

# Trusted AI Committee

## Overview

Below is an overview of the current discussion topics within the Trusted AI Committee. Further updates will follow as the committee work develops.

- Focus of the committee is on policies, guidelines, tooling and use cases by industry
- Survey and contact current open source Trusted AI related projects to join LF AI & Data efforts
- Create a badging or certification process for open source projects that meet the Trusted AI policies/guidelines defined by LF AI & Data
- Create a document that describes the basic concepts and definitions in relation to Trusted AI and also aims to standardize the vocabulary /terminology

## Mail List

Please self subscribe to the mail list here at <https://lists.lfai.foundation/g/trustedai-committee>.

Or email [trustedai-committee@lists.lfai.foundation](mailto:trustedai-committee@lists.lfai.foundation) for more information.

## Meetings

<https://lists.lfai.foundation/g/trustedai-committee/calendar>

Trusted AI Committee North America Monthly Meeting - 4th Thursday of the month, 10 PM Shenzhen China, 4 PM Paris, 10 AM ET, 7 AM PT USA (updated for daylight savings time as needed)

Zoom info : <https://zoom.us/j/7659717866>

We will starting ONE Monthly Call for Europe/Asia Time Zone, every 2nd Thursday of the month (Starting in November - time to be determined)

## Participants

Initial Organizations Participating: IBM, Orange, AT&T, Amdocs, Ericsson, TechM, Tencent

## Committee Chairs

Name	Region	Organization	Email Address	LF ID
Animesh Singh	North America	IBM	<a href="mailto:singhan@us.ibm.com">singhan@us.ibm.com</a>	
Souad Ouali	Europe	Orange	<a href="mailto:souad.ouali@orange.com">souad.ouali@orange.com</a>	
Jeff Cao	Asia	Tencent	<a href="mailto:jeffcao@tencent.com">jeffcao@tencent.com</a>	

## Committee Participants

Name	Organization	Email Address	LF ID
Ofer Hermoni	Amdocs	<a href="mailto:oferher@gmail.com">oferher@gmail.com</a>	
Mazin Gilbert	ATT	<a href="mailto:mazin@research.att.com">mazin@research.att.com</a>	
Alka Roy	Responsible Innovation Project	<a href="mailto:alka@responsibleproject.com">alka@responsibleproject.com</a>	
Mikael Anneroth	Ericsson	<a href="mailto:mikael.anneroth@ericsson.com">mikael.anneroth@ericsson.com</a>	
Alejandro Saucedo	The Institute for Ethical AI and Machine Learning	<a href="mailto:a@ethical.institute">a@ethical.institute</a>	
Jim Spohrer	IBM	<a href="mailto:spohrer@us.ibm.com">spohrer@us.ibm.com</a>	spohrer
Saishruthi Swaminathan	IBM	<a href="mailto:saishruthi.tn@ibm.com">saishruthi.tn@ibm.com</a>	
Susan Malaika	IBM	<a href="mailto:malaika@us.ibm.com">malaika@us.ibm.com</a>	sumalaika (but different email address)
Romeo Kienzler	IBM	<a href="mailto:romeo.kienzler@ch.ibm.com">romeo.kienzler@ch.ibm.com</a>	
Francois Jezequel	Orange	<a href="mailto:francois.jezequel@orange.com">francois.jezequel@orange.com</a>	
Nat Subramanian	Tech Mahindra	<a href="mailto:Natarajan.Subramanian@Techmahindra.com">Natarajan.Subramanian@Techmahindra.com</a>	
Han Xiao	Tencent	<a href="mailto:hanxiao@tencent.com">hanxiao@tencent.com</a>	
Wenjing Chu	Futurewei	<a href="mailto:chu.wenjing@gmail.com">chu.wenjing@gmail.com</a>	<a href="#">Wenjing Chu</a>
Yassi Moghaddam	ISSIP	<a href="mailto:yassi@issip.org">yassi@issip.org</a>	

## Assets

- All the assets being

#### Sub Categories

- Fairness: Methods to detect and mitigate bias in datasets and models, including bias against known protected populations

- Robustness: Methods to detect alterations/tampering with datasets and models, including alterations from known adversarial attacks

- Explainability: Methods to enhance understandability/interpretability by persona/roles in process of AI model outcomes/decision recommendations, including ranking and debating results/decision options

- Lineage: Methods to ensure provenance of datasets and AI models, including reproducibility of generated datasets and AI models

## Projects

Name	Github	Website
AI Fairness 360	<a href="https://github.com/IBM/AIF360">https://github.com/IBM/AIF360</a>	<a href="http://aif360.mybluemix.net/">http://aif360.mybluemix.net/</a>
Adversarial Robustness 360	<a href="https://github.com/IBM/adversarial-robustness-toolbox">https://github.com/IBM/adversarial-robustness-toolbox</a>	<a href="https://art-demo.mybluemix.net/">https://art-demo.mybluemix.net/</a>
AI Explainability 360	<a href="https://github.com/IBM/AIX360">https://github.com/IBM/AIX360</a>	<a href="http://aix360.mybluemix.net">http://aix360.mybluemix.net</a>

## Working Groups

[Trusted AI Principles Working Group](#)

[Trusted AI Technical Working Group](#)

## Meetings

Zoom info : Trusted AI Committee meeting - NA/Europe Meeting 4th Thursday of the month, 10 PM Shenzhen China, 4 PM Paris, 10 AM ET, 7 AM PT USA (updated for daylight savings time as needed)

<https://zoom.us/j/7659717866>


**How to Join:** Visit the [Trusted AI Committee Group Calendar](#) to self subscribe to meetings.

Or email [trustedai-committee@lists.lfaidata.foundation](mailto:trustedai-committee@lists.lfaidata.foundation) for more information.

**Meeting Content (minutes / recording / slides / other):**

Date	Agenda/Minutes
November 19, 2020	Recording ( <a href="#">video</a> ) Recording ( <a href="#">audio</a> ) Principles report =(R)REPEATS

22 Oct 2020	<p><a href="#">Recording</a> (audio)</p> <p><a href="#">Recording</a> (video)</p> <p>LFAI Trusted AI Committee Structure and Schedule: Animesh Singh</p> <p>Real World Trusted AI Usecase and Implementation in Financial Industry: Stacey Ronaghan</p> <p>AIF360 Update: Samuel Hoffman</p> <p>AIX360 Update: Vijay Arya</p> <p>ART Update: Beat Buesser</p>
October 1, 2020	<p><a href="#">Recording</a> (audio)</p> <p><a href="#">Recording</a> (video)</p> <p>Setting up Trusted AI TSC</p> <p>Principles Update</p> <p>Coursera Course Update</p> <p>Calendar discussion - Europe and Asia Friendly</p>
17 Sep 2020	<ul style="list-style-type: none"> <li>• <a href="#">Recording</a></li> <li>• Walkthrough of LFAI Trusted AI Website and github location of projects</li> <li>• Trusted AI Video Series</li> <li>• Trusted AI Course in collaboration with University of Pennsylvania</li> </ul>
20 Aug 2020	<ul style="list-style-type: none"> <li>• <a href="#">Recording</a></li> <li>• <b>Z-Inspection: A holistic and analytic process to assess Ethical AI - Roberto Zicari - University of Frankfurt, Germany</b></li> <li>• <b>Age-At-Home - the exemplar of TrustedAI, David Martin, Hacker in Charge at motion-ai.com</b></li> </ul>
06 Aug 2020	<ul style="list-style-type: none"> <li>• <b>Plotly Demo with SHAP and AIX360 - Xing Han, Plot.ly</b> <ul style="list-style-type: none"> <li>• Explain the Tips dataset with SHAP: <a href="#">Article</a>, <a href="#">Demo</a></li> <li>• Heart Disease Classification with AIX360: <a href="#">Post</a>, <a href="#">Demo</a></li> <li>• Community-made SHAP-to-dashboard API: <a href="#">Post</a></li> </ul> </li> <li>• <b>Swiss Digital Trust Label - short summary - Romeo Kienzler, IBM</b> <ul style="list-style-type: none"> <li>• Announced by Swiss President Doris Leuthard at WEF 2nd Sept. 2019</li> <li>• Geneva based initiative for sustainable and fair treatment of data</li> <li>• Among others, those companies are involved already Google, Uber, IBM, Microsoft, Facebook, Roche, Mozilla, <a href="#">Booking.com</a>, UBS, Credit Suisse, Zurich, Siemens IKRK, EPFL, ETH, UNO</li> <li>• <a href="#">Booking.com</a>, Credit Suisse, IBM, Swiss Re, SBB, Kudelski and Canton of Waadt to deliver pilot</li> </ul> </li> </ul>

23 Jul 2020	<p><b>Agenda</b></p> <ul style="list-style-type: none"> <li>• <b>Watson OpenScale and Trusted AI - Eric Martens, IBM</b></li> <li>• <b>LFAI Ethics Training course</b>  <a href="https://github.com/lfaiai-ethics-training/blob/master/ai-ethics-outline.md">https://github.com/lfaiai-ethics-training/blob/master/ai-ethics-outline.md</a></li> </ul> <div data-bbox="402 260 906 758" style="text-align: center;">  <p>LFAI TAIC 2020...deo zoom_0.mp4</p> </div> <ul style="list-style-type: none"> <li>• <b>recording:</b></li> <li>• <b>slack:</b> <a href="https://lfaifoundation.slack.com/archives/CPS6Q1E8G/p1595515808086900">https://lfaifoundation.slack.com/archives/CPS6Q1E8G/p1595515808086900</a></li> </ul>
09 Jul 2020	<p><b>Agenda</b></p> <ul style="list-style-type: none"> <li>• <b>Montreal AI Ethics Institute Presentation</b></li> <li>• <b>Status on Trusted AI projects in Open Governance</b></li> </ul>
28 May 2020	<p><b>Agenda</b></p> <ul style="list-style-type: none"> <li>• <b>Principles Working Group update - Susan Malaika</b></li> <li>• <b>Trusted AI Committee activities summarization for Governing board - Animesh Singh</b></li> </ul>
14 May 2020	<p><b>Agenda</b></p> <ul style="list-style-type: none"> <li>• <b>Swaminathan Chandrasekaran</b>, KPMG Managing Director would be talking about how they are working with practitioners in the field on their AI Governance and Trusted AI needs.</li> <li>• <b>Susan Malaika</b> from IBM will be giving an update from the Principles Working Group, and progress there.</li> </ul>
30 Apr 2020	<p><b>Agenda</b></p> <ul style="list-style-type: none"> <li>• <b>Saishruthi Swaminathan</b> to do a presentation on AI Transparency in Marketplace</li> <li>• <b>Francois Jezequel</b> to present on Orange Responsible AI initiative.</li> </ul>
16 Apr 2020	<p><b>Agenda</b></p> <ul style="list-style-type: none"> <li>• Andrew and Tommy did a deep dive in Kubeflow Serving and Trusted AI Integration</li> <li>• Principles Working Group discussion</li> </ul>
19 Mar 2020	<p><b>Proposed Agenda</b></p> <ul style="list-style-type: none"> <li>• AI for People is focused on the intersection of AI and Society with a lot of commonality with the focus areas of our committee. Marta will be joining to present their organization and what they are working on.</li> <li>• Proposal of Use Case to be tested by AT&amp;T using Apache Nifi and AIF360 (romeo)</li> </ul>
20 Feb 2020	<p><b>Proposed Agenda</b></p> <ul style="list-style-type: none"> <li>• Introduction to baseline data set for AI Bias detection (romeo)</li> <li>• Exemplar walk-through: retrospective bias detection with Apache Nifi and AIF360 (romeo)</li> <li>• Principles Working Group Status Update (Susan)</li> </ul>

23 Jan 2020	<p><b>Proposed Agenda</b></p> <ul style="list-style-type: none"> <li>• Discuss AIF360 work around SKLearn community (Samuel Hoffman, IBM Research demo)</li> <li>• Discuss "Many organizations have principles documents, and a bit of backlash - for not enough practical examples."</li> </ul> <p><b>Resource</b></p> <ul style="list-style-type: none"> <li>• Watch updates on production ML with Alejandro Saucedo done with Susan Malaika on the Cognitive Systems Institute call:</li> </ul> <p><b>Notes from the call:</b>  <a href="https://github.com/lfai/trusted-ai/blob/master/committee-meeting-notes/notes-20200123.md">https://github.com/lfai/trusted-ai/blob/master/committee-meeting-notes/notes-20200123.md</a></p>
12 Dec 2019	<p><b>Proposed Agenda</b></p> <ul style="list-style-type: none"> <li>• Meeting notes are now in GitHub here: <a href="https://github.com/lfai/trusted-ai/tree/master/committee-meeting-notes">https://github.com/lfai/trusted-ai/tree/master/committee-meeting-notes</a></li> <li>• Since we don't record and share our committee meetings should our committee channel in Slack be made private for asynchronous conversation outside these calls?</li> <li>• Introduction of MLOps in IBM Trusted AI projects</li> <li>• Design thinking around integrating Trusted AI projects in Kubeflow Serving</li> </ul> <p><b>Notes from the call:</b>  <a href="https://github.com/lfai/trusted-ai/blob/master/committee-meeting-notes/notes-20191212.md">https://github.com/lfai/trusted-ai/blob/master/committee-meeting-notes/notes-20191212.md</a></p>
14 Nov 2019	<p><b>Proposed Agenda:</b></p> <ul style="list-style-type: none"> <li>▪ Jim to get feedback from LFAI Board meeting</li> <li>▪ Romeo to demo AIF360 -Nifi Integration + feedback from his talk at OSS Lyon</li> <li>▪ Alka to present AT&amp;T Working doc</li> <li>▪ Discuss holiday week meeting potential conflicts (28 Nov - US Holiday, 26 Dec - Day after Christmas)</li> </ul> <p><b>Notes from call:</b>  <a href="https://github.com/lfai/trusted-ai/blob/master/committee-meeting-notes/notes-20191114.md">https://github.com/lfai/trusted-ai/blob/master/committee-meeting-notes/notes-20191114.md</a></p>
17 Oct 2019	<p><b>Attendees:</b>  Ofer, Alka, Francois, Nat, Han, Animesh, Jim, Maureen, Susan, Alejandro</p> <p><b>Summary</b></p> <ul style="list-style-type: none"> <li>• Animesh walked through the draft slides (to be presented in Lyon to LFAI governing board about TAIC)</li> <li>• Discussion of changes to make</li> <li>• Discussion of members, processes, and schedules</li> </ul> <p><b>Detail</b></p> <ul style="list-style-type: none"> <li>• Jim will put slides in Google Doc and share with all participants</li> <li>• Susan is exploring a slack channel for communications</li> <li>• Trust and Responsibility, Color, Icons to add Amdocs, Alejandro's Institute</li> <li>• Next call Cancelled (31 October) as many committee members will be at OSS EU and TensorFlow World</li> </ul> <p><b>Notes from the call:</b>  <a href="https://github.com/lfai/trusted-ai/blob/master/committee-meeting-notes/notes-20191017.md">https://github.com/lfai/trusted-ai/blob/master/committee-meeting-notes/notes-20191017.md</a></p>
03 Oct 2019	<p><b>Attendees:</b>  Animesh Singh (IBM), Maureen McElaney (IBM), Han Xiao (Tencent), Alejandro Saucedo, Mikael Anneroth (Eriksson), Ofer Hermoni (Am docs)</p> <p>Animesh will check with Souad Ouali to ensure Orange wants to lead the Principles working group and host regular meetings. Committee members on the call were not included in the email chains that occurred so we need to confirm who is in charge and how communication will occur.</p> <p>The Technical working group has made progress but nothing concrete to report.</p> <p>A possible third working group could form around AI Standards.</p> <p><b>Notes from the call:</b>  <a href="https://github.com/lfai/trusted-ai/blob/master/committee-meeting-notes/notes-20191003.md">https://github.com/lfai/trusted-ai/blob/master/committee-meeting-notes/notes-20191003.md</a></p>

19 Sep  
2019

**Attendees:**

Ibrahim. H, Nat .S, Animesh.S, Alka.R, Jim.S, Francios. J, Jeff. C, Maureen. M, Mikael. A, Ofer. H, Romeo.K

- Goals defined for the meeting:
- Working Group Names and Leads have been confirmed:

Principles, lead: Souad Ouali (Orange France) with members from Orange, AT&T, Tech Mahindra, Tencent, IBM, Ericsson, Amdocs, Technical, lead: Romeo Kienzler (IBM Switzerland) with members from IBM, AT&T, Tech Mahindra, Tencent, Ericsson, Amdocs, Orange.

Working groups will have a weekly meeting to make progress. First read out to LF AI governing board will be Oct 31 in Lyon France.

The Principles team will study the existing material from companies, governments, and professional associations (IEEE), and some up with set that can be shared with the technical team for feedback as a first step. We need to identify and compile the existing materials.

The Technical team is working on Acumos+Angel+AIF360 integration demonstration.

3. Possible Discussion about third working group

- Discussion about LFAI day in Paris
- More next steps
  - Will begin recording meetings in future calls.

**Notes from call:** <https://github.com/lfai/trusted-ai/blob/master/committee-meeting-notes/notes-20190919.md>