

Meeting of the LF AI & Data Technical Advisory Council (TAC)

September 21, 2023

 LF AI & DATA

Antitrust 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 Updegrove 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
- › Wiki: wiki.lfaidata.foundation
- › GitHub: github.com/lfaidata
- › Landscape: <https://landscape.lfaidata.foundation> or <https://l.lfaidata.foundation>
- › Mail Lists: <https://lists.lfaidata.foundation>
- › Slack: <https://slack.lfaidata.foundation>
- › Youtube: <https://www.youtube.com/channel/UCfasaeqXJBCAJMNO9HcHfbA>
- › 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
- › 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 (1 mins)
- › Approval of Minutes from previous meeting (2 mins)
- › AIDA new Sandbox project proposal (45 minutes)
- › Open Discussion

TAC Voting Members - Please note

Please ensure that you do the following to facilitate smooth procedural quorum and voting processes:

- Change your Zoom display name to include your First/Last Name, Company/Project Represented
 - example: Nancy Rausch, SAS
- State your First/Last Name and Company/Project when submitting a motion
 - example: First motion, Nancy Rausch/SAS

TAC Voting Members

Note: we still need a few designated backups specified on [wiki](#)

Member Company or Graduated Project	Membership Level or Project Level	Voting Eligibility	Country	TAC Representative	Designated TAC Representative Alternates
4paradigm	Premier	Voting Member	China	Zhongyi Tan	
Baidu	Premier	Voting Member	China	Jun Zhang	Daxiang Dong, Yanjun Ma
Ericsson	Premier	Voting Member	Sweden	Rani Yadav-Ranjan	
Huawei	Premier	Voting Member	China	Howard (Huang Zhipeng)	Charlotte (Xiaoman Hu), Leon (Hui Wang)
IBM	Premier	Voting Member	USA	Susan Malaika	Beat Buesser, Alexandre Eichenberger
Nokia	Premier	Voting Member	Finland	@Michael Rooke	@Jonne Soininen
OPPO	Premier	Voting Member	China	Jimmy (Hongmin Xu)	
SAS	Premier	Voting Member	USA	*Nancy Rausch	Liz McIntosh
ZTE	Premier	Voting Member	China	Wei Meng	Liya Yuan
Adversarial Robustness Toolbox Project	Graduated Technical Project	Voting Member	USA	Beat Buesser	Kevin Eykholt
Angel Project	Graduated Technical Project	Voting Member	China	Jun Yao	
Egeria Project	Graduated Technical Project	Voting Member	UK	Mandy Chessell	Nigel Jones, David Radley, Maryna Strelchuk, Ljupcho Palashevski, Chris Grote
Flyte Project	Graduated Technical Project	Voting Member	USA	Ketan Umare	
Horovod Project	Graduated Technical Project	Voting Member	USA	Travis Addair	
Milvus Project	Graduated Technical Project	Voting Member	China	Xiaofan Luan	Jun Gu
ONNX Project	Graduated Technical Project	Voting Member	USA	Alexandre Eichenberger	Andreas Fehlner, Prasanth Pulavarthi, Jim Spohrer
Pyro Project	Graduated Technical Project	Voting Member	USA	Fritz Obermeyer	
Open Lineage Project	Graduated Technical Project	Voting Member	USA	Awaiting confirmation from Project Lead	

Minutes approval

Approval of September 7, 2023 Minutes

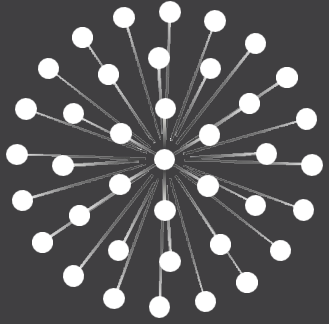
Draft minutes from the September 7 TAC call were previously distributed to the TAC members via the mailing list

Proposed Resolution:

- › That the minutes of the September 7 meeting of the Technical Advisory Council of the LF AI & Data Foundation are hereby approved.

AIDA New Project Proposal

Jean-Georges Perrin <jgp@jgp.ai>



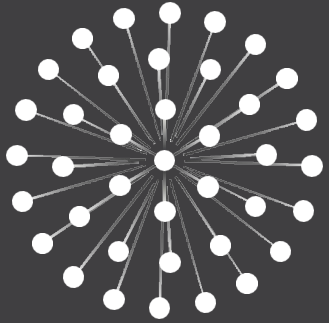
AI
DATA
ANALYTICS
USER GROUP

Artificial Intelligence, Data, and
Analytics
User Group

Open Data Contract Standard

v1.0.2

Jean-Georges Perrin • Co-founder & President • jgp@jgp.ai

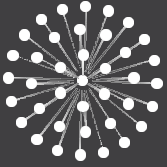


AIDA
USER GROUP

Artificial Intelligence, Data,
and Analytics
User Group

As **Artificial Intelligence is transforming society** and businesses, AIDA User Group has the ambition to **share knowledge, connect people, and provide** guidance through this next era of computing while embracing peace and the highest levels of **ethics**.

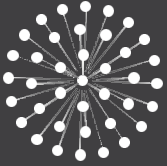
Join AIDAUG for free at <http://aidaug.org/join>



What is a data contract?

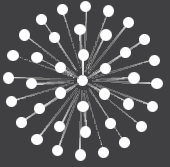
User Group

- Creates a link between data producers and data consumers.
- Creates a link between a logical representation of the data and its physical implementation.
1
3
- Describes meta meta data: rules, quality, and behavior.



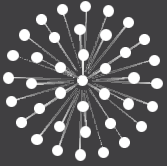
What are the problems we are trying to solve?

- Normalizing and keeping documentation relevant. Describing
- service-level expectations.
- Easing data & tools integration.
- Enabling data product thinking.



Benefits of a standard

- Boosting innovation.
- Simplifying integration.

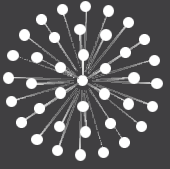


Who uses the Open Data Contract Standard?

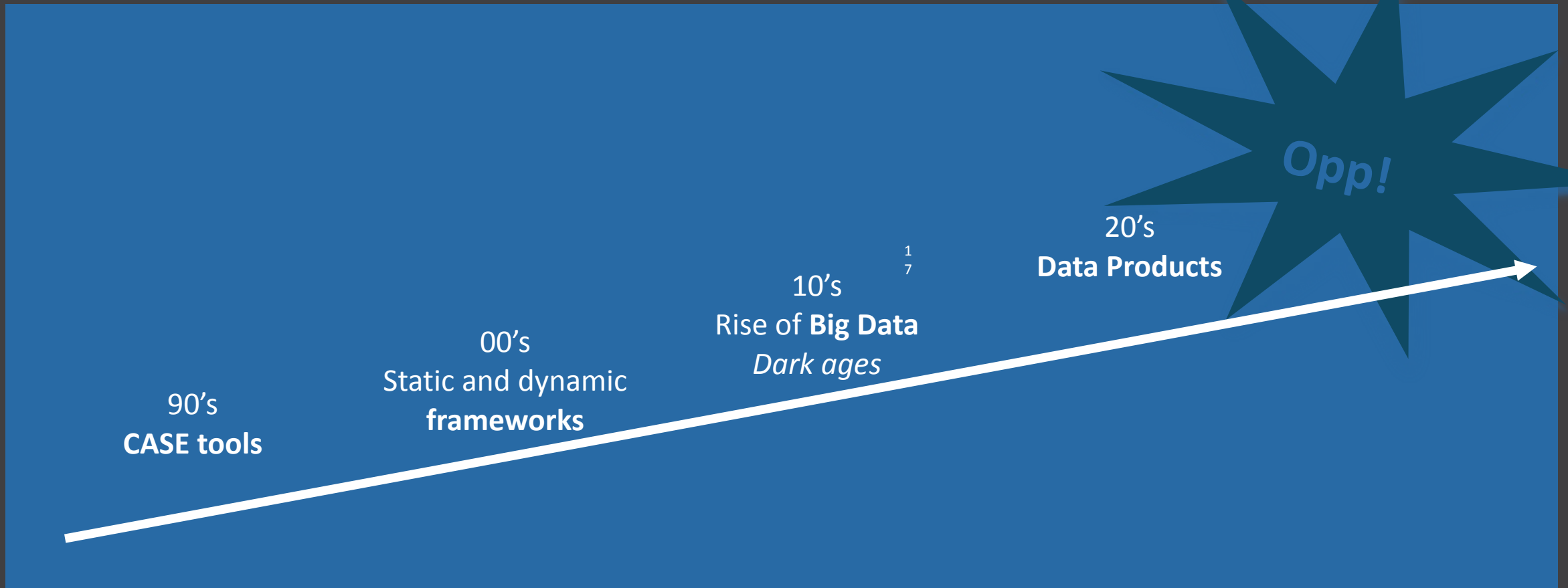
Various levels of maturity in:

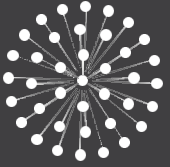
- Fortune 500: peer-to-peer payment leader, major cable company.
- Startups: several data-oriented startups in Europe and in NA (Among them: Raito). Mid-market companies: in the US & Europe (Among them: ProfitOptics,¹Innoboost). An active user group behind the standard (AIDA User Group).
- A community (Data Mesh Learning).

Active recruitment of users & vendors underway!



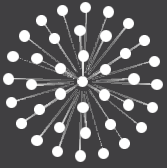
History of data-related development



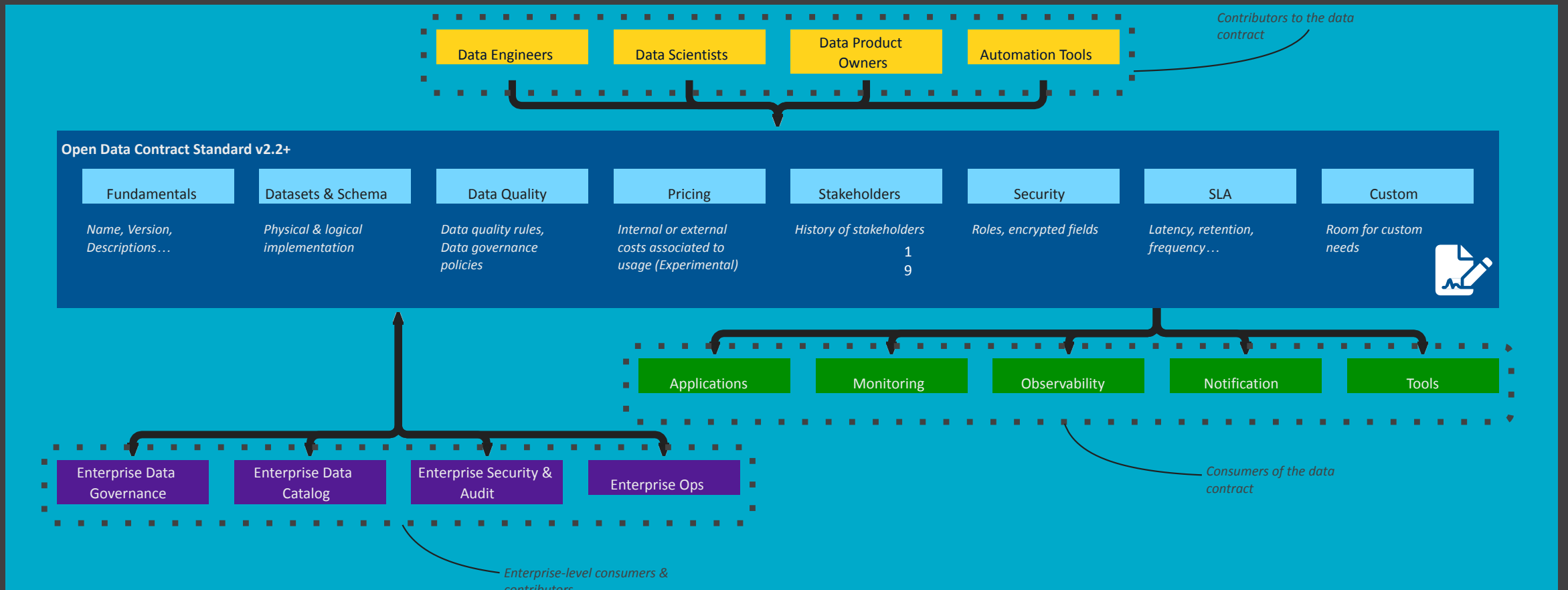


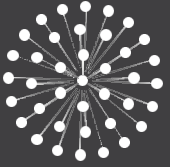
A popular standard

	<u>PayPal</u> (Original until v2.1)	<u>ODCS</u> (From v2.2)
Stars	341 18	15
Forks	38	40
Contributors	5	5



Architecture & Content





Implementation

- YAML file, both human -and computer-readable.
- Language agnostic.

- `table: tbl` Table name

`description:` Provides core payment metrics

`dataGranularity:` Aggregation on columns `txn_ref_dt`, `pmt_txn_id`

`columns:`

- `column: txn_ref_dt`
`businessName:` Transaction reference date

`logicalType:` date

Mapping from logical to physical world

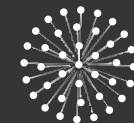
`physicalType:` date

`description:` Reference date for the transaction¹. Use this date in reports and aggregation rather than `txn_mystical_dt`, as it is slightly too mystical.²

`sampleValues:`

- 2022-10-03

- 2025-01-28

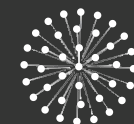


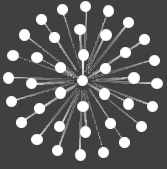
```
- table: tbl
  columns:
    - column: rcvr_cntry_code
      businessName: Receiver country code
      logicalType: string
      physicalType: varchar(2)
```

Multiple links to enterprise definitions

```
  authoritativeDefinitions:
```

```
    - url: https://collibra.com/asset/748f-71a5-4ab1-bda4-8c25
      type: Business definition
    - url: https://github.com/myorg/myrepo
      type: Reference Implementation
```

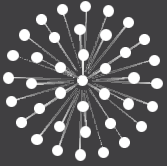




Here

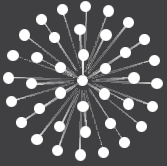
Evolution of this project

	Phase 1	Phase 2	Phase 3	Phase 4
Standard	Publish ODCS v2.x	Animate a working group Publish ODCS v3.x	Publish Data Product Standard v1.x Evolution of ODCS	Publish Data Mesh Standard v1.x Evolution of ODCS
Tools	ODCS parser	ODCS validator	Updated & new parsers & validators	Updated & new parsers & validators



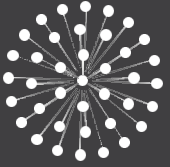
Communication plan

- 9/11 IBM TechXchange
- 9/20 Big Data London (Chat with exhibitors) 9/27 IBM TechTalk (Online APAC)
- 9/28 IBM TechTalk (Online Americas) 10/15 All Things Open (Raleigh, NC, USA) Q4: Articles
- 2024: Several conferences, articles, at least one **book** (Implementing Data Mesh, O'Reilly)
-



Roadmap post LF incubation

- Create a set of Open Standards, around:
 - Data Contracts.
 - Data Products.
 - Data Mesh.
- Host open-source tooling.
- Provide guidance around licensing & operations.



Request to the Linux Foundation

Requesting the support of the TAC to host the project in LF AI & Data as a Sandbox project.

Approval on AIDA Proposal

Proposed Resolution:

- › AIDA as a new Sandbox project of the LF AI & Data Foundation is hereby approved.

Upcoming TAC Meetings

 **DLF** AI & DATA

Upcoming TAC Meetings

- › October 5 – LMSYS project review / 20 mins (Lianmin Zheng) + Update on the Generative AI Commons from Matt White
- › October 21 – Open slot

Please note we are always open to special topics as well.

If you have a topic idea or agenda item, please send agenda topic requests to tac-general@lists.lfaidata.foundation

Open Discussion

 OLF AI & DATA

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>

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 with any questions about The Linux Foundation's policies or the notices set forth on this slide.