LFAI & Data

Webinar -
The Trusted AI Principles - Practical Examples

Trusted AI Committee - Principles Working Group (where you will find the slides and materials)

28 April 2021

LF AI & DATA

Agenda

Introduction to session (5 mins)
Practical examples in the context of the RREPEATS Principles (20 mins)

- Classification of Encrypted Traffic Application
  - Iman Akbari Azirani - University of Waterloo, Canada
  - Noura Limam - University of Waterloo, Canada
  - Bertrand Mathieu – Orange, France
- Rosae NLG Framework (an LF-AI project)
  - Natural Language Generation for AI decision explanation
  - Ludan Stoecklé, AI Lab at BNP Paribas CIB, France

Round-table – Applying Principles in a Corporation (20 mins)

- Calvin Lawrence, IBM, US
- Alejandro Saucedo, Seldon, Institute for Ethical AI & Machine Learning, UK
- Emilie Sirvent-Hien, Orange, France

Q/A from audience and closing (15 mins)

Session Host: Souad Ouali
Head of interoperators relationships Orange - Counsel / Responsable de relations inter opérateurs chez Orange - Conseil

The 8 LFAI Principles for Trusted AI – (R)REPEATS

Reproducibility	**Robustness**	**Equitability**	**Privacy**
Explainability	**Accountability**	**Transparency**	**Security**

The principles are of equal importance and value.

No principle is of higher priority than another.

The principles are related to each other.

LF AI & DATA

LFAI & Data

The Trusted AI Principles

Practical Examples

- Classification of Encrypted Traffic Application
- Natural Language Generation for AI Decision Explanation - Rosea NLG Framework

LF AI & DATA

Practical Examples Presenters

Classification of Encrypted Traffic Application

Iman Akbari Azirani, Artificial Intelligence x Cyber-security, University of Waterloo

Noura Limam, University of Waterloo, Canada

Bertrand Mathieu, Orange Labs, France

Natural Language Generation for AI Decision Explanation - Rosea NLG Framework

Ludan Stoecklé, CTO of Data & AI Lab BNP Paribas CIB. Author of RosaeNLG.

Introductions for Practical Examples Segment

Iman Akbari Azirani, University of Waterloo, Canada
Iman Akbari received his B.Sc. in 2018 from Sharif University, Iran, in Software Engineering. He is currently a researcher and graduate student at University of Waterloo, Canada. His research mostly revolves around the intersection of AI, cybersecurity and network management with a focus on automation and scalability.

Noura Limam, University of Waterloo, Canada
Noura Limam received her M.Sc. and Ph.D. degrees in computer science from the University Pierre & Marie Curie (Sorbonne University), France in 2002 and 2007, respectively. She is currently a research assistant professor of computer science at the University of Waterloo, Canada. She is an active researcher and contributor in the area of network and service management. Her current research interest is in applying ML to networking and network management problems including network security and traffic classification.

Bertrand Mathieu, Orange Labs, France
Bertrand Mathieu joined Orange Labs in 1994. He received the PhD degree from the University Pierre et Marie Curie in Paris, and the « Habilitation à Diriger des Recherches » from the University of Rennes. He is a senior researcher whose main research topics are programmable networks, QoS and QoE, transport protocols. He is currently using AI for providing solutions for network classification and troubleshooting of encrypted protocols. He is member of several conferences Technical Program Committee and an IEEE and SEE senior member.

Introductions for Practical Examples Segment

Ludan Stoecklé, AI Lab at BNP Paribas CIB, France
Ludan Stoecklé is a professional of the AI software industry.
He was the CTO of Yseop, a Natural Language Generation (NLG) solution editor, before building the technical team of Addventa, an AI consulting firm.
Today CTO of the AI Lab at BNP Paribas CIB, Ludan is developing and industrializing AI products internally.
Ludan is also Expert Professor at aivancity school for technology, business & society, and the original author of RosaeNLG, an open-source NLG library sandboxed by LF AI & Data Foundation.
Beyond his passion for NLG and running, Ludan is also known worldwide as a paperweight collector.

Practical Example: Classification of Encrypted Traffic

Bertrand Mathieu – Orange Labs Lannion, France

Noura Limam, Iman Akbari – University of Waterloo, Canada

28/04/2021

LF AI & Data event
The Trusted AI Principles – Practical Examples

orange™

LF AI & DATA

UNIVERSITY OF WATERLOO

Context & Objectives

Context

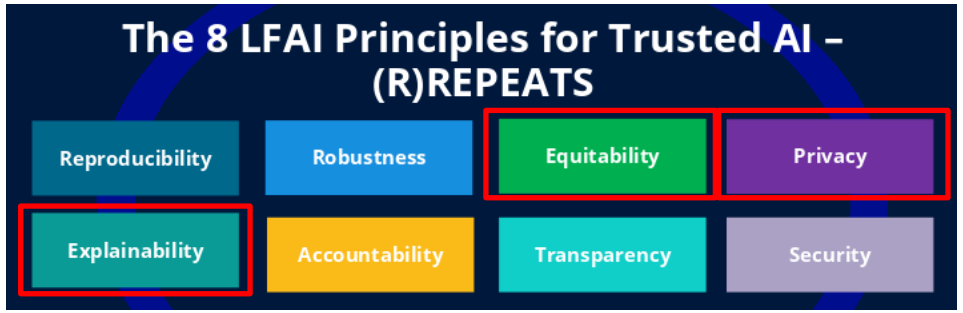
- Network traffic increasingly encrypted (HTTP2, HTTPS3/QUIC)
  - Today, 85% of Internet traffic is encrypted (almost 100% of Google's)
- Less « clear » information inside packets makes traffic classification a tricky task.
  - Still few header fields in « clear ». But for how long ?
- Deep Packet Inspection (DPI) solutions limited because of this encryption

Objectives

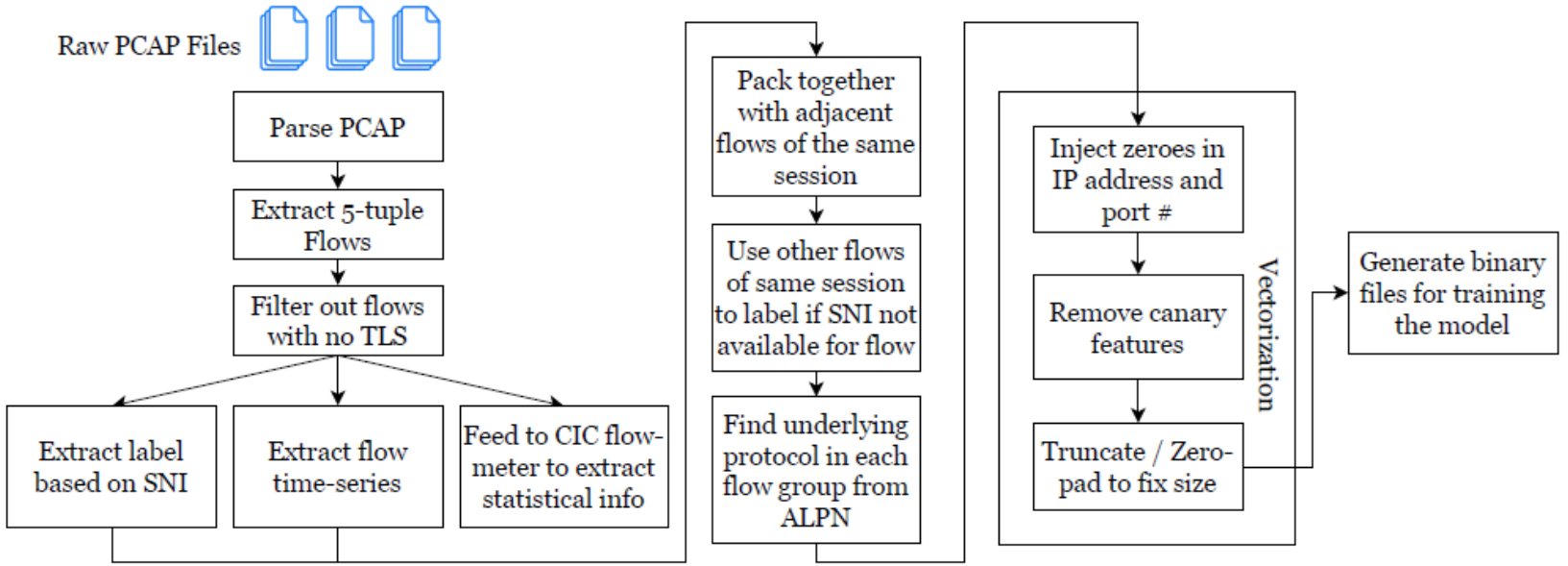
A network operator needs to know the use of its network : the main used services and applications, and the data volume

- To anticipate/prepare network evolutions, capacity planning and interconnections with main Internet actors
- To adapt its marketing offers
- To detect O-rating frauds
- For customers satisfaction

Trusted AI in mind



- Processing of the dataset for network traces anonymization
  - Keep only flows with useful packets for our classification
  - Truncate data payload
  - Remove temporal information of sessions : set to unix epoch time (1/1/70)
  - Remove end-users & servers network related information : IP addresses, security certificates information, information in protocol header
- Classification on a global level, not per user

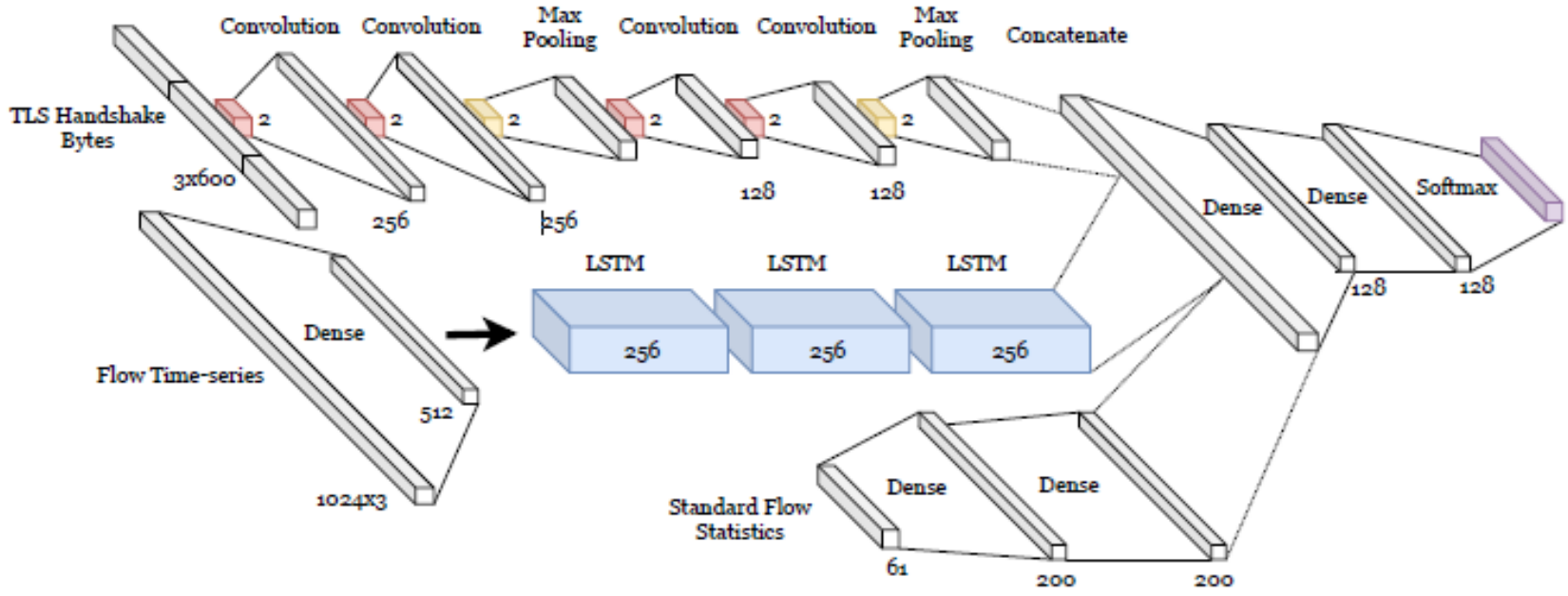


SNI (Server Name Indication)

ALPN (Application-Layer Protocol Negotiation)

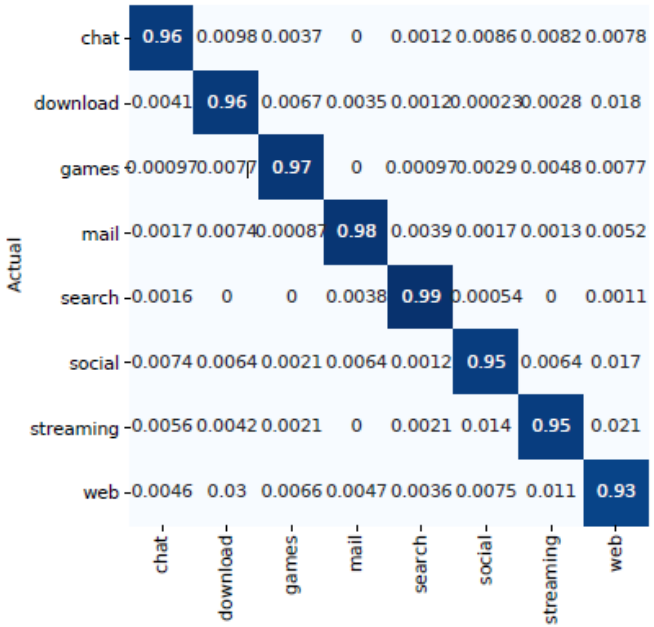
Solution

- AI-based solution; a deep neural network architecture composed of convolutional neural network (CNN) and a Long Short Term Memory network (LSTM) layers
- A 3-faceted model fed with 3 input data types, extracted/computed from network sessions
- Identify the service aka application category (video streaming, social network, web, etc.) and known applications (netflix, youtube, facebook, gmail, etc.) of one session based on flow-level data and characteristics

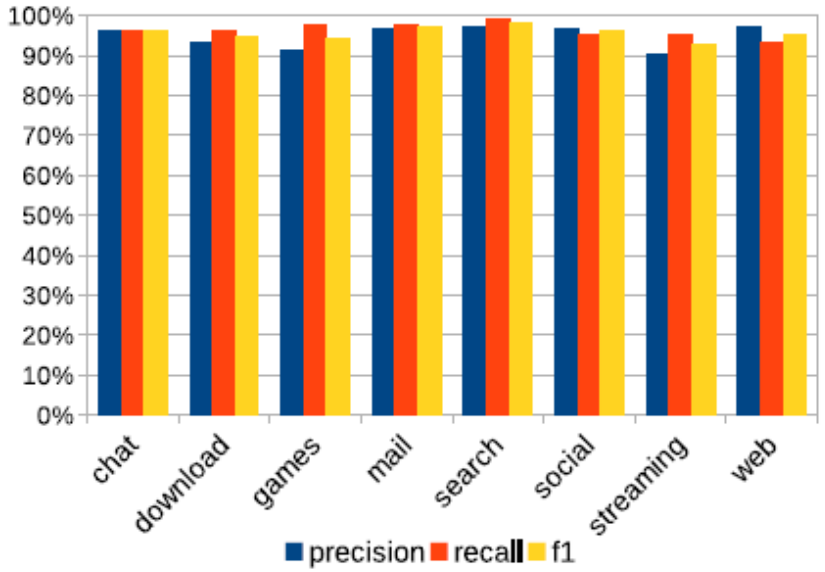


Evaluation

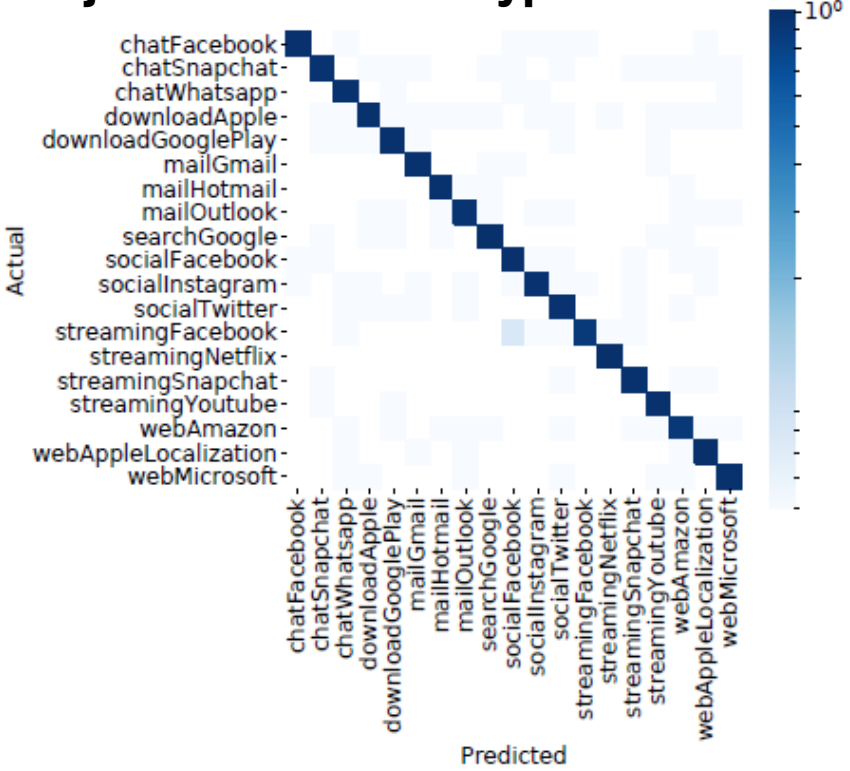
- Dataset collected on the French mobile Orange network
- About 95% of the HTTPS traffic can be correctly classified
- Protocol-agnostic features generalize for different classification objectives and encrypted web protocols



Services Confusion Matrix



Precision, Recall and F1-Score per service



Applications Confusion Matrix

Possible Requests for LF AI & Data Trusted AI Principles Working Group

- ❑ **Have a generic AI tool to anonymize network packets**
  - ❖ **Currently manually done via scripts**
- ❑ **Being able to make AI models run at network wire-speed, inside network equipment**
- ❑ **Tools to better explain and understand biases and decisions**
- ❑ **Robustness to adversarial attacks**

Thanks

Contacts:

bertrand2.mathieu@orange.com

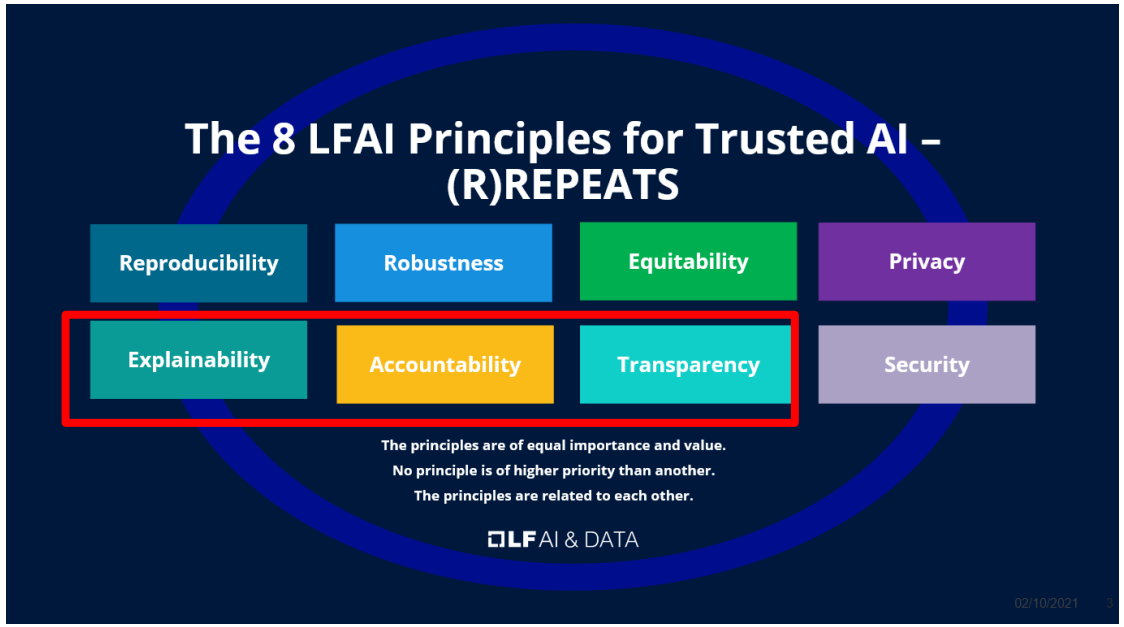
noura.limam@uwaterloo.ca

Practical Example: Natural Language Generation for AI decision explanation

Ludan Stoecklé, CTO of Data & AI Lab BNP Paribas CIB. Author of RosaeNLG.

How to explain a decision to a non-expert end user?

- Trusting **a prediction** (not Trusting a model)
- A **good explanation** must be:
  a. Understandable **for its target population**
  b. **Automated** to be scalable



Corporate and data scientist biases

- We, data scientists, we love:
  - Data
  - Tables
  - Dashboards
- Corporate environments favor data and complex dashboards
- Corporate power users have expertise and can understand dashboards

Corporate users and data scientists tend to believe that data and dashboards are a good way to convey information.

Is it true?

OECD PIAAC

Look at the graph about the number of births. Click to answer the question below.

During which period(s) was there a decline in the number of births? Click all that apply.

- ☐ 1957 - 1967
- ☐ 1967 - 1977
- ☐ 1977 - 1987
- ☐ 1987 - 1997
- ☐ 1997 - 2007

The following graph shows the number of births in the United States from 1957 to 2007. Data are presented every 10 years.

https://www.oecd.org/skills/piaac/Numeracy%20Sample%20Items.pdf

Welcome to the Real World!

- **Very simple line graphs are not understood by the general population**
- Non-experts do not understand statistics or data visualisations
- Studies show that 25% of adults cannot compare information presented in a simple table

Excluding large parts of the population of understandable explanations is not acceptable

“Non-experts users” are not an homogeneous category.
Who is your audience? What are their skills, what is their level of expertise?

https://ehudreiter.com/2017/10/02/non-experts-struggle-graphs/

Non-expert users prefer textual explanations

A Bank Customer wants to understand:

Why was my application rejected?
What can I improve to increase the likelihood my application is accepted?

https://aix360.mybluemix.net/explanation_cust#

Language can communicate background, context, caveats, and give advice.

Example #1

Your credit application has been denied.

The main reason is your existing debt, which is high. But when your $300/month car credit will be fully repaid in January 2022, your application will probably be accepted.

Example #2

Your credit application has been denied.

This is due to a couple of reasons. Even if your existing debt is rather low, some your internet bills in 2019 were not paid on time. Also, you have moved twice during the last 3 years. At last, you tend to use a large part of your available credit (an average of 80% during the last two years).

If you continue to pay your bills on time like you did in 2020, if you stop moving, and if you reduce the used part of your available credit (50% maximum is a good rule of thumb), your application is likely to be accepted in January 2022.

Also, if you reduce your missed payments, your credit charges will also be lower.

How?

- Producing such texts is impossible to do manually at scale and in real time

- Use NLG: Natural Language Generation
  - Transforms data to text, using textual templates
  - Used widely in production e.g. the financial industry

Computer-generated texts can be superior (from the reader's perspective) to human-written texts

Explainability pipeline using NLG

NLG can be used the end of an AI pipeline, to automate and convey expertise, explain and summarize situations, and communicate with end users.

1. Interpret the decision

When ML, using an explainability framework: AI Explainability 360, SHAP, LIME, etc.

2. "What to say"

- find key insights
- present them as a story
- **be sensitive to user goals and knowledge**

Filter data, apply static rules, simple statistics etc. - using Python for instance

3. "How to say"

Produce the texts, tell the story

Using textual templates and an **NLG engine**

1. *At design time, define what is a clear explanation for your users and your use case*
2. *Ask your users for feedback: “Was the explanation clear?”*
3. *Adapt to your real audience at runtime*

Introducing RosaeNLG

- RosaeNLG is an open source (Apache 2.0) Natural Language Generation (NLG) project
- Designed to be **developer and IT friendly**
- Supports multiple languages with linguistic resources: currently English, French, German, Italian and Spanish
- Extensive documentation https://rosaenlg.org
- An awesome logo!
- **LF AI & Data Sandbox project**

RosaeNLG.org

LF AI & DATA

SANDBOX PROJECT

LFAI & Data

The Trusted AI Principles

The Round Table

LF AI & DATA

Session Host :
Susan Malaika, Senior Technical Staff, IBM

Round Table Panelists – Applying Trusted AI Principles in a Corporation

Calvin Lawrence
CTO & Distinguished Engineer
Cognitive Solutions at IBM

Alejandro Saucedo
Engineering Director at Seldon
Chief Scientist at The Institute for Ethical AI

Emilie Sirvent-Hien
Responsible AI program manager at Orange

Introductions for Round Table Segment

Calvin Lawrence, IBM, US
Calvin Lawrence is a Distinguished Engineer and Chief Architect for Cognitive Computing and Innovation for IBM Global Markets. Calvin was raised in the Atlanta area in the US, where he grew up in the projects in an impoverished community that was riddled with crime. The contrast of having a sometimes-adversarial relationship with law enforcement, while also seeing the benefits of government assistance shaped his mindset and informs his commitment today to using tech for good.

Alejandro Saucedo, Seldon, Institute for Ethical AI & Machine Learning, UK
Alejandro Saucedo is the Director of Machine Learning Engineering at Seldon Technologies, where he leads large scale projects implementing open source and enterprise infrastructure for Machine Learning Orchestration and Explainability. Alejandro is also the Chief Scientist at the Institute for Ethical AI & Machine Learning, where he leads the development of industry standards on machine learning bias, adversarial attacks and differential privacy.

Emilie Sirvent-Hien, Orange, France
Emilie Sirvent-Hien received a MS degree from Telecom Paris in 2003. She joined Orange 15 years ago and she held several positions in different domains including network engineering, customer relationship and innovation and animation of a research program on privacy. She is currently leading Orange's research activities on Ethics and Responsible AI and is also involved in several french working group on Responsible AI in Impact AI association, Women and AI and on AI development certification.

References and Resources

- Trusted AI Committee - Principles Working Group (where you will find the slides and materials)
- [LF-AI] The Trusted-AI Principles document bit.ly/lfai-trustedai-principles
- [LF-AI Blog] LF AI & Data Announces Principles for Trusted AI
- [LF-AI Webinar] RREPEATS – An Introduction to the Principles for Trusted AI – Thoughts and Next Steps
- [ACM] ACM Principles for Algorithmic Transparency and Accountability https://www.acm.org/binaries/content/assets/publicpolicy/2017_usacm_statement_algorithms.pdf
- [EU] Ethics Guidelines for Trustworthy AI - High-Level Expert Group on Artificial Intelligence set up by the European Commission https://ec.europa.eu/futurium/en/ai-alliance-consultation
- [EUFeb2020] On Artificial Intelligence -A European approach to excellence and tru https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf
- [IEEE] Ethically Aligned Design, IEEE https://ethicsinaction.ieee.org/
- [DoD] AI Principles: Recommendations on the Ethical Use of Artificial Intelligence by the Department of Defense https://media.defense.gov/2019/Oct/31/2002204458/-1/-1/0/DIB_AI_PRINCIPLES_PRIMARY_DOCUMENT.PDF
- [OECD] Organisation for Economic Co-operation and Development https://www.oecd.org/going-digital/ai/principles/
- [SoA] State of the Art: Reproducibility in Artificial Intelligence Odd Erik Gundersen, Sigbjørn Kjensmo, Department of Computer Science Norwegian University of Science and Technology https://www.researchgate.net/publication/326450530_State_of_the_Art_Reproducibility_in_Artificial_Intelligence

Contributions

Principles Working Group Team:

- Souad Ouali (Orange)
- Jeff Cao (Tencent)
- Francois Jezequel (Orange)
- Sarah Luger (Orange)
- Susan Malaika (IBM)
- Alka Roy (The Responsible Innovation Project/ex-AT&T)
- Alejandro Saucedo (The Institute for Ethical AI / Seldon)
- Marta Ziosi (AI for People)

Thank you

Webinar -
The Trusted AI Principles – Tools & Techniques

When: Sep 15, 2021 10:00 AM Eastern Time (US and Canada)

Register in advance for this meeting:
https://zoom.us/meeting/register/tJUpc-GoqjwtGtH_xVM6KiqCc-sPlcGRmQB1

After registering, you will receive a confirmation email containing information about joining the meeting.
.

The Trusted AI Principles - Practical Examples

Head of interoperators relationships Orange - Counsel / Responsable de relations inter opérateurs chez Orange - Conseil

Join us at 10am US Eastern on April 28 to meet with Souad Ouali Chair of the Trusted AI Principles Working Group at the LF-AI & Data - to hear about the application of the RREPEATS Principles to two practical examples:

- Classification of Encrypted Traffic Application, Iman Akbari Azirani & Noura Limam, University of Waterloo ; Bertrand Mathieu, Orange Labs, France
- Rosae NLG Framework (an LF-AI project) - Ludan Stoecklé, CTO of Data & AI Lab BNP Paribas CIB. Author of RosaeNLG

The session will also include a round table discussion with Calvin Lawrence CTO & Distinguished Engineer Cognitive Solutions at IBM, Alejandro Saucedo, Engineering Director at Seldon, Chief Scientist at The Institute for Ethical AI, Emilie Sirvent-Hien, Responsible AI program manager at Orange

Iman Akbari Azirani, Artificial Intelligence x Cyber-security, University of Waterloo

Noura Limam, University of Waterloo, Canada

Bertrand Mathieu, Orange Labs, France

Ludan Stoecklé, CTO of Data & AI Lab BNP Paribas CIB. Author of RosaeNLG

Session at 10am US Eastern April 28
Register :
https://zoom.us/meeting/register/tJwldu-urTIiE9xVfY07SedHw2-jjhyHb5NC

Calvin Lawrence
CTO & Distinguished Engineer
Cognitive Solutions at IBM

Alejandro Saucedo
Engineering Director at Seldon
Chief Scientist at The Institute for Ethical AI

Emilie Sirvent-Hien
Responsible AI program manager at Orange