

# LF AI & Data

Ibrahim Haddad, PhD

# Introduction to the Linux Foundation

The Linux Foundation's goal is to create the greatest shared technology investment in history by enabling open collaboration across companies, developers and users.

We are the nonprofit organization of choice to build ecosystems that accelerate open source technology development and commercial adoption on a global scale.

# We are behind some of the most critical projects in the world

## Security



## Networking



## Cloud



## Automotive



## Blockchain



## Edge/IoT



## Web



## AI



## Film



## CI/CD



## Energy



## Hardware



## Standards



# Building large, sustainable ecosystems requires collective resources

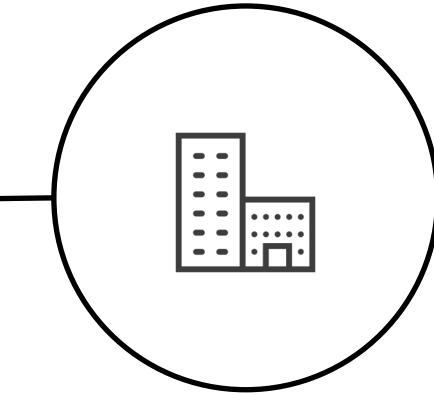
**2,000+**

Members From  
41 Countries



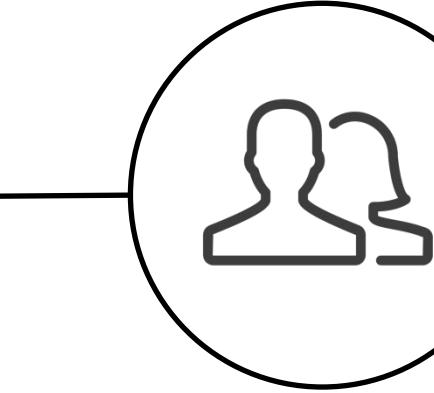
**100%**

of Fortune 100  
Tech & Telecom



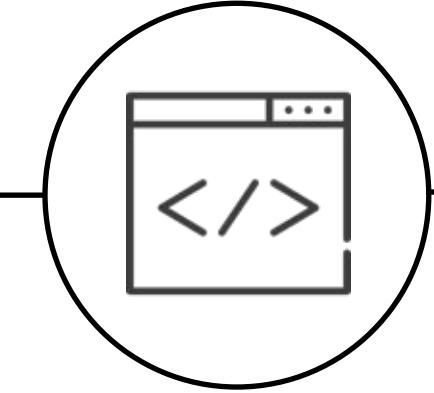
**242,960+**

Developers  
Contributing Code



**730+**

Critical Open Source  
Projects



**\$54.1B**

Shared  
Value



## The Linux Foundation is a critical part of modern technology

**31.2M**  
Lines of Code  
Added Weekly

**20M**  
Lines of Code  
Removed Weekly

**236,040**  
Contributing  
Developers

**19,442**  
Contributing  
Companies

**12,290**  
Repositories

**10.4M**  
Commits

**1.1M**  
Pull Requests

**735,760**  
Builds Monitored

**815,110**  
Logged Issues

**4.6B**  
Container  
Downloads

**5.1M**  
Chat  
Messages  
Sent

**4.4M**  
Email  
Messages  
Sent

**10,334**  
Scanned  
Repositories

**236,499**  
Vulnerabilities  
Detected

**21,597**  
Recommended  
Fixes

**13,802**  
Vulnerabilities  
Fixed

**31,496**  
CLA  
Contributors

**26,998**  
Community  
Meetings

# For an Enterprise, Open Source is about Business Strategy

# Open Source Strategic Impact



- Accelerate the development of open solutions
- Provide an implementation to an open standard



- Commoditize a market
- Reduce prices of non-strategic software assets
- Share development costs



- Drive demand by building an ecosystem for products & services



- Partner with others
- Engage customers
- Strengthen relationships with common goals

# Increase market opportunity

- If a certain part of our stack became a commodity and every player adopted it, where can we get first mover and industry leader advantages to expand our market opportunity?
- Are there open, industry adopted southbound or northbound APIs to allow for interoperability and/or data egress/ingress?
- Where can we increase customer value with offerings of additional data services, analytics, management, and security? What industry open source base stack can facilitate this?

# Take market share through de facto standards

- Where in the stack can we disrupt a market leading proprietary implementation with a standardized open source implementation that creates a market opportunity for us?
- Which peers, partners and suppliers, do I need to involve to ensure an open source effort becomes adopted industry-wide?
- Where can a customer or competitor use open source to create a de facto standard that forces us to change our product development roadmap and R&D/IT investments?

# Decrease costs

- Where is there opportunity to externalize commodity R&D through shared open source development?
- Where is the client shouldering duplicative costs in non differentiating vendor solutions and how we can we use open source to move our relationship up the stack while removing duplicative R&D?

**Strategy is Built Upon Category Creation  
through EcoSystem Formation and  
Industry-Wide Standardization on  
Open Source Software**

Open Source enables companies to reshape their industry value chain in ways that are both neutral (to create **Industry Platforms**) and that drive their own competitive goals (through active **Stakeholdership**)

Old World “Standards” =  
Specifications with  
Divergent  
Implementations

**New World Standards =**  
**Industry-wide adoption**  
**of shared technology**  
**development and critical**  
**deployment alongside**  
**specifications to become a**  
**de facto platform**

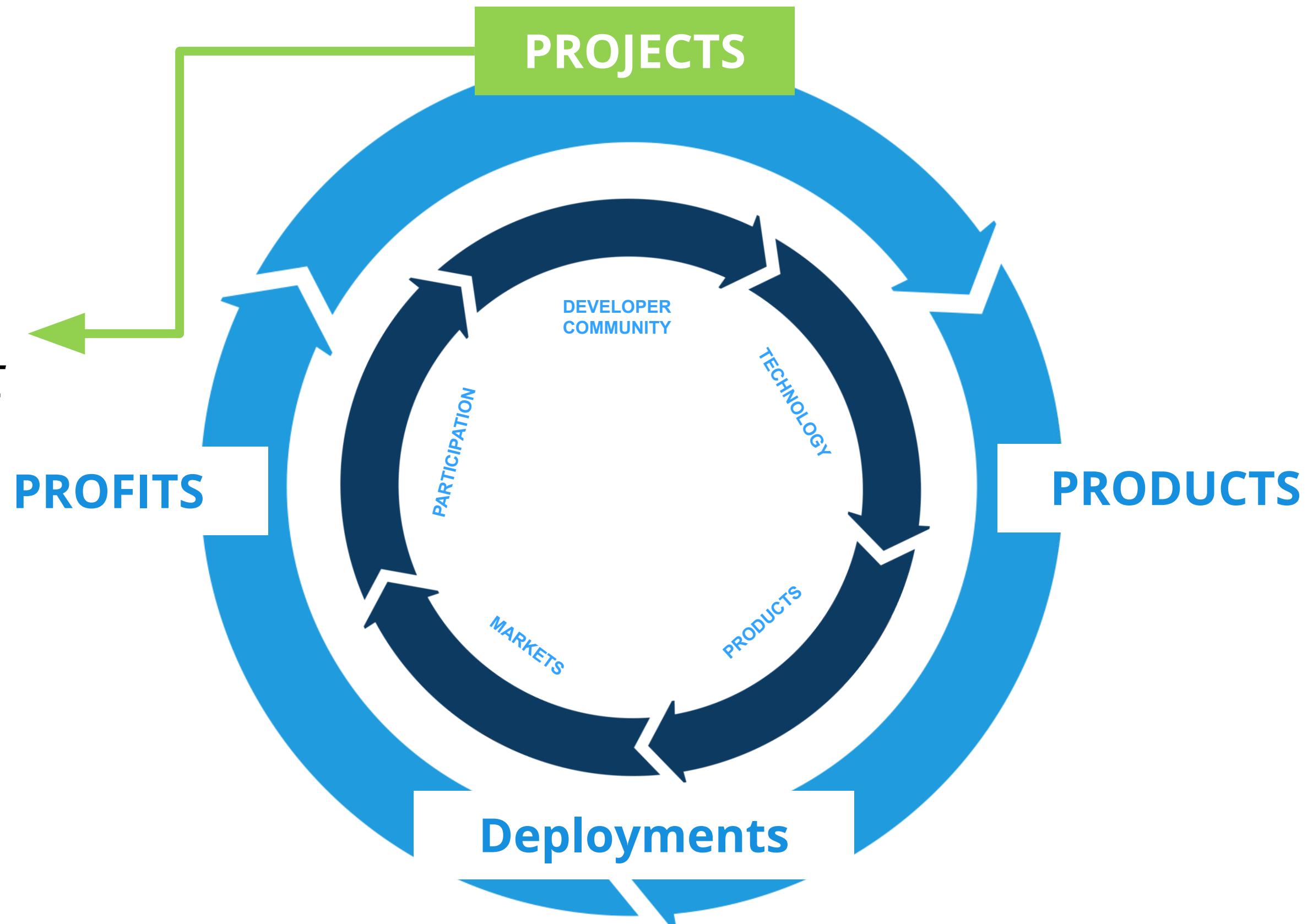
**Stakeholders** invest in active, multi-year  
**Membership** to see hundreds of millions,  
sometimes billions, in return by collecting  
value from **Industry Platforms**

“You are either at the table or on the table.”  
—Anonymous Executive



# Open Source Journey – First Came Projects

*Successful Open Source Development depends on the complete life cycle of projects, products that market will adopt and deploy*



# Our Role Has Been Recognized Alongside Tech Titans

- › “This category looks at those companies, associations and projects that have inspired development and IT shops to build upon the work they have created, and recognizes them for their leadership as we begin to create a digital world we only could have dreamed about a generation ago.
- › SD Times Influencers: Apple, Facebook, Google, IBM, Intel, Microsoft, GitHub, Netflix, Red Hat, Slack, *The Linux Foundation*

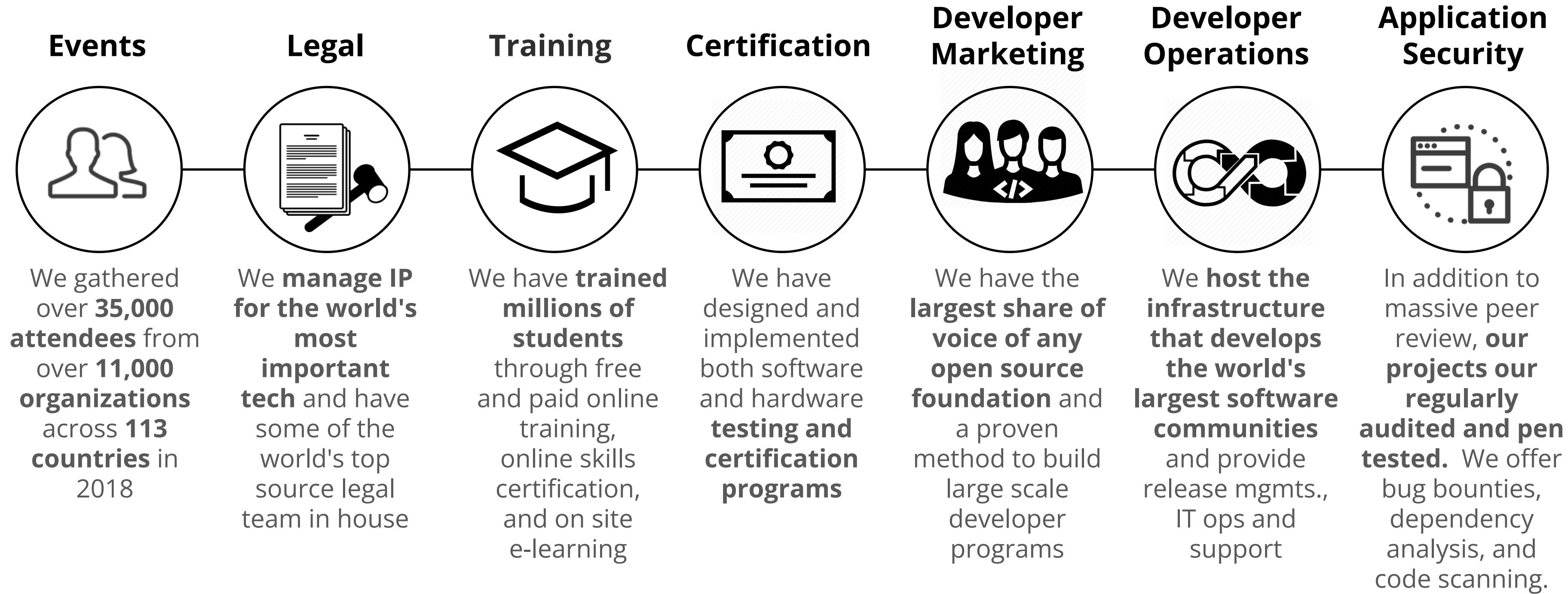
In 2019, The Linux Foundation is the Only Non-Profit Recognized For Three Consecutive Years For Its Critical Role in the Industry



# The Linux Foundation Platform assists projects in 5 key areas

<b>Governance and Membership</b>	<ul style="list-style-type: none"><li>• Project technical governance, policies, etc.</li><li>• Ongoing business development and membership recruitment</li><li>• Membership management</li></ul>
<b>Development Methodology and Processes</b>	<ul style="list-style-type: none"><li>• Technical decision making</li><li>• Project life cycle</li><li>• Release processes</li></ul>
<b>Infrastructure</b>	<ul style="list-style-type: none"><li>• CI/CD infrastructure using open source best practices</li><li>• Release engineering, DevOps</li><li>• Security and reliability</li></ul>
<b>Ecosystem Development</b>	<ul style="list-style-type: none"><li>• Evangelism, marketing and outreach initiatives</li><li>• Events bringing developers, users and solution providers together</li><li>• Training for developers and administrators, establishing professional certifications</li></ul>
<b>IP and Assets Management</b>	<ul style="list-style-type: none"><li>• Code provenance</li><li>• Trademark management</li><li>• IP Policy, license scanning, IP defenseas</li></ul>

# Building large, sustainable ecosystems require collective resources



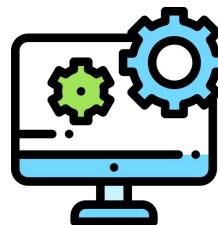
# We enable organizations to become open source leaders

The path to leadership in enterprise open source



# We help organizations navigate the open source ecosystem

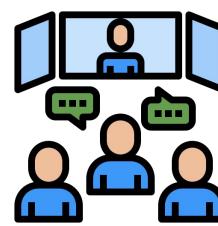
Our services help companies manage open source strategically.



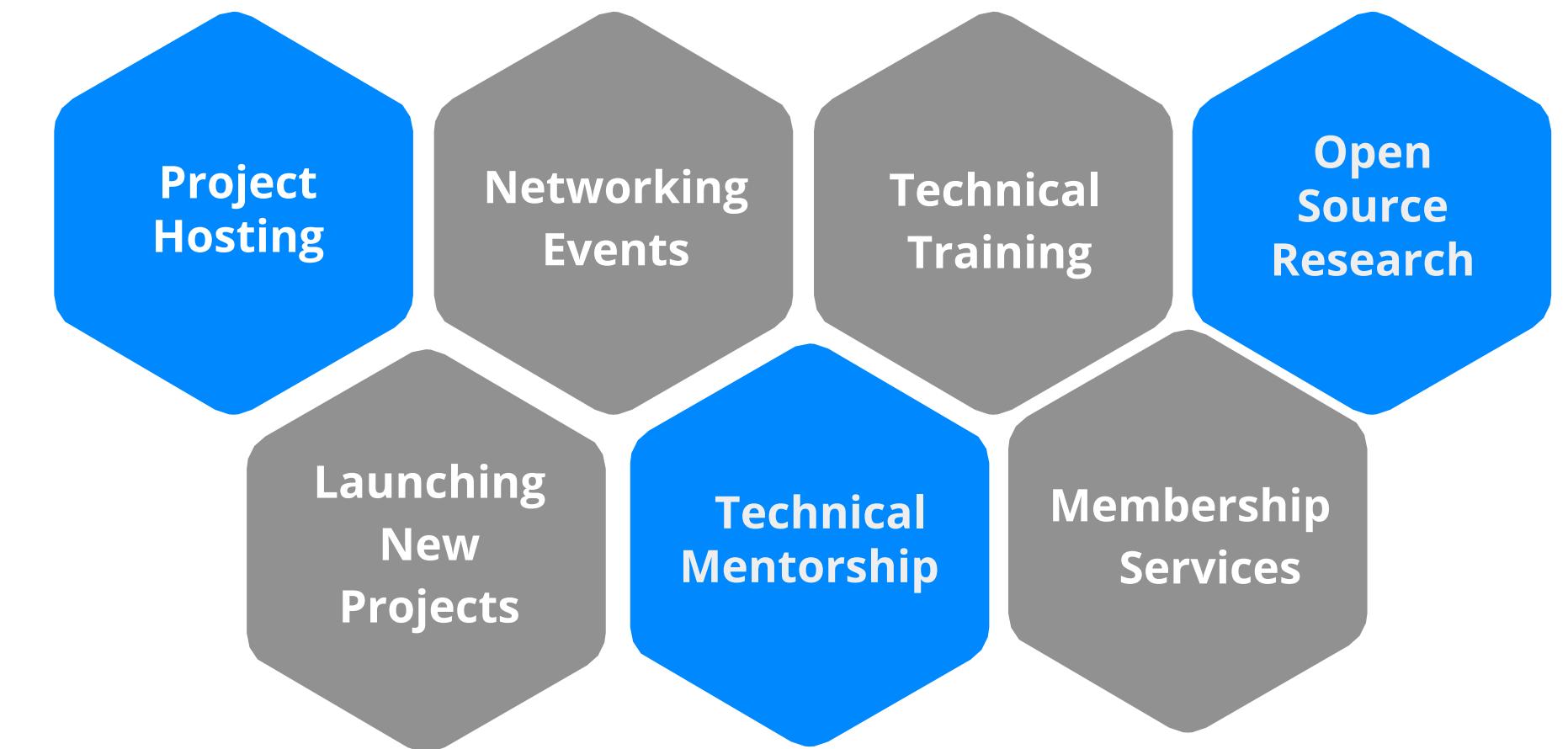
We host the world's most important projects.



Our events are the place where top developers and users of open source meet.



Our training programs meet the demand for OSS skills from the source.





# LF AI & Data Foundation

Ibrahim Haddad, PhD  
Executive Director

Open source libraries, frameworks, platforms, tools, are a two-way street: they make AI accessible to everyone, and companies benefit from a community of other contributors helping accelerate open AI applied research.

The barrier to entry in AI is lower than ever before, thanks to open source software.

# Open source benefits for AI & Data

## FAIRNESS

Methods to detect and mitigate bias in datasets and models, e.g., bias against known protected populations

## ROBUSTNESS

Methods to detect alterations and tampering with datasets and models, e.g., modifications from known adversarial attacks

## EXPLAINABILITY

Methods to enhance persona's or role's ability to understand and interpret AI model outcomes, decisions, and recommendations, e.g., ranking and debating results and options

## LINEAGE

Methods to ensure the provenance of datasets and AI models, e.g., reproducibility of generated datasets and AI models

## AVAILABILITY

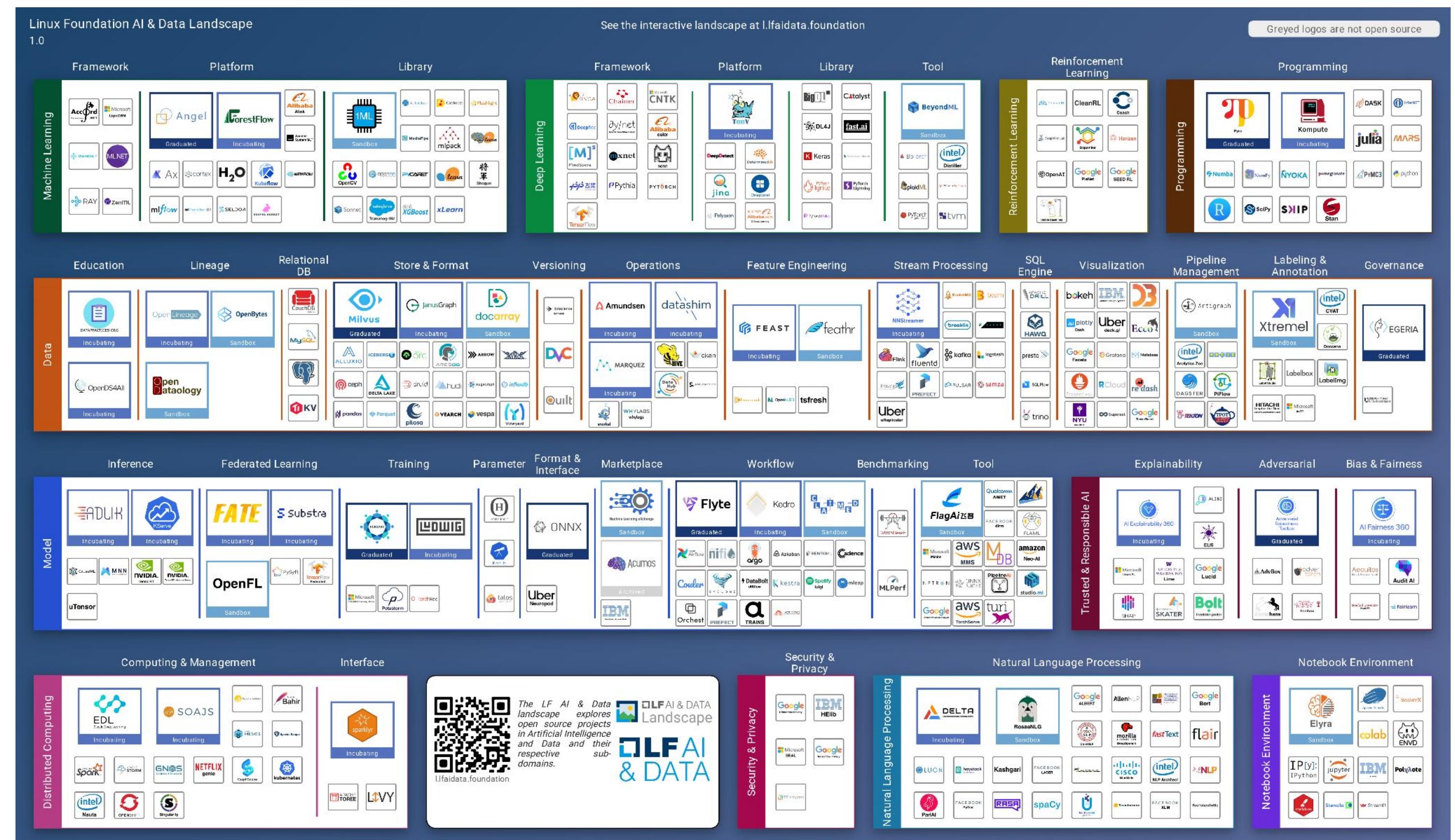
Open source data-specific licenses make data freely accessible for use without mechanisms of control

## GOVERNABILITY

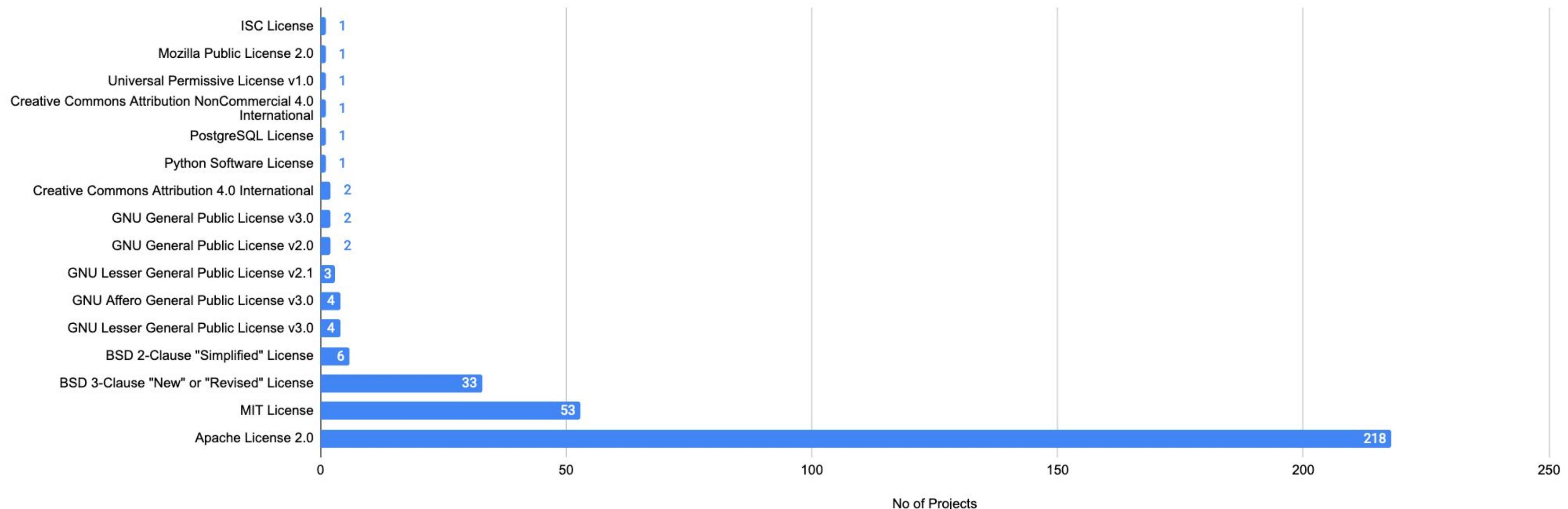
A governance structure and tools to clean, sort, tag, trace, and govern data and datasets

# A growing ecosystem

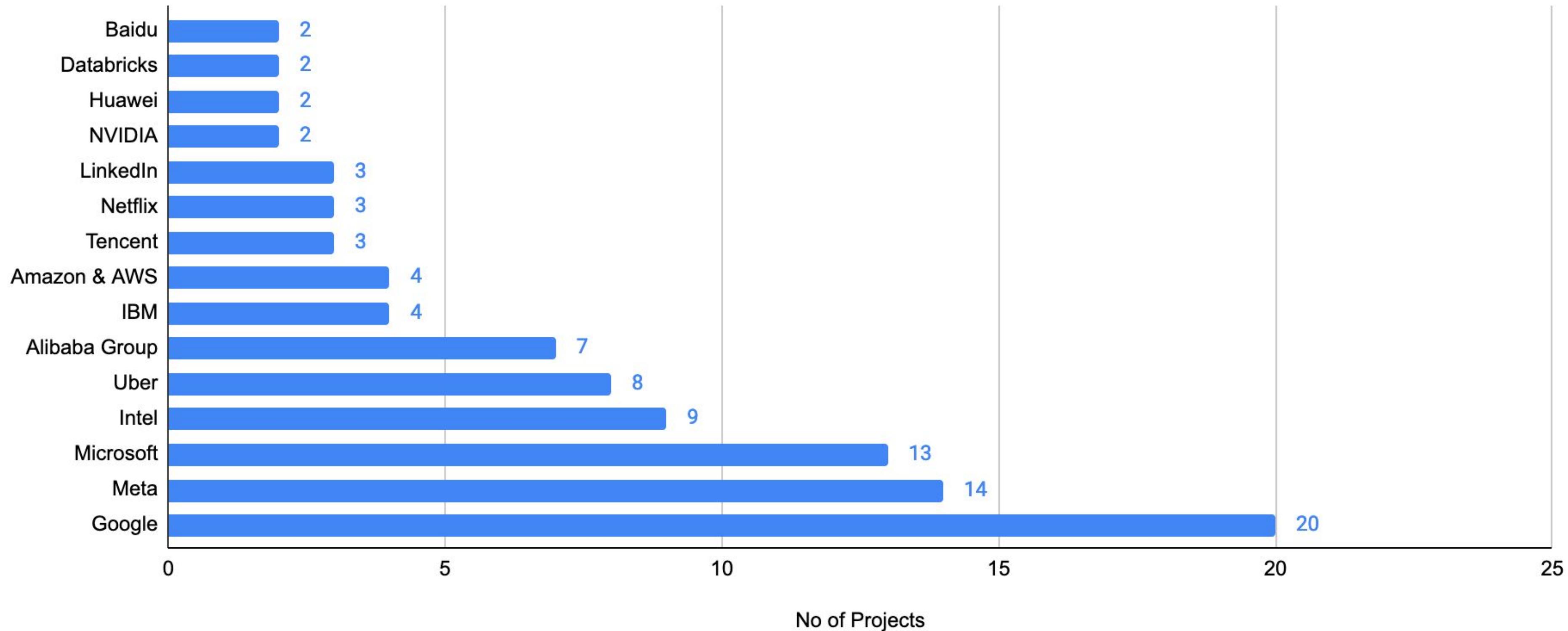
<https://l.lfaidata.foundation>



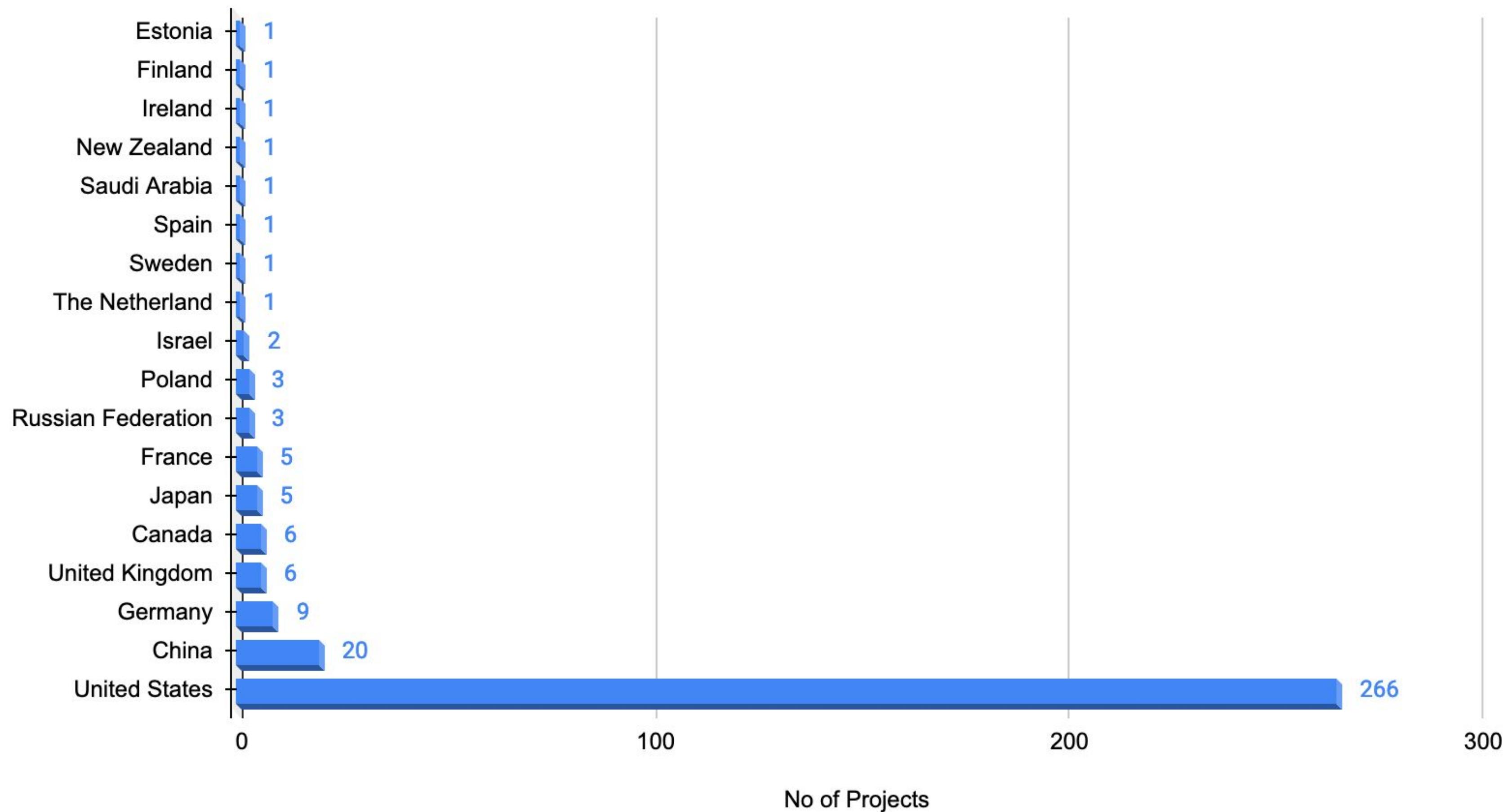
## No of Projects per License



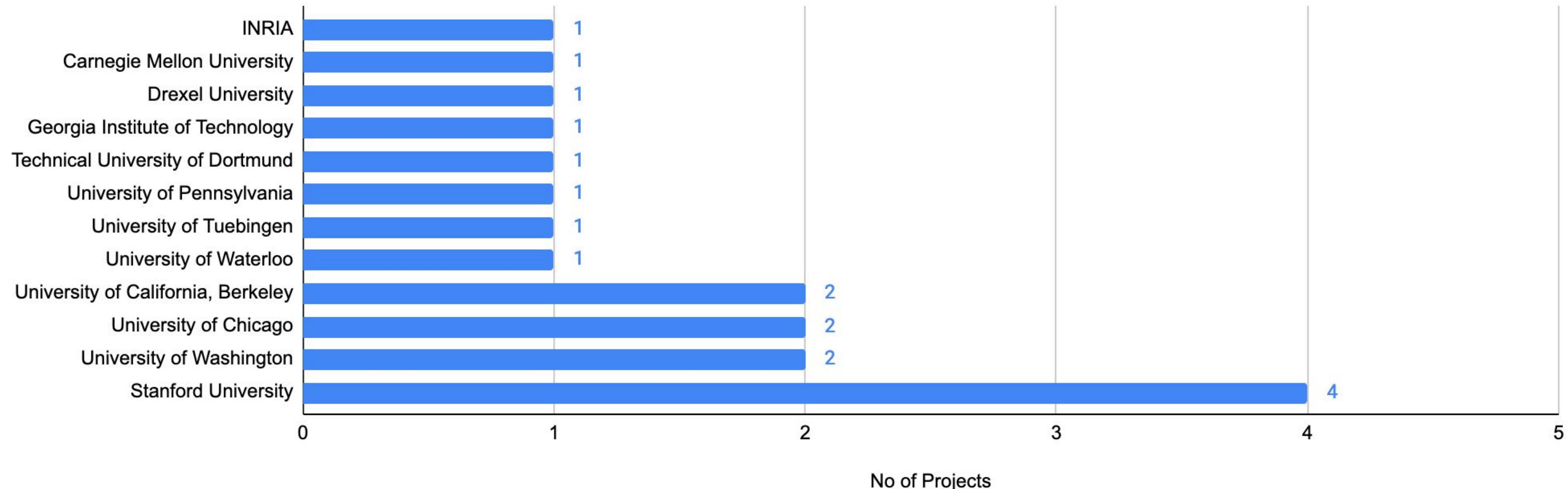
## No of Projects per Company



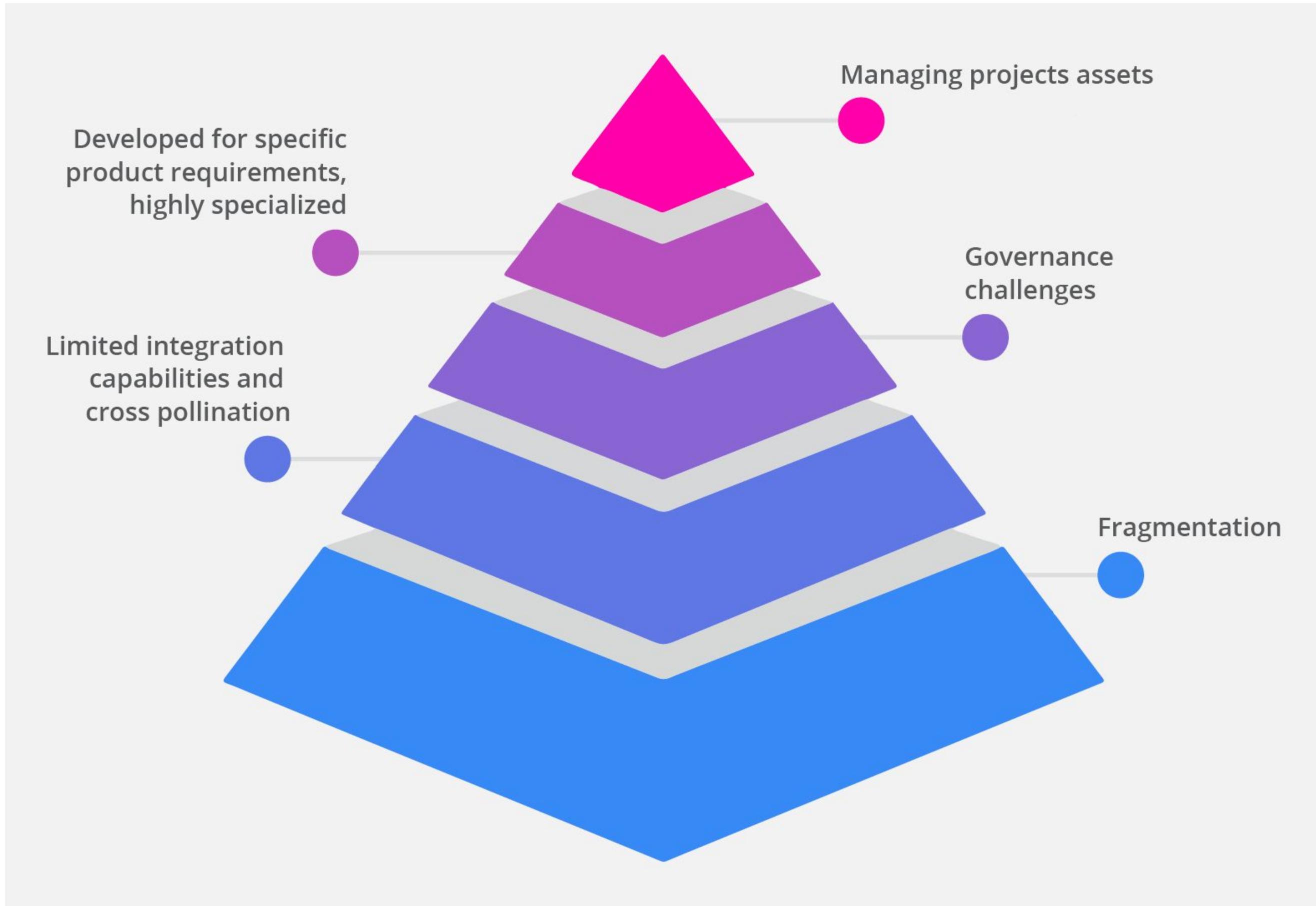
## No of Projects per Originating Country



## Founding Universities



# But with challenges...



# Who is LF AI & Data?

The LF AI & Data is a global not for profit foundation that hosts critical components of the global AI & Data technology infrastructure. It brings together the world's top developers, end users, and vendors to identify and contribute to the projects and initiatives that address industry challenges for the benefit of all participants.

# Our Mission

Our mission is to build and support an open AI community, and drive open source innovation in the AI, ML, DL and Data domains by enabling collaboration, sharing best practices, supporting development efforts, and the creation of new opportunities for all the members of the community.

# Key stats - unprecedented momentum

50  
Member  
Organization

46  
Technical  
Projects

6  
Active  
Committees

20+  
Technical  
Integrations

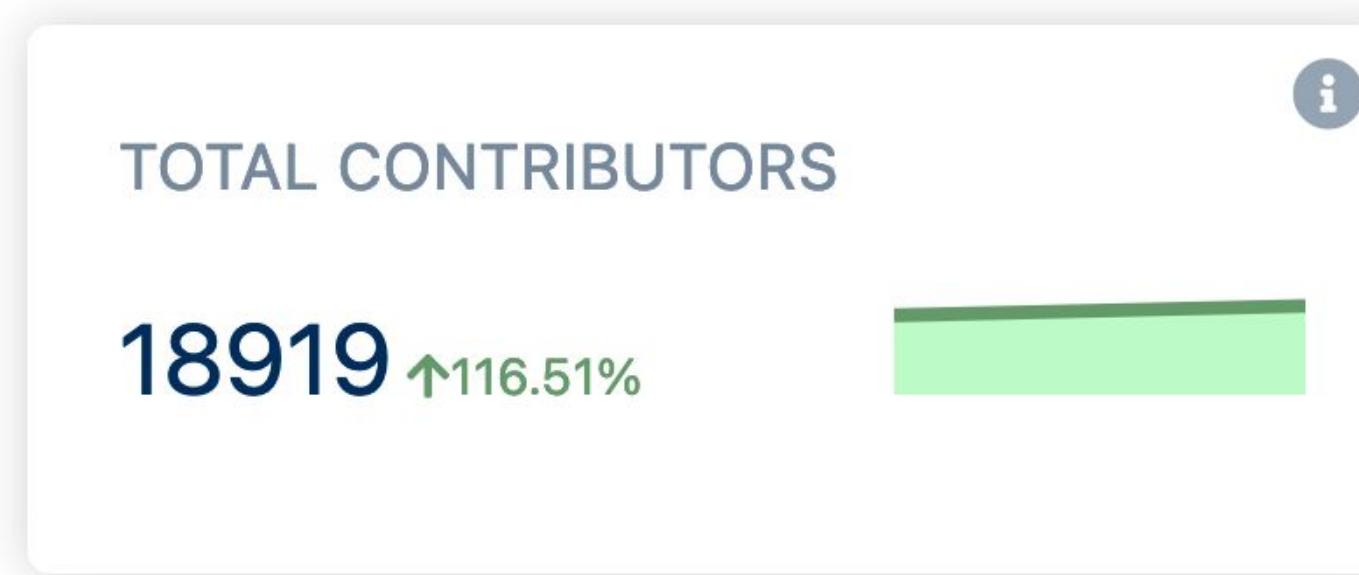
18,500+  
Contributors

188,000+  
GitHub Pull  
Requests

500+  
Active  
Organizations

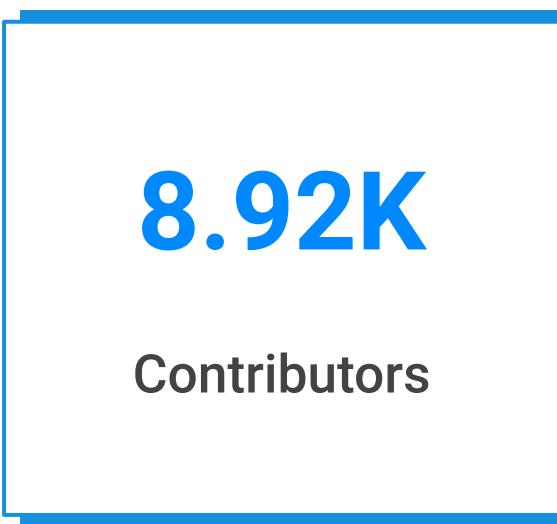
Feb 23, 2023

Source: [LFX Insights](#)



Dec 31, 2020

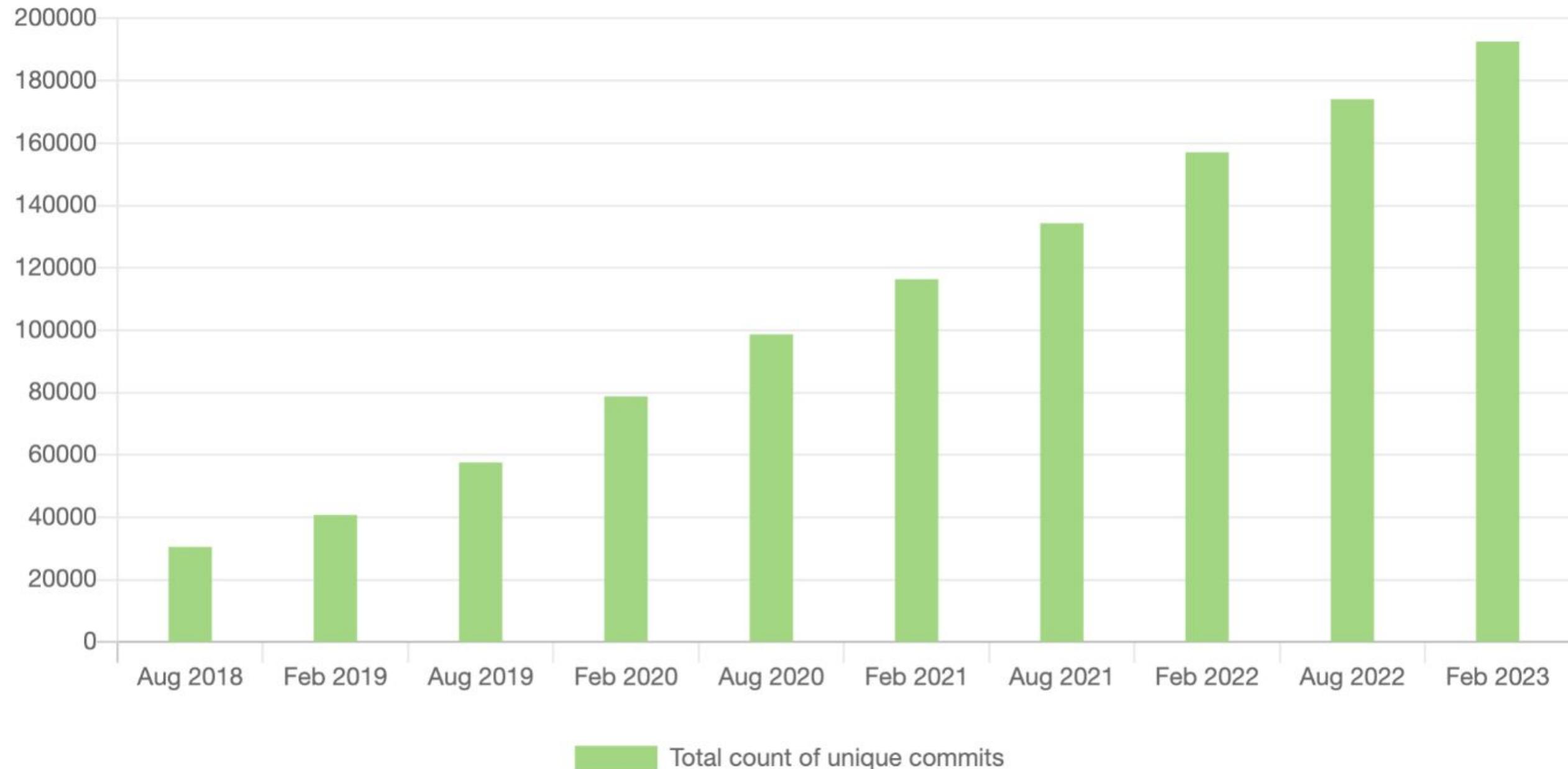
Source: LFX Insights



*Contributors performing code activity (commits/PRs/changesets) or submitting and resolving bugs are code contributors.*

# Commit Growth

Feb 23, 2023, Source: [LFX Insights](#)



There has been a  
**growth** of **531.29%**  
in the total commits  
during the last 5  
Years.

# Members

 <b>AI for people</b> AI for People	 <b>aivancy</b> PARIS-CACHAN Avancy School for Technology, Business & Society Paris-Cachan	 <b>Alibaba</b> Alibaba Cloud (Singapore) MCap: \$209.5B Alibaba Cloud Private LTD	 <b>AMBIANIC</b> Astronomer Astronomer	 <b>BAAI</b> 智源研究院 Funding: \$202.9M	 <b>BANQUE DE FRANCE EUROSYSTEME</b> Banque de France	 <b>BASIC AI</b> BasicAI	 <b>Baidu</b> Baidu Baidu Netcom Science Technology Co., Ltd.	 <b>PRIMITIVE HUB</b> Primitive Technology	 <b>BROADCOM</b> Broadcom Corporation	 <b>CHAITANYA BHARATHI INSTITUTE OF TECHNOLOGY</b> Chaitanya Bharathi Institute of Technology	 <b>CLOUDERA</b> Cloudera, Inc. Cloudera	
 <b>databricks</b> Databricks Inc. Databricks	 <b>DiDi</b> DiDi USA DiDi	 <b>元灵数智 DMetaSoul</b> DMetaSoul	 <b>ERICSSON</b> Ericsson Ericsson	 <b>Ersilia</b> Ersilia Open Source Initiative Ersilia Open Source Initiative	 <b>Institute for Biomedical Law &amp; Ethics</b> Ersilia Institute for Leadership Development	 <b>FURIOSA</b> FuriosaAI, Inc.	 <b>GALGOTIAS UNIVERSITY</b> Galgotias University	 <b>GRAVITI</b> GRAVITI TECHNOLOGIES INC Graviti	 <b>OPPO</b> Guangdong OPPO Mobile Telecommunications Corp., Ltd. Oppo	 <b>HST</b> High School Technology Services	 <b>HUAWEI</b> Huawei Technologies Co., Ltd. Huawei Technologies	 <b>ING</b> ING Group
 <b>IBM</b> International Business Machines Corporation IBM	 <b>The International Society of Service Innovation Professionals</b> ISSIP	 <b>jina</b> Jina AI Limited Jina AI Limited	 <b>RICE KEN KENNEDY INSTITUTE</b> Ken Kennedy Institute Ken Kennedy Institute	 <b>MAIEI</b> Montreal AI Ethics Institute Montreal AI Ethics Institute	 <b>NYU</b> New York University New York University	 <b>NOKIA</b> Nokia Corporation Nokia	 <b>OpenAI</b> OpenAI 新一代人工智能开源开放平台	 <b>OpenUK</b> OpenUK	 <b>PennState Great Valley</b> Penn State Great Valley - Funding: \$13.4M The Pennsylvania State University Penn State University	 <b>PSIT Kanpur</b> Premier Singh Institute Of Technology Premier Singh Institute Of Technology	 <b>precisely</b> Precisely Holdings, LLC Precisely	 <b>QuantUniversity, LLC</b> QuantUniversity Quantitative Insights
 <b>IBM   Red Hat</b> Red Hat, Inc. Red Hat	 <b>R Studio</b> RStudio PBC RStudio	 <b>SAHYADRI COLLEGE OF ENGINEERING &amp; MANAGEMENT</b> Sahyadri College of Engineering & Management Sahyadri College of Engineering and Management	 <b>sas</b> SAS Institute Inc. SAS	 <b>shopen</b> Shanghai OpenSource Information Technology Association Shanghai OpenSource Information Technology Association	 <b>4Paradigm</b> The 4th Paradigm Technology Co., Ltd. Paradigm4, Inc.	 <b>ETL</b> The Institute for Ethical AI & Machine Learning The Institute for Ethical AI & Machine Learning	 <b>UNION</b> Union	 <b>UNSW SYDNEY</b> University of New South Wales University of New South Wales	 <b>W UNIVERSITY OF WASHINGTON</b> University of Washington University of Washington	 <b>ULT</b> Université Libre de Tunis Université Libre de Tunis	 <b>vmware</b> VMware, Inc. VMware	 <b>X PRIZE</b> XPRIZE XPRIZE
 <b>zilliz</b> Zilliz	 <b>ZTE</b> ZTE Corporation ZTE											

# Foundation Structure and Governance

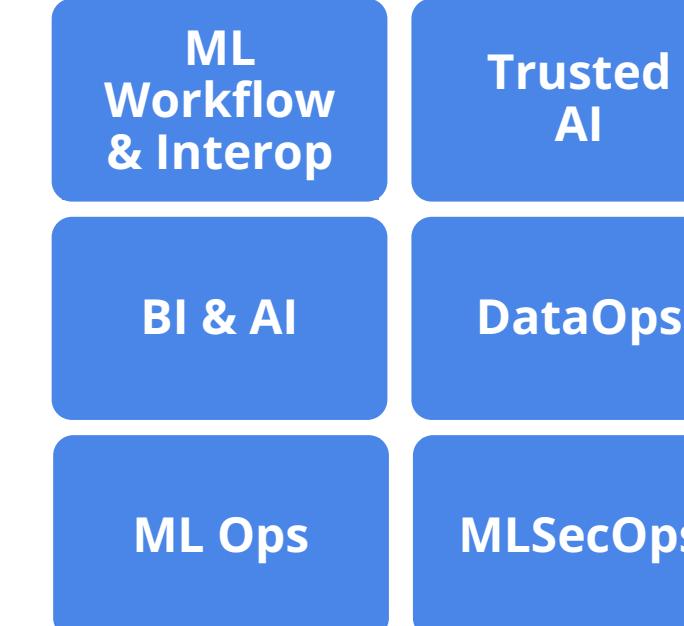
## Foundation Governance

(Funding effort)

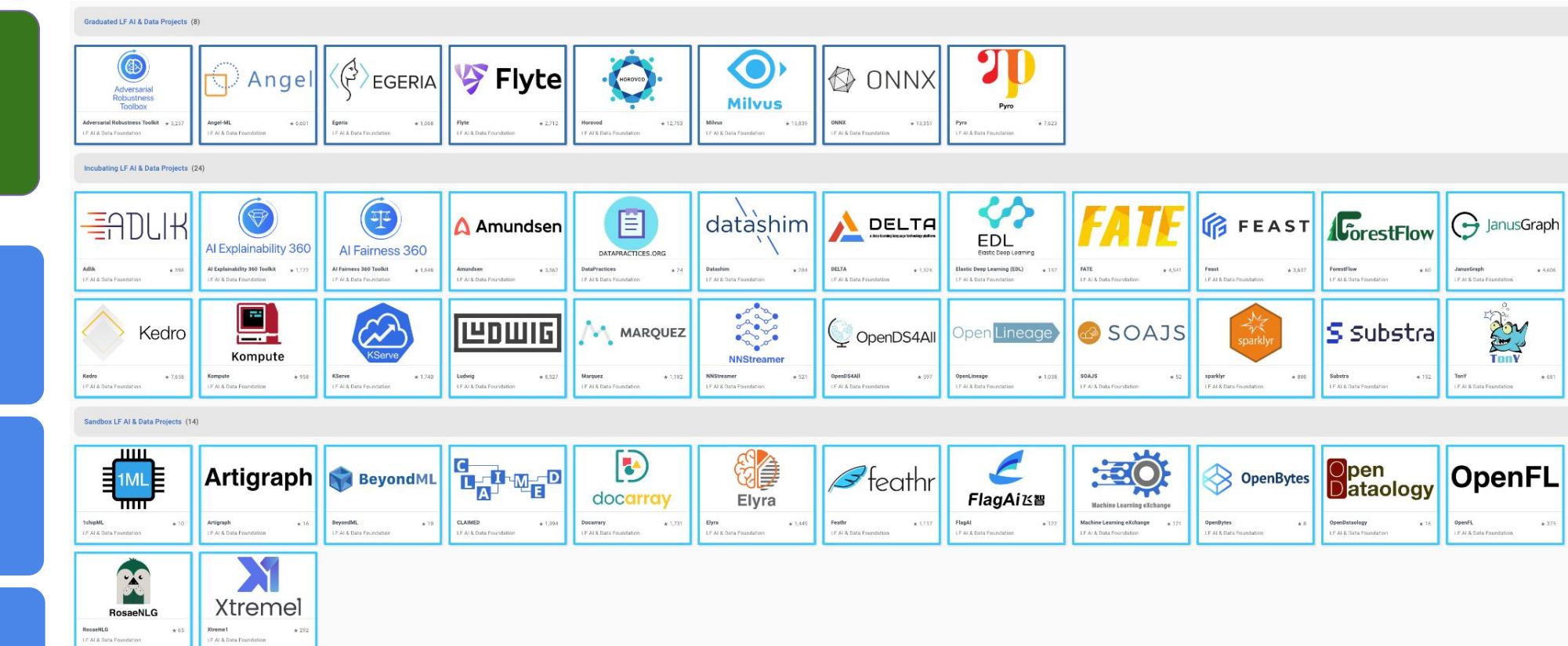


## Technical Coordination

### Technical Advisory Council



## Hosted Projects



# Hosted Projects

Graduated LF AI & Data Projects (8)

Adversarial Robustness Toolkit \* 3,257 LF AI & Data Foundation  
Angel-ML \* 6,601 LF AI & Data Foundation  
EGERIA \* 1,068 LF AI & Data Foundation  
Flyte \* 2,712 LF AI & Data Foundation  
Horovod \* 12,753 LF AI & Data Foundation  
Milvus \* 13,839 LF AI & Data Foundation  
ONNX \* 13,351 LF AI & Data Foundation  
Pyro \* 7,623 LF AI & Data Foundation

Incubating LF AI & Data Projects (24)

ADLIK \* 598 LF AI & Data Foundation  
AI Explainability 360 Toolkit \* 1,172 LF AI & Data Foundation  
AI Fairness 360 Toolkit \* 1,848 LF AI & Data Foundation  
Amundsen \* 3,562 LF AI & Data Foundation  
DataPractices.org \* 24 LF AI & Data Foundation  
datashim \* 284 LF AI & Data Foundation  
DELTA \* 1,506 LF AI & Data Foundation  
EDL Elastic Deep Learning \* 157 LF AI & Data Foundation  
FATE \* 4,541 LF AI & Data Foundation  
FEAST \* 3,637 LF AI & Data Foundation  
ForestFlow \* 60 LF AI & Data Foundation  
JanusGraph \* 4,605 LF AI & Data Foundation  
Kedro \* 7,658 LF AI & Data Foundation  
Kompute \* 958 LF AI & Data Foundation  
KServe \* 1,740 LF AI & Data Foundation  
LUDWIG \* 8,527 LF AI & Data Foundation  
MARQUEZ \* 7,182 LF AI & Data Foundation  
NNStreamer \* 521 LF AI & Data Foundation  
OpenDS4All \* 597 LF AI & Data Foundation  
OpenLineage \* 1,038 LF AI & Data Foundation  
SOAJS \* 52 LF AI & Data Foundation  
sparklyr \* 880 LF AI & Data Foundation  
Substra \* 152 LF AI & Data Foundation  
TonY \* 681 LF AI & Data Foundation

Sandbox LF AI & Data Projects (14)

1ML \* 10 LF AI & Data Foundation  
Artigraph \* 16 LF AI & Data Foundation  
BeyondML \* 10 LF AI & Data Foundation  
CLAIMED \* 1,094 LF AI & Data Foundation  
docarray \* 1,731 LF AI & Data Foundation  
Elyra \* 1,449 LF AI & Data Foundation  
feathr \* 1,117 LF AI & Data Foundation  
FlagAI 飞智 \* 122 LF AI & Data Foundation  
Machine Learning eXchange \* 121 LF AI & Data Foundation  
OpenBytes \* 8 LF AI & Data Foundation  
OpenDataology \* 16 LF AI & Data Foundation  
OpenFL \* 375 LF AI & Data Foundation  
RosaeNLG \* 65 LF AI & Data Foundation  
Xtremel \* 292 LF AI & Data Foundation

# Industry Leaders Host Their Projects in LF AI & Data



Bloomberg



Herron Tech



Tencent 腾讯



Uber

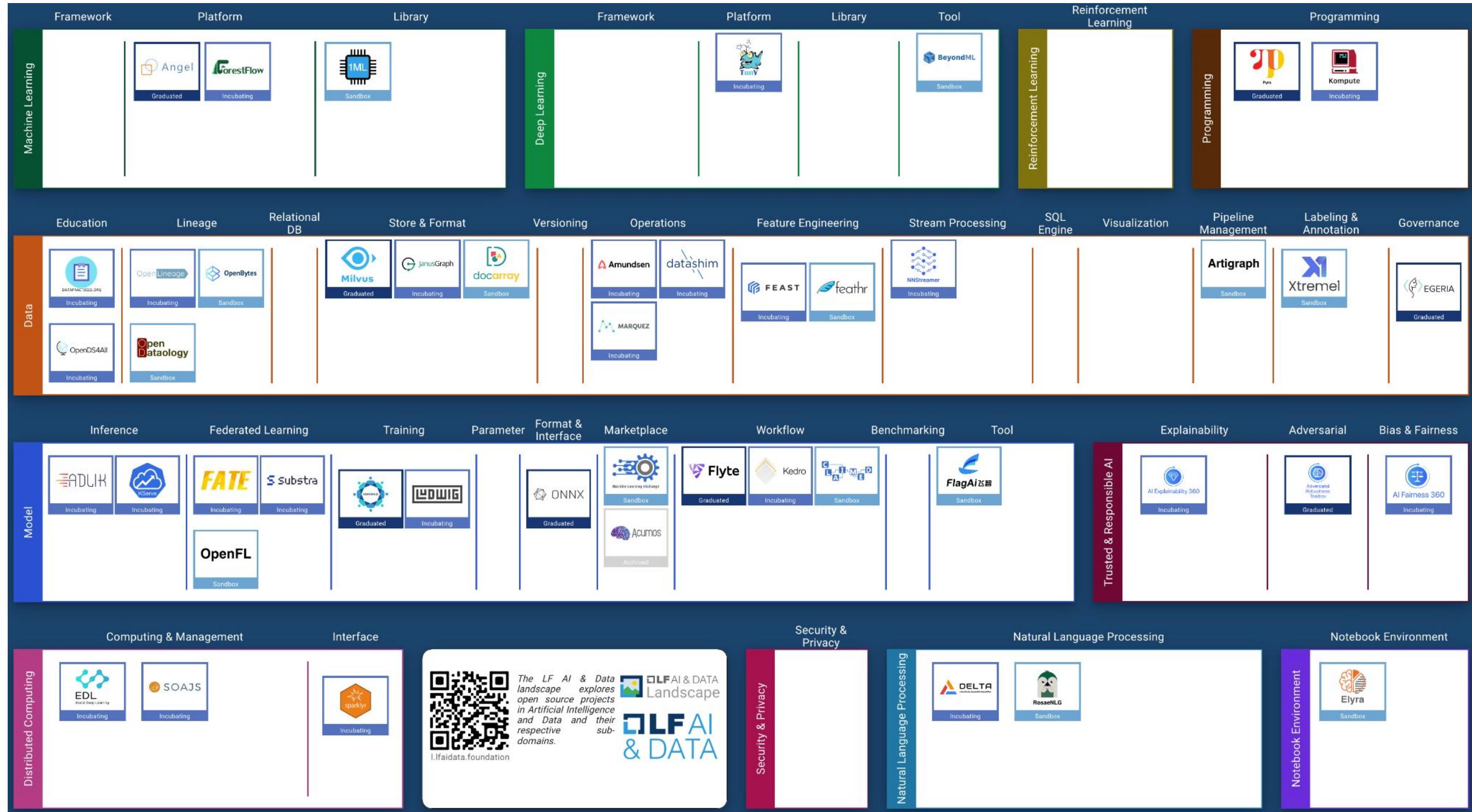
wework



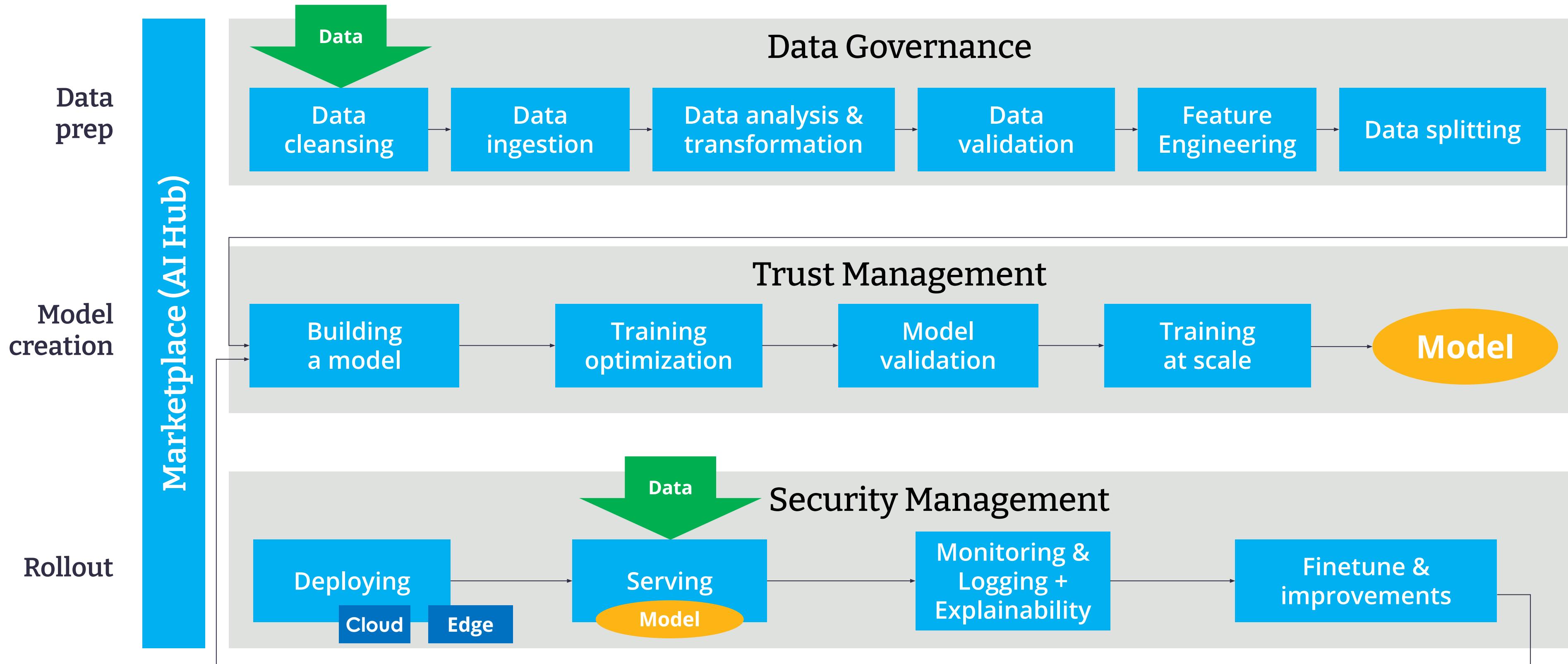
ZTE

# Over 12% of the ecosystem's key projects depend on LF AI & Data

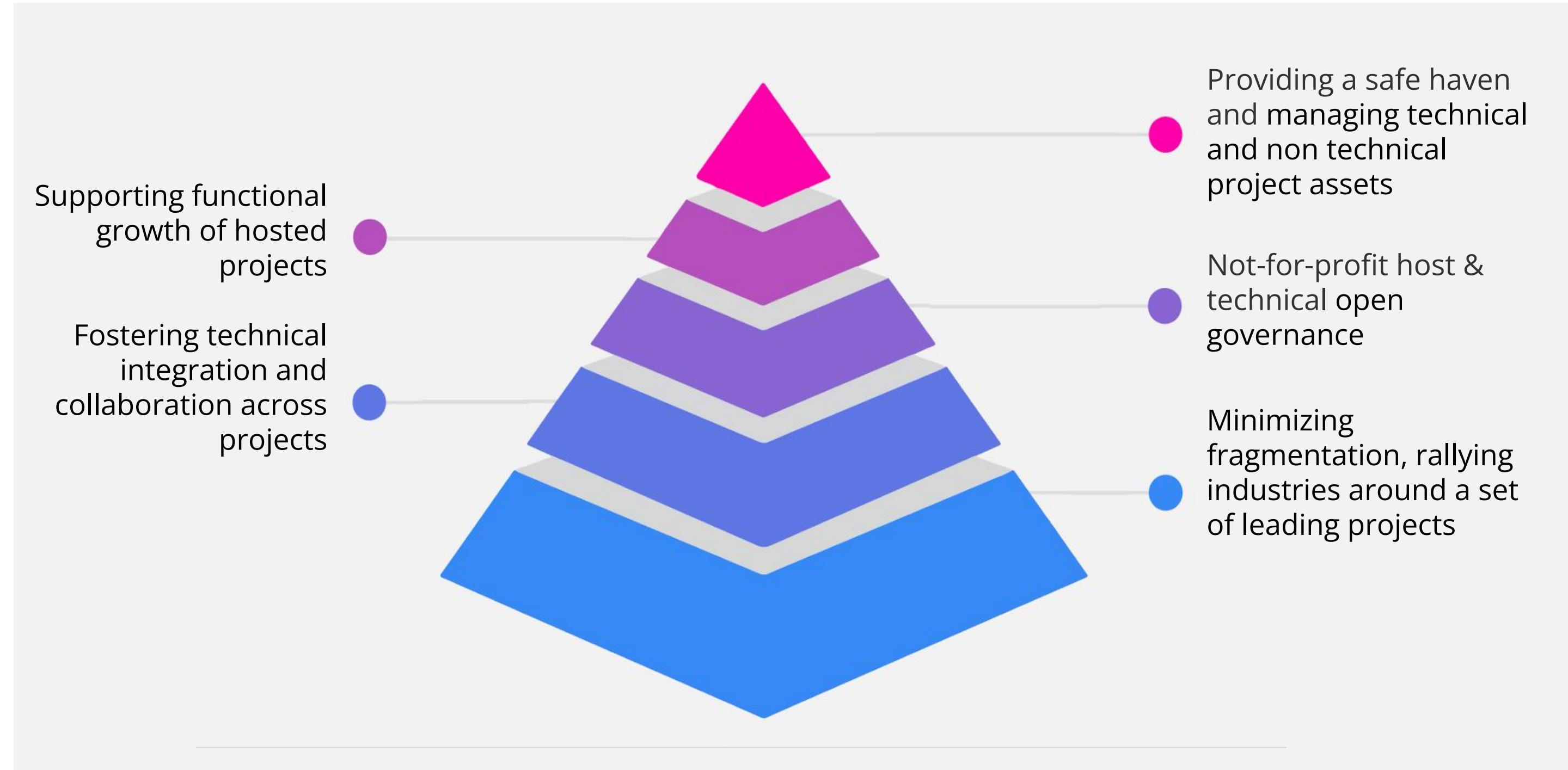
for their open governance, safe haven for their assets, infrastructure, and enabling marketing, legal and event services, and staff that is eager to help and support the communities of these various projects



# LF AI & Data provides a leading technical project for every phase of your ML Workflow



# LF AI & Data is addressing key ecosystem challenges



# LF AI & DATA TIMELINE 2022

## EVENTS



INFO@LFAIDATA.FOUNDATION

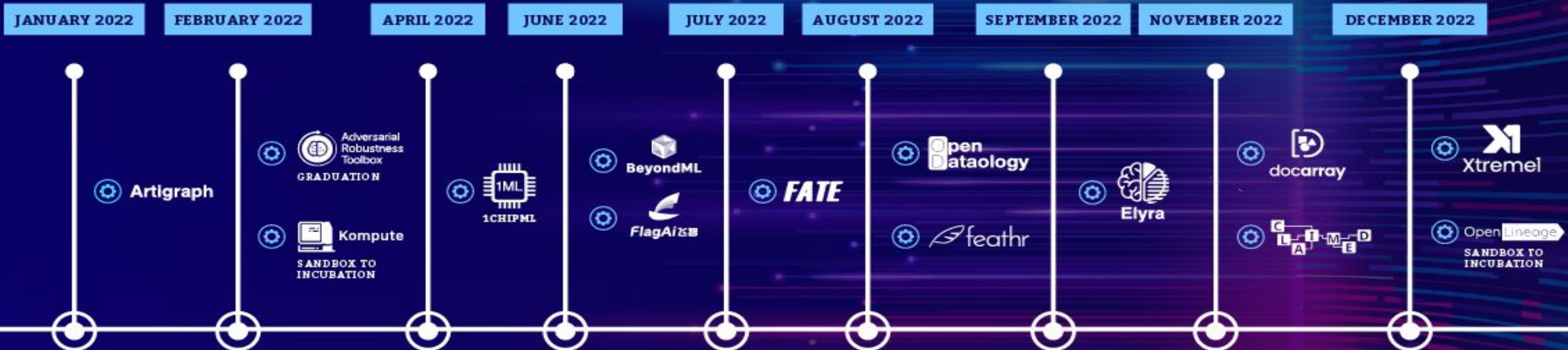
TECHNICAL PROJECTS

FOUNDATION

COMMUNITY

# OLFAI & DATA TIMELINE 2022

## PROJECTS



INFO@LFAIDATA.FOUNDATION

TECHNICAL PROJECTS

FOUNDATION

COMMUNITY

# Services available to hosted projects

<b>NEUTRAL HOSTING</b> <p>A neutral home for an open source project increases the willingness of developers from software companies, startups, academia, and elsewhere to collaborate, contribute, and become committers.</p>	<b>DEDICATED STAFF</b> <p>Projects have access to full-time staff (executive director, program manager, project coordinator) who cultivate the maturity and adoption of open source AI and data projects</p>	<b>TRAINING AND CERTIFICATION</b> <p>We develop training classes and, through the Linux Foundation, can execute and launch certification programs in support of hosted projects.</p>
<b>EVENTS MANAGEMENT</b> <p>Events are part of LF AI &amp; Data's core strategy to help projects build a community and accelerate knowledge-sharing and integration. Many LF AI &amp; Data projects have their own events.</p>	<b>DEV-FOCUSED OPERATION</b> <p>Services include IT infrastructure, release management, IT ops, support, security audits, and a host of tools (FOSSA, LastPass, Slack, Synk, Zoom, etc.).</p>	<b>MENTORSHIP</b> <p>Members of the LF AI &amp; Data technical advisory committee and leaders of graduated projects are available to support and mentor new projects.</p>
<b>MARKET SERVICES</b> <p>We offer a wide range of marketing services to increase project awareness, project adoption, and the number of contributors.</p>	<b>DESIGN AND AESTHETICS</b> <p>Our in-house team provides graphic design resources for new logos, websites, and website refreshes or enhancements.</p>	<b>PROGRAM MANAGEMENT</b> <p>We have decades of experience in program management of open source projects. We bring best practices to all LF AI &amp; Data hosted projects.</p>
<b>LFX PLATFORM EXPERIENCE</b> <p>This Linux Foundation product offers a set of integrated tools for project insights, security, easy contributor license agreements, crowdfunding, member engagement, and more.</p>		

# Strong Presence in China

Premier  
Members



OPPO

ZTE

General  
Members



OpenI 启智  
新一代人工智能开源开放平台



ZILLIZ



Hosted Projects



FATE



Companies  
hosting projects



Tencent 腾讯



ZTE

Board Chair



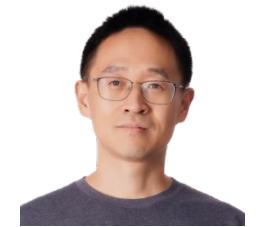
Junping Du  
(Huawei)

General  
Members Board  
Representative



Charles Xie  
(Zilliz)

Outreach Chair



Jerry Tan  
(4paradigm)

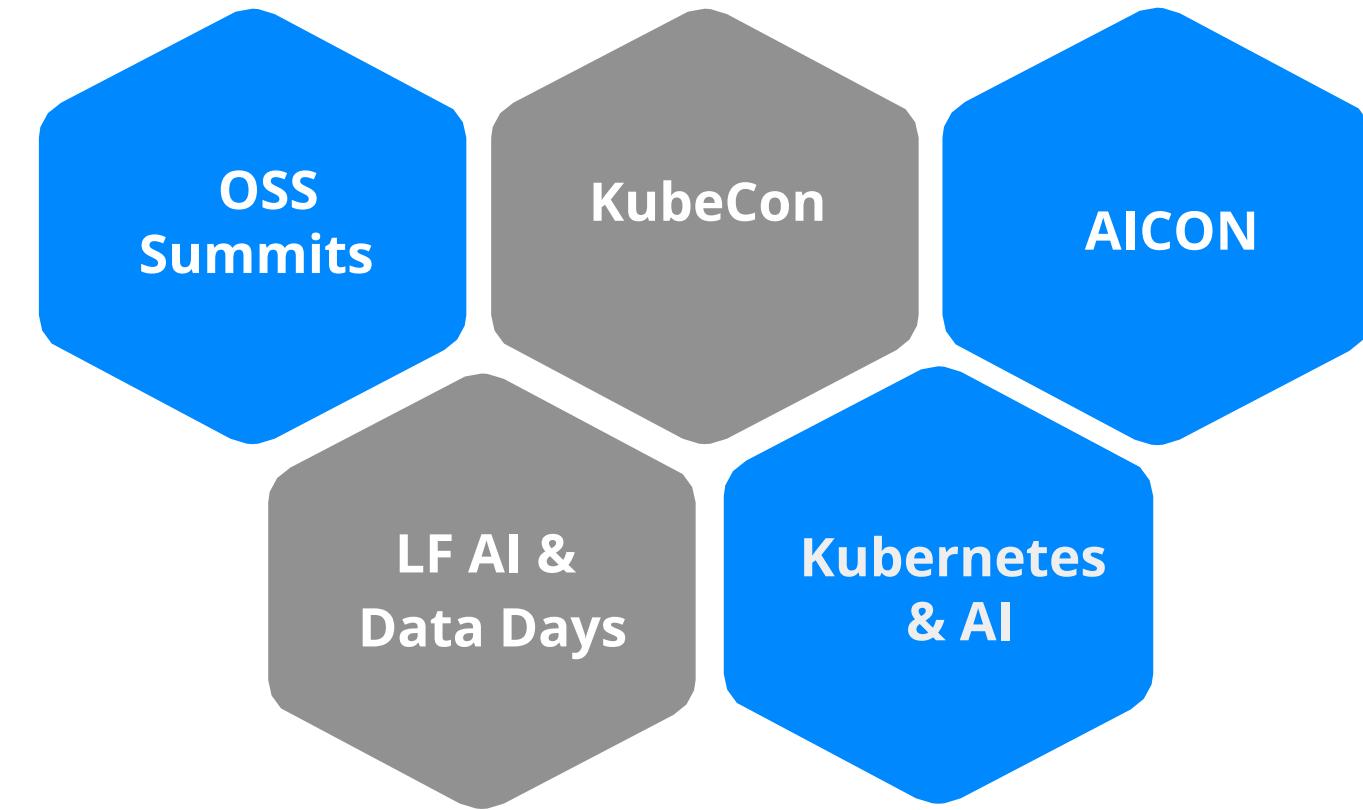
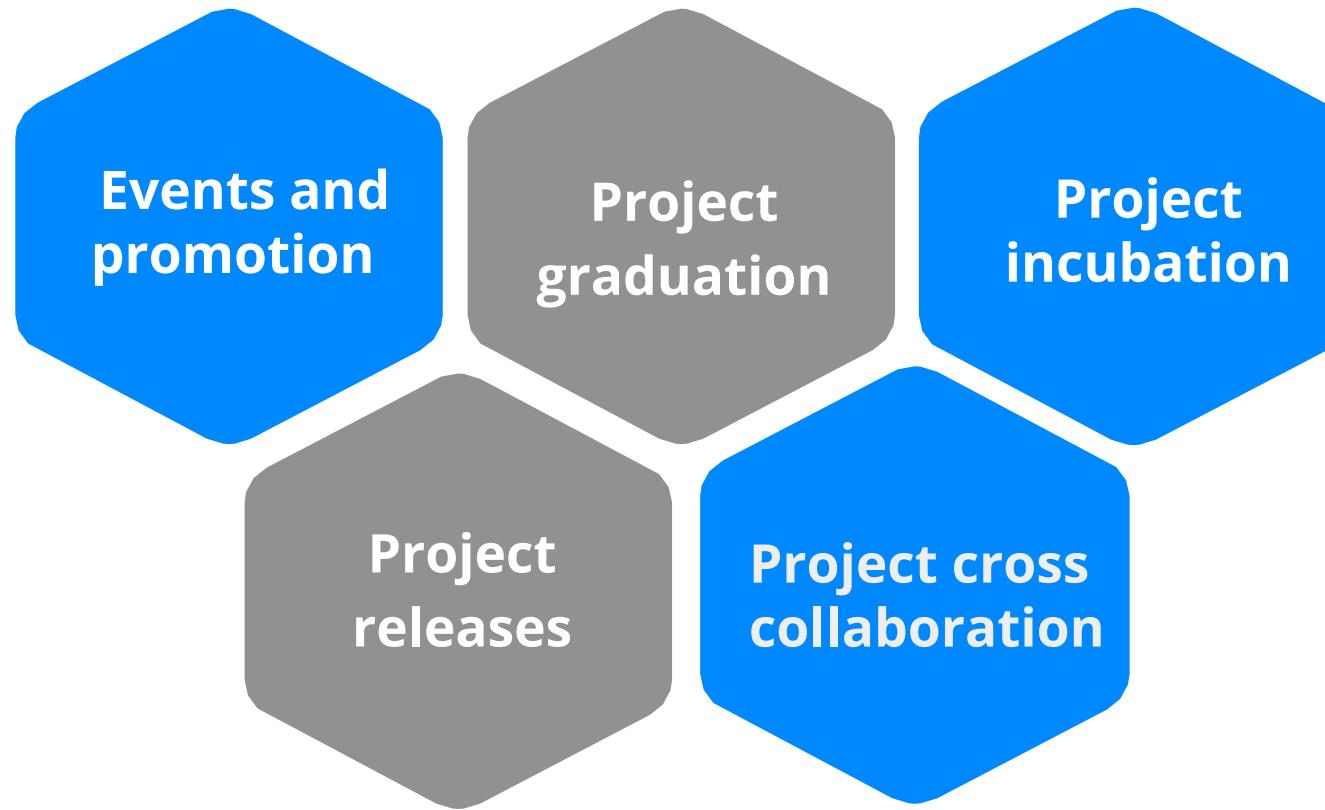
# Focused Developer & Community Marketing Effort



**MarComm services to support community and ecosystem engagement**



**Events and community ecosystem building services**



# Hosting Process

# Technical Advisory Committee

- The TAC serves a coordination role:
  - Votes on new projects joining LF AI & Data, as well as on promoting projects across incubation stages
  - Facilitates communication and fosters collaboration across hosted technical projects
  - Communicates needs and requirements of the projects to the Governing Board
  - Onboards new projects, assists in progression of existing projects, and reviews projects annually
  - Defines and maintains the technical vision for the LF AI & Data Foundation
  - Creates a conceptual architecture for the projects, aligning projects, promoting, removing or archiving projects
  - Defines common practices to be implemented across LF AI & Data projects
- Meetings via conference calls take place every 2 weeks, are recorded and open to the general public
  - <https://wiki.lfai.foundation/pages/viewpage.action?pageId=7733341>

# Project Incubation levels

## Sandbox

- 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).

## Incubation

Sandbox requirements plus:

- Have 2+ 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.
- Have at least 300 stars on GitHub.
- Have achieved and maintained a Core Infrastructure Initiative Best Practices Silver Badge.
- Have the affirmative vote of the TAC.

## Graduate

Incubation requirements plus:

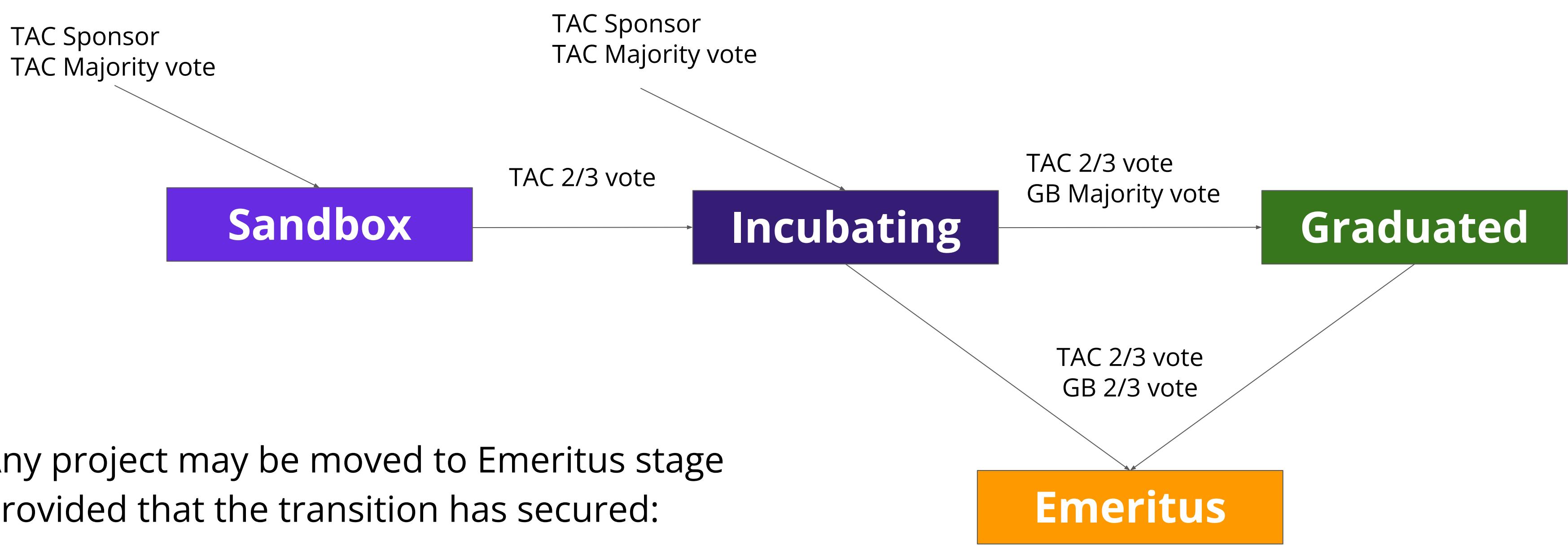
- Have a healthy number of code contributions coming from at least five organizations.
- Have reached a minimum of 1000 stars on GH.
- Have achieved and maintained a Core Infrastructure Initiative Best Practices Gold Badge.
- Have demonstrated a substantial ongoing flow of commits and merged contributions 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 TAC

# Transitioning from Incubation to Graduation

- The TAC will undertake an annual review of all projects
- Projects in the Sandbox Stage are generally expected to move to Incubation within 12 months from joining the Foundation following an evaluation by the TAC committee
- Projects in the Incubation Stage are generally expected to move to Graduation within 12-18 months from joining the Foundation following an evaluation by the TAC committee

Projects can be provided with an extension of time in their stage (up to the discretion of the TAC)

# Project Lifecycle



- › TAC affirmative  $\frac{2}{3}$  vote
- › GB affirmative  $\frac{2}{3}$  vote

# Process for proposing a project for hosting in LF AI & Data

1. Decide on a date to present to the TAC and request incubation
2. Ensure that your project implements these [recommendations](#)
3. Submit a formal request to incubate the project via a [GH](#) PR
4. Prepare deck and share with ED about 10 days prior to the presentation
5. Present to the TAC and get approval
6. Onboard the project with the LF AI & Data team and integrate the project with our service
7. Announce the project becoming hosted in LF AI & Data

# Incubation Benefits by Levels

## Sandbox

- Neutral hosting of the project's trademark and assets
- Announcement of the project hosting and social promotion.
- Right to refer to the project as a "LF AI & Data Foundation Project."
- An initial and regularly scheduled license scan of the project's codebase with results reported to the project's mailing list.
- Basic infrastructure support that includes mailing lists, wiki space, slack channel, etc.
- Basic marketing, communication, and PR support, limited to major announcements.
- Access to Foundation staff who are eager to help and support the project.

## Incubation

Sandbox benefits plus:

- Appointment of an existing TAC member that will act as a sponsor of the project and provide recommendations regarding governance best practices.
- Create the project's artwork, website and other required creative work.
- Advanced marketing, PR and communication support that includes project promotion via blog posts, social media and web site; support with white papers and posters, and any type of marketing collateral designated to promote the project.
- Zoom bridge to be used by the project's technical steering committee.
- Basic events support and promotion.

## Graduate

Incubation benefits plus:

- Graduate projects are eligible to receive support as determined by the Governing Board.
- LF AI & Data blog announcement or similar announcing the graduation of the project including promotion activities.
- Graduation stage projects gain a voting seat on the TAC.
- Advanced events support and promotion that includes event website, registration, support with sponsors, logistics, etc.
- Advanced IT infrastructure support.
- Additional ecosystem development opportunities, such as conformance programs, training, etc.

# LF AI & Data Membership Benefits

LF AI & Data is positioned to bolster cross-company collaboration and interoperability with our neutral IP zone, rich contributor programs, and most importantly, the trust end users place in us.

# Organizations join LF AI & Data to take an active role in supporting the growth and evolution of the Open Source AI & Data ecosystem

**Support our mission,** programs, and the community by helping fund services that hosted projects rely on

**Host projects with us** and benefit from the services & support we offer our hosted projects to increase their adoption and footprint in the ecosystem

**Demonstrate thought leadership** by participating in a wide reaching networking and marketing programs

**Provide technical leadership** to the community via the TAC, its various sub-committees and efforts

Be part of **defining and maintaining technologies** that are at the forefront of the industry

**Receive greater insight** into the foundation's strategy, efforts, projects and initiatives through increased engagement with the ED, staff and committees' leads

# LF AI & Data Membership Benefits



Marketing  
Amplification &  
Brand Awareness



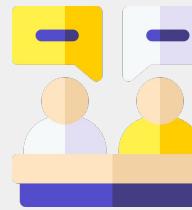
Community  
Engagement



Thought & Tech  
Leadership Across  
Key Technologies

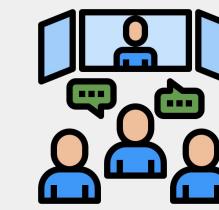
# Marketing amplification and brand awareness

**Broaden your reach and awareness in the community with LF AI & Data marketing programs.**  
**As a member you can participate in:**



## LF AI & Data Outreach Committee

Participate in the marketing committee monthly to engage with your peers in the cloud native space.



## LF AI & Data Summit

LF AI & Data organizes its own community's open source developer summits focused on open source AI and Data technologies.



## LF AI & Data Online Programs

Showcase your organization's open source technology by educating new and existing community members about best practices, trends, and new technologies



## Public Relations & Analyst Relations Support

LF AI & Data supports members with analyst reports and research highlights.

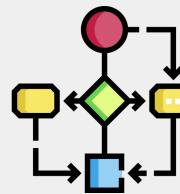


## LF AI & Data Blog

Showcase your thought-leadership and industry commentary, as well as share technical walkthroughs for LF AI & Data projects here.

# Community Engagement

**LF AI & Data is a constellation of open source projects. Our members leverage many efforts to engage with our project's ecosystems and share their stories. As a member you can participate in:**



## ML Workflow and Interoperability

Design a reference architecture for end-to-end ML Workflow, provide a reference implementation using hosted projects, and facilitate integration and interop across projects.



## BI & AI

Integrate the power of AI and BI to make it CI (Cognitive Intelligence) by combining the speed machines accelerate (AI) with the direction intuited by human insight (BI).



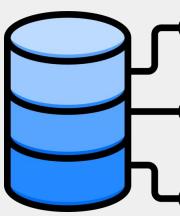
## Trusted AI

A global group working on policies, guidelines, tools and use cases by industry to ensure the development of trustworthy AI systems and processes to develop them continue to improve over time.



## LF AI & Data Day

A 1-day event organized multiple times per year focused on specific technical projects and collaboration ongoing in LF AI & Data



## DataOps

Share best practices, and bring technology awareness around DataOps across industries and enable collaboration across LF AI & Data member companies, projects, and external participants.

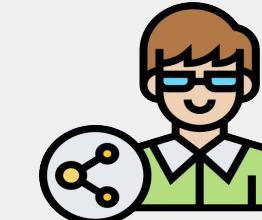
# Thought leadership

**Members of the LF AI & Data can network and help shape the open source AI and Data market.**  
**As a member you can participate in:**



## Governing Board Participation

Participate in elections to serve on the Governing Board to oversee the vision of LF AI & Data and work with the TAC.



## Technical Advisory Council Leadership

LF AI & Data TAC provides technical leadership to the open source AI and Data community, accepts projects into incubation, graduates projects and launches new efforts and collaborations.



## Engagement with other Foundations and Industry Consortia

Participate in establishing collaborations with other umbrella foundations under the LF and broadly with other industry consortia organizations.



## Interactive Landscape Placement

Our landscape is a comprehensive view of all open source AI and Data critical projects in the ecosystem. Anyone looking to adopt open source AI and Data projects comes here to review what technologies they should assess and adopt.

# LF AI & Data Annual Dues

	Not Yet a Linux Foundation Member	Existing Linux Foundation Member
Premier Member	\$120,000	\$100,000
General Member	5,000 employees +: \$45,000 2,000 - 4,999: \$30,000 500 - 1,999: \$25,000 Up to 499 employees: \$10,000	5,000 employees +: \$25,000 2,000 - 4,999: \$15,000 500 - 1,999: \$10,000 Up to 499 employees: \$5,000
Associate Member		Free <small>Limited to academic and nonprofit institutions respectively and requires approval by the Governing Board</small>

# Premier Membership

Highest tier of membership – For organizations contributing heavily to open source AI and Data, and bringing their own projects to be hosted at the Foundation. They work in concert with the Foundation team members. These companies want to take the most active role in enabling the open source AI and Data ecosystem.

## Premier members are eligible to:

(Enjoy all the benefits of General level, plus;)

- **Hold one (1) guaranteed seat** on the LF AI & Data Governing Board + one alternate (1) representative
- **Appoint one (1) voting representative in any subcommittees** or activities of the LF AI & Data Governing Board
- Receive greater **insight into LF AI & Data strategy and projects** through engagement with the LF AI & Data leadership team. Premier level members have the unique opportunity to **customize their experience** with LF AI & Data. The team will make themselves available to help achieve your strategic goals. We can help with guidance in open source contributions, new market creation, and/or open source project donation. **Have ideas? Just ask!**
- Enjoy most prominent placement in displays of membership including website, landscape and marketing materials.
- Create an individualized press release upon membership announcement with the LF AI PR team.

# General Membership

Targeted for organizations that want to support of LF AI & Data and our mission. Organizations that join at the General level are deeply committed to using open source technology, helping LF AI & Data grow, voicing the opinions of their customers, and giving back to the community.

## General members are eligible to:

- **Participate in elections** between **other General members** to appoint one (1) representative to the **LF AI & Data Governing Board**. Three (3) total General representatives will be elected to represent all General members. Voice your opinions amongst the leaders in the industry and help determine the strategic direction of LF AI & Data.
- Receive **greater insight into LF AI & Data strategy and project roadmaps** through increased engagement with the LF AI & Data Executive Director and staff.
- Create an announcement upon membership announcement with the LF AI & Data PR team.
- Participate in all Marketing, Community, Thought Leadership opportunities.
- Opportunity to **host “LF AI & Data Day”**, including on-demand webinars and livestream.
- Receive **discounts** on LF AI & Data event sponsorships.
- Demonstrate your support for LF AI & Data by displaying your logo on the LF AI & Data website, landscape and in marketing materials.

# Associate Membership

 **AI & DATA**  
ASSOCIATE MEMBER

The Associate membership is a free complimentary membership limited to academic and non-profit institutions.

## Associate members are eligible to:

- Participate in all Marketing, Community, and Thought Leadership opportunities.
- Identify your organization as a member and display your logo on the LF AI & Data website, landscape and in marketing materials.
- Feature your organization in the quarterly new members announcement.
- Receive discounts on LF AI & Data events' sponsorship.

### Specific opportunities for associate members:

- Practical experience working with open source foundations
- Visibility into and within the ecosystem
- Internships opportunities with commercial members
- Collaboration opportunities
- Publication opportunities based on R&D conducted with projects
- Open source working experience for all participants
- Access to LF AI & Data events

# Getting Involved in LF AI & Data

# Getting involved with technical projects

Participate in the development: Review and submit patches, report bugs, request features, test, etc.

Contribute to project documentation

Join the projects' mailing lists and participate in the discussions

Attend developer events for LF AI & Data projects

Provide your testing and deployment feedback via appropriate project channel

Start a local User Group Meetup

# Getting involved with the Technical Advisory Council

Support TAC leadership in inviting speakers to present for the LF AI & Data technical community

Share success stories, opportunities and challenges in relation to open source AI with the broader technical community

Support technical leadership for integration and harmonization efforts with other open source projects

Support TAC in evaluating new projects for incubating; recommend new projects

Participate in the ML Workflow effort aiming to provide a reference workflow and support integration across LF AI & Data projects

Attend TAC Bi-weekly calls, participate in the discussions

Identify opportunities for collaboration on common interests and initiatives, seek input from peers

Support TAC in hosting and sponsoring intra-project and inter-project developer events

Support TAC Chair working with the GB to highlight opportunities and needed resources in support of hosted projects

Participate in the effort to enable collaboration with other LF umbrella foundations and external communities

# Getting involved with the Outreach Committee

Promotion of project updates, releases, and news via LF AI & Data social media accounts

Marketing and PR support for demos at meetups and events

Contribute to the LF AI & Data landscape, promote in talks

Identify speaking opportunities and help secure speakers from the LF AI & Data community

Attend Outreach Committee meetings and participate in ongoing activities

Host vendor neutral content via LF AI blog site

Coordination at events, speaking proposals, booth attendance, demos, etc.

Help secure user stories about LF AI & Data based deployments

Publish use cases, case studies, white papers, and deployment insights

Get support for artwork, web site, content creation, etc., for LF AI & Data projects

Volunteer to host an LF AI & Data Day and developer events

Support LF AI & Data marketing and PR staff

# Getting involved with the LF AI & Data Committees: BI & AI, DataOps, ML Workflow & Interop

Join the committee as a representative of your company

Contribute to the landscape, promote in talks

Host your related projects in the LF AI & Data Foundation

Join the conference calls and contribute to ongoing efforts

Coordination at events, speaking proposals, booth attendance, demos, etc.

Help secure user stories about LF AI & Data based deployments

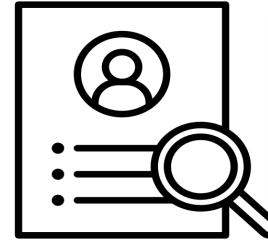
Promote the work of the committee, invite your collaborators to participate

Subscribe to the mailing list and participate in discussions

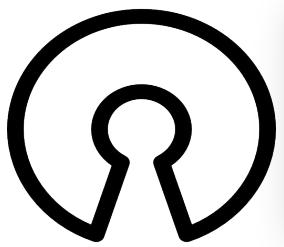
Invite prominent researchers and developers in the space to speak to the community

# Challenges and Trends

# Global Trends



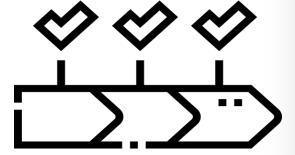
Talent  
Shortage



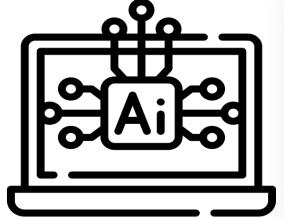
Rise of OSPOS



Software  
security



Software  
supply chain



Infusing AI  
and ML

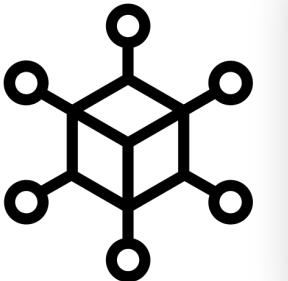


Ethical AI

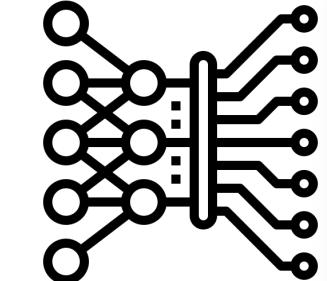
# AI Challenges and Opportunities



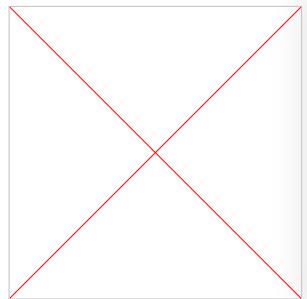
Trusted and  
Responsible AI



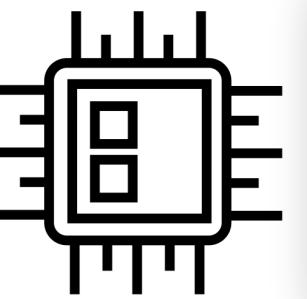
AI on edge  
devices



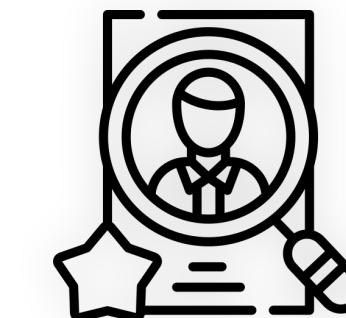
Federated  
learning



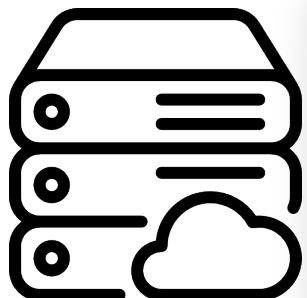
Social Impact



Specialized  
Hardware



Availability of  
Skilled Talent



Scalability,  
Efficiency



Policies,  
Regulations &  
Standards



User  
Acceptance,  
Building Trust

# Data Challenges and Opportunities

Security

Privacy

Quality

Lineage

Deriving  
value

Integration

Storage

Ethical

Licensing

Labeling

Quantity

Diversity

Governance

Accessibility

Regulations,  
Legislations

Standards

Join us, become a member!

Incubate your project

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

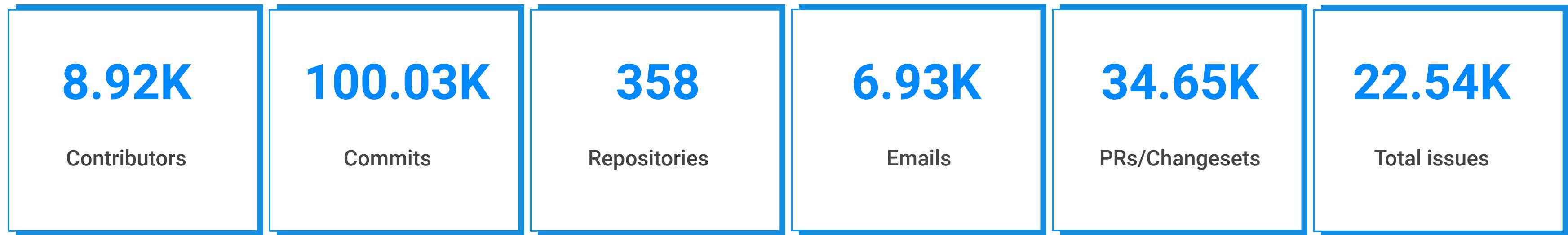


blog

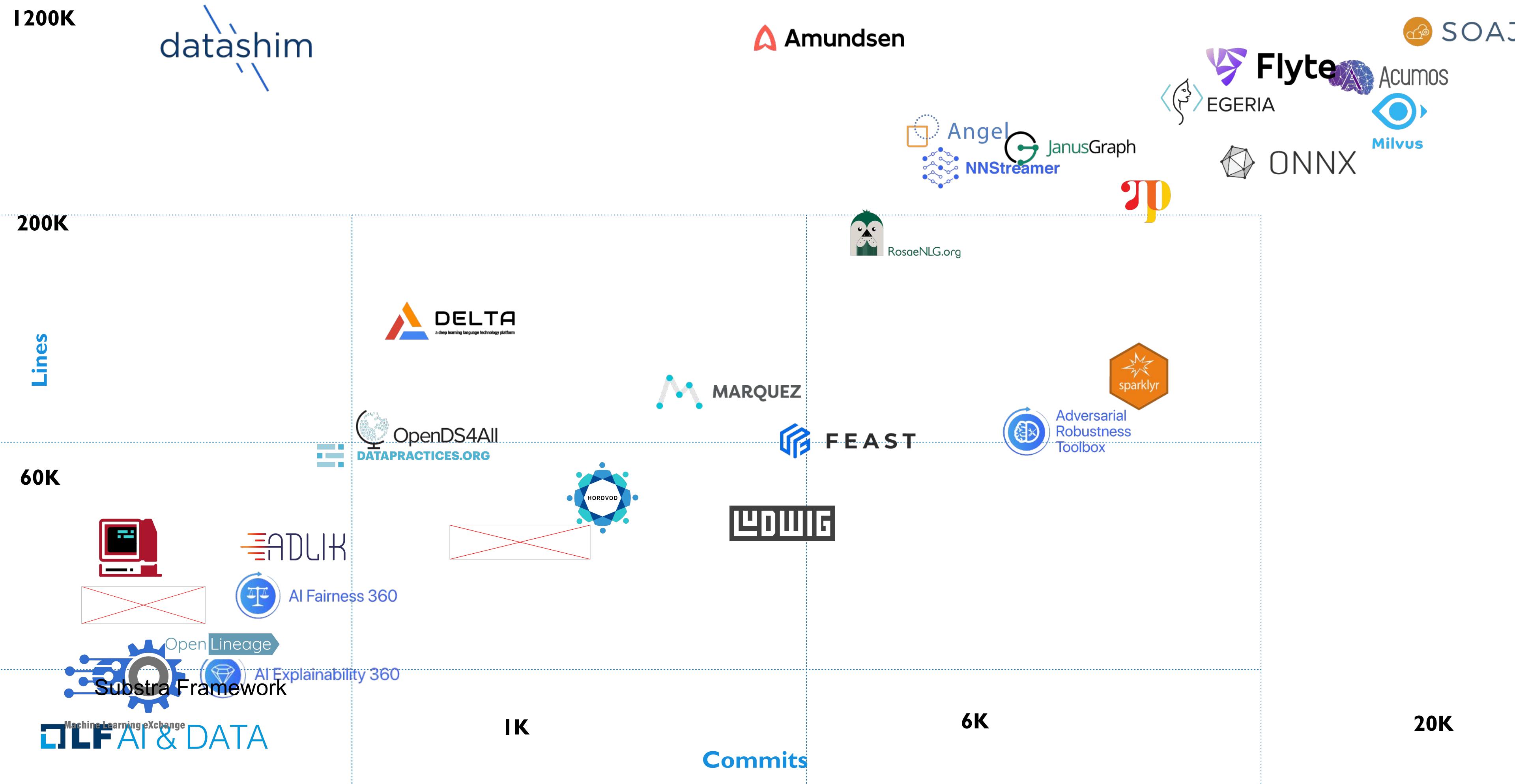
OLD FOR REFERENCE ONLY

# Active and growing developer community

Jan 1- Dec 31, 2020



# Data Coverage (June 7, 2021) - Line of Code and Commits



# Data Coverage (June 7, 2021) - Stars and Contributors

11K



ONNX



GitHub  
Stars

0.4K

MLF AI & DATA  
Machine Learning eXchange

0.3K

0.7K

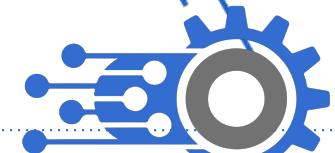
6.2K

Contributors



Substra Framework

datashim



OpenLineage



EGERIA



Flyte



AI Explainability 360



DELTA

a deep learning language technology platform



AI Fairness 360



MARQUEZ



RosaeNLG.org



ADLIK



NNStreamer



SOAJS



Acumos



LUDWIG



Angel



Adversarial  
Robustness  
Toolbox



Milvus



JanusGraph

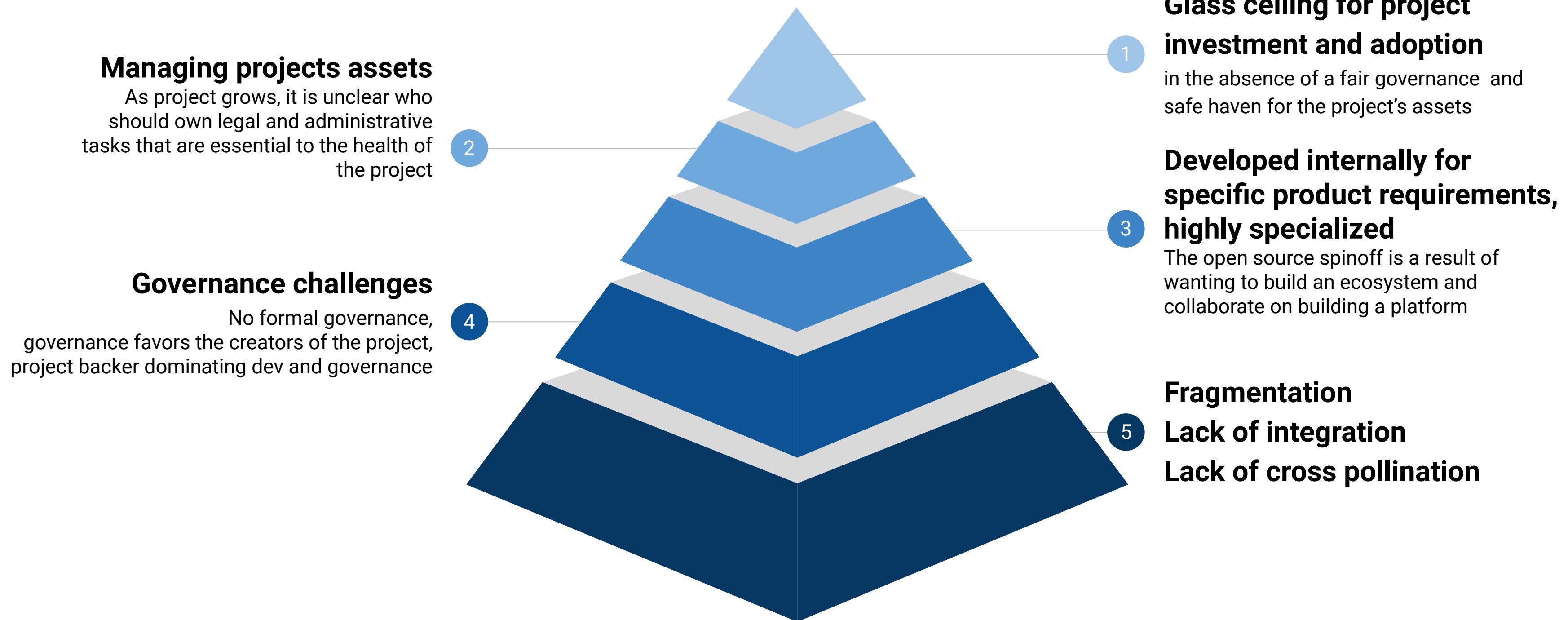


FEAST



sparklyr

# But with challenges



# Why host your project in LF AI & Data?

<b>Neutral hosting</b> <p>A neutral home for an open source project increases the willingness of developers from software companies, startups, academia, and elsewhere to collaborate, contribute, and become committers.</p>	<b>Dedicated staff</b> <p>Projects have access to full-time staff (executive director, program manager, project coordinator) who cultivate the maturity and adoption of open source AI and data projects.</p>	<b>Training and certification</b> <p>We develop training classes and, through the Linux Foundation, can execute and launch certification programs in support of hosted projects.</p>	<b>Events management</b> <p>Events are part of LF AI &amp; Data's core strategy, to help projects build a community and accelerate knowledge-sharing and integration. Many LF AI &amp; Data projects have their own events.</p>
<b>Dev-focused operation</b> <p>Services include IT infrastructure, release management, IT ops, support, security audits, and a host of tools (FOSSA, LastPass, Slack, Synk, Zoom, etc.).</p>	<b>Mentorship</b> <p>Members of the LF AI &amp; Data technical advisory committee and leaders of graduated projects are available to support and mentor new projects.</p>	<b>Market services</b> <p>We offer a wide range of marketing services to increase project awareness, project adoption, and number of contributors.</p>	<b>Legal services</b> <p>We help projects navigate licensing requirements, IP regimes, trademark management, compliance scans, export control filings, and developer certificate of origin or contributor license agreement integration with GitHub, etc.</p>
<b>Design and aesthetics</b> <p>Our in-house team provides graphic design resources for new logos, websites, and website refreshes or enhancements.</p>	<b>Program management</b> <p>We have collectively decades of experience in program management of open source projects. We bring best practices to all LF AI &amp; Data hosted projects.</p>	<b>LFX platform experience</b> <p>This Linux Foundation product offers a set of integrated tools for project insights, security, easy collaborator license agreements, crowdfunding, member engagement, and more.</p>	