

# Technical Advisory Council Meeting

October 22, 2020

 THE **LINUX** FOUNDATION

 **LF AI**

# Antitrust Policy Notice

- › 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 Useful Links

Web site: [lfai.foundation](https://lfai.foundation)  
Wiki: [wiki.lfai.foundation](https://wiki.lfai.foundation)  
GitHub: [github.com/lfai](https://github.com/lfai)  
Landscape: [landscape.lfai.foundation](https://landscape.lfai.foundation) or [l.lfai.foundation](https://l.lfai.foundation)  
Mail Lists: <https://lists.lfai.foundation>

LF AI Logos: <https://github.com/lfai/artwork/tree/master/lfai>

LF AI Presentation Template:

[https://drive.google.com/file/d/1eiDNJvXCqSZHT4Zk\\_-czASlz2GTBRZk2/view?usp=sharing](https://drive.google.com/file/d/1eiDNJvXCqSZHT4Zk_-czASlz2GTBRZk2/view?usp=sharing)

Events Page on LF AI Website: <https://lfai.foundation/events/>

Events Calendar on LF AI Wiki (subscribe available):

<https://wiki.lfai.foundation/pages/viewpage.action?pageId=12091544>

Event Wiki Pages: <https://wiki.lfai.foundation/display/DL/LF+AI+Foundation+Events>

# Agenda

- › Roll Call (3 minutes)
- › Approval of Minutes (3 minutes)
- › Invited Presentation: Open Lineage - Datakin (15 minutes)
- › Invited Presentation: IDA - Salesforce, IBM (20 minutes)
- › Upcoming TAC Meeting (4 minutes)
- › LF AI General Updates (5 minutes)
- › Open Discussion (15 minutes)
- ›

# TAC Voting Members - additions coming Nov 5th

Board Member	Contact Person	Email
AT&T	Anwar Atfab	<a href="mailto:anwar@research.att.com">anwar@research.att.com</a>
Baidu	Daxiang Dong	<a href="mailto:dongdaxiang@baidu.com">dongdaxiang@baidu.com</a>
Ericsson	Rani Yadav-Ranjan	<a href="mailto:rani.yadav-ranjan@ericsson.com">rani.yadav-ranjan@ericsson.com</a>
Huawei	Huang Zhipeng	<a href="mailto:huangzhipeng@huawei.com">huangzhipeng@huawei.com</a>
IBM	Susan Malaika	<a href="mailto:malaika@us.ibm.com">malaika@us.ibm.com</a>
Nokia	Jonne Soininen	<a href="mailto:jonne.soininen@nokia.com">jonne.soininen@nokia.com</a>
Tech Mahindra	Nikunj Nirmal	<a href="mailto:nn006444@techmahindra.com">nn006444@techmahindra.com</a>
Tencent	Bruce Tao	<a href="mailto:brucetao@tencent.com">brucetao@tencent.com</a>
Zilliz	Jun Gu	<a href="mailto:jun.gu@zilliz.com">jun.gu@zilliz.com</a>
ZTE	Wei Meng	<a href="mailto:meng.wei2@zte.com.cn">meng.wei2@zte.com.cn</a>
Graduate Project	Contact Person	Email
Acumos	Nat Subramanian	<a href="mailto:natarajan.subramanian@techmahindra.com">natarajan.subramanian@techmahindra.com</a>
Angel	Bruce Tao	<a href="mailto:brucetao@tencent.com">brucetao@tencent.com</a>
Horovod	Travis Addair	<a href="mailto:taddair@uber.com">taddair@uber.com</a>
ONNX	Jim Spohrer (Chair of TAC)	<a href="mailto:spohrer@us.ibm.com">spohrer@us.ibm.com</a>

# Approval of Sept 24th TAC Minutes

Draft minutes from the September 24<sup>th</sup> TAC call were previously distributed to the TAC members via the mailing list

## **Proposed Resolution:**

- › That the minutes of the September 24<sup>th</sup> meeting of the Technical Advisory Council of the LF AI Foundation are hereby approved

# Invited Presentation - Open Lineage (Datakin)



# Open Lineage (Datakin)

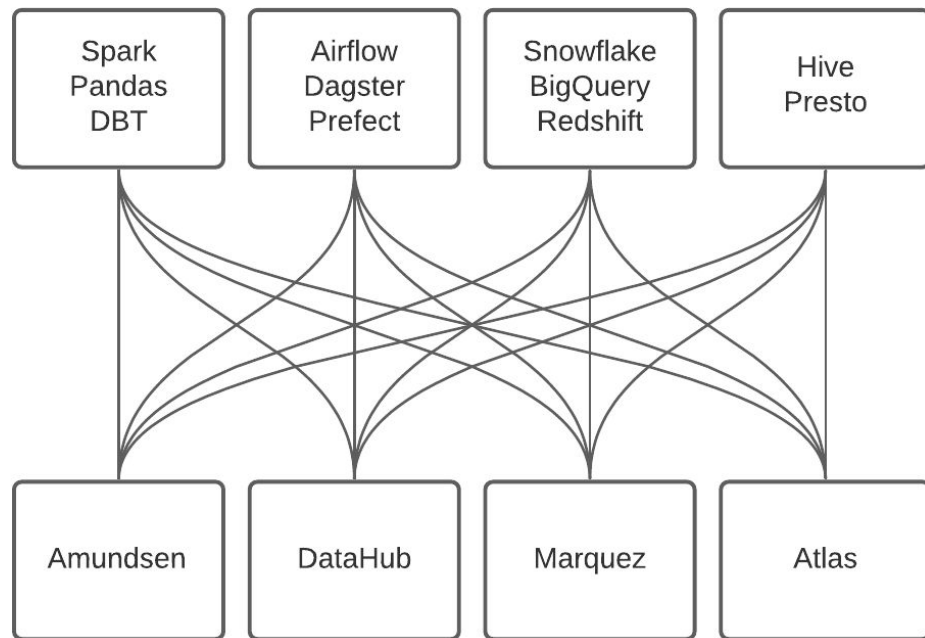
Open Lineage is a proposed standard to simplify data and AI projects integration (new version breakage) and instrumentation (jobs/dataset/runs entities) in end-to-end workflows.

## **Presenter:**

- › Julien Le Dem (Datakin)

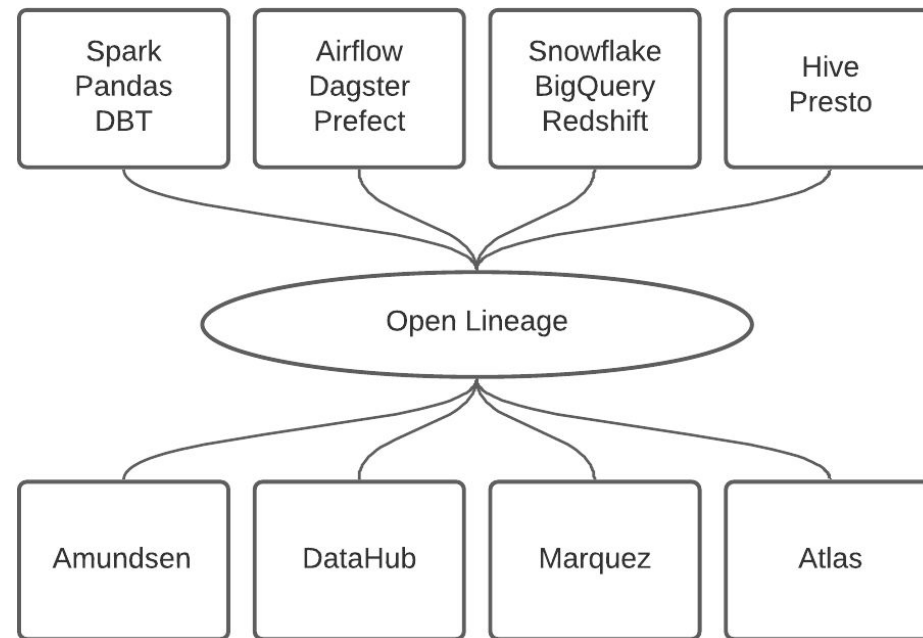
# Problem

Today:



- Duplication of effort: Each project has to instrument all jobs
- Integrations are external and can break with new versions

With an open lineage standard:



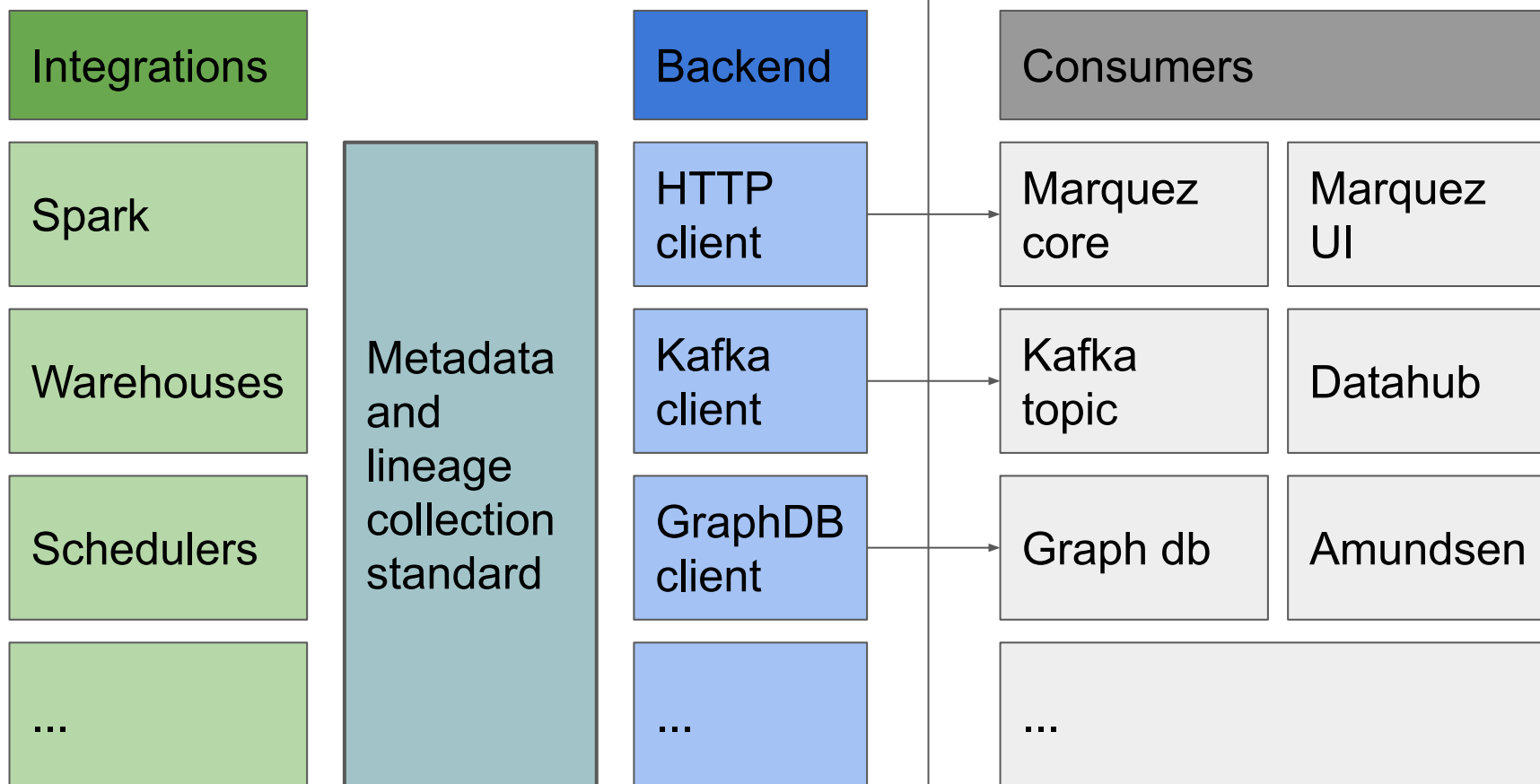
- Effort of integration is shared
- Integration can be pushed in each project: no need to play catch up

# Purpose

- Open standard for metadata and lineage collection
- Instrument jobs as they are running
- Define a generic model of job/dataset/runs entities
- Consistent naming strategies for jobs and datasets
- Define specific facets that can enrich those entities

# Open Lineage scope

# Not in scope



# Plan

0) Agree on coordination / project management

Then:

- 1) Agree on protocols
- 2) Agree on core model
- 3) Agree on lifecycle
- 4) Agree on individual facets

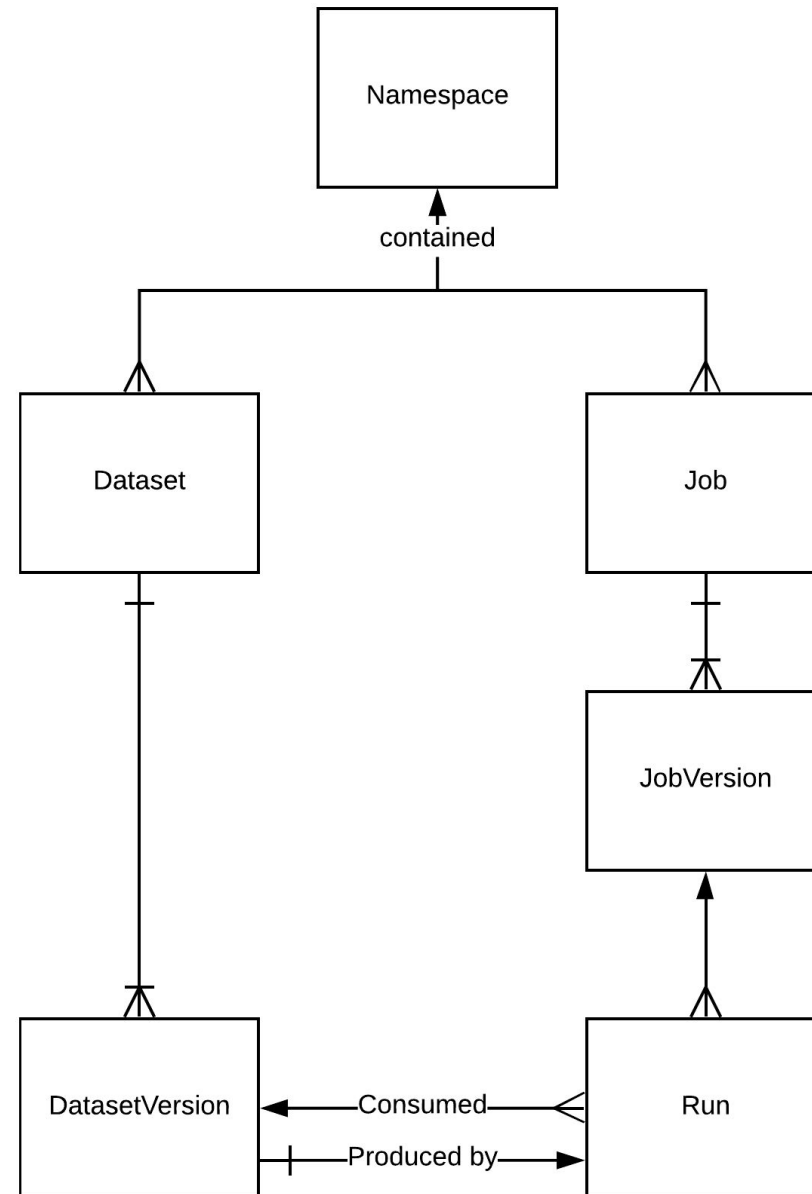
# 0) Consensus building / Project management

- Use git to document the spec:
  - Model
  - Lifecycle
  - Facet schema
- Github issues for individual discussion points
- PRs to evolve the spec.

# 1) Protocol

- Asynchronous events
  - UUID for identifying a run and correlate events
- Configurable backend
  - Kafka
  - Http
  - ...

## 2) Core Model





## 2) Core Model

Consistent naming:

- Jobs:

Example: scheduler.job.task

- Datasets:

Example: instance.schema.table

# 3) Lifecycle

- Create run id
  - Send current version of job and inputs
- Run start
  - Send plan/profile info
- Run end
  - Send output Dataset version updates

## 4) Facets

### **Dataset:**

- Stats
- Schema
- Version
- Column level lineage
- ...

### **Job:**

- Source code
- Dependencies
- params
- Source control
- Query plan
- Query profile

### **Run:**

- Logical time
- Batch id

# Invited Presentation - IDA (Salesforce, IBM)

# IDA (Salesforce, IBM)

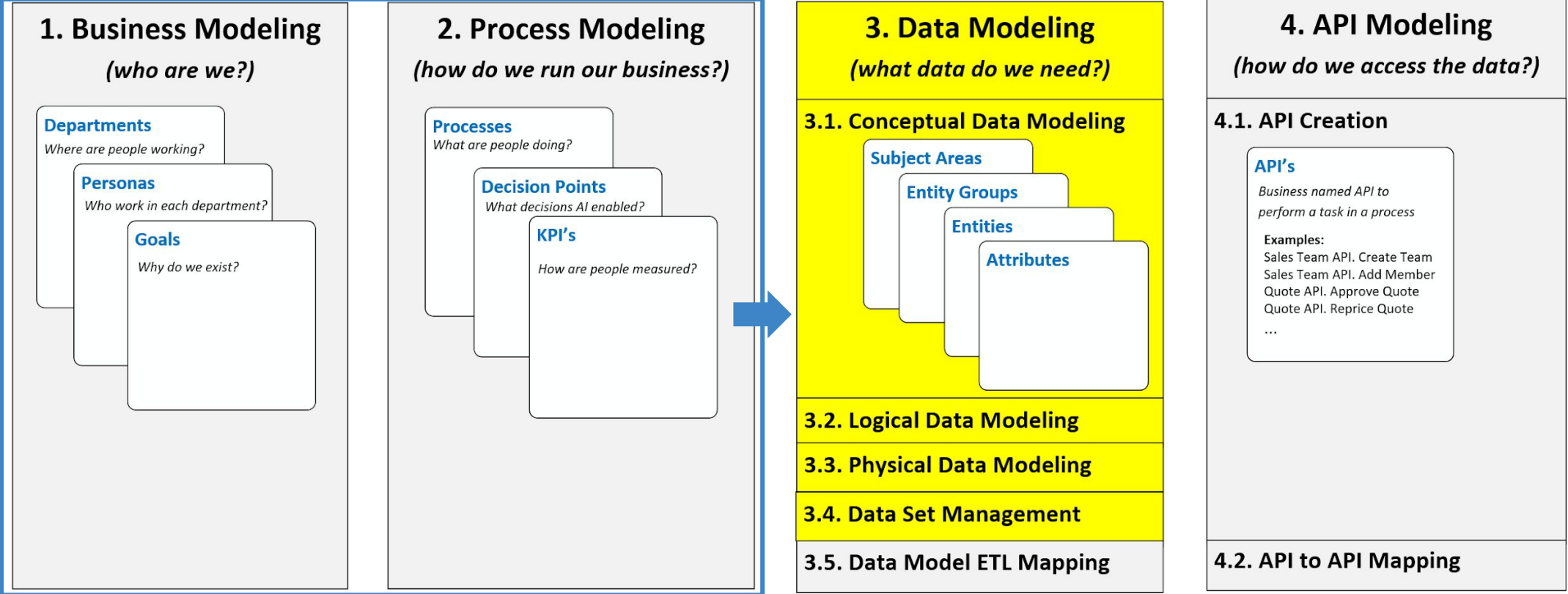
Vision: Overall vision of filling a major gap in the industry for creating, sharing data model and related functions: data governance, predictive analytics and other AI.

## Presenters:

- › Param Singh, Senior Director Product Management Salesforce.
- › Maryia Rakina, Product Manager, IDA, Db2 UI & UX, IBM Data & AI.
- › Nathan Horne, Architect Salesforce Platform.
- › Lars Martinssen, Chief Data Modeler of Salesforce

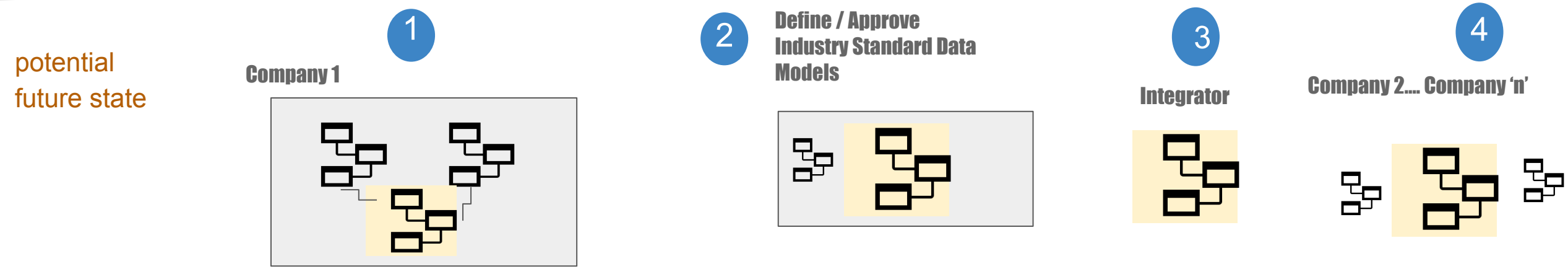
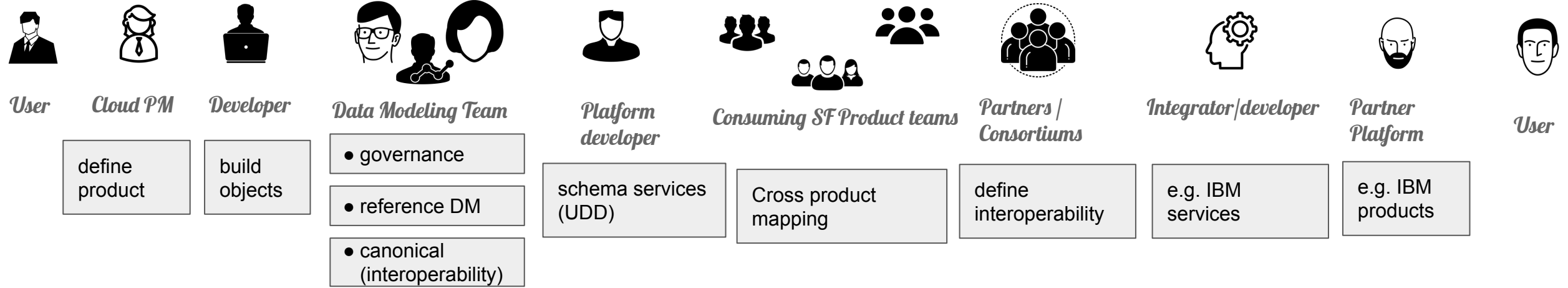
# Information Management relative to Data Modeling & platform

We focus on IM discipline #3 - Data Modeling



Salesforce platform (ORM, DB)

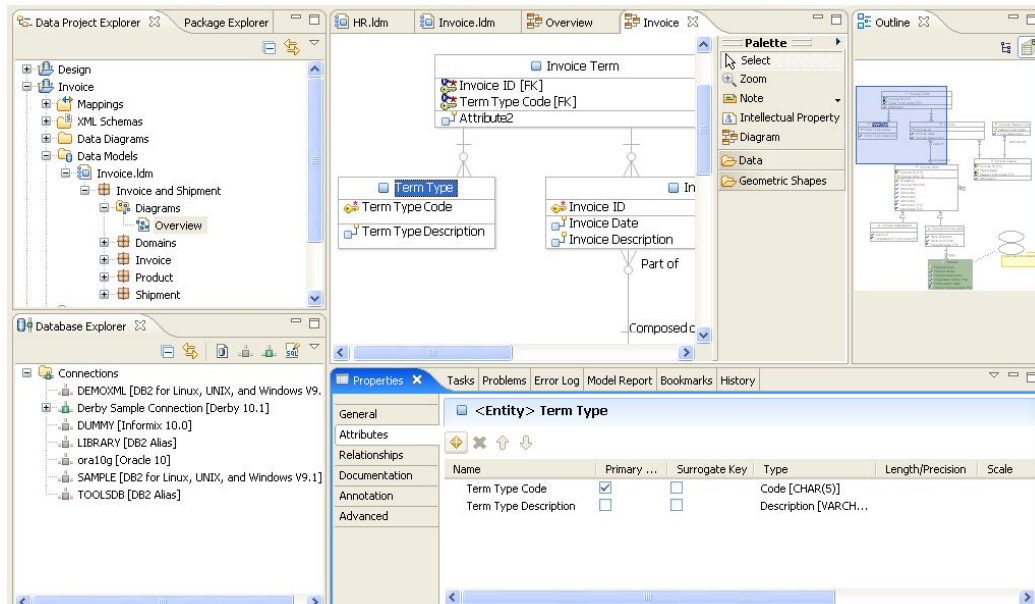
# Data Modeling Toolchain



# InfoSphere Data Architect – Data Modeling

***A collaborative, data design solution to discover, model, relate, and standardize diverse data assets.***

- Design and manage enterprise data models
- Enforce model conformance to enterprise standards
- Leverage industry data models for best practices
- Optimize existing investments



- Data Models
  - Logical
  - Physical
  - Dimensional
  - Glossary
  - Mapping
- Heterogeneous Database support
- Reverse Engineering
- Visualize Design
- Model Validation
- Naming Standards
- Team collaboration



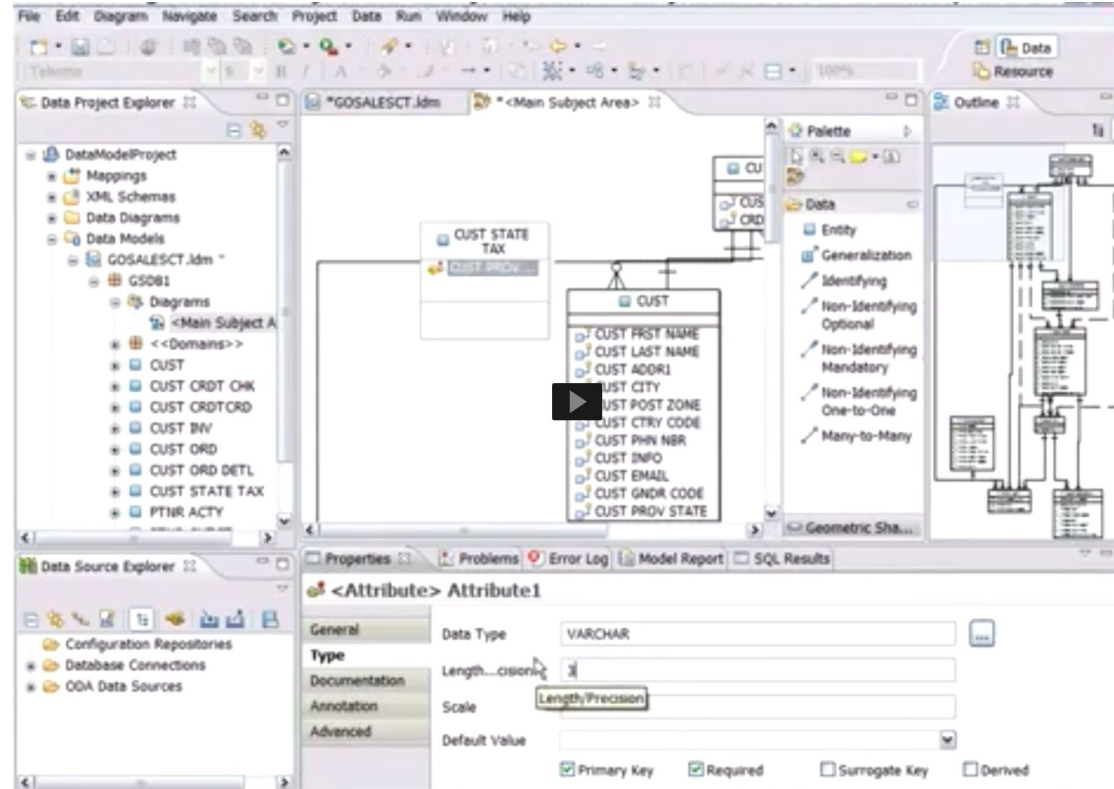
# Areas for Community Collaboration - open source IDA

Logical Model -  
meta model, export  
JSON

Example: Salesforce, IDA  
customers.

Physical Models -  
support for DBs,  
Cloud

Example: DB vendors, Cloud  
DB vendors



Plugins- Import,  
Export

Example: Tooling vendor  
(UML import, DB, AI)

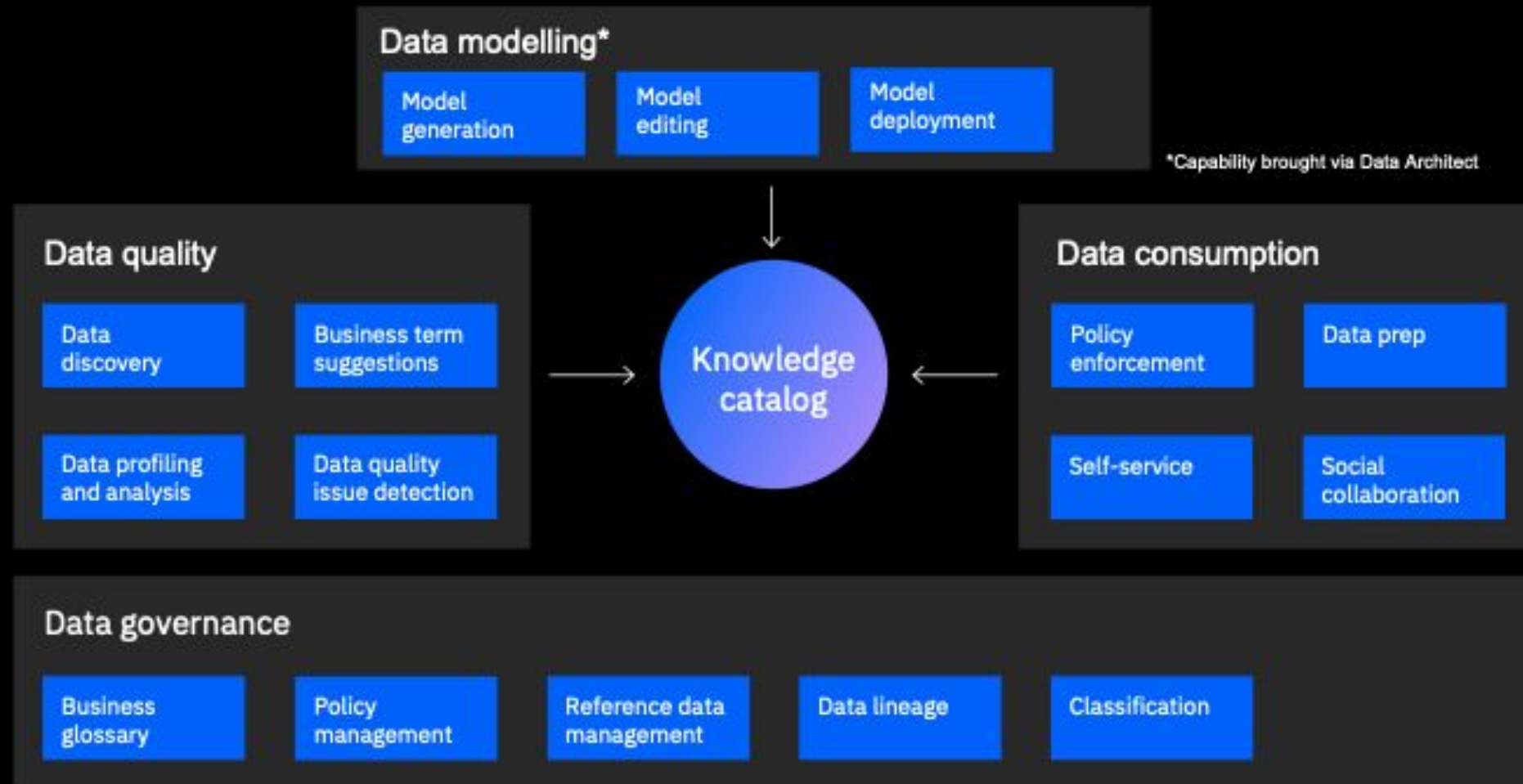
Multi-user,  
Collaboration

Example: multi-user  
collaboration in the cloud,  
internally and externally

Core Capabilities - Diagramming/Layout, Meta model

# IBM End to End, Value chain diagram

## All capabilities in a single experience





## Support for the Initiative

1. Current IDA customers
2. Salesforce and IBM
3. Linux Foundation members

# LF AI General Updates

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

The LF AI landscape explores open source projects in the domains of artificial intelligence, machine learning, and deep learning.
   
  
[l.fai.foundation](https://l.fai.foundation)

Security & Privacy	Security & Privacy

Natural Language Processing	Natural Language Processing

Education	Education




















# Projects

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

## Graduated LFAI Projects (5)

 <b>Acumos</b> ★ 11 LF Artificial Intelligence Foundation	 <b>Angel-ML</b> ★ 6,003 LF Artificial Intelligence Foundation	 <b>Egeria</b> ★ 254 LF Artificial Intelligence Foundation	 <b>Horovod</b> ★ 10,182 LF Artificial Intelligence Foundation	 <b>ONNX</b> ★ 9,155 LF Artificial Intelligence Foundation
---	---	---	--	---

## Incubating LFAI Projects (17)

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

# 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

# Looking to host a project with LF AI

Hosted project stages and life cycle:

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

Offered services for hosted projects:

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

Contact:

Jim Spohrer (TAC Chair) and Ibrahim Haddad (ED, LF AI)



# Promoting Upcoming Project Releases

We promote project releases via a blog post and on LF AI [Twitter](#) and/or [LinkedIn](#) social channels

For links to details on upcoming releases for LF AI hosted projects visit the [Technical Project Releases wiki](#)

If you are an LF AI hosted project and would like LF AI to promote your release, reach out to [pr@lfai.foundation](mailto:pr@lfai.foundation) to coordinate in advance (min 2 wks) of your expected release date.

## Note on quorum

As LF AI is growing, we now have 14 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 the Outreach Committee

# Events

- › Upcoming Events
  - › Visit the [LF AI Events Calendar](#) or the [LF AI 2020 Events wiki](#) for a list of all events
  - › To participate visit the [LF AI 2020 Events wiki page](#) or email [info@lfai.foundation](mailto:info@lfai.foundation)
- › Please consider holding virtual events
  - › To discuss participation, please email [events@lfai.foundation](mailto:events@lfai.foundation)

# Upcoming Events

## **LF AI Booth at ONES – Sept 28-30**

September 28 - September 30

LF AI Foundation will have a booth at Open Networking & Edge

## **LF AI Booth at OSS EU – Oct 26-28**

October 26 - October 28

LF AI Foundation will have a booth at Open Source Summit Europe (OSS EU)

## **“AI/ML/DL presented by LF AI Foundation” Track at OSS EU – Oct 26-28**

October 26 - October 28

"AI/ML/DL presented by LF AI Foundation" Track at Open Source Summit Europe (OSS EU)

## **LF AI Mini Summit at OSS EU – Virtual – Oct 29**

October 29 @ 12:00 am

LF AI Foundation will hold a Mini Summit at Open Source Summit EU (OSS EU)

# LF AI PR/Comms

- › Please follow LF AI 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 Blog](#)
- › Open call to publish project/committee updates or other relevant content on the [LF AI Blog](#)
- › To discuss more details on participation or upcoming announcements, please email [pr@lfai.foundation](mailto:pr@lfai.foundation)

# Call to Participate in Ongoing Efforts

# Trusted AI

- › **Leadership:**  
Animesh Singh (IBM), Souad Ouali (Orange), and Jeff Cao (Tencent)
- › **Goal:** Create policies, guidelines, tooling and use cases by industry
- › **Github:**  
<https://github.com/lfai/trusted-ai>
- › **Wiki:**  
<https://wiki.lfai.foundation/display/DL/Trusted+AI+Committee>
- › **To participate:**  
<https://lists.lfai.foundation/g/trustedai-committee/>
- › **Next call:** Bi-weekly on Thursdays at 7am PT, subscribe to group calendar on wiki  
<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
- › **Wiki:**  
<https://wiki.lfai.foundation/display/DL/ML+Workflow+Committee>
- › **To participate:**  
<https://lists.lfai.foundation/g/mlworkflow-committee>
- › **Next call:** Every 4 weeks on Thursdays at 7:00 am PT, subscribe to group calendar on wiki  
<https://wiki.lfai.foundation/pages/viewpage.action?pageId=18481242>



# Launching an 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.lfai.foundation/g/aiethics-training>

# Upcoming TAC Meetings

# Upcoming TAC Meetings

- › **November 5:** RosaeNLG
- › **November 19 :** TBD

Please send agenda topic requests to [tac-general@lists.lfai.foundation](mailto:tac-general@lists.lfai.foundation)

# TAC Meeting Details

- › To subscribe to the TAC Group Calendar, visit the wiki: <https://wiki.lfai.foundation/x/XQB2>
- › 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

# Legal Notices

- › 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](mailto:legal@linuxfoundation.org) with any questions about The Linux Foundation's policies or the notices set forth on this slide.