

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

June 15, 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)
- › Ethical AI project (40 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 - Please note

- › TAC members must attend consistently to maintain their voting status
- › After 2 absences voting members will lose voting privileges
- › Voting privileges will only be reinstated after attending 2 meetings in a row

TAC Voting Members

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

with the new attendance and voting rule adopted by the TAC, the number of eligible voting members varies per week and is tracked [here](#). Please see meeting minutes on the [TAC wiki](#) to

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)
Nokia	Premier	Voting Member	Finland	@ Michael Rooke	@ Jonne Soininen
OPPO	Premier	Voting Member	China	Jimmy (Hongmin Xu)	
SAS	Premier	Voting Member	USA	*Nancy Rausch	JP Trawinski
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	

Minutes approval

Approval of June 1, 2023 Minutes

Draft minutes from the June 1 TAC call were previously distributed to the TAC members via the mailing list

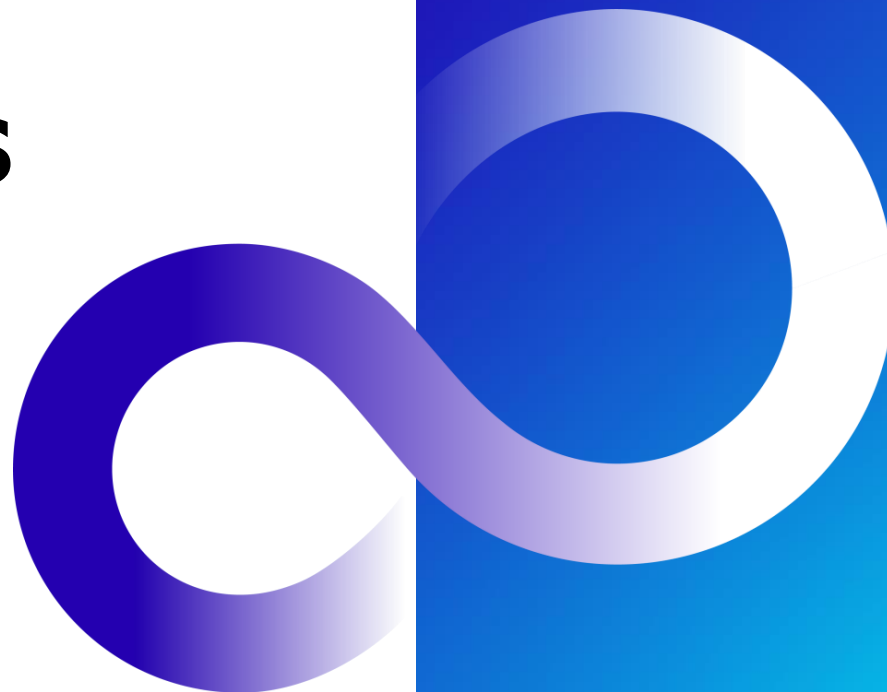
Proposed Resolution:

- › That the minutes of the June 1 meeting of the Technical Advisory Council of the LF AI & Data Foundation are hereby approved.

Intersectional Fairness (ISF) Technology

June 15, 2023

AI Trust Research Centre
Fujitsu Research, Fujitsu Limited



Why contribute ISF technology to Linux Foundation?

- Neutral holding ground
 - Vendor-neutral, Not for profit
- Growing community
 - Opportunities to collaborate with other hosted projects
 - Increase visibility of project through LF ecosystem
 - Increase contributors by converting new & existing users
- Open Governance model
 - Transparent and open governance model
 - Instill trust in contributors and adopters in the management of the project
 - Neutral management of projects' assets by the foundation

Problem

- Fairness in ML (Machine Learning) has got a great deal of attention in recent years.
- A variety of fair ML methods have been developed to mitigate discriminatory bias against sensitive attributes (e.g., gender, race) in a wide range of scenarios.
 - Examples of the methods: Adversarial debiasing, Equalized odds, Reject option-based classification, and Massaging, which AIF360 and other popular OSS projects for ML fairness toolkit contain.

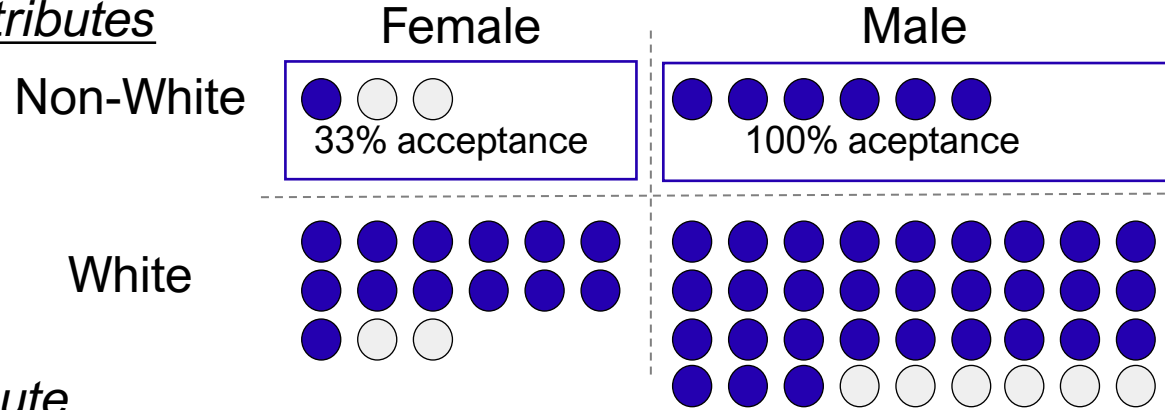
OSS projects for ML fairness toolkit	License	Repository
Google What-if- tool	Apache-2.0	https://github.com/PAIR-code/what-if-tool
Microsoft FairLean	MIT	https://github.com/fairlearn/fairlearn
AIF360	Apache-2.0	https://github.com/Trusted-AI/AIF360

- However, most of them only consider bias for a *single* sensitive attribute.

Problem: Loss of fairness in intersection

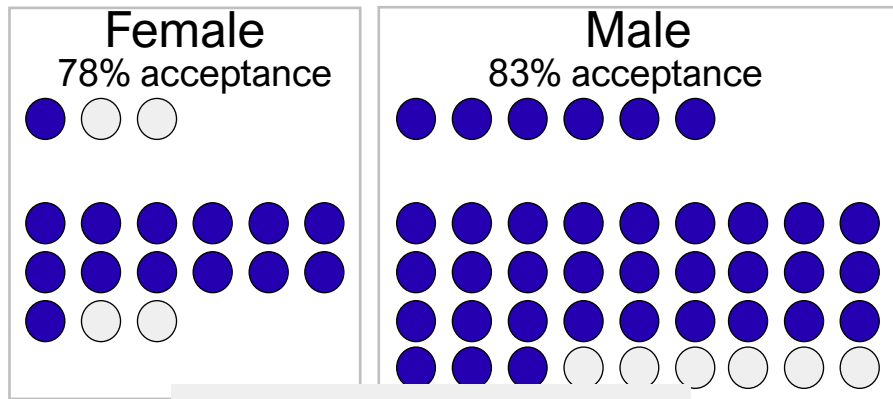
- Combining multiple sensitive attributes causes unfairness even if each single attribute causes no unfairness

In case you treat combined attributes (i.e., an intersection)

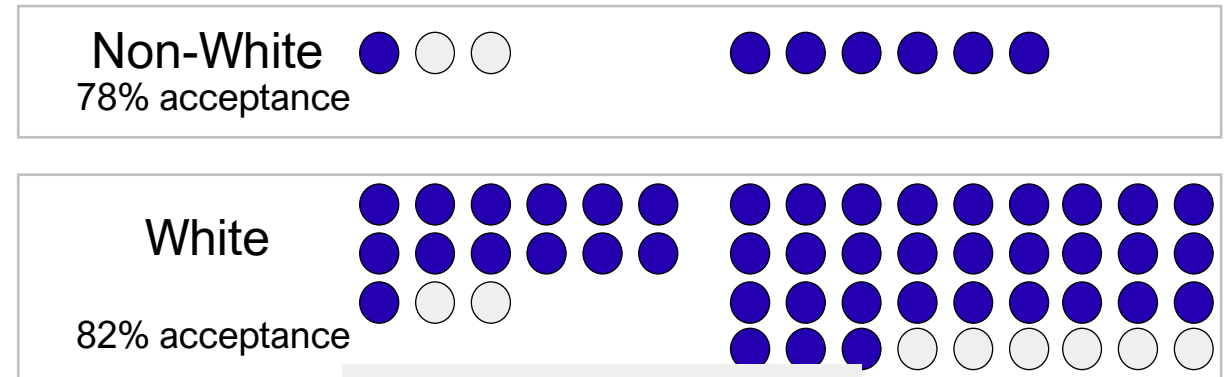


→ $DI=0.3 < 0.8$
(Unfair)

In case you treat a single attribute



→ $DI=0.9 > 0.8$ (fair)



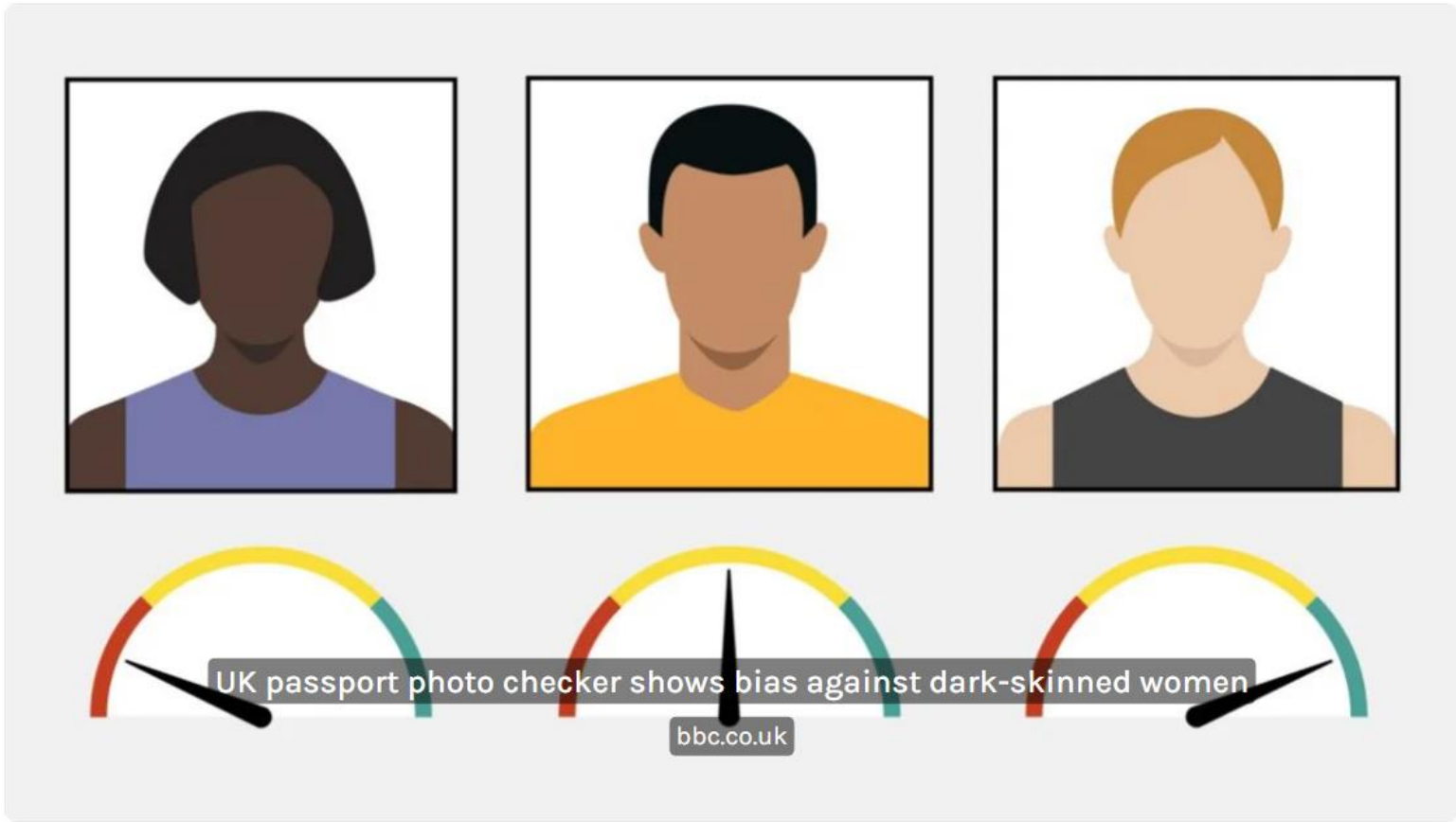
→ $DI=0.9 > 0.8$ (fair)

● Accepted applicant
○ Rejected applicant

DI (Disparate Impact) : Ratio of the acceptance rate between groups; $DI > 0.8$ means "fair"

Practical example

UK passport photo checker shows bias against dark-skinned women



Conventional methods

- ✓ can detect intersectional fairness violation, but
- ✗ cannot conform to intersectional fairness.



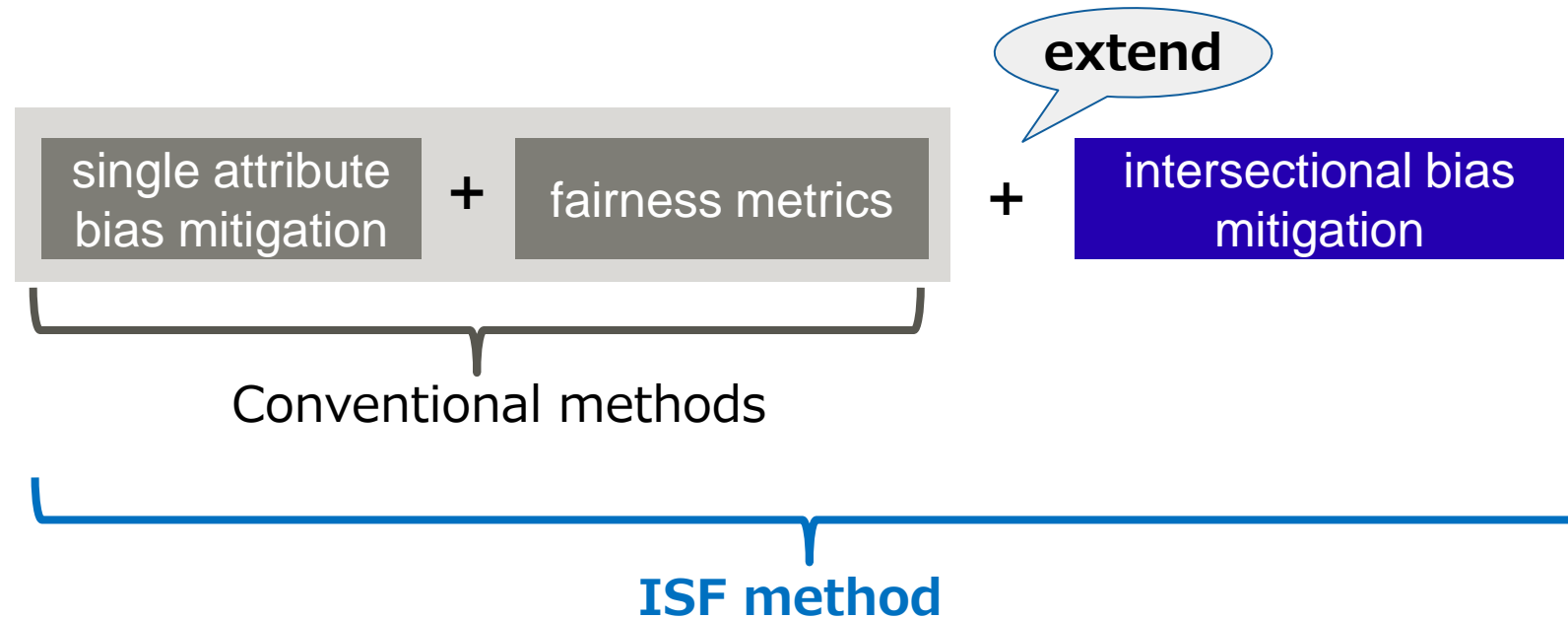
Our goal is to

- ✓ detect intersectional fairness violation, and
- ✓ *conform to intersectional fairness*

Intersectional Fairness (ISF) Technology

- ISF method

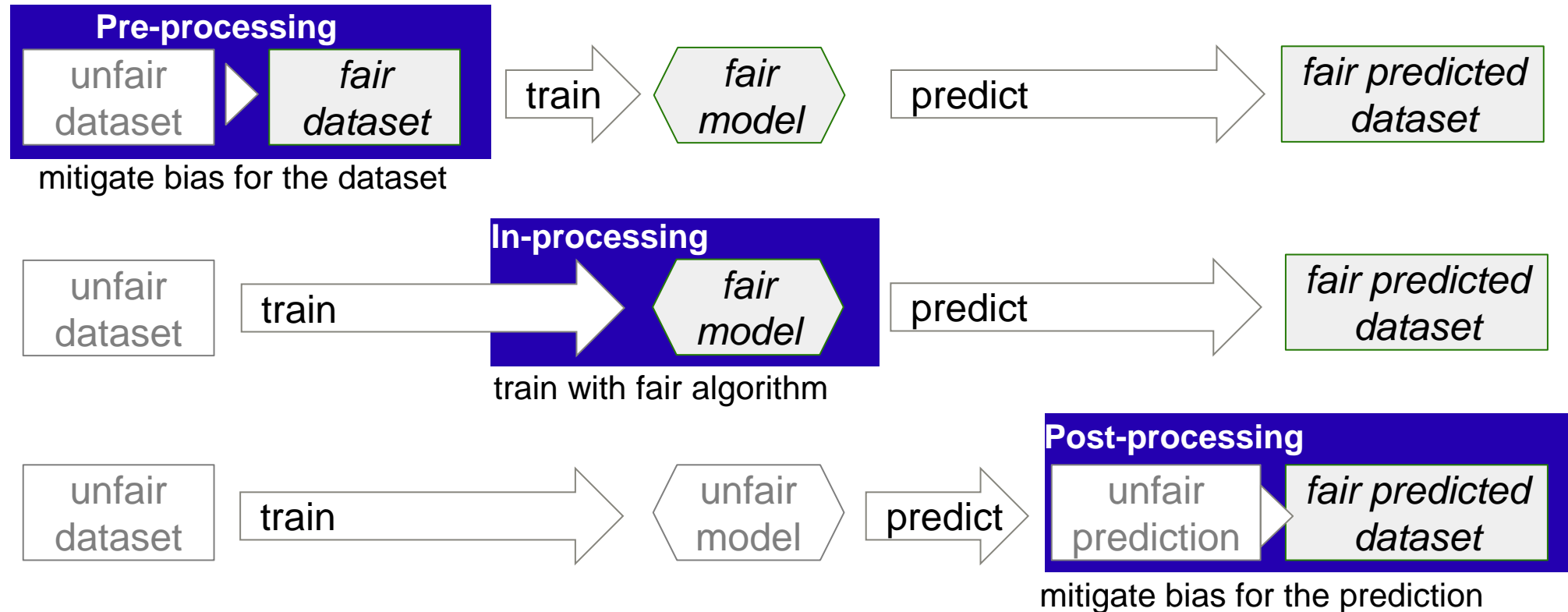
- extends a conventional method to enable intersectional bias mitigation
- inherits bias mitigation approaches and fairness metrics from conventional methods



- focuses on binary classification tasks for now

Applicable bias mitigation methods

All three types of mitigation methods can be used: pre-, in- & post-processing



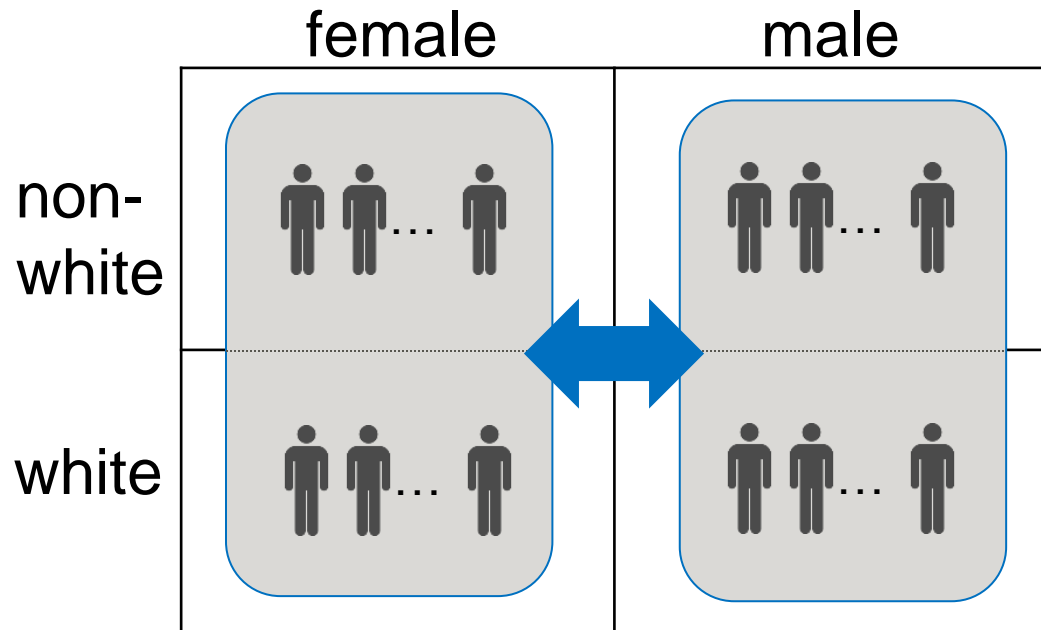
- Select a mitigation method (and metrics) that fits your use-case scenario

- To learn an ML and predict with it, call ISF APIs with parameters specifying a mitigation method.
 - (Todo: List key operations in Jupyter Notebook and the visualized result)

- In python (3.7)
- Referenced libraries:
 - numpy
 - pandas
 - sklearn
 - tempeh
 - matplotlib
 - tensorflow
 - protobuf
 - fairlearn
 - cvxpy
 - aif360

Technical details

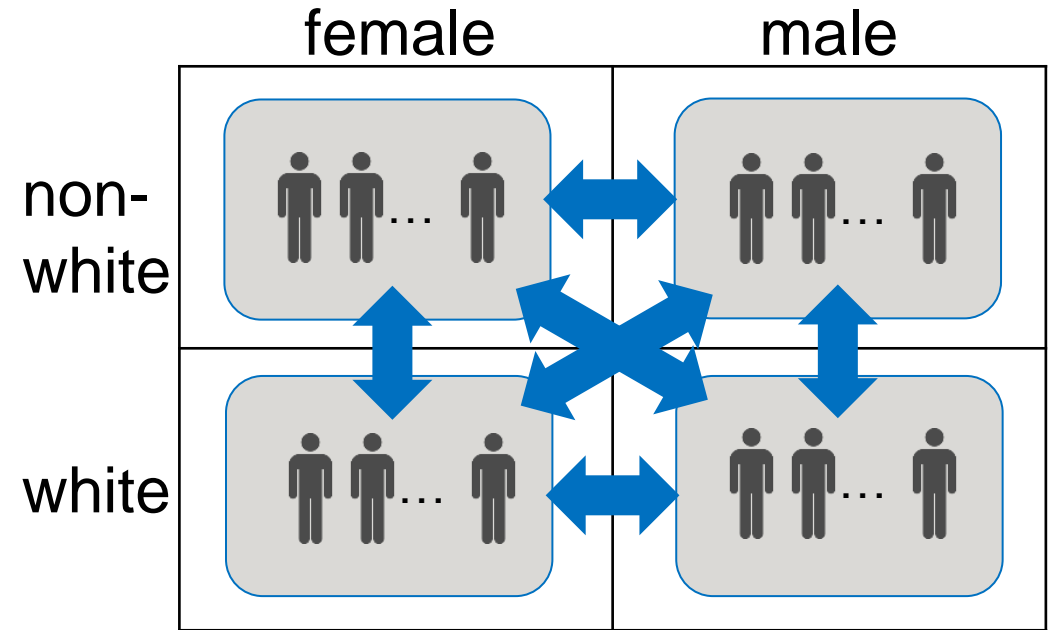
Conventional methods



mitigate bias between two groups

(in case bias is mitigated for gender)

ISF method

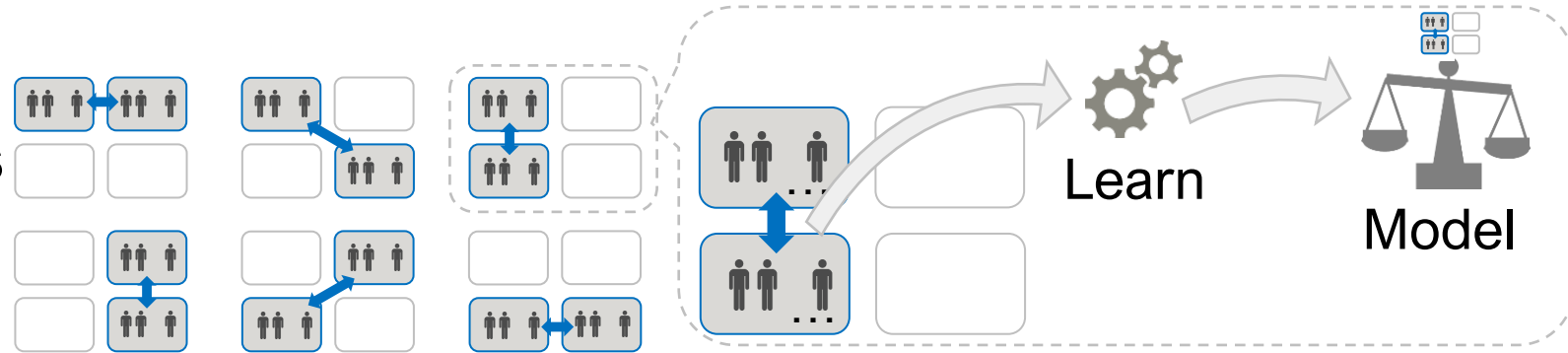


**mitigates bias between
all subgroup pairs
using conventional methods**

Detail procedure (for in-process)

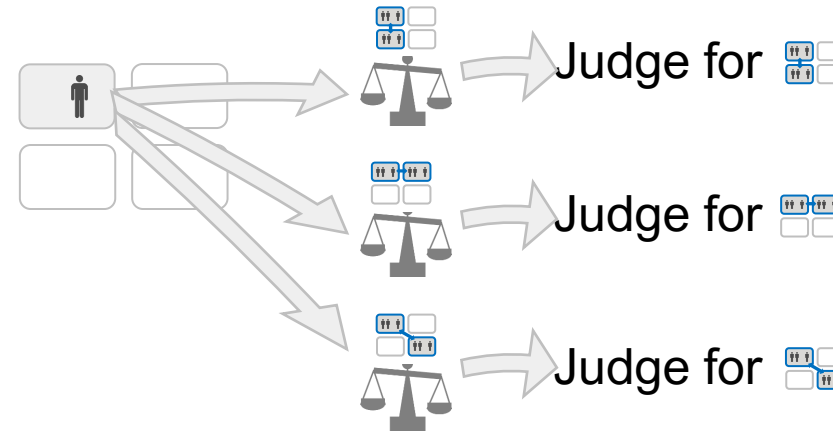
○ Training

- Obtain prediction models for all pair of subgroups

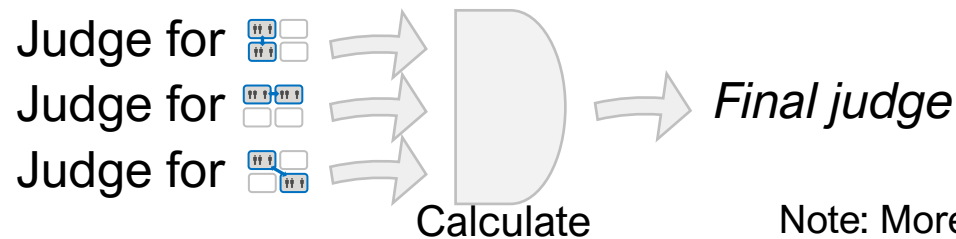


○ Prediction

- Predict judgements only for pairs related to the target



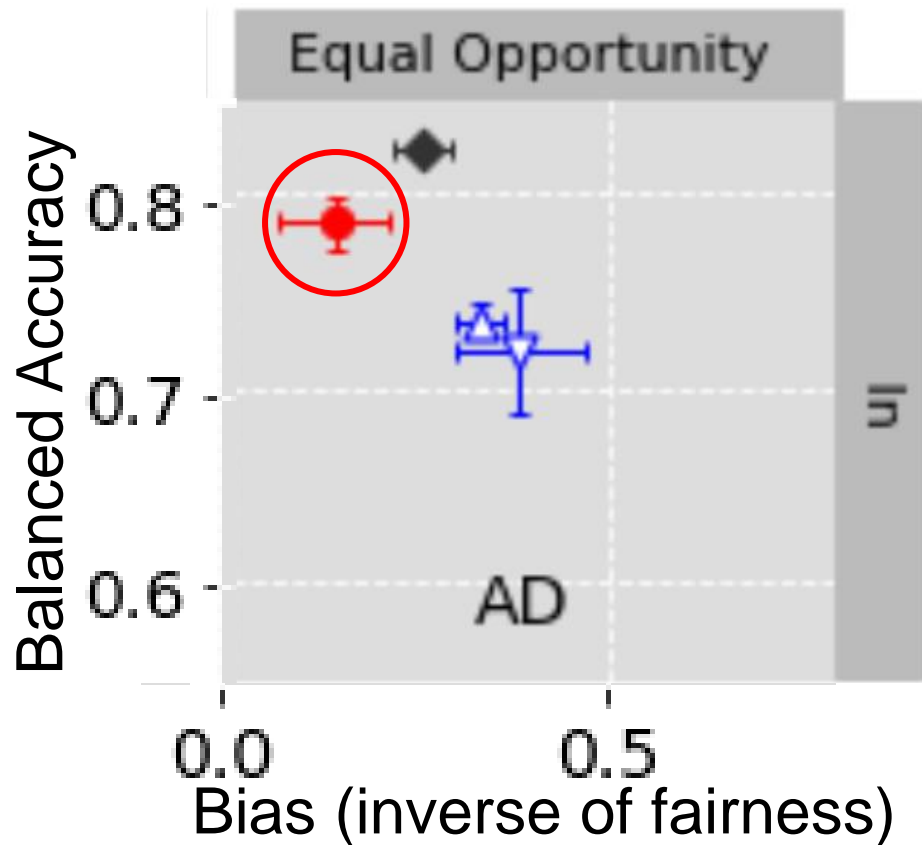
- Merge these judgements



Note: More precise procedure in appendix

Effect: Intersectional fairness improves fairness & accuracy

Intersectional fairness has the best fairness across all results, with not much accuracy loss.



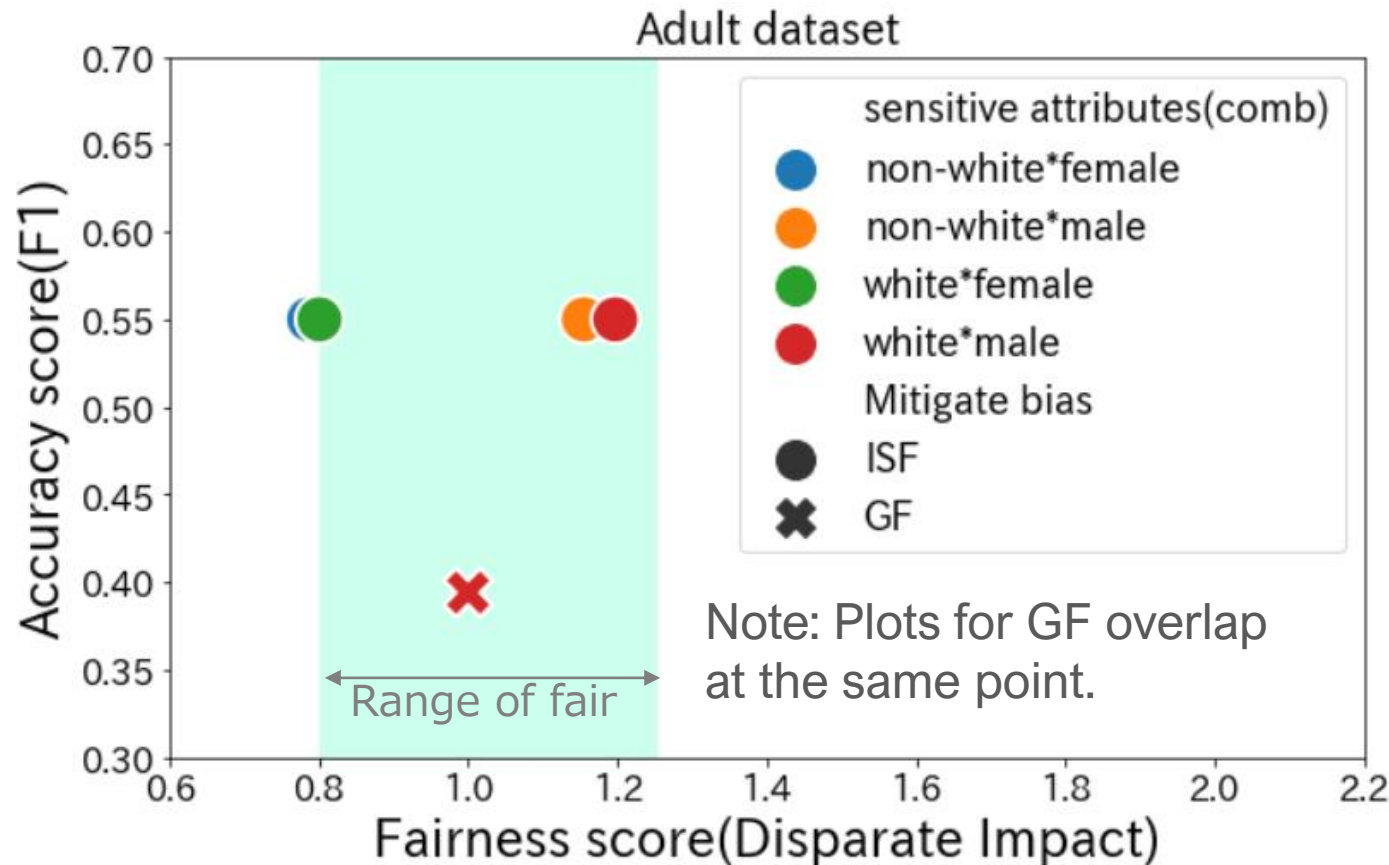
Top left is better

- ◆ Plain classifier (Logistic Regression)
 - △▽ Conventional method for gender and race, respectively
 - ⊙ Our method (extends a method for each cell)
- MS:** Messaging
AD: Adversarial debiasing
ROC: Reject option-based classification
EO: Optimal Equalized Odds/Opportunity Predictor

Comparison with GerryFair

To compare fairness with the existing same-purposed method GerryFair,

ISF has less accuracy loss, while both techniques have better fairness scores.



GerryFair: Kearns, M., et al. (2018). *Preventing fairness gerrymandering: Auditing and learning for subgroup fairness*. In International conference on machine learning. PMLR.

Potential co-operation and future

- AI Fairness 360 (AIF360) project
 - develops a toolkit “AI Fairness 360” to help detect and mitigate bias in machine learning models throughout the AI application lifecycle.
- By incorporating ISF code in AIF360,
 - the choice of the methods increases for applicators (users);
 - applicators can use a method for intersectional fairness that they need; and
 - applicators need not install ISF separately while retaining consistency among libraries.

Technical matters:

- Explainability: to explain how to mitigate bias
- Improve the trade-offs between accuracy and fairness (with data augmentation)

Activities:

- Publicize ISF as an individual project
- Promote activities of the LF and ISF
- Merge it with or collaborate with AIF360
- Support for more practical use cases

The loss of fairness in intersections in sensitive attribute combinations can cause problems in practical ML uses.

ISF helps solve for making MLs fair.

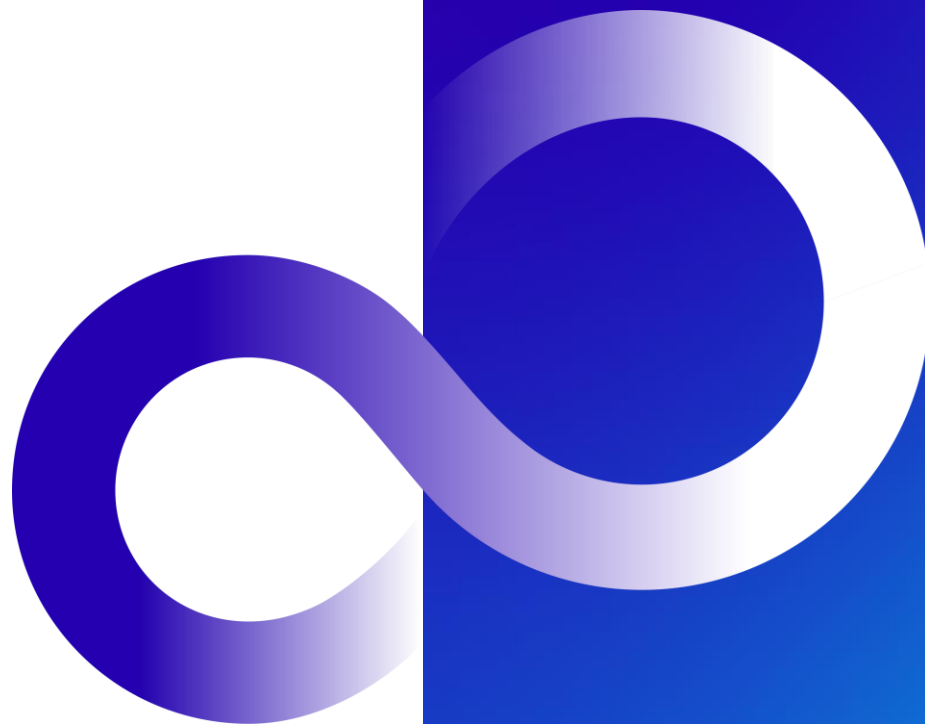
We believe joining ISF to the LF AI & Data and collaborating with community members accelerates the improvement of intersectional fairness methods. That also matches the mission of the LF AI & Data.

Therefore, we would sincerely like to host ISF with the LF AI & Data.

We seek your support in welcoming ISF as a sandbox project in LF AI & Data.

We're happy to address any questions you may have now.

Thank you



Upcoming TAC Meetings

 **DLF** AI & DATA

Upcoming TAC Meetings

- › June 29 - Microsoft recommenders project
- › July 13 - OPPO new sandbox project ShaderNN

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

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/j/430697670>

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.