

LF AI & Data

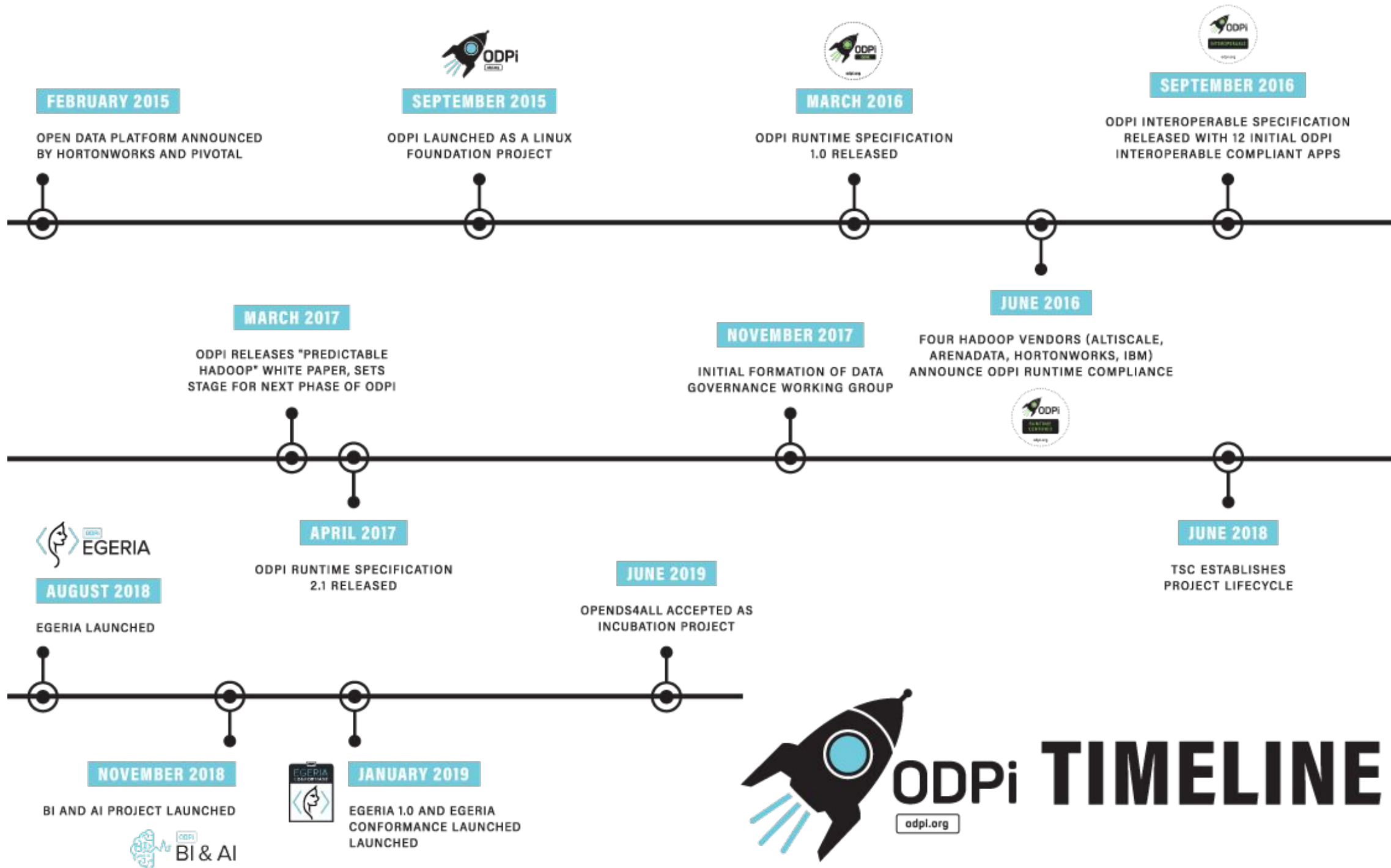
October 27, 2020

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

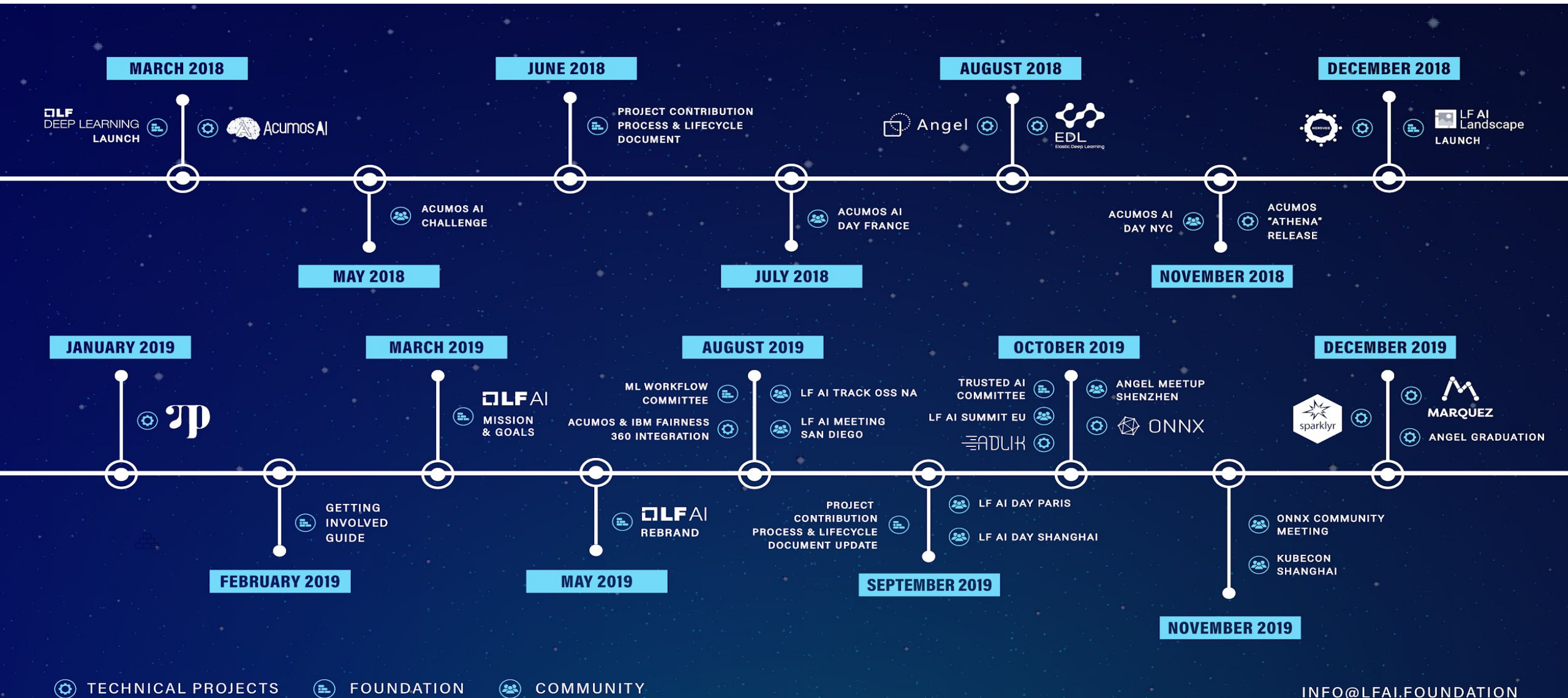
 LF AI & DATA

LF AI & ODPI joining efforts under LF AI & Data





LF AI 2018 - 2019 Timeline



LF AI 2020 Timeline

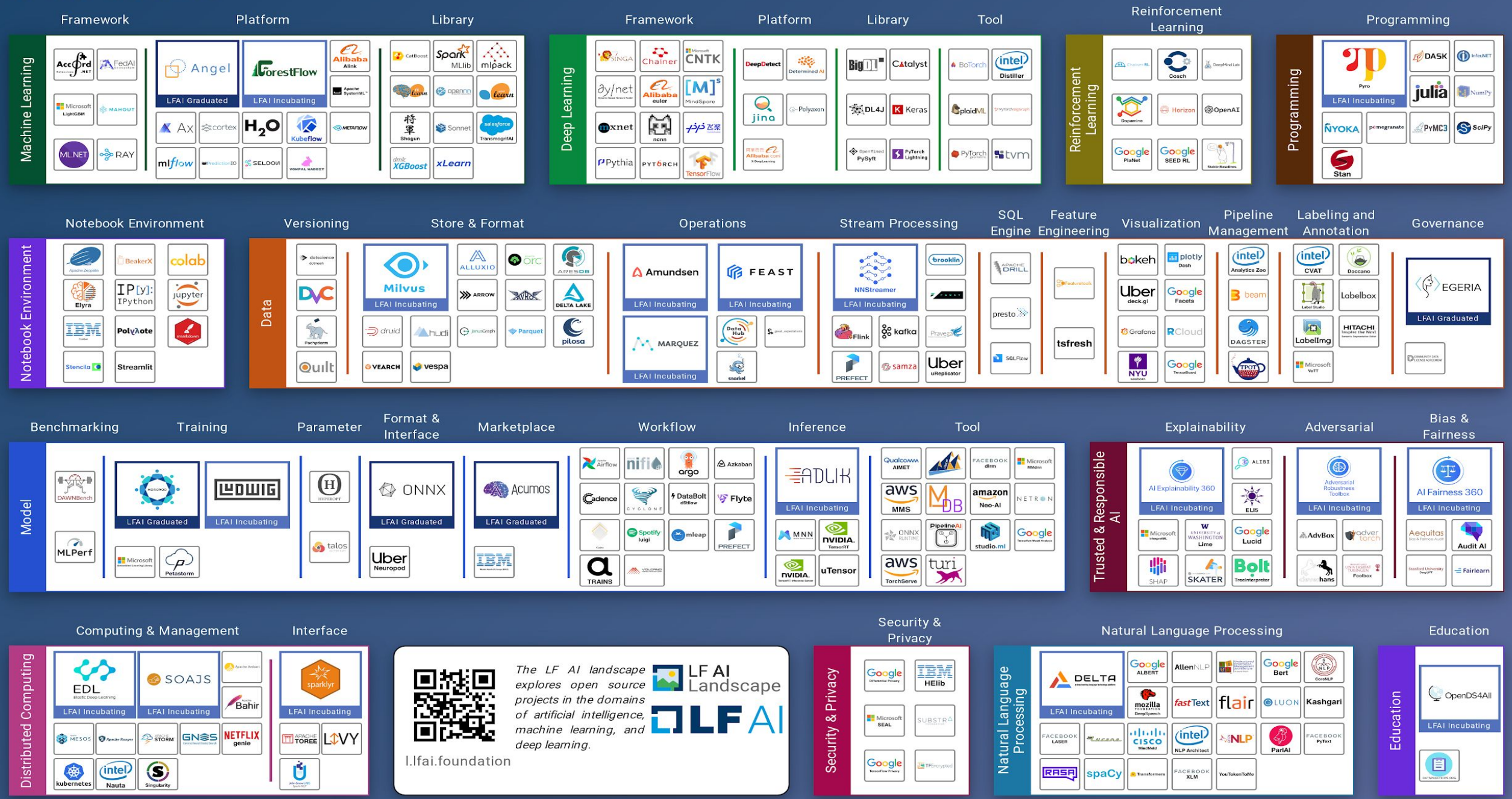


LF AI and ODPi Combining Efforts in AI and Data

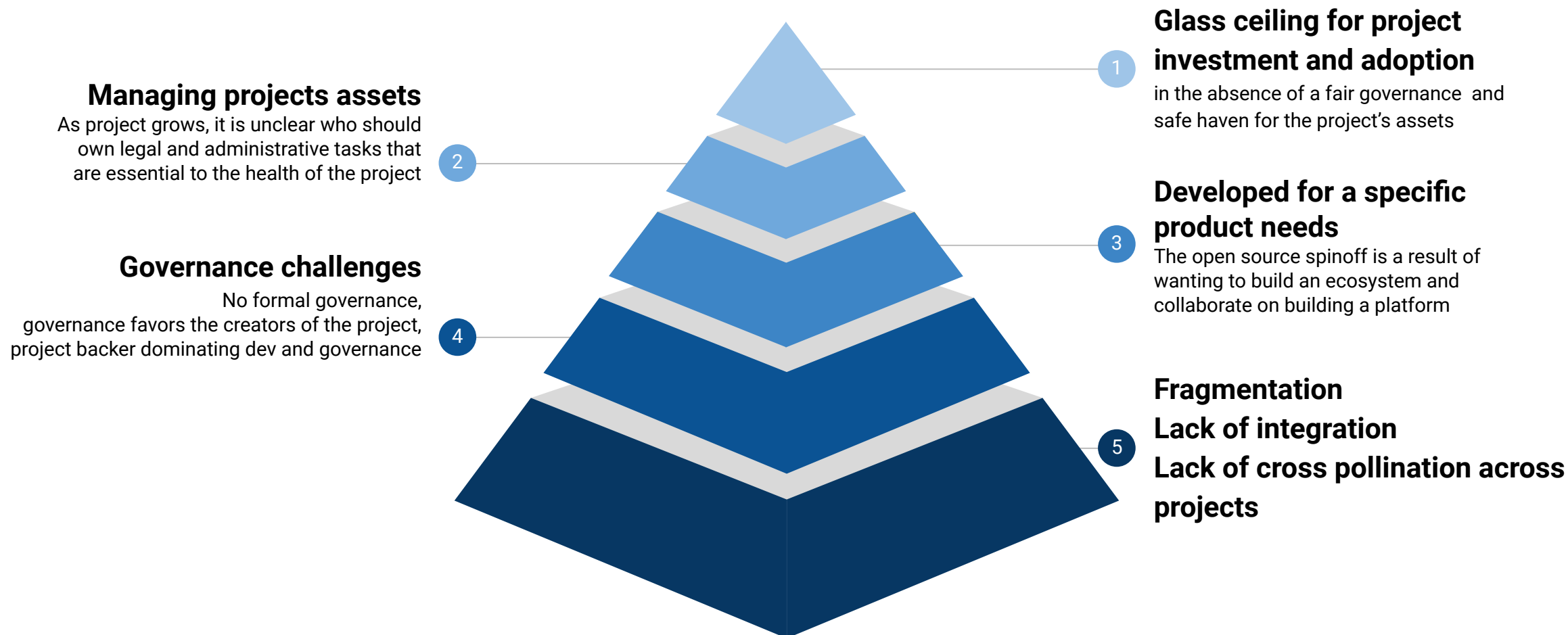
Combining efforts will:

- › Bring developers/data scientists and their projects under a single roof, served by a single Technical Advisory Council and several committees (Trusted AI, BI & AI, ML Workflow and InterOp), to work together towards building the open source AI & Data ecosystem and accelerating development and innovation.
- › Provide unified guidance for end users on tools, interoperability, integration, standards, and the future of AI, Data and Analytics as its use continues to grow in every industry.
- › Enables close collaboration, integration, and interoperability across projects - hosted within the same umbrella foundation - which is a proven, industry best practice recipe for building strong open ecosystems.
- › Allow greater efficiency for members across the various services we offer to hosted projects in the AI and Data domains.

Ecosystem Overview



Open Source AI & Data - Ecosystem Challenges



We all are aware of the advantages of open source ...

**Access to
source code**

**Ability to
customize**

**Rapid
development**

Peer review

**Improved
quality and
security**

**Reduce dev
time and TTM**

**Flexible
licensing model**

**Community
driven
innovation**

**Influence
roadmap via
participation**

**Transparent
technical
discussions**

But, Open Source offers *unique* benefits to AI & Data

Fairness

Methods to detect and mitigate bias in datasets and models, including bias against known protected populations

Robustness

Methods to detect alterations/tampering with datasets and models, including alterations from known adversarial attacks

Explainability

Methods to enhance understandability and interpretability by persona/roles in process of AI model outcomes/decision recommendations, incl ranking and debating results and decision options

Lineage

Methods to ensure provenance of datasets and AI models, including reproducibility of generated datasets and AI models

Open Data

Methods to clean, sort, tagg and track provenance and a governance structure for doing these things

LF AI & DATA

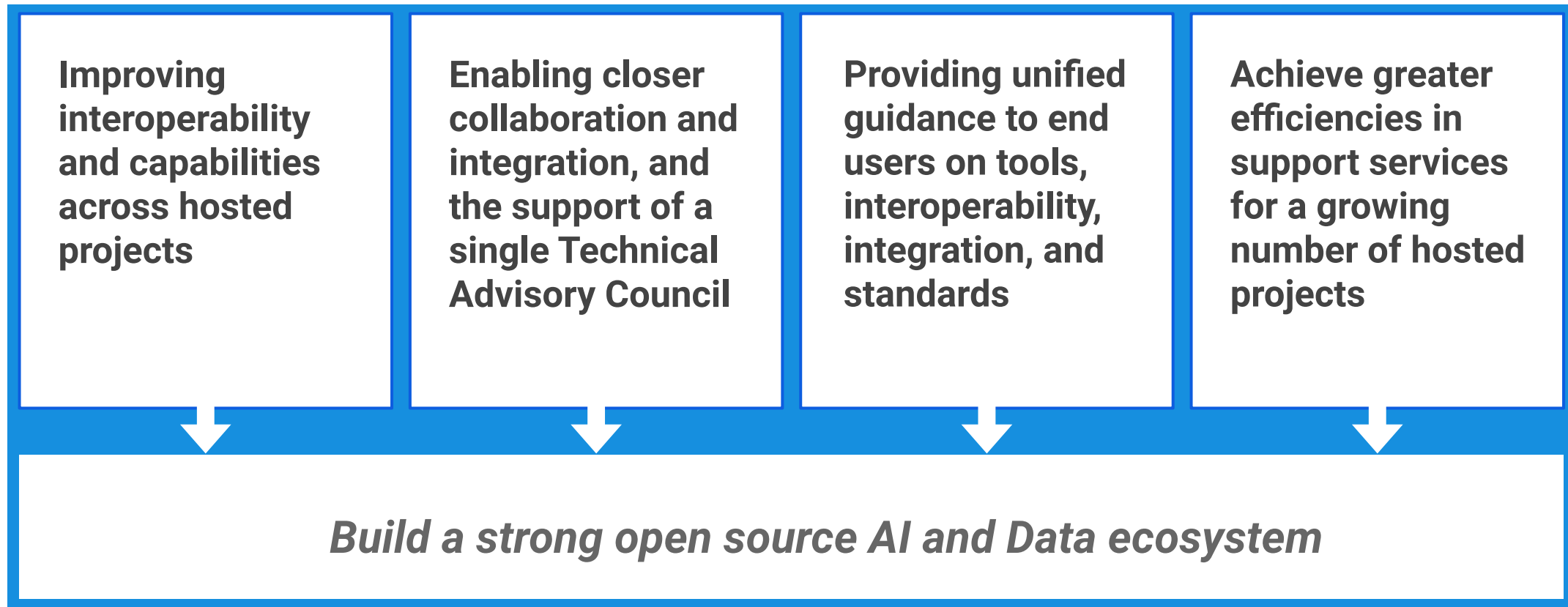
LF AI & Data is an LF umbrella foundation that brings together the LF AI and ODPi communities to further support collaboration and the acceleration open source innovation in AI, ML, DL and Data.

LF AI & DATA

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

Motivations for Harmonization



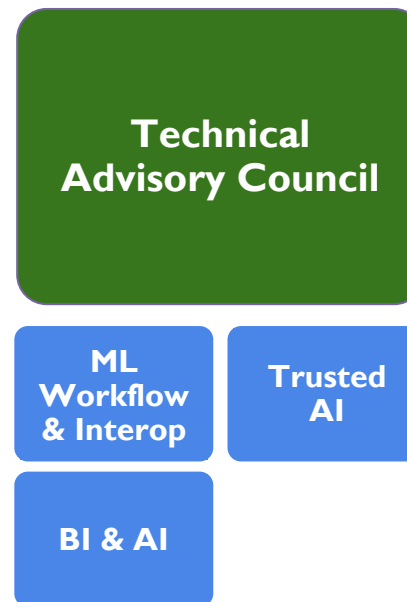
Structure and Governance

Foundation Governance

Funding effort to support hosted technical projects

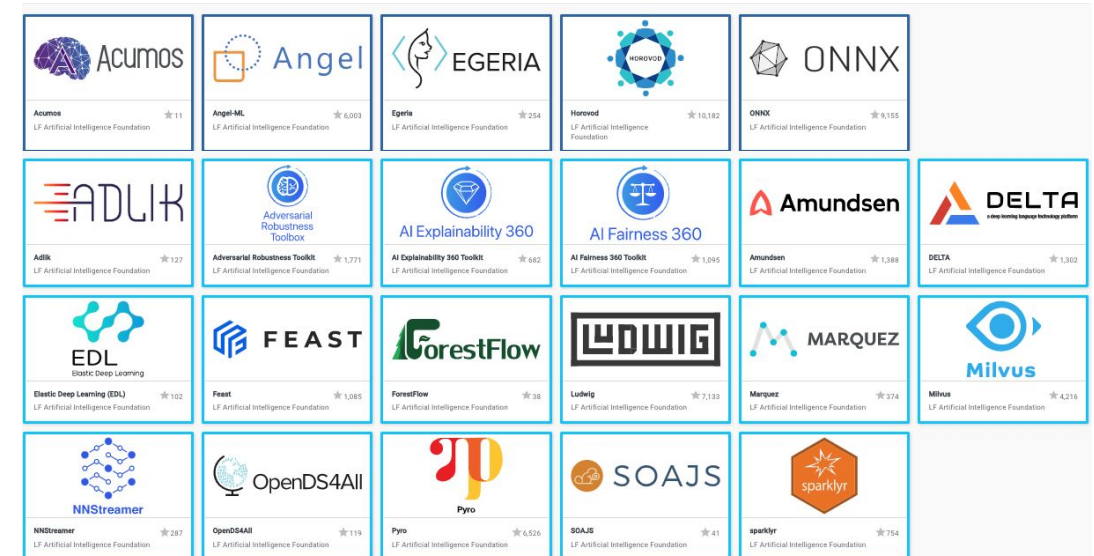


Technical Coordination, New Initiatives



Hosted Projects

Projects have their own independent technical governance



Major Efforts

Neutral Environment

Provide an neutral environment and an open governance that fosters collaboration and cross-pollination across projects and communities



Vendor neutral

Not-for profit

Open governance

IP clarity

Harmonization, Interoperability

Increase efforts to harmonize AI/ML/DL open source projects and reduce fragmentation; increase interoperability among technical projects



TAC
ML Workflow & Interop

Cross Foundation Collaboration

Integration efforts

Trusted and Responsible AI

Become the home of open source trusted and responsible AI, both at the level of technical projects and also at the level of principles and guidelines



Trusted AI Committee

3 Core Projects

edX Training

Data

Develop the only vendor neutral, open source standards to enable best practices for data governance, connectivity, business intelligence, and analytics.



BI & AI Committee

Large number of Data hosted projects

Model Marketplace

Host open source models and provide supporting tools; publish best practices; offer a marketplace for AI/ML/DL models



Open Source Model marketplace

Funding and Awareness

Provide a funding model for key projects

Increase awareness via a number of marketing and outreach activities

Training courses published in edX



Funding model

LF AI & Data Days, LF AI & Data Summit

Services: Marketing, Legal, Events and PM

Members

<https://landscape.lfai.foundation/format=members>

Premier	 AT&T	 Baidu 百度	 ERICSSON	 HUAWEI	 IBM	 NOKIA
	 sas	 Tech Mahindra	 Tencent 腾讯	 ZILLIZ Reinvent Data Science	 ZTE	
General	 ALPHABRAVO.IO	 CLOUDERA	 databricks	 DiDi	 Herron Tech	 index analytics
	 inwinSTACK	 orange	 precisely	 redhat	 R Studio	 XENONSTACK
Associate	 AI for people	 aivancity PARIS-CACHAN	 AMBIANIC	 ETL	 The National Strategy of Service Innovation Professionals	 MAIEI
	 NYU	 Open 启智 新一代人工智能开发平台	 PennState Great Valley	 PSIT Kanpur	 shopen 上海开放科学	


















Projects

<https://landscape.lfai.foundation/format=card-mode&project=hosted>

Graduated LFAI Projects (5)

 Acumos ★ 11 LF Artificial Intelligence Foundation	 Angel ★ 6,003 LF Artificial Intelligence Foundation	 EGERIA ★ 254 LF Artificial Intelligence Foundation	 Horovod ★ 10,182 LF Artificial Intelligence Foundation	 ONNX ★ 9,155 LF Artificial Intelligence Foundation
---	--	---	--	---

Incubating LFAI Projects (17)

 ADLIK ★ 127 LF Artificial Intelligence Foundation	 Adversarial Robustness Toolbox ★ 1,771 LF Artificial Intelligence Foundation	 AI Explainability 360 ★ 682 LF Artificial Intelligence Foundation	 AI Fairness 360 ★ 1,095 LF Artificial Intelligence Foundation	 Amundsen ★ 1,388 LF Artificial Intelligence Foundation	 DELTA ★ 1,302 LF Artificial Intelligence Foundation	 EDL ★ 102 Elastic Deep Learning LF Artificial Intelligence Foundation	 FEAST ★ 1,085 LF Artificial Intelligence Foundation	 ForestFlow ★ 38 LF Artificial Intelligence Foundation
 Ludwig ★ 7,133 LF Artificial Intelligence Foundation	 MARQUEZ ★ 374 LF Artificial Intelligence Foundation	 Milvus ★ 4,216 LF Artificial Intelligence Foundation	 NNStreamer ★ 287 LF Artificial Intelligence Foundation	 OpenDS4All ★ 119 LF Artificial Intelligence Foundation	 Pyro ★ 6,526 LF Artificial Intelligence Foundation	 SOAJS ★ 41 LF Artificial Intelligence Foundation	 sparklyr ★ 754 LF Artificial Intelligence Foundation	

What does it mean to be a Foundation project?

- › **A neutral home for the project** increases the willingness of developers from organizations to collaborate, contribute, and become committers
- › **Endorsement by members of the LF AI & Data Technical Advisory Committee (TAC)** is an independent signal of the quality of your project
- › **Open and neutral governance model**
- › **Program and project management services**
- › **Access to full-time staff** who are eager to assist your project
- › **Access to marketing and communications services** to support community and ecosystem engagement
- › **Access to a full range of events services** to help the project build and grow a community around it
- › **Legal support** ranging from hosting the trademark in a neutral organization to CLA and DCO system integration with GitHub, regular compliance scans, etc.
- › **Presence across multiple geographies** including China and Japan

Companies hosting projects in LF AI & Data

<https://landscape.lfai.foundation/format=hosting>

amazon

AT&T

Baidu 百度

DiDi

DREAMWORKS

FACEBOOK

Herron Tech

IBM

lyft

Microsoft

gojek

R Studio

SAMSUNG

Tech Mahindra

Tencent 腾讯

Uber

wework

ZILLIZ
Reinvent Data Science

ZTE

Why these top companies host with LF AI & Data?

The Linux Foundation is the leading commons for community assets

- The Linux Foundation (LF) is a recognized and respected brand and can activate communications across tech, business, and social media channels.
- The LF is the largest open source foundation, providing management of community assets for over 400 projects:
 - Over 3,400 source code repositories.
 - Over 700 trademark registrations and applications, plus hundreds of unregistered marks.
 - Over 700 project domains and DNS records.
 - Over 12,000 Contributor License Agreements signed for project communities that decide to require one.
- The LF offers neutral sponsorship for managing fundraising.
 - Over 1,600 unique member company relationships.
- Our communities raised over \$100 million in 2019 via 6,500+ invoices.
 - Our communities spent funds in 180,000+ separate transactions in 2019.
 - The LF reports its audited financials to its board.

Why incubate with us?

Building large, sustainable ecosystems requires collective resources



Events - We gathered over 45,000 attendees from over 12,000 organizations across 113 countries in 2019.



Legal - We manage IP for the world's most important tech and have one of the world's top open source legal teams in house.



Training - We trained millions of students (free and paid online), online skills certification, and on-site e-learning.



Certification - We designed and executed software and hardware testing and certification programs.



Developer Marketing - We have the largest share of voice of any open source foundation and a proven method to build large scale developer programs.



Developer Operations - We host the infrastructure that develops the world's largest software communities and provide release management, IT ops, and support.



Application Security- our projects are regularly audited and pen tested. We offer bug bounties, dependency analysis, and code scanning.

Strong LF AI & Data Marketing/PR Effort

37 announcements between LF AI and ODPI in 2019

27 announcements between LF AI and ODPI in 1H 2020

- › Events announcements and promotion
- › Project release
- › Project incubation
- › New members joining
- › Project cross collaboration
- › Invited posts

LF AI provides marketing services to support community and ecosystem engagement.

- › Content ranging from the LF AI blog, white papers, posters, presentations, and more
- › Oversee blog editorial calendars and writing, editing, and/or identifying and funding freelancers to develop content
- › Help develop the project website and other required creative work

Events and Community

- › **LF AI & Data Summit** - co-located event with major LF conferences + Booth
- › **LF AI & Data Track at Open Source Summit** (NA, EU, China) - Supporting 18-20 talks over 3 days + Booth
- › **LF AI & Data Days** - a whole day (in person and virtual) hosted in collaboration with an LF AI & Data member with audience ranging from 50-250. Multiple/year.
- › Talks at other leading industry events and regional events (All Things Open, Data Con LA, Big Data Conference Europe)
- › Project specific events
- › Sponsorship and booth at KubeCon
- › Strong presence & event participation in China



Incubating Projects

- › TAC committee is responsible to voting in new incubation projects.
- › The TAC meets bi-weekly and generally booked 4-6 weeks in advance.
- › [Proposing projects for incubation](#)

Joining LF AI & Data is Easy!



General Inquiries:

info@lfaidata.foundation

 **LF** AI & DATA