

# LF AI & Data

March 16, 2021

Ibrahim Haddad, Ph.D.  
Jacqueline Cardoso  
John Mertic

Executive Director, LF AI & Data  
Head of Program Management  
Director of Technical Projects

[Ibrahim@LinuxFoundation.org](mailto:Ibrahim@LinuxFoundation.org)  
[jzcardoso@linuxfoundation.org](mailto:jzcardoso@linuxfoundation.org)  
[jmertic@linuxfoundation.org](mailto:jmertic@linuxfoundation.org)

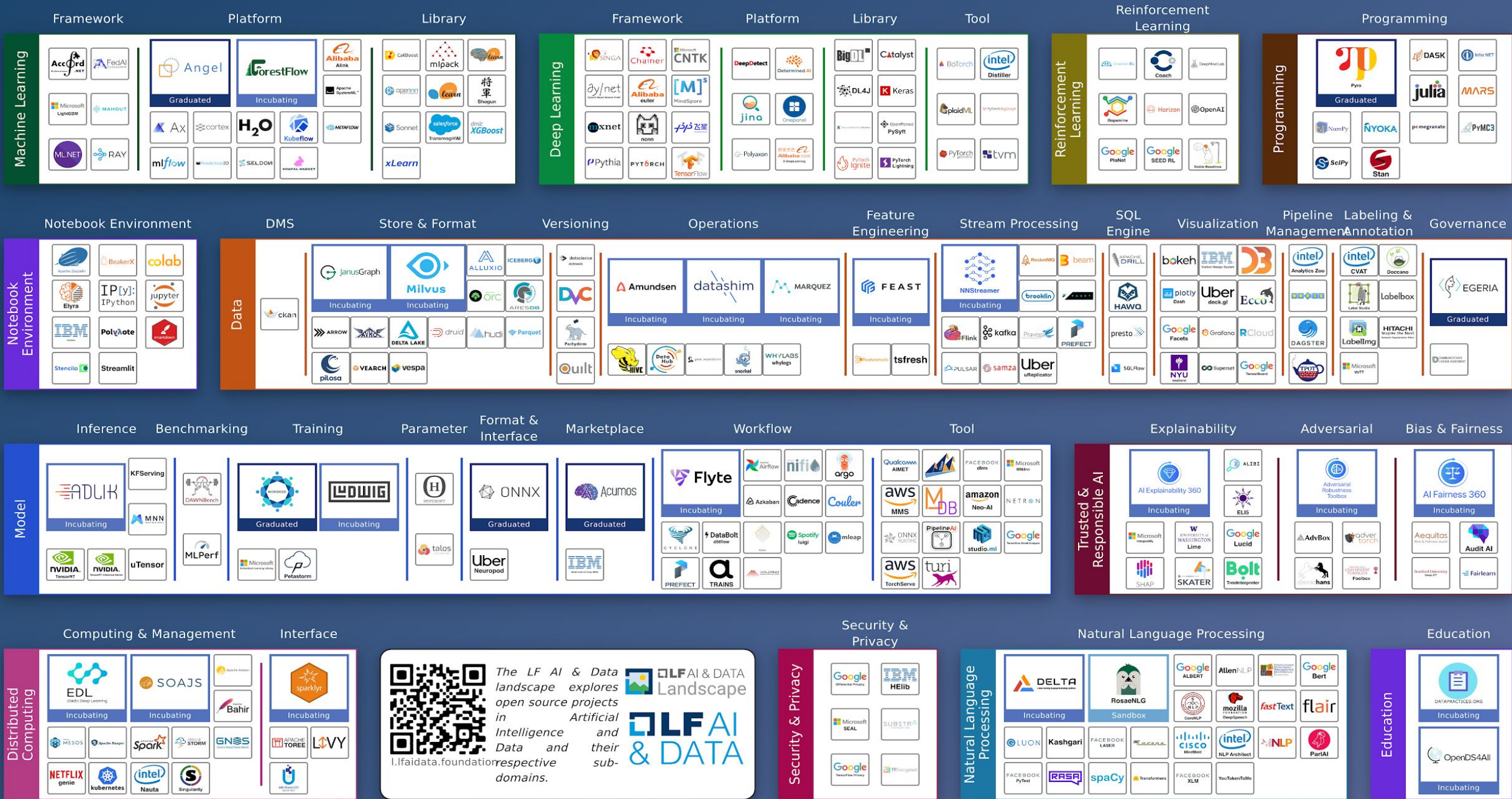
 LF AI & DATA

# LF AI & Data

Collaborating to support an open and 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

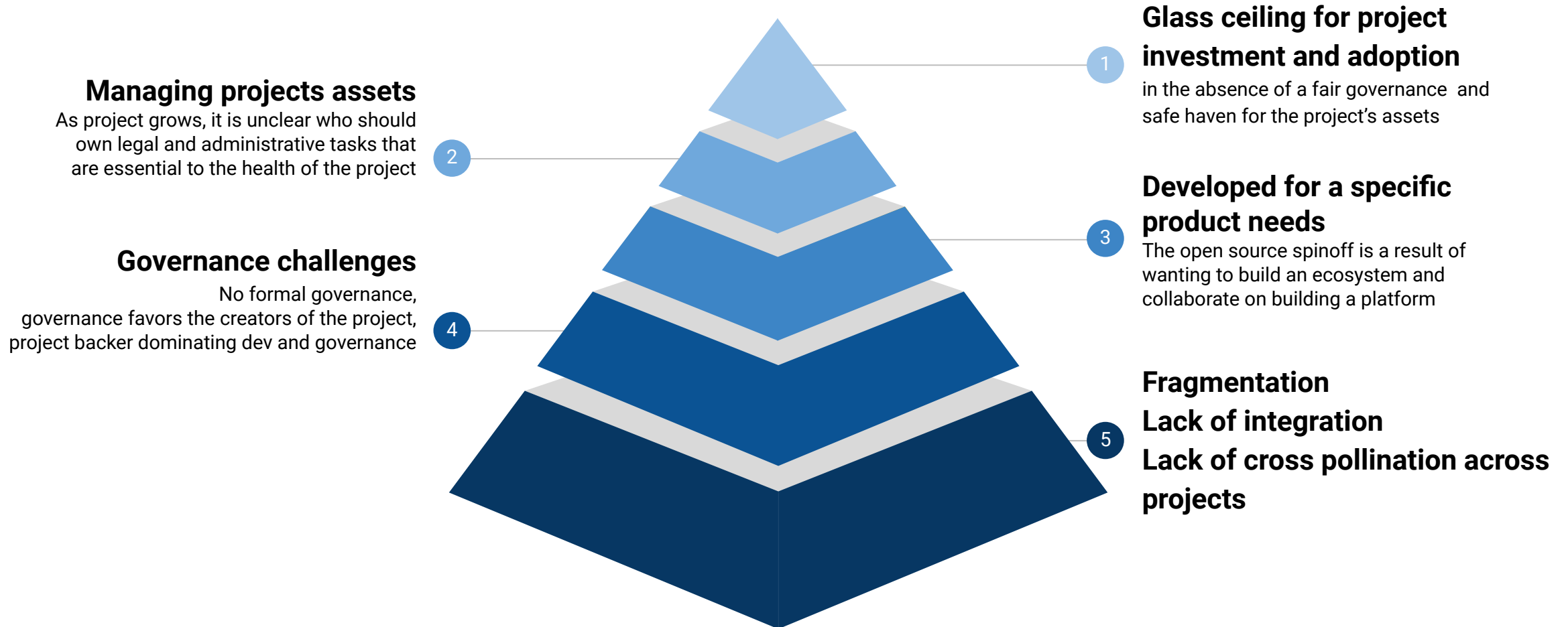
<https://lfaidata.foundation>

# Ecosystem Overview





# Open Source AI & Data - Ecosystem Challenges



# Enter LF AI & Data

 **LF** AI & DATA

# What is LF AI & Data?

- › Non-profit
- › Open source collaborative effort focused on AI & Data
- › Hosted at the Linux Foundation
- › Focused on supporting open source AI & data projects and accelerating innovation and their trajectory to adoption

# 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

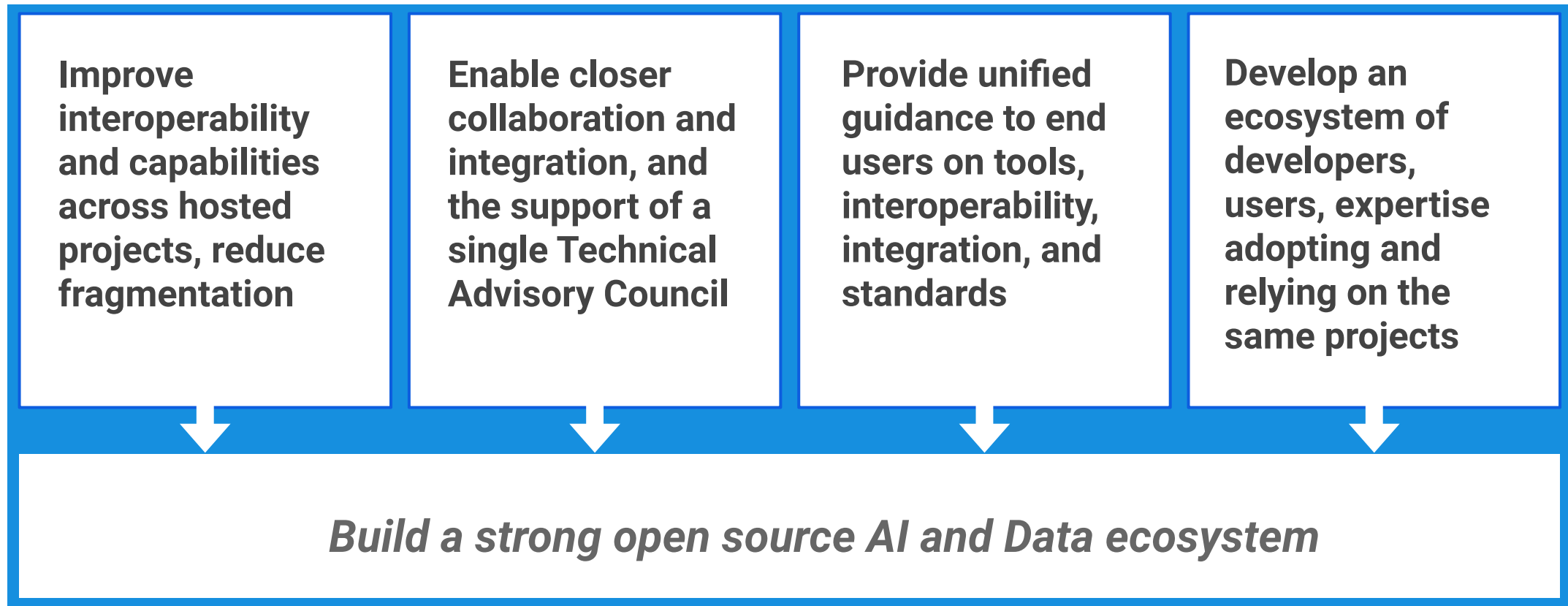
Methods to ensure provenance of datasets and AI models, including reproducibility of generated datasets and AI models

## Open Data

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



# Goals of LF AI & Data



# Why is LF AI & Data important?

1

## OPEN & NEUTRAL

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

2

## HARMONIZATION & INTEROPERABILITY

Increase efforts to harmonize AI & Data open source projects and reduce fragmentation; increase interoperability among technical projects

3

## TRUSTED & 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

4

## ECOSYSTEM DEVELOPMENT

Fund key projects; Increase awareness via a number of marketing and outreach activities; Training courses published in edX

5

## DATA

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

# LF AI & Data Growth

- › Launched in 03/2018 with 9 members and 1 hosted project
- › +3 projects in 2018
- › +5 projects in 2019
- › +15 projects, +16 members in 2020
- › Currently supported by 46 member companies
- › >2000 subscribers to our main mailing list

# Very active and growing developer community

Jan 1- Dec 31, 2020

8.92K

Contributors

217K

Contributions

100.03K

Commits

358

Repositories

6.93K

Emails

34.65K

PRs/Changesets

2.08K

Slack messages

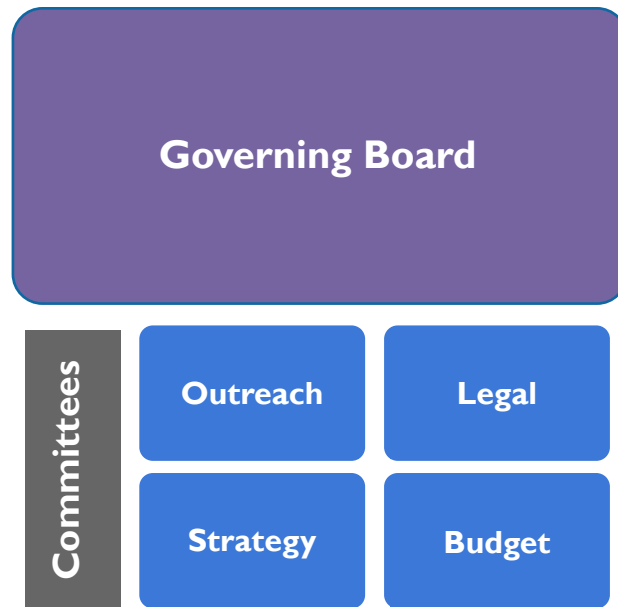
22.54K

Total issues

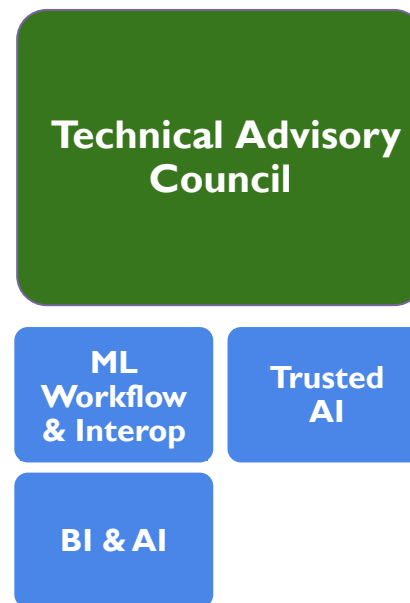
# Structure and Governance

## Foundation Governance

Funding effort to support hosted technical projects

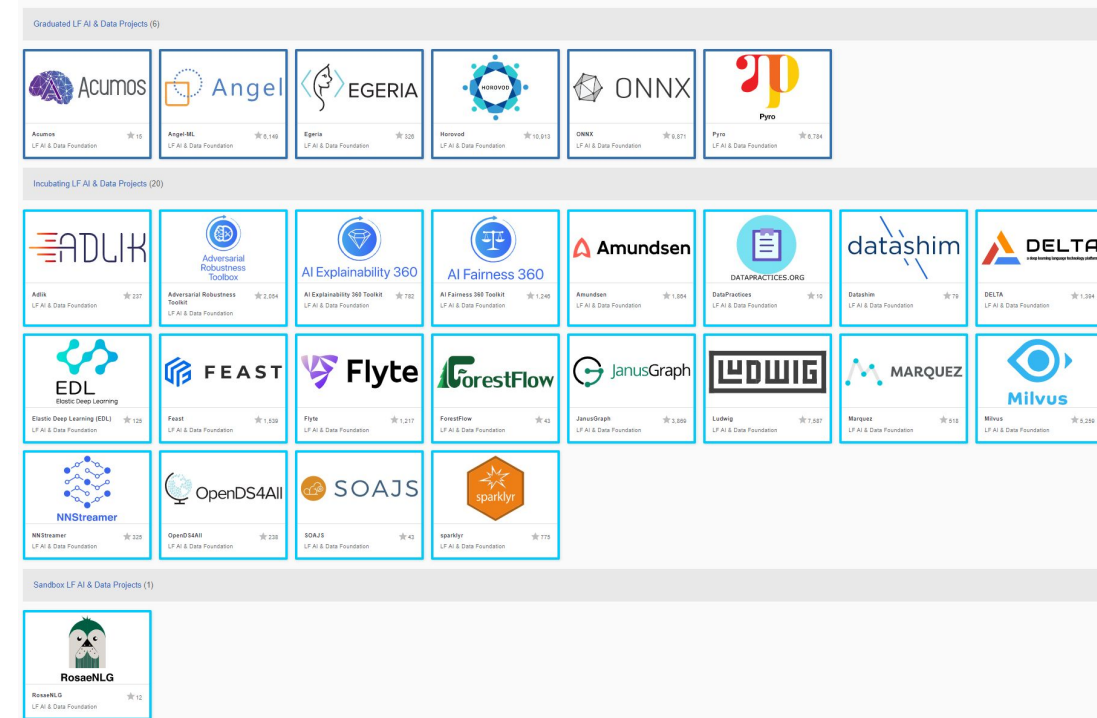


## Technical Coordination, New Initiatives































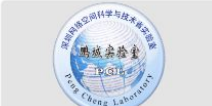
















## Hosted Projects

Projects have their own independent technical governance



# Members

<https://landscape.lfai.foundation/card-mode?project=company>







 <p><b>AI for people</b> AI For People</p>	 <p><b>aivancity</b> PARIS-CACHAN</p>	 <p><b>AMBIANIC</b> Ambianic</p>	 <p><b>AT&amp;T</b> AT&amp;T MCap: \$213.88B</p>	 <p><b>Baidu</b> Baidu MCap: \$78.32B</p>	 <p><b>BROADCOM</b> Broadcom</p>	 <p><b>CLOUDERA</b> Cloudera MCap: \$4B</p>	 <p><b>D2IQ</b> D2iq Funding: \$247.25M</p>	 <p><b>databricks</b> Databricks Funding: \$1.9B</p>
 <p><b>Didi</b> Didi Funding: \$21.25B</p>	 <p><b>ERICSSON</b> Ericsson MCap: \$40.21B</p>	 <p><b>GALGOTIAS UNIVERSITY</b> Galgotias University</p>	 <p><b>Herron Tech</b> Herron Tech</p>	 <p><b>High School Technical Services</b> High School Technical Services</p>	 <p><b>HUAWEI</b> Huawei Technologies</p>	 <p><b>IBM</b> IBM MCap: \$111.53B</p>	 <p><b>index analytics</b> Index Analytics</p>	 <p><b>ING</b> ING Group</p>
 <p><b>Institute of Ethical AI and Machine Learning</b> The Institute for Ethical AI &amp; Machine Learning</p>	 <p><b>The International Society of Service Innovation Professionals</b> International Society of Service Innovation Professionals</p>	 <p><b>inwinSTACK</b> inwinSTACK</p>	 <p><b>RICE KEN KENNEDY INSTITUTE</b> Ken Kennedy Institute - Rice University Ken Kennedy Institute</p>	 <p><b>MAIEI</b> Montreal AI Ethics Institute</p>	 <p><b>NYU</b> New York University Funding: \$5M</p>	 <p><b>NOKIA</b> Nokia MCap: \$21.88B</p>	 <p><b>OpenI 启智</b> 新一代人工智能开源开放平台</p>	 <p><b>OpenUK</b> OpenUK</p>
 <p><b>orange</b> Orange MCap: \$31.33B</p>	 <p><b>Peng Cheng Laboratory</b> Peng Cheng Laboratory</p>	 <p><b>PennState Great Valley</b> Pennsylvania State University Funding: \$13.4M</p>	 <p><b>PSIT Kanpur</b> Pranveer Singh Institute Of Technology</p>	 <p><b>precisely</b> Precisely Holdings, LLC Funding: \$2.18M</p>	 <p><b>redhat</b> Red Hat MCap: \$111.53B</p>	 <p><b>R Studio</b> RStudio</p>	 <p><b>sas</b> SAS</p>	 <p><b>shopen</b> 上海开源信息技术协会 Shanghai Open Source Information Technology Association</p>
 <p><b>Tech Mahindra</b> Tech Mahindra</p>	 <p><b>Tencent 腾讯</b> Tencent Holdings MCap: \$794.03B</p>	 <p><b>UNIVERSITY of WASHINGTON - TACOMA</b> University of Washington</p>	 <p><b>UL T</b> Université Libre de Tunis</p>	 <p><b>vmware</b> VMware MCap: \$60.22B</p>	 <p><b>XENONSTACK</b> Xenonstack</p>	 <p><b>XPRIIZE</b> XPRIIZE Foundation Funding: \$1.75M</p>	 <p><b>ZILLIZ</b> ZILLIZ Funding: \$53M</p>	 <p><b>ZTE</b> ZTE MCap: \$20.21B</p>























# Projects

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


## Graduated LF AI & Data Projects (6)

 <b>Acumos</b> LF AI & Data Foundation ★ 15	 <b>Angel</b> Angel-ML LF AI & Data Foundation ★ 8,149	 <b>EGERIA</b> Egeria LF AI & Data Foundation ★ 328	 <b>HOROVOD</b> Horovod LF AI & Data Foundation ★ 10,913	 <b>ONNX</b> LF AI & Data Foundation ★ 9,871	 <b>Pyro</b> LF AI & Data Foundation ★ 8,784
--	--	---	--	--	---

## Incubating LF AI & Data Projects (20)

 <b>ADLIK</b> Adlik LF AI & Data Foundation ★ 237	 <b>Adversarial Robustness Toolbox</b> LF AI & Data Foundation ★ 2,054	 <b>AI Explainability 360</b> AI Explainability 360 Toolkit LF AI & Data Foundation ★ 782	 <b>AI Fairness 360</b> AI Fairness 360 Toolkit LF AI & Data Foundation ★ 1,248	 <b>Amundsen</b> LF AI & Data Foundation ★ 1,864	 <b>DATA PRACTICES.ORG</b> DataPractices LF AI & Data Foundation ★ 10	 <b>datashim</b> Datashim LF AI & Data Foundation ★ 78	 <b>DELTA</b> a deep learning language technology platform LF AI & Data Foundation ★ 1,384
 <b>EDL</b> Elastic Deep Learning LF AI & Data Foundation ★ 126	 <b>FEAST</b> Feast LF AI & Data Foundation ★ 1,539	 <b>Flyte</b> LF AI & Data Foundation ★ 1,217	 <b>ForestFlow</b> LF AI & Data Foundation ★ 43	 <b>JanusGraph</b> LF AI & Data Foundation ★ 3,809	 <b>LUDWIG</b> Ludwig LF AI & Data Foundation ★ 7,587	 <b>MARQUEZ</b> LF AI & Data Foundation ★ 518	 <b>Milvus</b> LF AI & Data Foundation ★ 5,259
 <b>NNStreamer</b> LF AI & Data Foundation ★ 328	 <b>OpenDS4All</b> LF AI & Data Foundation ★ 238	 <b>SOAJS</b> LF AI & Data Foundation ★ 43	 <b>sparklyr</b> LF AI & Data Foundation ★ 775				

## Sandbox LF AI & Data Projects (1)

 <b>RosaeNLG</b> RosaeNLG LF AI & Data Foundation ★ 12
--

# 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](#)

# Companies hosting projects in LF AI & Data

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

amazon

AT&T

Baidu 百度

data.world

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 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.

# 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

# Strong Marketing/PR/Community Effort

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

**37 announcements in 2019**

**61 announcements in 2020**

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

**LF AI provides event and community ecosystem building services.**





# Joining LF AI & Data is Easy!



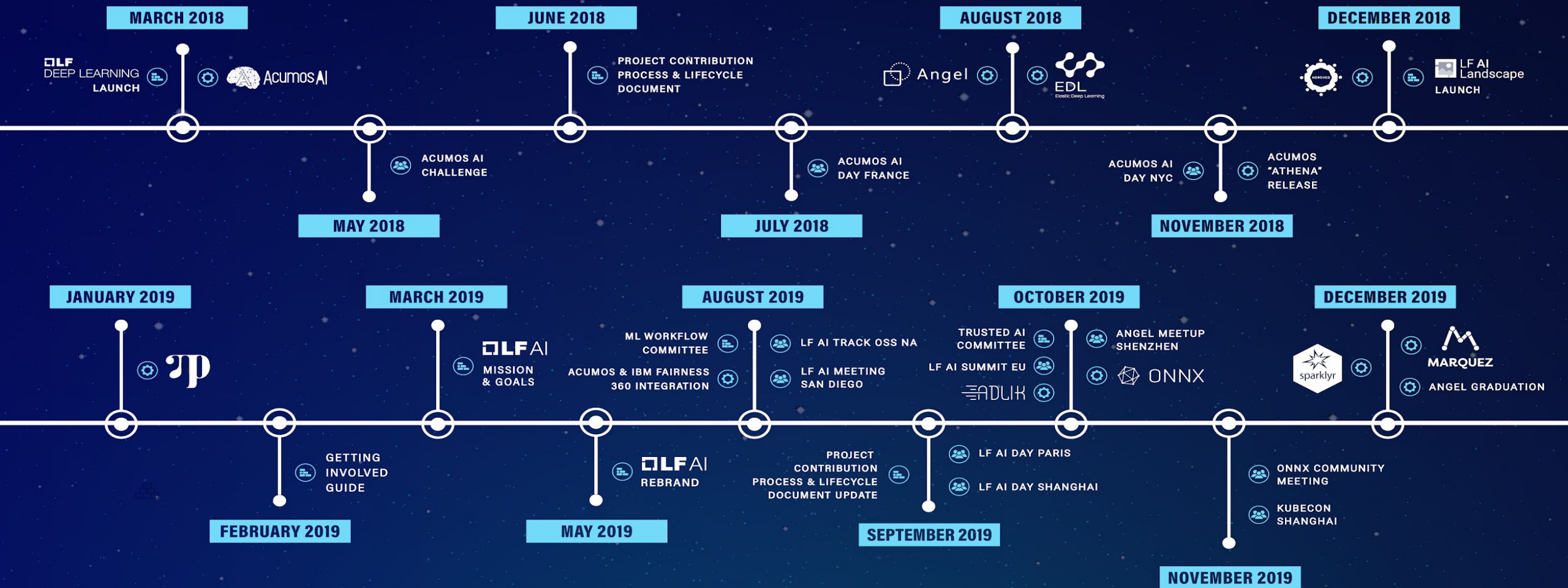
General Inquiries:

[info@lfaidata.foundation](mailto:info@lfaidata.foundation)

 **LF** AI & DATA

# Backup

# LF AI TIMELINE 2018-2019



# LF AI & DATA TIMELINE 2020-2021

