

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 efforts
- Create a badging or certification process for open source projects that meet the Trusted AI policies/guidelines defined by LF AI
- 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 for more information.

Meetings

Zoom info : Trusted AI Committee meeting - Alternate Thursday's, 10 PM Shenzen China, 4 PM Paris, 10 AM ET, 7 AM PT USA (updated for daylight savings time as needed)

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

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	singhan@us.ibm.com	
Souad Ouali	Europe	Orange	souad.ouali@orange.com	
Jeff Cao	Asia	Tencent	jeffcao@tencent.com	

Committee Participants

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

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	https://github.com/IBM/AIF360	http://aif360.mybluemix.net/
Adversarial Robustness 360	https://github.com/IBM/adversarial-robustness-toolbox	https://art-demo.mybluemix.net/
AI Explainability 360	https://github.com/IBM/AIX360	http://aix360.mybluemix.net

Working Groups

Trusted AI Principles Working Group

Trusted AI Technical Working Group

Meetings


Zoom info : Trusted AI Committee meeting - Alternate Thursday's, 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.lfai.foundation for more information.

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

Date	Agenda/Minutes
06 Aug 2020	<ul style="list-style-type: none"> • Swiss Digital Trust Label - short summary - Romeo Kienzler, IBM <ul style="list-style-type: none"> • Announced by Swiss President Doris Leuthard at WEF 2nd Sept. 2019 • Geneva based initiative for sustainable and fair treatment of data • Among others, those companies are involved already Google, Uber, IBM, Microsoft, Facebook, Roche, Mozilla, Booking.com, UBS, Credit Suisse, Zurich, Siemens IKRK, EPFL, ETH, UNO • Booking.com, Credit Suisse, IBM, Swiss Re, SBB, Kudelski and Canton of Waadt to deliver pilot • Z-Inspection: A holistic and analytic process to assess Ethical AI - Roberto Zicari - University of Frankfurt, Germany

23 Jul 2020	<p>Agenda</p> <ul style="list-style-type: none"> • Watson OpenScale and Trusted AI - Eric Martens, IBM • LFAI Ethics Training course https://github.com/lfaiai-ethics-training/blob/master/ai-ethics-outline.md <div data-bbox="396 260 894 758" style="text-align: center;">  <p>LFAI TAIC 2020...deo zoom_0.mp4</p> </div> <ul style="list-style-type: none"> • recording: • slack: https://lfaifoundation.slack.com/archives/CPS6Q1E8G/p1595515808086900
09 Jul 2020	<p>Agenda</p> <ul style="list-style-type: none"> • Montreal AI Ethics Institute Presentation • Status on Trusted AI projects in Open Governance
28 May 2020	<p>Agenda</p> <ul style="list-style-type: none"> • Principles Working Group update - Susan Malaika • Trusted AI Committee activities summarization for Governing board - Animesh Singh
14 May 2020	<p>Agenda</p> <ul style="list-style-type: none"> • Swaminathan Chandrasekaran, KPMG Managing Director would be talking about how they are working with practitioners in the field on their AI Governance and Trusted AI needs. • Susan Malaika from IBM will be giving an update from the Principles Working Group, and progress there.
30 Apr 2020	<p>Agenda</p> <ul style="list-style-type: none"> • Saishruthi Swaminathan to do a presentation on AI Transparency in Marketplace • Francois Jezequel to present on Orange Responsible AI initiative.
16 Apr 2020	<p>Agenda</p> <ul style="list-style-type: none"> • Andrew and Tommy did a deep dive in Kubeflow Serving and Trusted AI Integration • Principles Working Group discussion
19 Mar 2020	<p>Proposed Agenda</p> <ul style="list-style-type: none"> • 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. • Proposal of Use Case to be tested by AT&T using Apache Nifi and AIF360 (romeo)
20 Feb 2020	<p>Proposed Agenda</p> <ul style="list-style-type: none"> • Introduction to baseline data set for AI Bias detection (romeo) • Exemplar walk-through: retrospective bias detection with Apache Nifi and AIF360 (romeo) • Principles Working Group Status Update (Susan)

<p>23 Jan 2020</p>	<p>Proposed Agenda</p> <ul style="list-style-type: none"> • Discuss AIF360 work around SKLearn community (Samuel Hoffman, IBM Research demo) • Discuss "Many organizations have principles documents, and a bit of backlash - for not enough practical examples." <p>Resource</p> <ul style="list-style-type: none"> • Watch updates on production ML with Alejandro Saucedo done with Susan Malaika on the Cognitive Systems Institute call: <p>Notes from the call: https://github.com/lfaif/trusted-ai/blob/master/committee-meeting-notes/notes-20200123.md</p>
<p>12 Dec 2019</p>	<p>Proposed Agenda</p> <ul style="list-style-type: none"> • Meeting notes are now in GitHub here: https://github.com/lfaif/trusted-ai/tree/master/committee-meeting-notes • 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? • Introduction of MLOps in IBM Trusted AI projects • Design thinking around integrating Trusted AI projects in Kubeflow Serving <p>Notes from the call: https://github.com/lfaif/trusted-ai/blob/master/committee-meeting-notes/notes-20191212.md</p>
<p>14 Nov 2019</p>	<p>Proposed Agenda:</p> <ul style="list-style-type: none"> ▪ Jim to get feedback from LFAI Board meeting ▪ Romeo to demo AIF360 -Nifi Integration + feedback from his talk at OSS Lyon ▪ Alka to present AT&T Working doc ▪ Discuss holiday week meeting potential conflicts (28 Nov - US Holiday, 26 Dec - Day after Christmas) <p>Notes from call: https://github.com/lfaif/trusted-ai/blob/master/committee-meeting-notes/notes-20191114.md</p>
<p>17 Oct 2019</p>	<p>Attendees: Ofer, Alka, Francois, Nat, Han, Animesh, Jim, Maureen, Susan, Alejandro</p> <p>Summary</p> <ul style="list-style-type: none"> • Animesh walked through the draft slides (to be presented in Lyon to LFAI governing board about TAIC) • Discussion of changes to make • Discussion of members, processes, and schedules <p>Detail</p> <ul style="list-style-type: none"> • Jim will put slides in Google Doc and share with all participants • Susan is exploring a slack channel for communications • Trust and Responsibility, Color, Icons to add Amdocs, Alejandro's Institute • Next call Cancelled (31 October) as many committee members will be at OSS EU and TensorFlow World <p>Notes from the call: https://github.com/lfaif/trusted-ai/blob/master/committee-meeting-notes/notes-20191017.md</p>
<p>03 Oct 2019</p>	<p>Attendees: Animesh Singh (IBM), Maureen McElaney (IBM), Han Xiao (Tencent), Alejandro Saucedo, Mikael Anneroth (Eriksson), Ofer Hermoni (Amdocs)</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>Notes from the call: https://github.com/lfaif/trusted-ai/blob/master/committee-meeting-notes/notes-20191003.md</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 Acuomos+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>