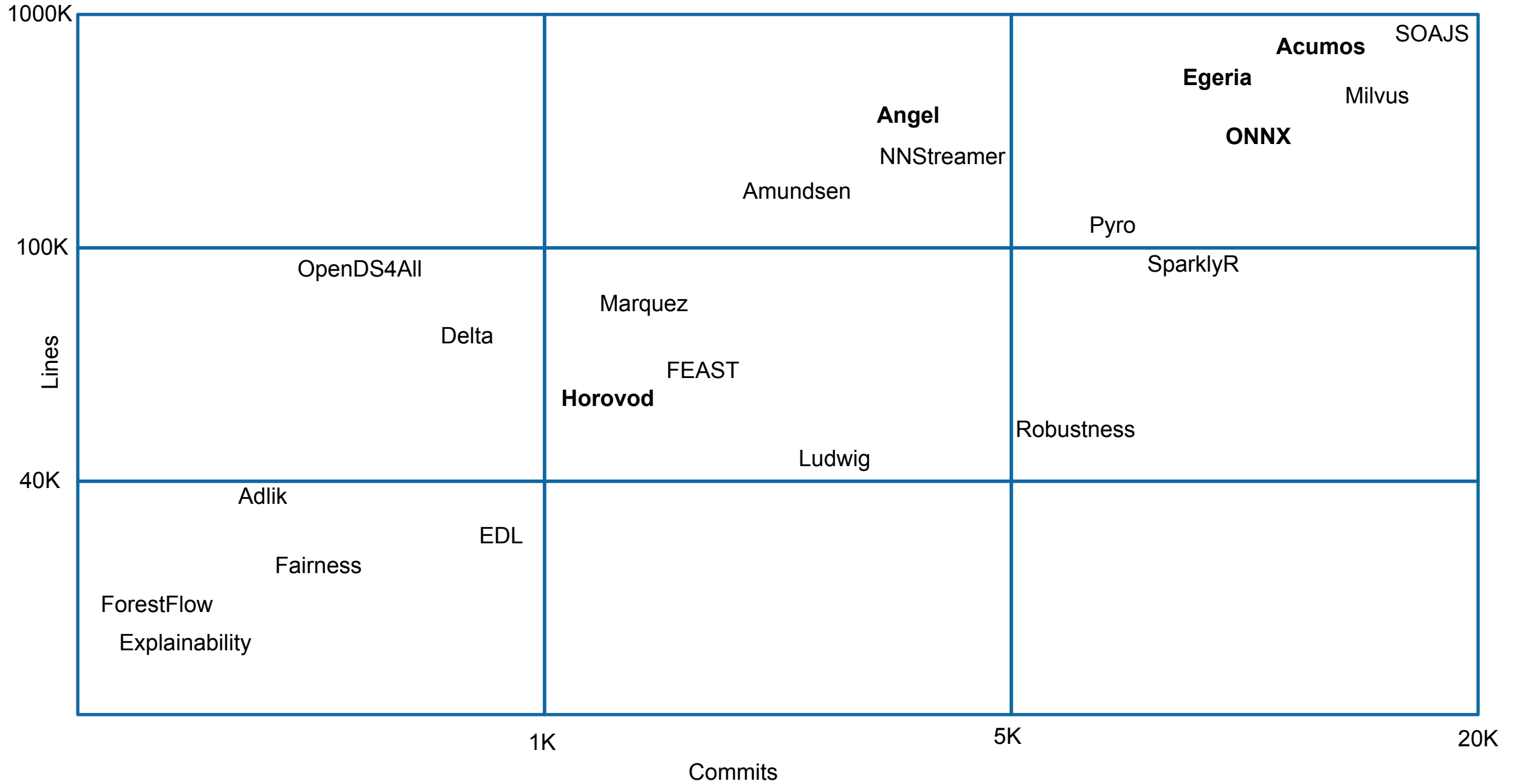
Italics = invited project presentation

Getting to know the projects more

Data from November 23, 2020 – Stars and Contributors



Data from November 23, 2020 – Lines of Code and Commits



Looking to host a project with LF AI & Data

- › Hosted project stages and life cycle:

<https://lfaidata.foundation/project-stages-and-lifecycle/>

- › Offered services for hosted projects:

<https://lfaidata.foundation/services-for-projects/>

- › Contact:

Jim Spohrer (TAC Chair) and Ibrahim Haddad (ED, 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@lfai.foundation to coordinate in advance (min 2 wks) of your expected release date.

Note on quorum

As LF AI & Data is growing, we now have 16 voting members on the TAC.

TAC representative - please ensure you attend the bi-weekly calls or email Jacqueline/Ibrahim to designate an alternate representative when you can not make it.

We need to ensure quorum on the calls especially when we have items to vote on.

Updates from Outreach Committee

Upcoming 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

› Please consider holding virtual events

To discuss participation, please email events@lfaidata.foundation

Upcoming Events

<https://lfaidata.foundation/events/>

- **March 24, 2021 - ONNX Community Virtual Meetup**
 - a. **Wednesday @ 5:00 pm - 8:00 pm PT USA**
Thursday @ 8:00am - 11am China Time
[LF AI Day: ONNX Community Virtual Meetup – March 2021](#)
(Virtual - Free - Asia-friendly time – Host Ti Zhou - Baidu)
- **August 4-6, 2021 - OSS North America, Vancouver**
 - a. **Mini-Summit, Booth, Track**
- **Sept 29 - Oct 1, 2021 - OSS Europe, Dublin**
 - a. **Mini-Summit, Booth, Track**

LF AI PR/Comms

- › Please follow LF AI & Data on [Twitter](#) & [LinkedIn](#) and help amplify news via your social networks - Please retweet and share!
 - › Also watch for news updates via the tac-general mail list
 - › View recent announcement on the [LF AI & Data Blog](#)
- › 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@lfaidata.foundation

Call to Participate in Ongoing Efforts

 **OLF** AI & DATA

Trusted AI

- › **Leadership:**
Animesh Singh (IBM), Souad Ouali (Orange), and Jeff Cao (Tencent)
- › **Goal:** Create policies, guidelines, tooling and use cases by industry
- › **Slack conversation channel:**
#trusted-ai-committee
<https://lfaifoundation.slack.com/archives/CPS6Q1E8G>
- › **Github:**
<https://github.com/lfai/trusted-ai>
- › **Wiki:**
<https://wiki.lfai.foundation/display/DL/Trusted+AI+Committee>
- › **Email lists:**
<https://lists.lfaidata.foundation/g/trustedai-committee/>
- › **Next call:** Monthly alternating times
<https://wiki.lfai.foundation/pages/viewpage.action?pageId=12091895>

ML Workflow & Interop

- › **Leadership:**
Huang “Howard” Zhipeng (Huawei)
- › **Goal:**
Define an ML Workflow and promote cross project integration
- › **Slack conversation channel:**
#ml-workflow
<https://lfaifoundation.slack.com/archives/C011V9VSMQR>
- › **Wiki:**
<https://wiki.lfaidata.foundation/pages/viewpage.action?pageId=10518537>
- › **Email lists:**
<https://lists.lfaidata.foundation/g/mlworkflow-committee>
- › **Next call:** Monthly check calendar/slack
<https://wiki.lfai.foundation/pages/viewpage.action?pageId=18481242>

BI & AI

- › **Leadership:**
Cupid Chan (Index Analytics)
- › **Goal:** Identify and share industry best practices that combine the speed of machine learning with human insights to create a new business intelligence and better strategic direction for your organization.

- › **Slack conversations channel:**
#bi-ai-committee
<https://lfaifoundation.slack.com/archives/C01EK5ND073>
- › **Github:**
<https://github.com/odpi/bi-ai>
- Wiki:**
<https://wiki.lfaidata.foundation/pages/viewpage.action?pageId=35160417>
- Email lists:**
<https://lists.lfaidata.foundation/g/biai-discussion>
- Next call:** Monthly community call TBD

Ongoing effort to create AI Ethics Training

Initial developed course by the LF: Ethics in AI and Big Data - published on edX platform:

<https://www.edx.org/course/ethics-in-ai-and-big-data>

The goal is to build 2 more modules and package all 3 as a professional certificate - a requirement for edX

- › **To participate:**
<https://lists.lfaidata.foundation/g/aiethics-training>

Upcoming TAC Meetings

Upcoming TAC Meetings (Tentative)

- › Jan 28th: Seldon
- › Feb 11: MARS (Alibaba)
- › Feb 25: FLYTE (Lyft)
- › Mar 11: Streams (IBM)
- › Mar 25: Substra
- › April 8: Adlik (ZTE)
- › April 22: TBD
- › May 6: All project updates

›

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

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>

Open Discussion

Mission

To build and support an open community and a growing ecosystem of open source AI, data and analytics projects, by accelerating innovation, enabling collaboration and the creation of new opportunities for all the members of the community

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.

