

# Meeting of the Technical Advisory Council (TAC)

January 28, 2021

 **DLF** AI & DATA

# Anti-Trust Policy

- › Linux Foundation meetings involve participation by industry competitors, and it is the intention of the Linux Foundation to conduct all of its activities in accordance with applicable antitrust and competition laws. It is therefore extremely important that attendees adhere to meeting agendas, and be aware of, and not participate in, any activities that are prohibited under applicable US state, federal or foreign antitrust and competition laws.
- › Examples of types of actions that are prohibited at Linux Foundation meetings and in connection with Linux Foundation activities are described in the Linux Foundation Antitrust Policy available at <http://www.linuxfoundation.org/antitrust-policy>. If you have questions about these matters, please contact your company counsel, or if you are a member of the Linux Foundation, feel free to contact Andrew Updegrave of the firm of Gesmer Undergone LLP, which provides legal counsel to the Linux Foundation.

# Recording of Calls

## **Reminder:**

TAC calls are recorded and available for viewing on the [TAC Wiki](#)

# Reminder: LF AI & Data Useful Links

- › Web site: [lfaidata.foundation](https://lfaidata.foundation)
- › Wiki: [wiki.lfaidata.foundation](https://wiki.lfaidata.foundation)
- › GitHub: [github.com/lfaidata](https://github.com/lfaidata)
- › Landscape: <https://landscape.lfaidata.foundation> or <https://l.lfaidata.foundation>
- › Mail Lists: <https://lists.lfaidata.foundation>
- › Slack: <https://slack.lfaidata.foundation>
- ›
- › LF AI Logos: <https://github.com/lfaidata/artwork/tree/master/lfaidata>
- › LF AI Presentation Template:  
[https://drive.google.com/file/d/1eiDNJvXCqSZHT4Zk\\_-czASlz2GTBRZk2/view?usp=sharing](https://drive.google.com/file/d/1eiDNJvXCqSZHT4Zk_-czASlz2GTBRZk2/view?usp=sharing)
- ›
- › Events Page on LF AI Website: <https://lfaidata.foundation/events/>
- › Events Calendar on LF AI Wiki (subscribe available):  
<https://wiki.lfaidata.foundation/pages/viewpage.action?pageId=12091544>
- › Event Wiki Pages: <https://wiki.lfaidata.foundation/display/DL/LF+AI+Data+Foundation+Events>

# Agenda

- › Roll Call (2 mins)
- › Approval of Minutes (2 mins)
- › Welcome new Members (2 minutes)
  - › Associate Member: Galgotias University (India)
  - › General Member:VMWare
- › Discussion and Vote (15 minutes)
  - › Project Stages (Ibrahim Haddad)
- › Invited Presentation (20 minutes)
  - › Sedna - KubeEdge SIG AI (Howard Huang (Huawei), Kevin Wang)
- › Invited Presentation (15 minutes)
  - › Cloud Information Model (Justin Digrazia (Salesforce), Joaquin Prado (LF))
- › LF AI General Updates (2 minutes)
- › Open Discussion (2 minutes)

# TAC Voting Members

\* = still need backup specified on [wiki](#)

Board Member	Contact Person	Email
AT&T	Anwar Atfab	<a href="mailto:anwar@research.att.com">anwar@research.att.com</a>
Baidu	Ti Zhou	<a href="mailto:zhouti@baidu.com">zhouti@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>
SAS	Nancy Rausch	<a href="mailto:nancy.rausch@sas.com">nancy.rausch@sas.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>
Egeria	Mandy Chessell	<a href="mailto:mandy_chessell@uk.ibm.com">mandy_chessell@uk.ibm.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 January 14th, 2021 Minutes

Draft minutes from the January 14<sup>th</sup> TAC call were previously distributed to the TAC members via the mailing list

## **Proposed Resolution:**

- › That the minutes of the January 14<sup>th</sup> meeting of the Technical Advisory Council of the LF AI & Data Foundation are hereby approved.

# Welcome new associate member!

## Galgotias University

LF AI & Data Member Company - Associate





# Welcome new general member!

vmware

LF AI & Data Member Company - General



# Discussion and Vote

## New Project Stages Document

Presenter (January 14th TAC)

- Ibrahim Haddad

# LF AI & Data Project Lifecycle Document

Ibrahim Haddad, Ph.D.  
Executive Director, LF AI & Data  
[Ibrahim@LinuxFoundation.org](mailto:Ibrahim@LinuxFoundation.org)

 LF AI & DATA

# Background

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

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

Current version dates May 2018.

# Revisiting the Document

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

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

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

# Key Updates Introduced to the Document

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

# 1. Introducing Sandbox Stage

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

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



## 2. Improving requirements to incubate projects

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

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



# 3. Improving requirements to graduate projects

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

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

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

# 5. Elaborating on the Archive Stage projects to eliminate ambiguities

To archive a project:

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

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

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

Reactivating an Archived Project

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

# 6. Adding information on the Annual Review of projects

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

# Existing Projects

No impact.

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

# Request for Feedback

Call for comment is open until Friday Jan 22nd.

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

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

# Joining LF AI & Data is Easy!



General Inquiries:

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

 LF AI & DATA

# Approval of New Project Stages Document

Draft project stages document were previously distributed to the TAC members via the mailing list as well as presented to TAC on Jan. 14th.

## **Proposed Resolution:**

- › TAC approves the new project stages documents..



# Next Steps

LF AI & Data Governing Board will vote to approve.

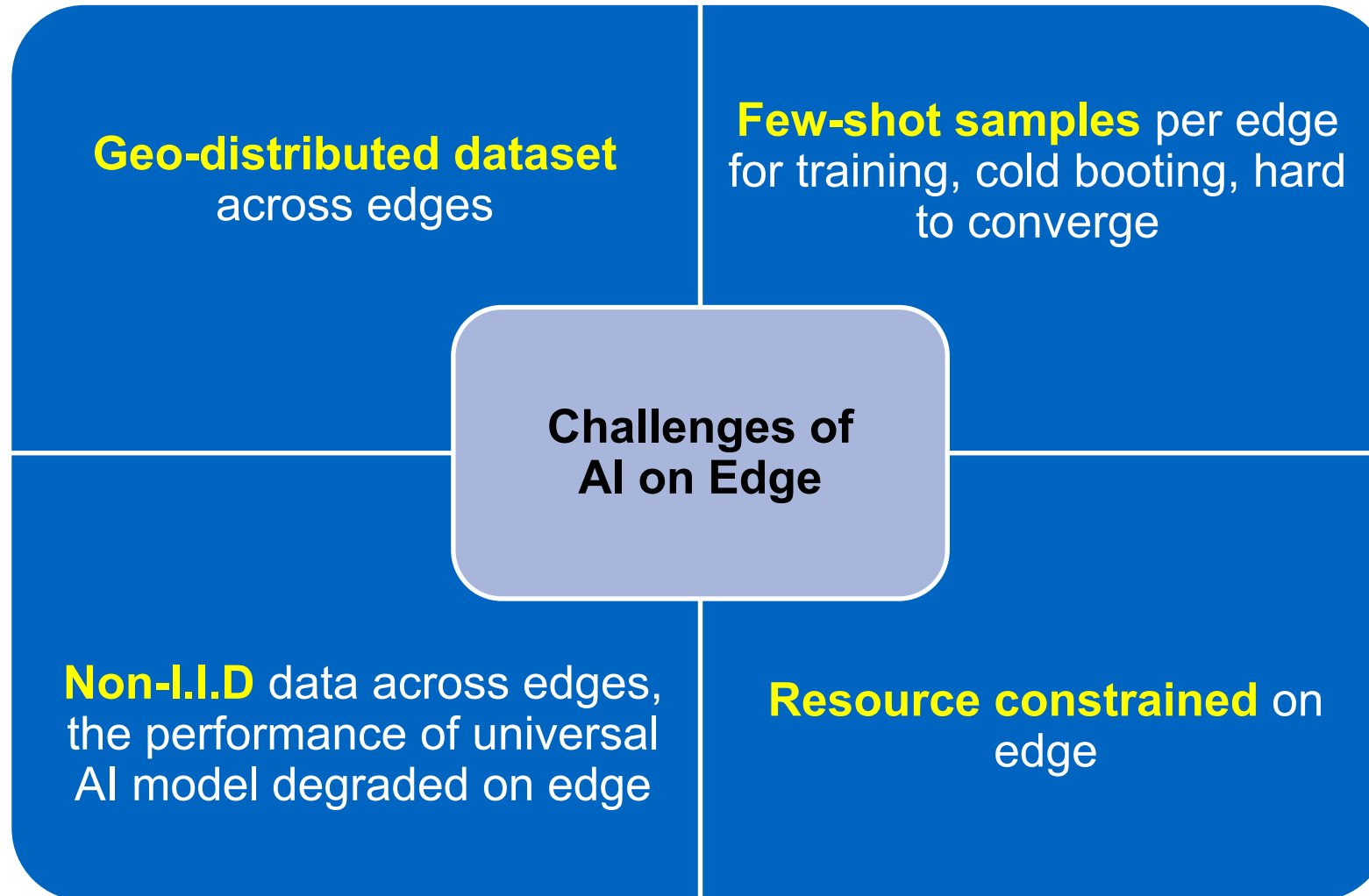
# Invited Presentation - Sedna - KubeEdge SIG AI

Howard Huang (Huawei)  
Kevin Wang

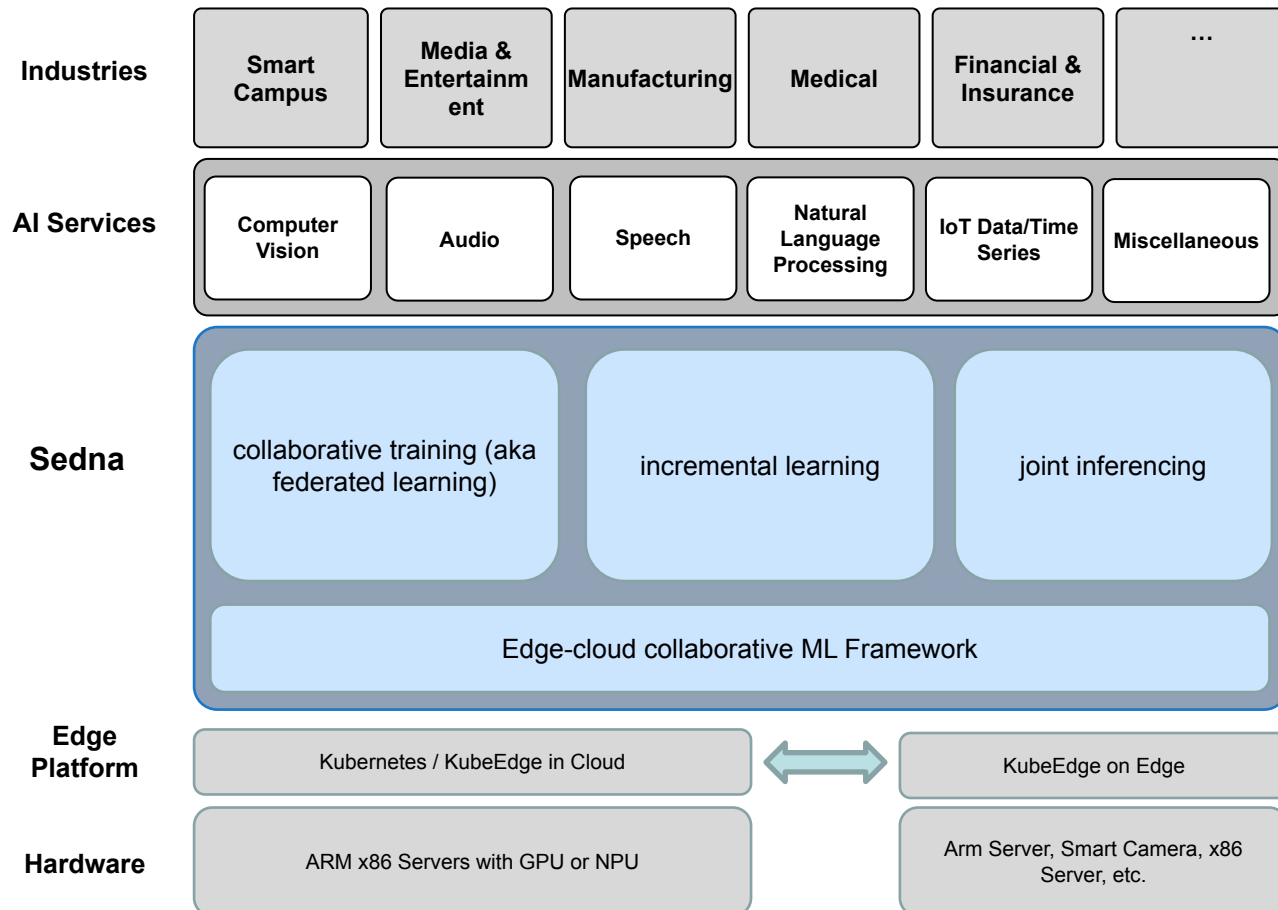
# Sedna Intro

KubeEdge SIG AI

# Challenges of AI on Edge



# Sedna Project

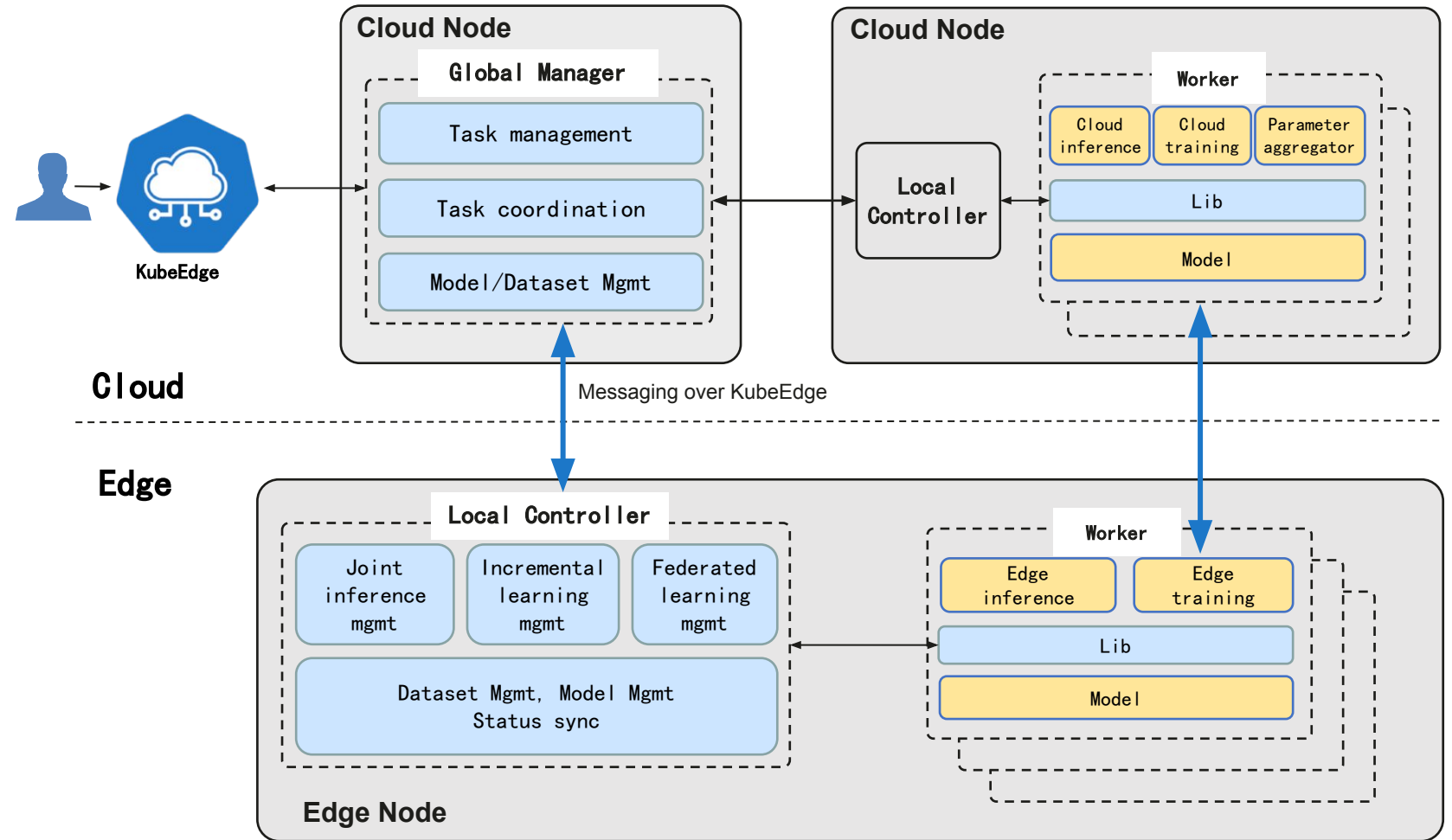


- What to propose:
  1. an edge-cloud collaborative ML framework based on KubeEdge
  2. with embed collaborative training and joint inferencing algorithm, which can
  3. work with existing AI frameworks e.g., TensorFlow, PyTorch, MindSpore, etc.
- 3 Features:
  1. joint inferencing
  2. incremental learning
  3. collaborative training (aka federated learning)
- Targeting Users:
  1. Domain-specific AI Developers: build and publish edge-cloud collaborative AI services/functions easily
  2. Application Developers: use edge-cloud collaborative AI capabilities.
- *We are NOT:*
  1. to re-invent existing ML framework, i.e., TensorFlow, PyTorch, MindSpore, etc.
  2. to re-invent existing edge platform, i.e., KubeEdge, etc.
  3. to offer domain/application-specific algorithms, e.g., facial recognition, text classification, etc.



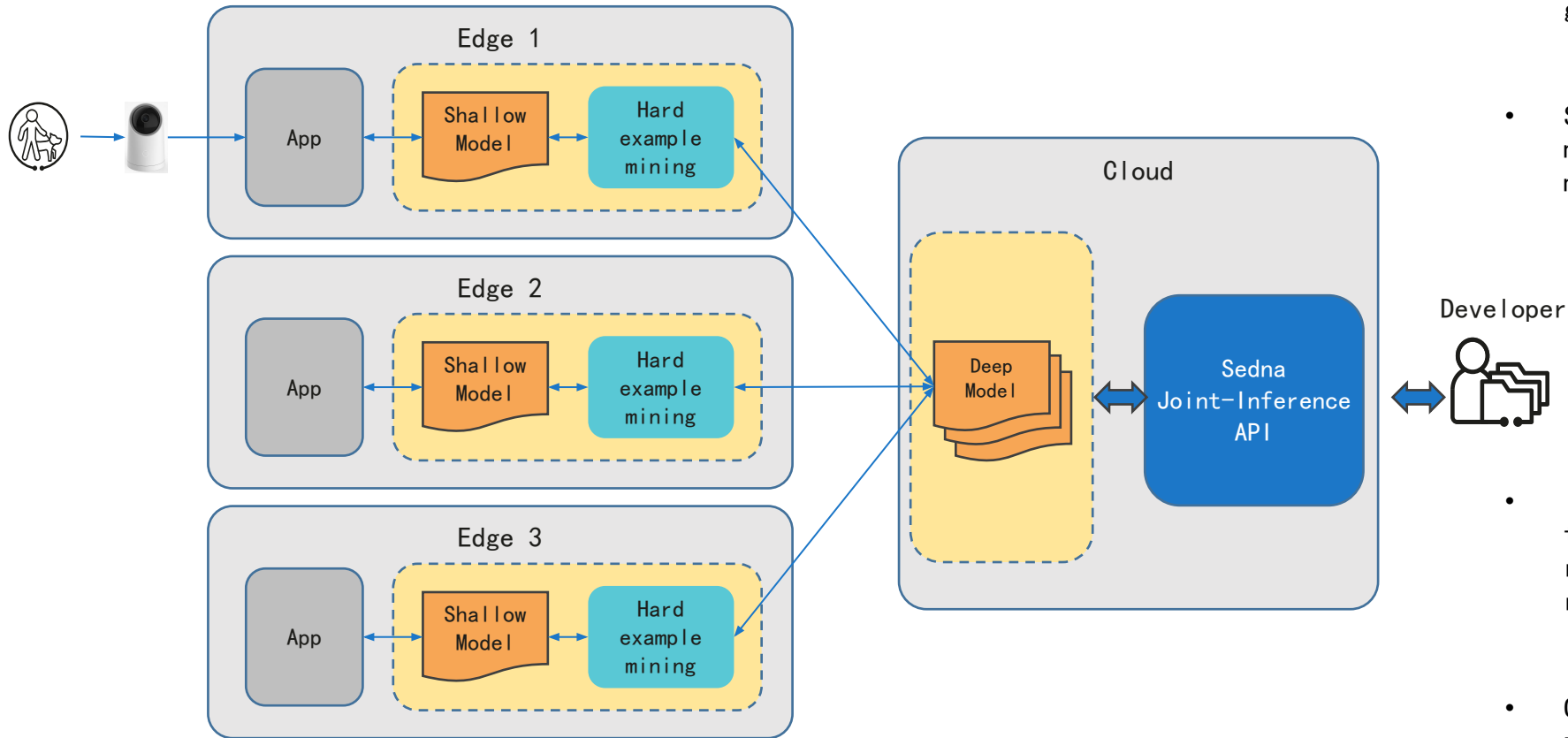
# Sedna Architecture

- **Workers:**
  - Inferencing or training, based on existing ML framework;
  - Launch on demand, as containers or functions;
  - Different workers for different features;
  - Could run either on edge or cloud.
- **Lib:**
  - Expose the Edge AI features to applications, i.e. training or inference programs.
- **GlobalManager**
  - Unified edge-cloud synergy AI task management
  - Cross edge-cloud synergy management and collaboration
  - Central Configuration Management
- **LocalController**
  - Local process control of edge-cloud synergy AI tasks
  - Local general management: model, dataset, and status synchronization



# Edge-cloud Collaborative JOINT INFERENCE

Improve the inference performance, when edge resources are limited.



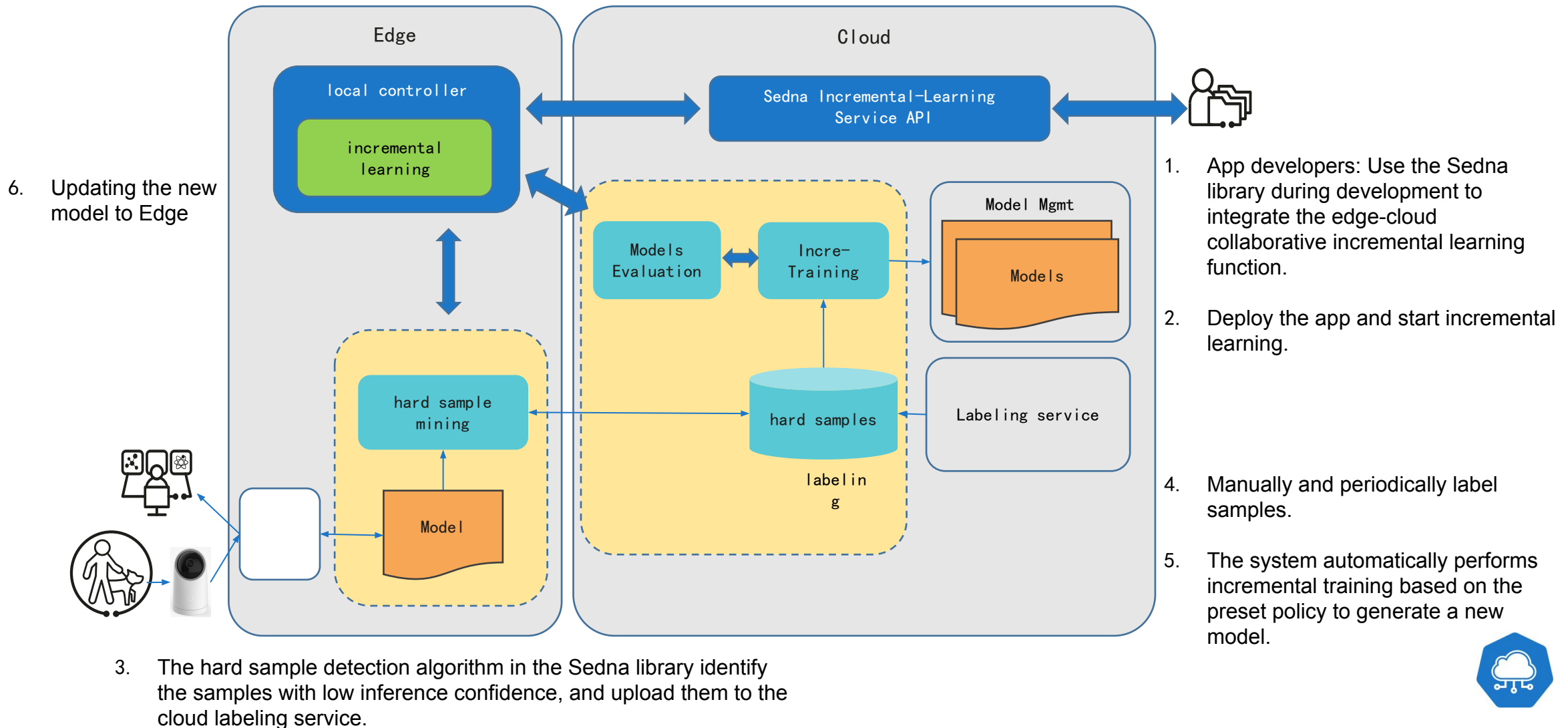
- AI developer: provides training data to generate deep and shallow models.
- Service developers: invoke collaboration models through the library and deploy the models to the edge.

- Inference based on the shallow model on the edge side. If the confidence requirement is met, the result is returned.
- Otherwise, the data is sent to the cloud for deep model inference.



# Edge-cloud Collaborative INCREMENTAL LEARNING

The more models are used, the smarter they are.

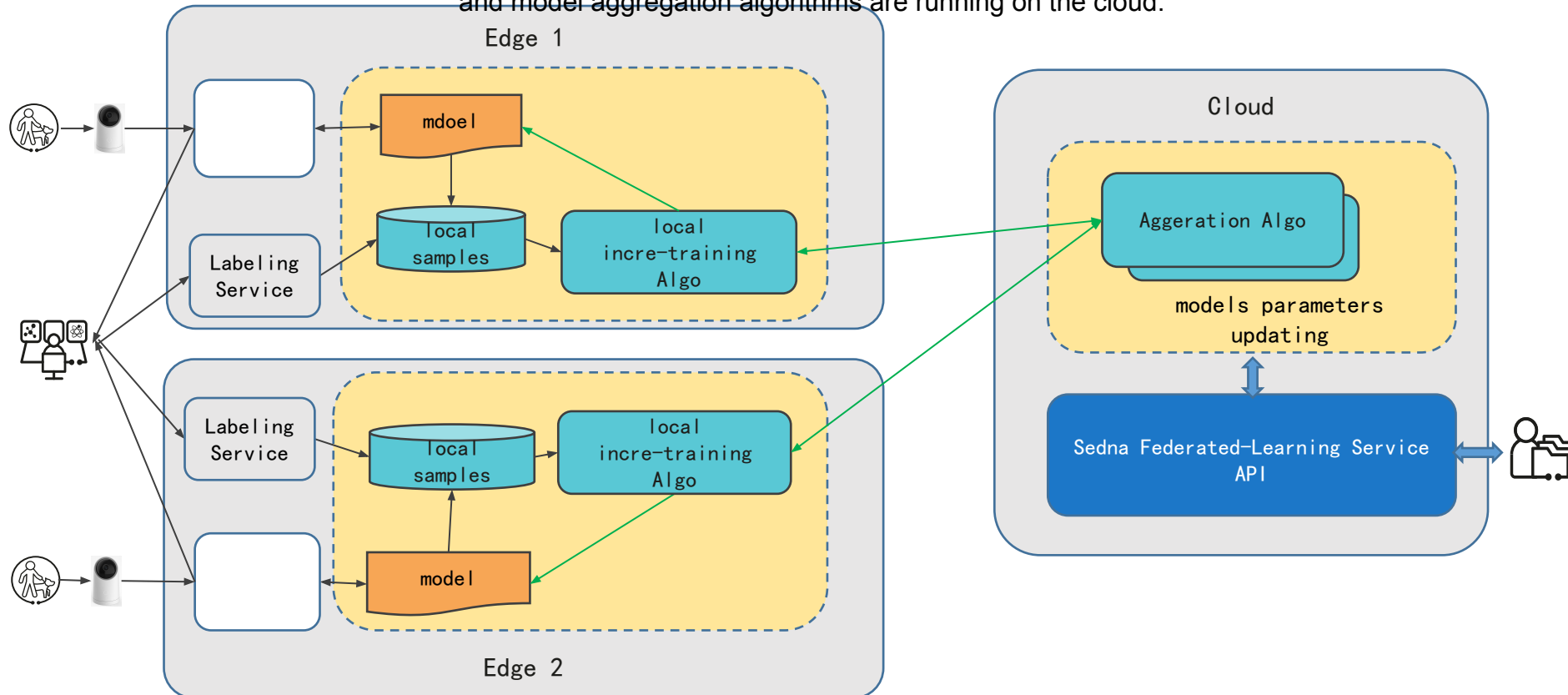




# Edge-cloud collaborative FEDERATED LEARNING

Raw data is not transmitted out of the edge, and the model is generated by knowledge aggregation.

3. Multi-task detection: Divide non-IID sample sets and work with the cloud to identify similar tasks.
4. Local training: Model parameters are uploaded to the cloud, and the cross-edge transferring and model aggregation algorithms are running on the cloud.



1. Developer: Import the Sedna library and develop the edge-cloud collaborative federated learning program.
2. Start the federated learning task and deploy the training program to the edge.



# JOINT INFERENCE

## Developer perspective

- Design Objectives: Try not to change the existing code of developers and do not require developers to learn new frameworks, reducing learning costs.
- How To Use:
- Importing the Sedna library: Developers use the familiar ML framework (such as TensorFlow) to import the Sedna library (solar\_corona library in the figure).
- JOINT INFERENCE: Replace the original load model object part, configure and generate the edge-cloud synergy model, and the background automatically generates a large model on the cloud. Developers do not need to change other parts of the code.

## JOINT INFERENCE code example (Based on TensorFlow)

```
import hilens
import solar_corona

def pre_fun():
    """
    Preprocessing function may be, for example, image rotation or resize.
    """
    return

def post_fun():
    """
    Post-processing function may be, for example, a post-processing module NMS in the object detection framework.
    """
    return

def run():
    # endpoint may be: 1. Services started by users' own models 2. Existing cloud services
    big_model_endpoint = solar_corona.joint_inference.get_big_model_endpoint() # deep model on cloud
    ibt = solar_corona.joint_inference.IBT(upload_ratio=0.5) # "transfer to cloud" algorithm
    model_path = solar_corona.context.get_model_path()

    # Configure the edge-cloud model, including:
    # the local path of the model,
    # local pre-processing,
    # post-processing methods,
    # cloud migration algorithm,
    # cloud migration endpoint.
    model = solar_corona.load_model(model_path, pre_fun, post_fun,
                                   cloud_offload_algorithm=ibt,
                                   big_model_endpoint=big_model_endpoint)

    # Service parameters settings.
    camera_address = solar_corona.context.get_parameters('ip_camera_address')
    camera = hilens.VideoCapture(camera_address)

    # Service related code.
    while True:
        image = read_one_frame_from_camera(camera)
        predictions = model.predict(image)

if __name__ == "__main__":
    run()
```

# FEDERATED LEARNING

## Developer perspective

- Design Objectives: Try not to change the existing code of developers and do not require developers to learn new frameworks, reducing learning costs.
- How To Use:
- Importing the Sedna library: Developers use the familiar ML framework (such as TensorFlow) to import the Sedna library (solar\_corona library in the figure).
- FEDERATED LEARNING: Import the local training loss function, optimizer, and the collaborative\_train function from the solar\_corona library.

FEDERATED LEARNING code example  
(Based on TensorFlow)

```
def main():
    tf.random.set_seed(22)

    # load dataset.
    (x, y) = solar_corona.load_train_dataset()
    (x_test, y_test) = solar_corona.load_eval_dataset()
    x, x_test = normalize(x, x_test)

    # read parameters from deployment config.
    epochs = solar_corona.context.get_parameters('epochs')
    batch_size = solar_corona.context.get_parameters('batch_size')
    aggregation_algorithm = solar_corona.context.get_parameters('aggregation_algorithm')

    train_loader = tf.data.Dataset.from_tensor_slices((x, y))
    train_loader = train_loader.map(prepare_cifar).shuffle(50000).batch(batch_size)
    test_loader = tf.data.Dataset.from_tensor_slices((x_test, y_test))
    test_loader = test_loader.map(prepare_cifar).shuffle(10000).batch(batch_size)
    model = VGG16([32, 32, 3])

    # you can use the loss/metric function defined in keras
    # loss = keras.losses.CategoricalCrossentropy(from_logits=True)
    # metric = keras.metrics.CategoricalAccuracy()
    # also you can use the loss/metric function defined in solar_corona
    loss = solar_corona.losses.ADifferentCategoricalCrossentropy()
    metric = solar_corona.metrics.ADifferentCategoricalAccuracy()
    # again, you can use either.
    # optimizer = keras.optimizers.Adam(learning_rate=0.0001)
    optimizer = solar_corona.optimizers.ADifferentAdam(learning_rate=0.0001)

    model = solar_corona.collaborative_training.fit(train_loader=train_loader,
                                                    test_loader=test_loader,
                                                    model=model,
                                                    loss=loss,
                                                    metric=metric,
                                                    optimizer=optimizer,
                                                    batch_size=batch_size,
                                                    epochs=epochs,
                                                    # config the aggregation algorithm
                                                    aggregation_algorithm=aggregation_algorithm)

    # Save the model based on the config.
    solar_corona.save_model(model)

if __name__ == '__main__':
    main()
```

Thank you

# Invited Presentation - Cloud Information Model

Justin Digrazia (Salesforce)  
Joaquin Prado (Linux Foundation)



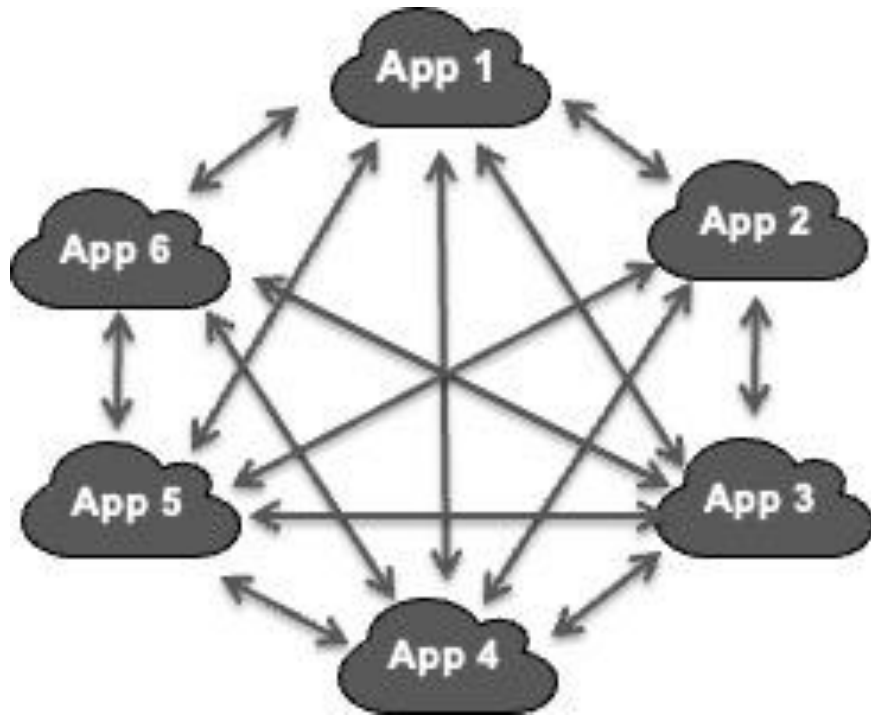
# Cloud Information Model

Facilitating Cloud Integration



Customers expect their environments to “just work” together out of the box...

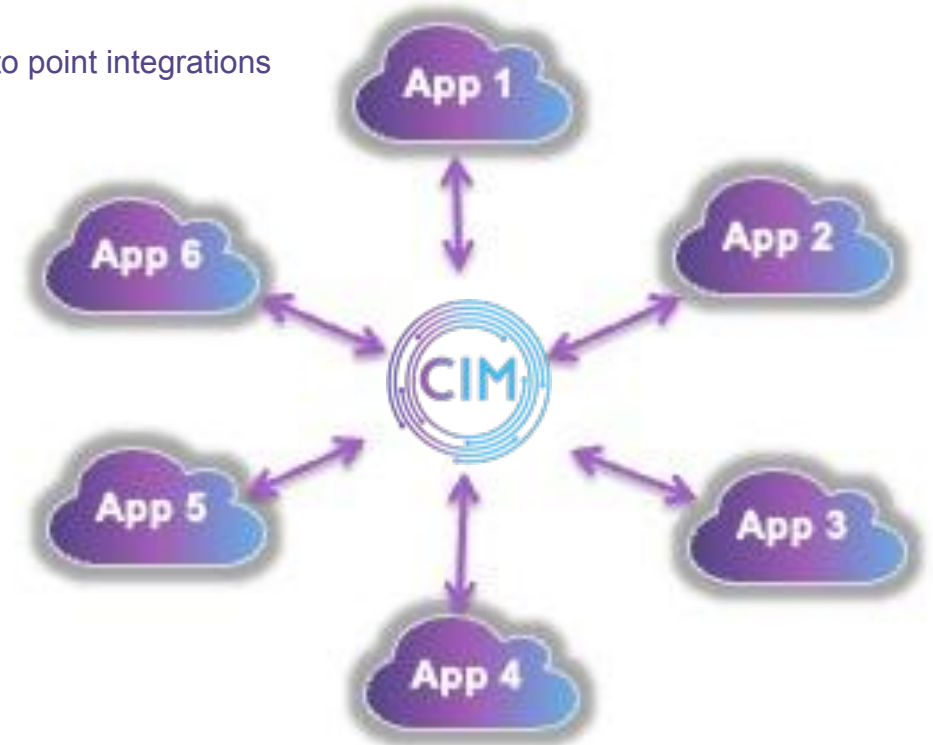
Before



15 point to point integrations

After

6 point to point integrations



...but the reality is more complicated.

# Challenges without a shared data model



## More Costly

Due to custom integrations



## Uncertain ROI

Due to longer development time



## Inability to Scale

Business functions become silos





What if you could utilize a shared data model for all of your enterprise systems?





# Introducing Cloud Information Model

The first open standard enabling faster and easier integration of cloud applications

## Anyone can join

Developed as an open standard...

## Flexible meta-model

Modular design, conceptual and schema view, extensible and customizable

**Build reliable integrations** Utilize multiple formats available to communicate between different systems

## Innovate faster

Adopt the modern standard for any use case and reduce time-to-value

## Standards Compliant

W3C Stack, Relational Mappings, additional mappings available



# Why use CIM?

CIM facilitates interoperability



- Achieve faster time to value through interoperability.
- Create seamless, customized experiences across cloud-native applications.
- Enable a cross-system view of data and reduce time to build and maintain integrations.

# Our Work Program



# How we work

CIM is organized into components including subject areas, entity groups, entities, and attributes.



## Subject area:

A major business concept identified by the CIM consortium, such as Party. Each subject area contains one or more entity groups.



## Entity group:

A logical grouping of related entities within a subject area, such as Account. Each entity group contains one or more entities.



## Entity:

A unique object that an organization collects information about, such as an Account Contact. An entity is analogous to a standard database table.









## Attribute:

A unique characteristic of an entity, such as Account Id or Contact Email. An attribute is analogous to a standard database field within a table.

# Data Model Domains











- Data models are grouped by business process in order from left to right.
- On the Subject Area/Domain Taxonomy slide, you will find a list of horizontal data models under each subject area group header.
- The vertical extensions are data models which extend the model for industry verticals while aligning to the subject area group above.

 <b>Setup</b>	 <b>Hire</b>	 <b>Produce</b>	 <b>Market</b>	 <b>Sell</b>
Defines who you deal with, for example customer supplier and seller.	Activities related to setting up your business. For example, internal business unit and worker.	Handling of material that you will be buying, moving and selling. For example, product and inventory product.	Activities used to promote your product. For example, marketing campaign and web store.	Activities used to sell your product. For example, creation of quotes and opportunities.

 <b>Serve</b>	 <b>Fulfill</b>	 <b>Interact</b>	 <b>Finance</b>	 <b>Analyze</b>
Activities to provide support for a product sold or serviced. For example, case or a survey.	Activities you perform to fulfill an order to a customer. For example, shipment and return order.	Activities to track engaging with either end users or other systems.	Activities to trace financial information in the company. For example, payment, invoice and expense report.	Activities related to analyzing data e.g. analyze patterns, analyze product use, data movement, data changes, customer satisfaction.



# Subject Area/Domain Taxonomy CIM Release Plan

 Setup	 Hire	 Produce	 Market	 Sell	 Service	 Fulfill	 Interact	 Finance	 Analyze
Software Host	Job Application	Supplier	<b>Party</b>	Price Book	AI Assistant	Fulfillment Order	Engagement	Budget	AI Model
Software Tenant	Employee	<b>Product</b>	Party Resolution	Shopping Cart	Asset	<b>Shipment</b>	Conversation	Invoice	AI Application
Software User	Compensation	Inventory Received	Privacy Consent	Quote	Asset Subscription	Return Order	Appointment	<b>Payment Method</b>	IoT Device Use
Software App.	Training	Inventory Product	Market Audience	Contract	Web Content	Work Order	Software Event	<b>Payment</b>	Data Lineage
Software Test	Location	Inventory Transfer	Campaign	Opportunity	Case	Work Resource	Data Connector	Credit Memo	Blockchain
Software Service	Work Territory	Electronic Media	Promotion	Opportunity Forecast	Task	Work Forecast	Data Movement	Financial Ledger	Survey
Soft. Batch Job	Work Report	Purchase Order	Trade Event	<b>Sales Order</b>	Event		Loyalty Journey	Account Forecast	Loyalty Journal
IoT Device	Biz Process	Sales Agreement	Ad Buy	Loyalty Program				Calendar	
Data Model	Biz Continuity		Web Site	Competitor				Tax Policy	
Available	Coming in Spring '21								

# Mappings, Dataflow & Compatibility





# What does it mean to be CIM compatible?



## Enabling CIM Utilization

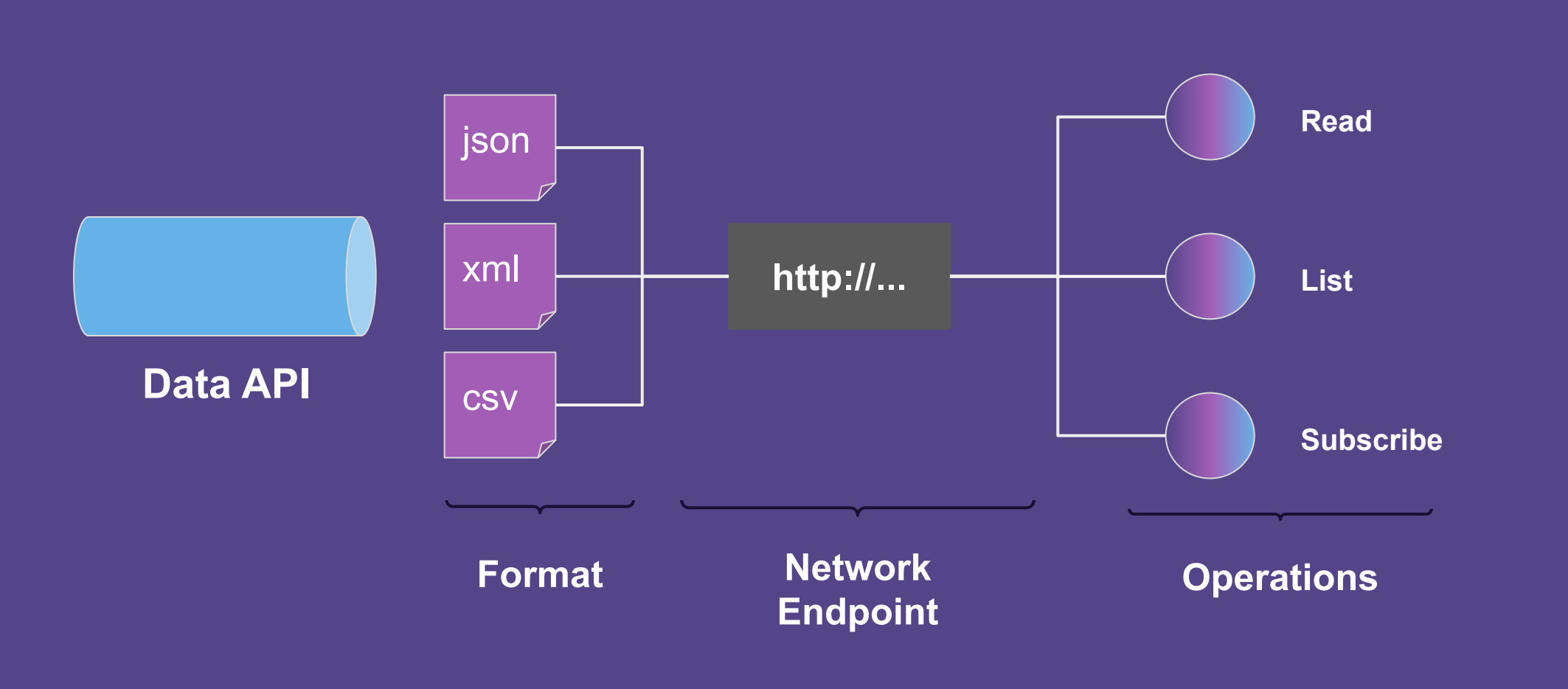
CIM availability in frameworks, builders or design tools



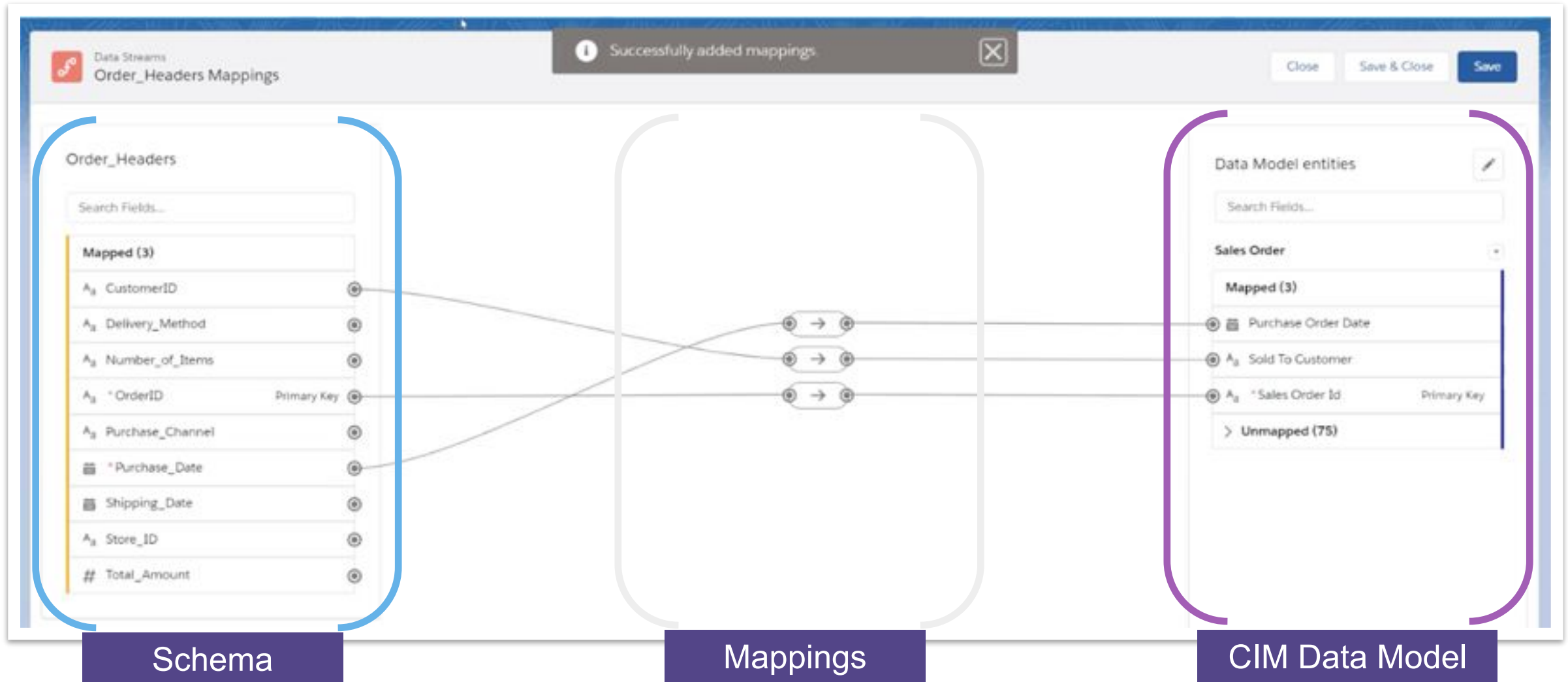
## CIM Alignment/Mapping

Existing application data models mapped to the CIM standard

# Data Flow Between Systems



# Open Discussion – Goals of “CIM Compatibility”



# Use Cases



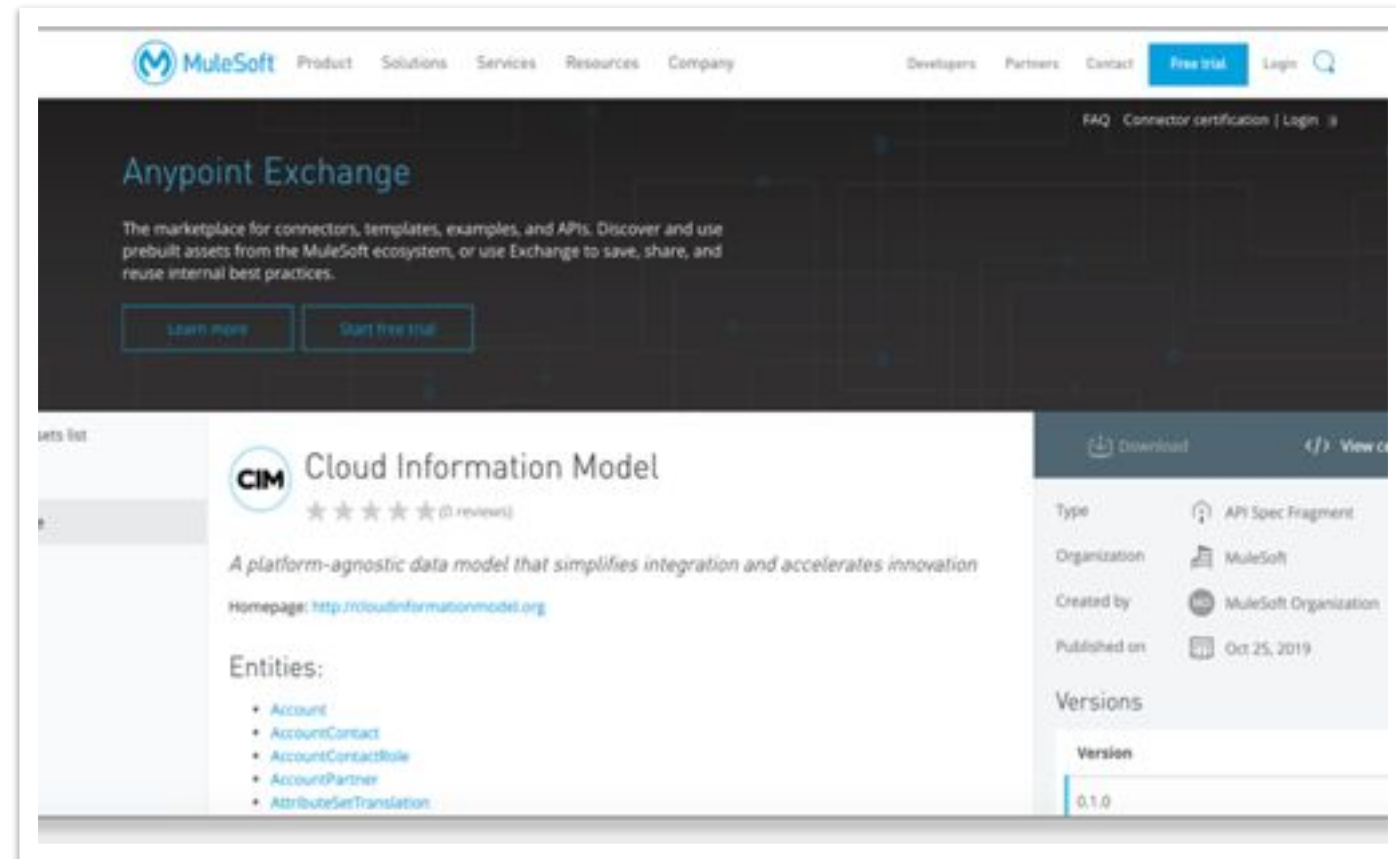




# CIM in MuleSoft

Easily create CIM compatible APIs and integrations

- Available in Anypoint Exchange
- CIM is supported across the entire Anypoint Platform (e.g. Mule Flows, DataSense, metadata)
- Utilize in API Designer to model CIM compatible APIs
- Automated C360 integration for CIM compatible sources





Join us and help define  
communication across  
clouds!



<https://cloudinformationmodel.org/>

[info@cloudinformationmodel.org](mailto:info@cloudinformationmodel.org)

# LF AI & Data - General Updates

 LF AI & DATA



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




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







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







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

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

[l.fai.foundation](https://l.fai.foundation)

Machine Learning	Framework	Platform	Library	Framework	Platform	Library	Tool	Reinforcement Learning	Programming
		 Angel LF AI & Data	 ForestFlow LF AI & Data						 PyTorch LF AI & Data

Notebook Environment	Notebook Environment	Versioning	Store & Format	Operations	Stream Processing	SQL Engine	Feature Engineering	Visualization	Pipeline Management	Labeling and Annotation	Governance
			 Milvus LF AI & Data	 Amundsen LF AI & Data  FEAST LF AI & Data  MARQUEZ LF AI & Data Incubating	 NNStramer LF AI & Data						 EGERIA LF AI & Data

Model	Benchmarking	Training	Parameter	Format & Interface	Marketplace	Workflow	Inference	Tool	Explainability	Adversarial	Bias & Fairness
		 H2O LF AI & Data	 LUDWIG LF AI & Data	 ONNX LF AI & Data	 Acumos LF AI & Data		 ADLIX LF AI & Data		 AI Explainability 360 LF AI & Data	 Adversarial Robustness Toolbox LF AI & Data	 AI Fairness 360 LF AI & Data

Distributed Computing	Computing & Management	Interface	 <p>The LF AI &amp; Data landscape explores open source projects in Artificial Intelligence and Data and their respective sub-</p> 				Security & Privacy	Natural Language Processing	Education
	 EDL LF AI & Data	 SOAJS LF AI & Data	 sparklyr LF AI & Data					 DELTA LF AI & Data	 DATAFACTICES.ORG LF AI & Data  OpenDS4AI LF AI & Data Incubating



# Suggested Additions

## Project Key

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

## Programming

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

## Notebooks

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

## Security & Privacy

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

## Distributed Computing

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

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

## Data

*Versioning*  
[Pachyderm](#) (\*)

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

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

*Stream Processing*

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

*Relational DB*

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

*SQL Engine*  
[Presto](#) (\*)

*Visualization*

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

## Data

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

*Feature Engineering*  
[Tsfresh](#)

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

*Pipeline Management*  
[Beam](#)

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

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

## Machine Learning

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

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

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

## Deep Learning

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

*Library*  
[Keras](#)

## Reinforcement Learning

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

## Model

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

*Benchmarking*  
[MLPerf](#)

*Training*  
[Horovod](#) (\*)

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

*Format & Interface*  
[ONNX](#)

*Marketplace*  
[MAX](#) (\*)

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

[Airflow](#) (\*)  
[Nifi](#) (\*)  
[Argp](#) (\*)  
[MLeap](#) (\*)  
[Volcano](#) (\*)

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

## Trusted AI

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

*Bias & Fairness*  
[AI Fairness 360](#)

*Adversarial Attacks*  
[Adversarial Robustness Toolbox](#)

## Natural Language Processing

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

*Education*  
[OpenDS4All](#)

## 2020 TAC Meetings Summary

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

(Entity)\* = incubating vote

**\*\* bold = graduate vote**

*Italics = invited project presentation*

## 2021 TAC Meetings Pipeline Summary

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

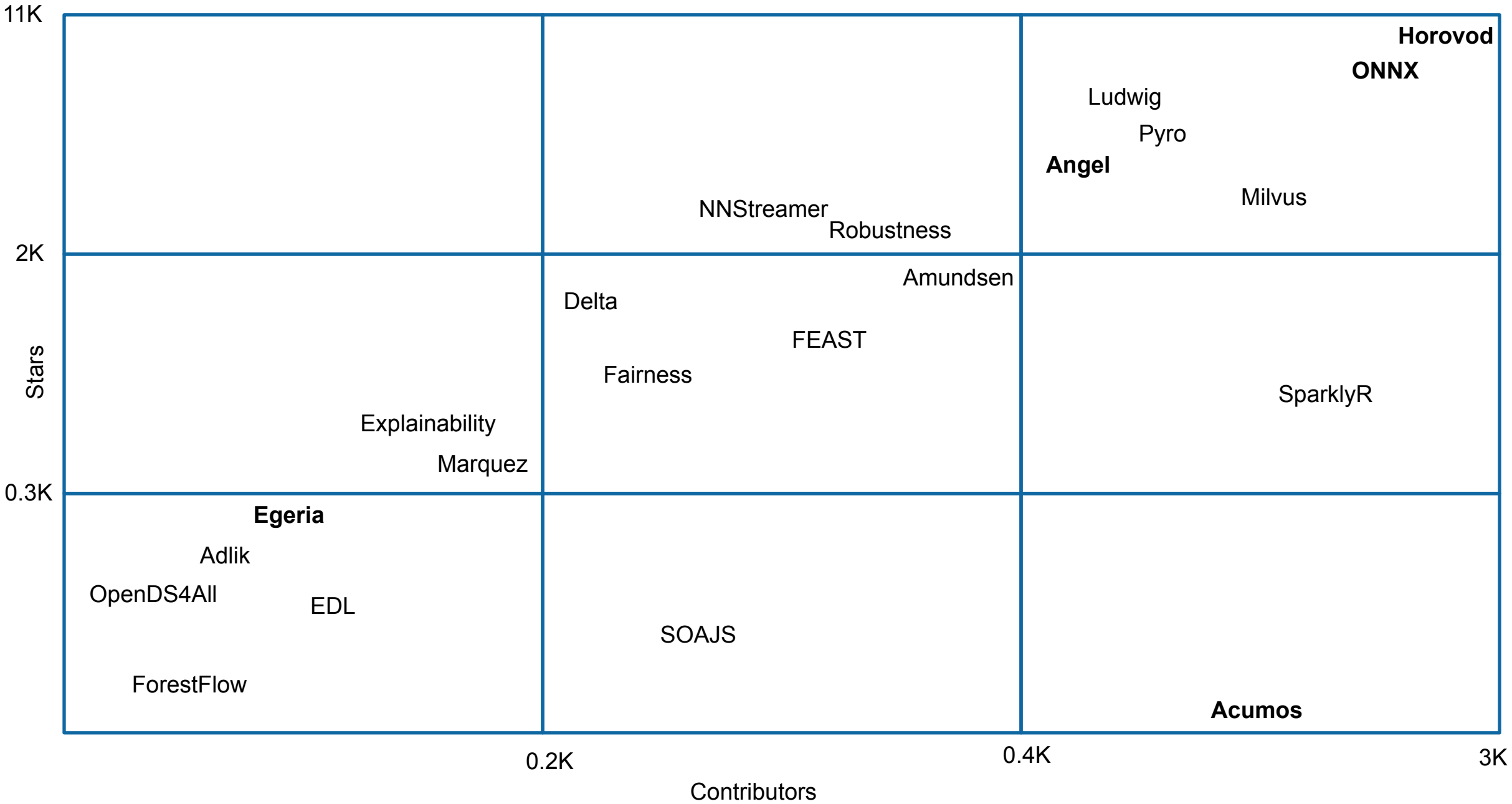
(Entity)\* = incubating vote

\*\* **bold** = graduate vote

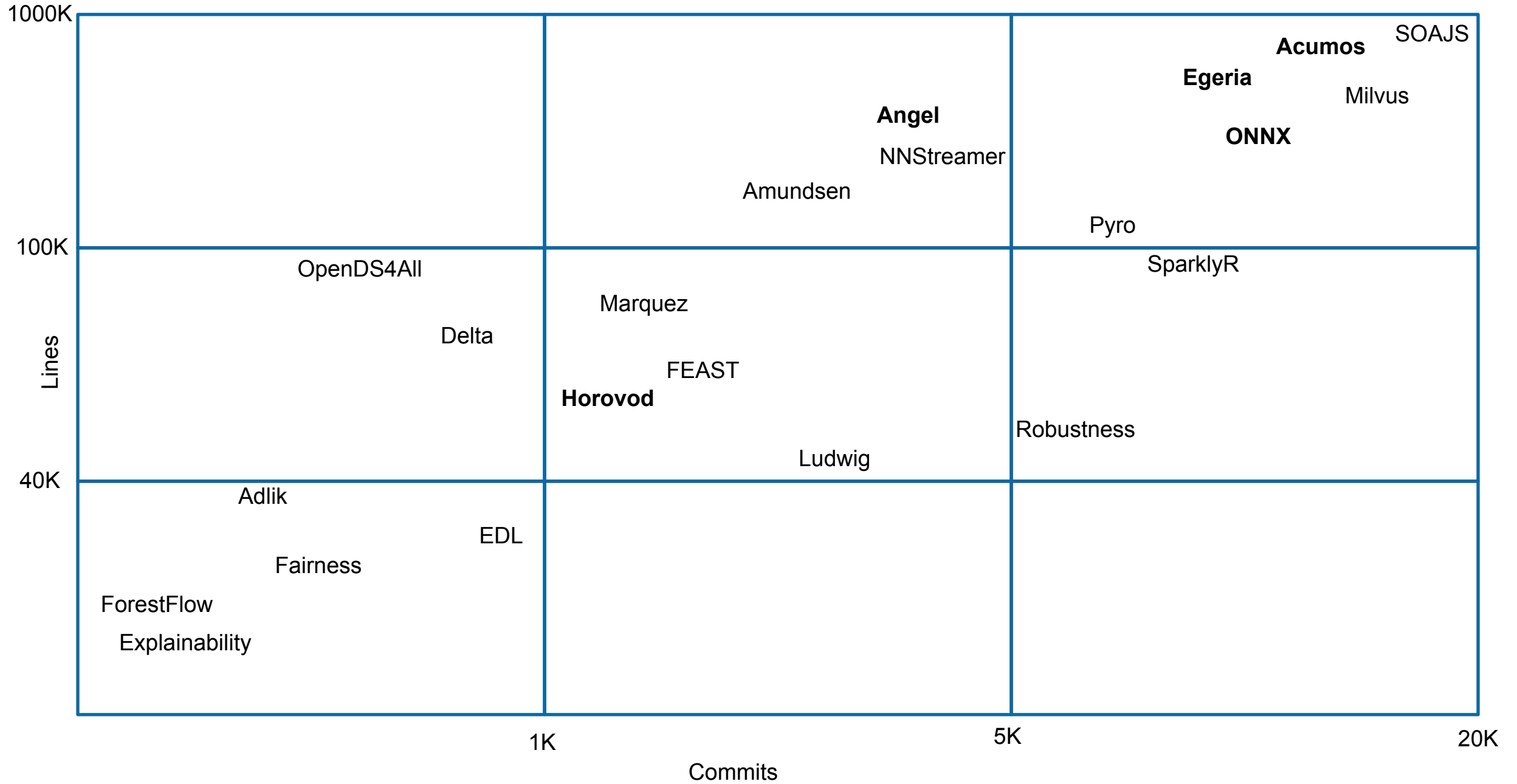
*Italics* = invited project presentation

Getting to know the projects more

Data from November 23, 2020 – Stars and Contributors



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





# Looking to host a project with LF AI & Data

- › Hosted project stages and life cycle:

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

- › Offered services for hosted projects:

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

- › Contact:

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

# Promoting Upcoming Project Releases

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

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

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

# Note on quorum

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

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

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

# Updates from Outreach Committee

# Upcoming Events

- › Upcoming Events

- › Visit the [LF AI & Data Events Calendar](#) or the [LF AI & Data 2021 Events wiki](#) for a list of all events
- › To participate visit the [LF AI & Data 2021 Events wiki page](#) or email [info@lfaidata.foundation](mailto:info@lfaidata.foundation)

› Please consider holding virtual events

To discuss participation, please email [events@lfaidata.foundation](mailto:events@lfaidata.foundation)

# Upcoming Events

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

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

# LF AI PR/Comms

- › Please follow LF AI & Data on [Twitter](#) & [LinkedIn](#) and help amplify news via your social networks - Please retweet and share!
  - › Also watch for news updates via the tac-general mail list
  - › View recent announcement on the [LF AI & Data Blog](#)
- › Open call to publish project/committee updates or other relevant content on the [LF AI & Data Blog](#)
- › To discuss more details on participation or upcoming announcements, please email [pr@lfaidata.foundation](mailto:pr@lfaidata.foundation)

# Call to Participate in Ongoing Efforts

 **OLF** AI & DATA



# Trusted AI

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

# ML Workflow & Interop

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

# BI & AI

- › **Leadership:**  
Cupid Chan (Index Analytics)
- › **Goal:** Identify and share industry best practices that combine the speed of machine learning with human insights to create a new business intelligence and better strategic direction for your organization.
  
- › **Slack conversations channel:**  
**#bi-ai-committee**  
<https://lfaifoundation.slack.com/archives/C01EK5ND073>
- › **Github:**  
<https://github.com/odpi/bi-ai>
- Wiki:**  
<https://wiki.lfaidata.foundation/pages/viewpage.action?pageId=35160417>
- Email lists:**  
<https://lists.lfaidata.foundation/g/biai-discussion>
- Next call:** Monthly community call TBD

# Ongoing effort to create AI Ethics Training

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

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

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

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

# Upcoming TAC Meetings

# Upcoming TAC Meetings (Tentative)

- › Feb 11: MARS (Alibaba)
- › Feb 25: FLYTE (Lyft)
- › Mar 11: TBD
- › Mar 25: Substra Foundation
- › April 8: Adlik (ZTE)
- › April 22: TBD
- › May 6: All project updates
- ›

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

# TAC Meeting Details

- › To subscribe to the TAC Group Calendar, visit the wiki: <https://wiki.lfaidata.foundation/x/cQB2>
- › Join from PC, Mac, Linux, iOS or Android: <https://zoom.us/j/430697670>
- › Or iPhone one-tap:
  - › US: +16465588656,,430697670# or +16699006833,,430697670#
- › Or Telephone:
  - › Dial(for higher quality, dial a number based on your current location):
  - › US: +1 646 558 8656 or +1 669 900 6833 or +1 855 880 1246 (Toll Free) or +1 877 369 0926 (Toll Free)
- › Meeting ID: 430 697 670
- › International numbers available: <https://zoom.us/u/achYtcw7uN>

# Open Discussion

# Mission

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



# Legal Notice

- › The Linux Foundation, The Linux Foundation logos, and other marks that may be used herein are owned by The Linux Foundation or its affiliated entities, and are subject to The Linux Foundation's Trademark Usage Policy at <https://www.linuxfoundation.org/trademark-usage>, as may be modified from time to time.
- › Linux is a registered trademark of Linus Torvalds. Please see the Linux Mark Institute's trademark usage page at <https://lmi.linuxfoundation.org> for details regarding use of this trademark.
- › Some marks that may be used herein are owned by projects operating as separately incorporated entities managed by The Linux Foundation, and have their own trademarks, policies and usage guidelines.
- › TWITTER, TWEET, RETWEET and the Twitter logo are trademarks of Twitter, Inc. or its affiliates.
- › Facebook and the “f” logo are trademarks of Facebook or its affiliates.
- › LinkedIn, the LinkedIn logo, the IN logo and InMail are registered trademarks or trademarks of LinkedIn Corporation and its affiliates in the United States and/or other countries.
- › YouTube and the YouTube icon are trademarks of YouTube or its affiliates.
- › All other trademarks are the property of their respective owners. Use of such marks herein does not represent affiliation with or authorization, sponsorship or approval by such owners unless otherwise expressly specified.
- › The Linux Foundation is subject to other policies, including without limitation its Privacy Policy at <https://www.linuxfoundation.org/privacy> and its Antitrust Policy at <https://www.linuxfoundation.org/antitrust-policy>, each as may be modified from time to time. More information about The Linux Foundation's policies is available at <https://www.linuxfoundation.org>.
- › Please email [legal@linuxfoundation.org](mailto:legal@linuxfoundation.org) with any questions about The Linux Foundation's policies or the notices set forth on this slide.



