NOKIA

Responsible AI for Telecom

10-13 October 2022

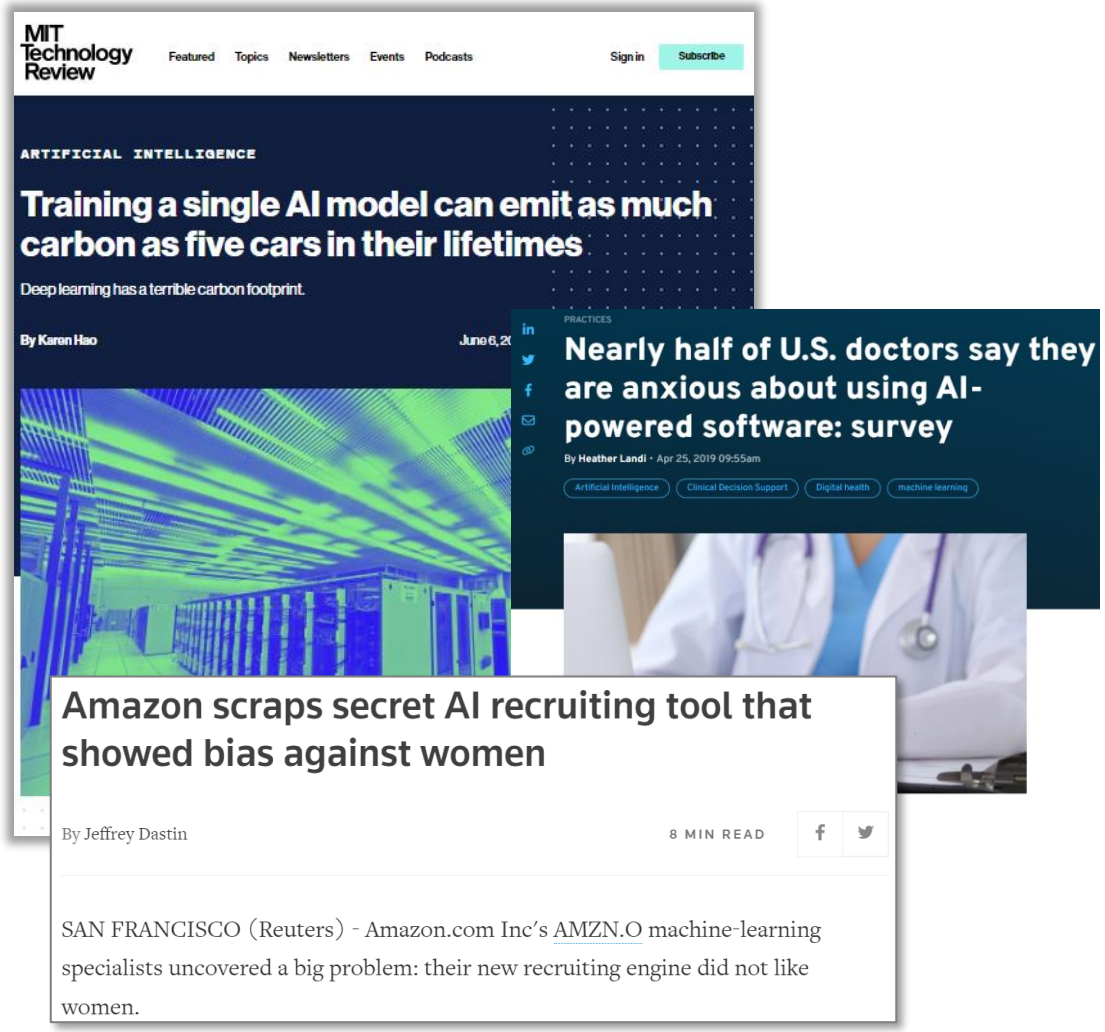
Anne Lee

Bell Labs Fellow

Introduction

Responsible AI = Trustworthy AI = Ethical AI

Why is responsibility in the AI field required?



- Harmful results from detrimental usage of AI
- Examples:
  - Sustainability
  - Trust
  - Reinforcing harmful biases and behaviors impacting fairness

Governments / Regulations Perspectives

- North America
  - ✓ USA Memorandum M-21-06 “Guidance for Regulation of AI Applications” with guidance to Federal agencies in the development of regulatory approaches to the use of AI in the private sector. (2020)
  - ✓ Canada bill BLL C-27 “Artificial Intelligence and Data Act” (2022)
- South America
  - ✓ Brazil Bill created legal framework for AI (2021)

- Europe
  - ✓ EU Commision Regulatory Framework for AI/ML Systems to be trustworthy based on the risk level of use cases. (2021)
  - ✓ UK National AI Strategy intends to establish ‘the most trusted and pro-innovation system for AI governance in the world’ (10 year strategy)
- Africa
  - ✓ The Future Society is developing 3 national AI strategies – for Rwanda, Tunisia, and Ghana – tailored towards developing AI governance frameworks, policies and AI ethical guidelines. (2022)

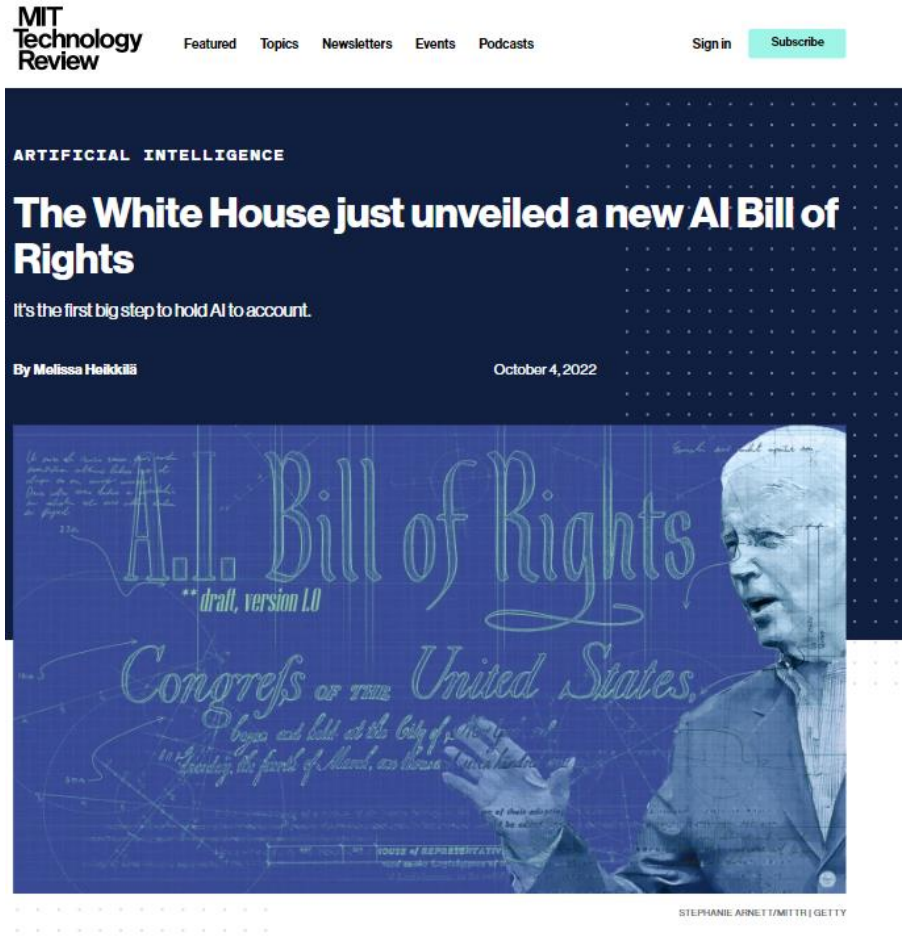
- Asia
  - ✓ China CAICT – an influential govt thinktank of MIIT - released white paper on Trustworthy AI. (2021)
  - ✓ India NITI Aayog – the public policy thinktank – released an approach document “Principles for Responsible AI”. (2021)
- Australia
  - ✓ Australia Artificial Intelligence Ethics Framework guides businesses and govts to responsibly design, develop, and implement AI. (2019)
  - ✓ New Zealand govt is developing an approach for ethical adoption of AI focusing on building AI ecosystem on foundation of trust. (2022)

Breaking News…

Oct 4, 2022: The White House wants Americans to know - The age of AI accountability is coming.

President Joe Biden has today unveiled a new AI Bill of Rights, which outlines five protections Americans should have in the AI age.

1. You should be protected from unsafe or ineffective systems.
2. You should not face discrimination by algorithms and systems should be used and designed in an equitable way.
3. You should be protected from abusive data practices via built-in protections and you should have agency over how data about you is used.
4. You should know that an automated system is being used and understand how and why it contributes to outcomes that impact you.
5. You should be able to opt out, where appropriate, and have access to a person who can quickly consider and remedy problems you encounter.



Addressing responsibility in AI

RAII Certification Beta

The world's first independent, accredited certification program of its kind.

Developed under the Global AI Action Alliance for the World Economic Forum (WEF), along with a diverse community of leading experts, RAII Certification is based on objective assessments of fairness, bias, explainability, and other concrete metrics of responsibly built AI systems.

- Responsible AI Institute
  - The World's First Independent, Accredited Certification Program for Responsible AI
- Early tools to help address responsibility
  - Microsoft Responsible AI toolbox
  - Google Tensorflow Responsible AI toolkit
  - IBM AI Fairness 360
  - PwC Responsible AI toolkit
  - And more…

AI and Responsibility in Telecom

- The three perspectives of AI and Responsibility for Telecom
- The six components of Responsible AI for Telecom
  1. Transparency / Explainability
  2. Sustainability
  3. Safe / Secure / Reliable
  4. Privacy
  5. Fairness
  6. Accountability
- Telecom standards
- Conclusion

The three perspectives of AI and Responsibility

Making AI responsible

Using AI to make a Telecom system responsible

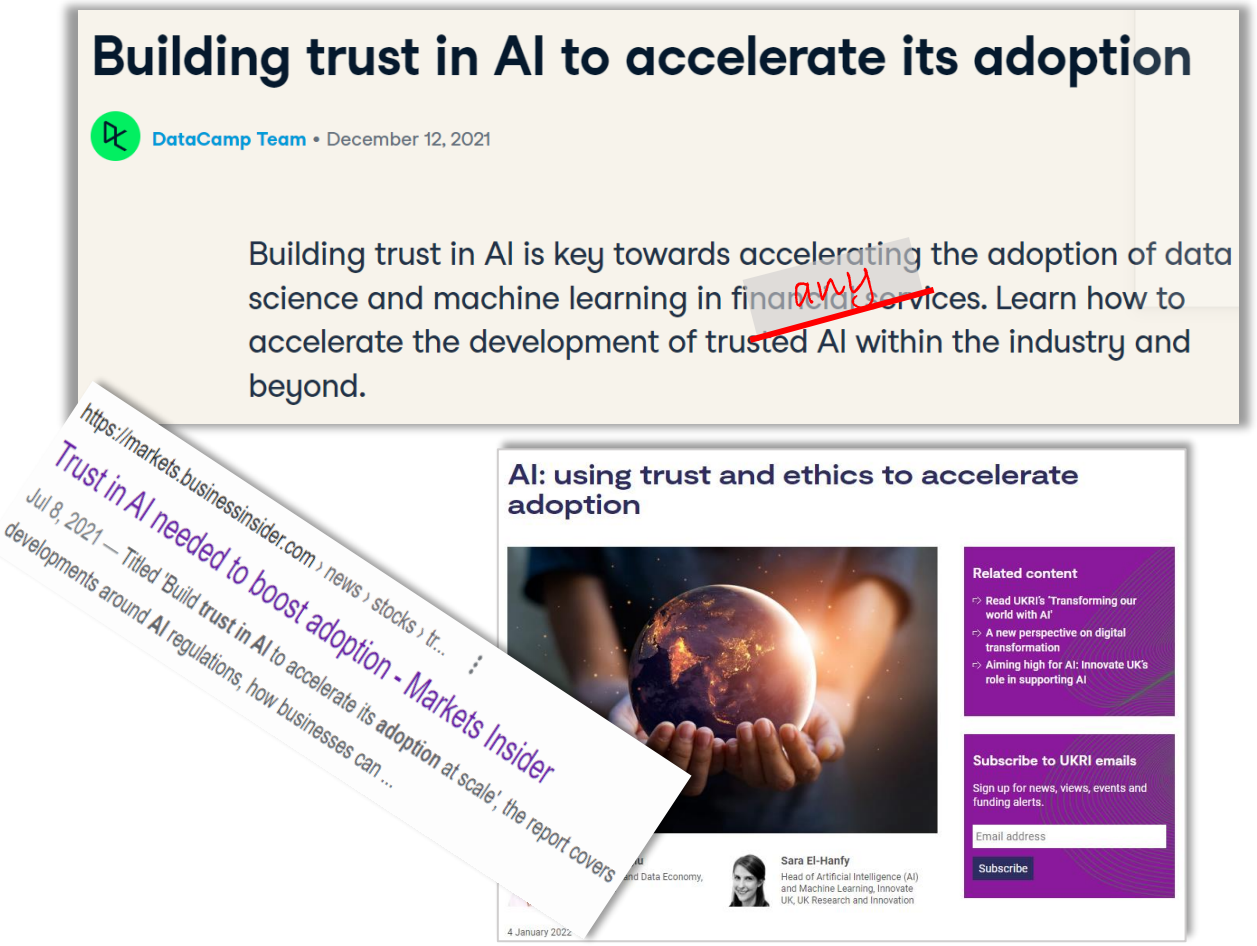
Ensuring AI that is used in a Telecom system is responsible

Transparency & explainability

- Transparency leads to trust in AI results and higher usage by engineers and executives.
  - Correlation does not equal causation
  - AI needs to evolve beyond correlation for predictions, e.g. leverage causal inferencing, use XAI toolkits
- Explainability (XAI) can help enable transparency.

- Explainability helps AI engineers better architect, feature engineer, train and re-train AI models.
  - Identifies the features most influencing predictions
  - Includes use of counterfactuals – What Ifs?
  - Optimizes AI model metrics – accuracy, precision, recall, F1, etc.

  - Explainability enables faster root cause analysis



Sustainability – reduce AI carbon footprint

Location

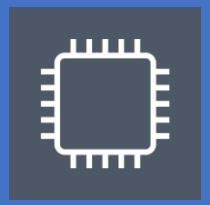
- Train AI models in colder regions, reducing cooling needs

Energy Sources

- Use renewable energy for training

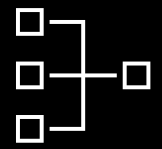
Chipsets

- Use neuromorphic chips – cut energy by factor of 16



Algorithm Design

- Sparse networks
- One-shot, few-shot learning

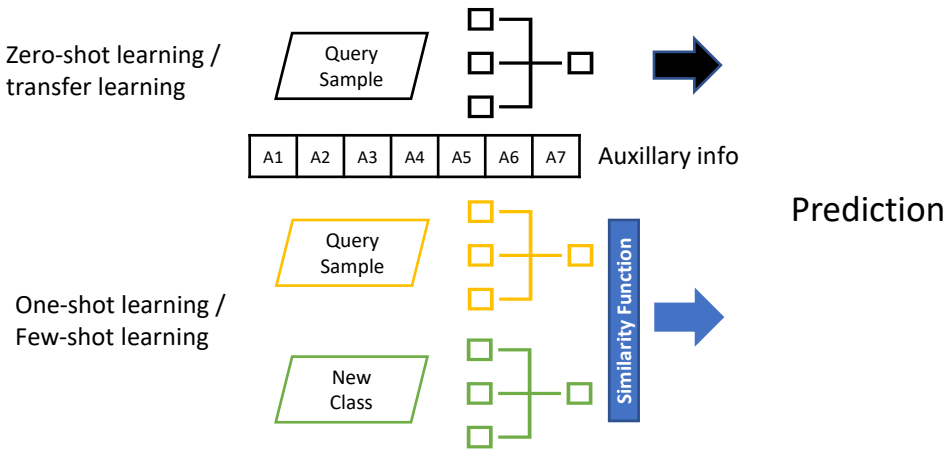


AI Inferencing

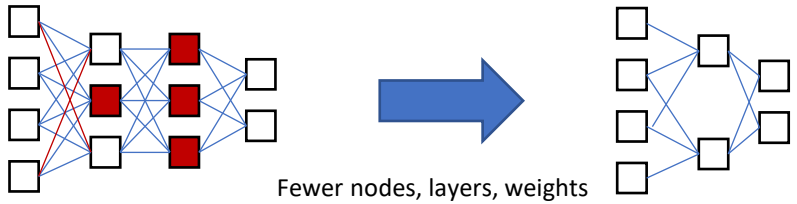
- Model compression, e.g. via knowledge distillation, etc.

Sustainability – algorithm methods

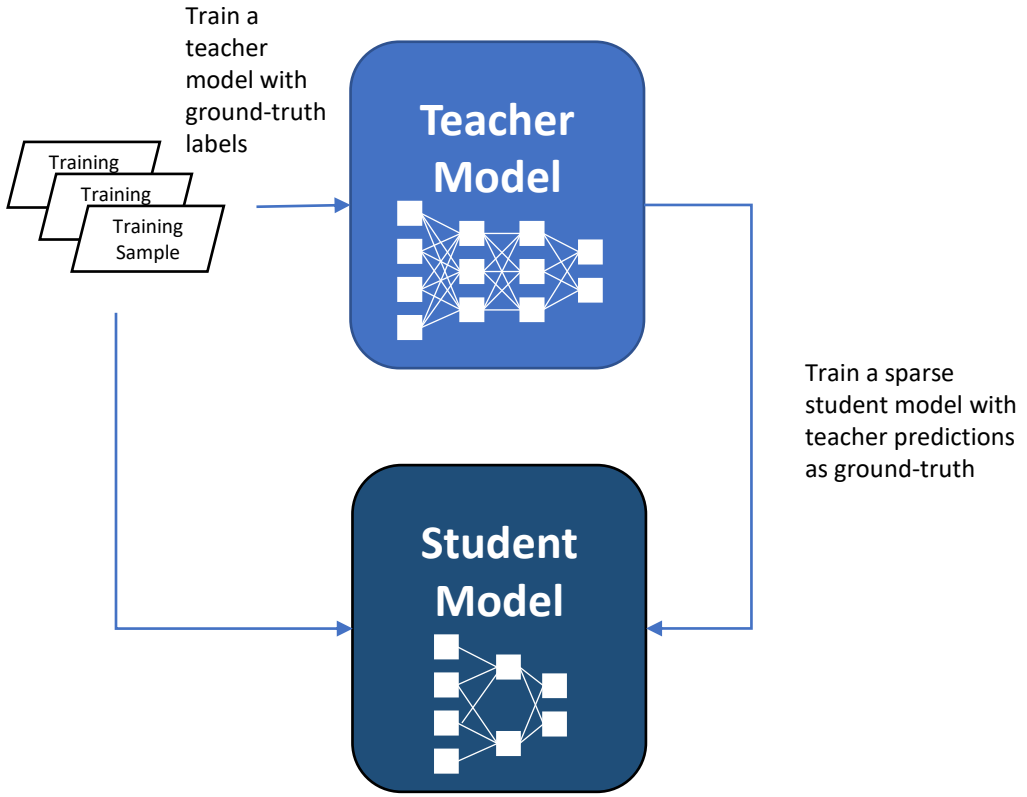
AI Algorithm Design and Training



Sparse Networks / Sparse Models



AI Inference

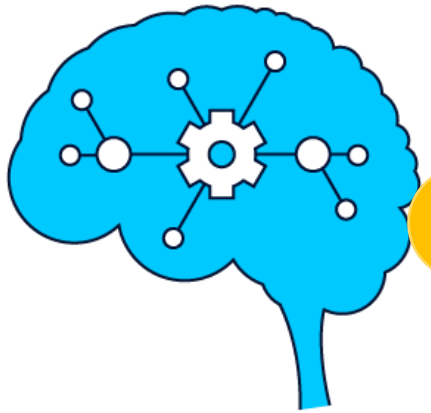


Model compression via Knowledge Distillation

Sustainability – use AI to reduce carbon footprint of Telecom systems

1. Predict demand and turn off components in the Radio, Transport, and Core when not needed.

2. Minimize RF transmission power output while meeting QoS targets across active sessions.

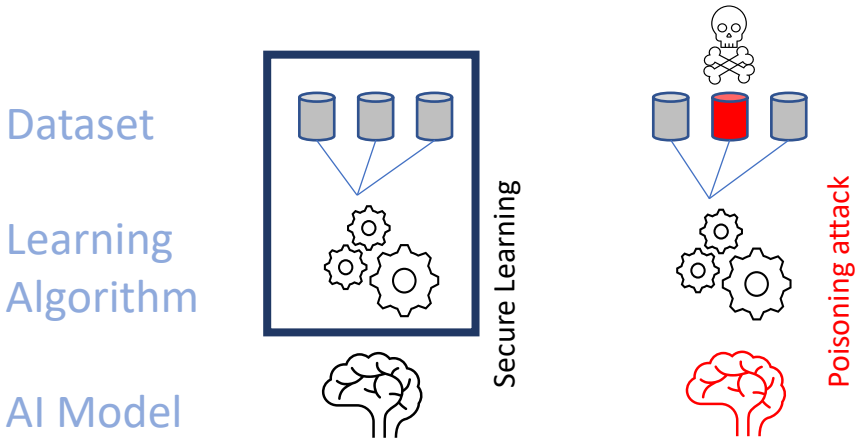


3. Load balance while minimizing number of components required.

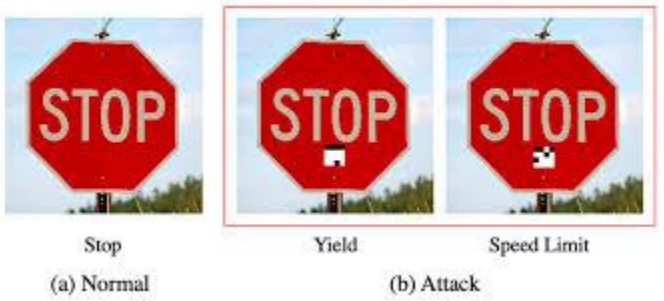
4. Network planning and deployment to minimize energy needs.

Safe, secure, reliable

- Secure AI model training - protect against poisoning the training data

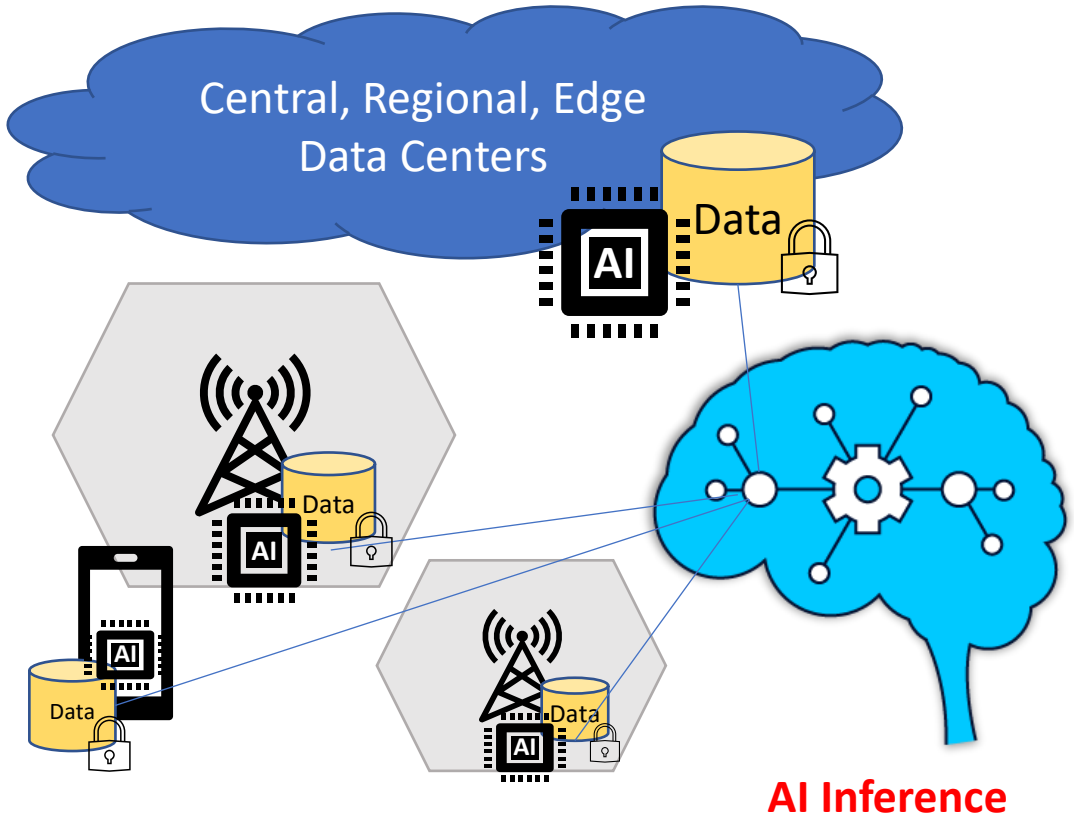


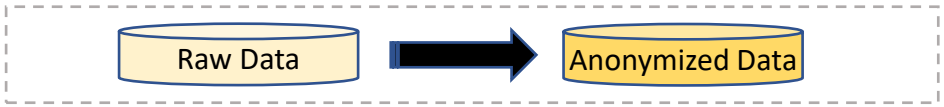
- Secure AI model inferencing - protect against poisoning the input data for inferencing



AI systems must be *safe*, performing as intended (*reliable*), and *resistant* and *resilient* to being compromised by unauthorized parties (*secure*)

Privacy

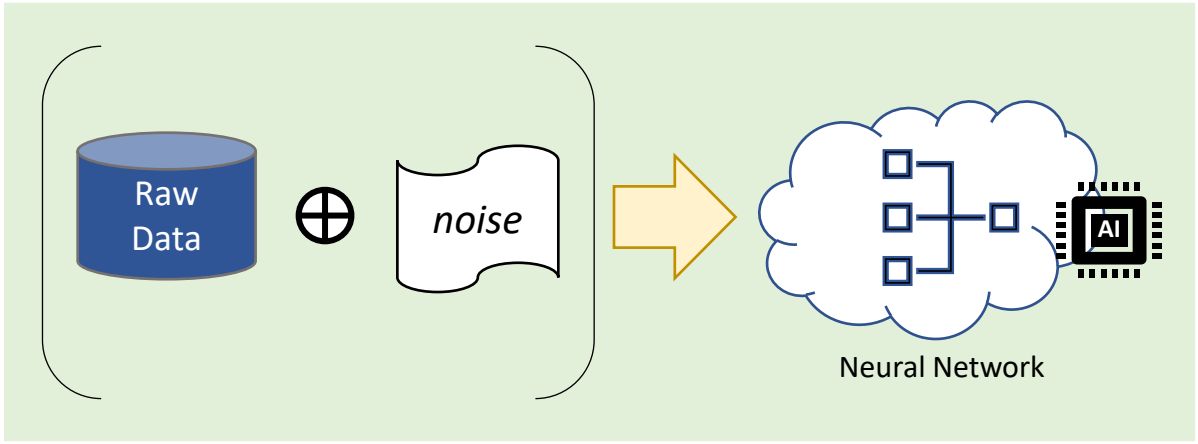


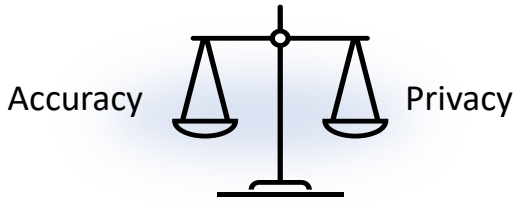


- What: privacy regulations
  - GDPR, HIPAA, etc.
- Why: give impacted people agency over their data
- How: privacy approaches
  - Data privacy in training and inference – secured and/or encrypted data collection and storage
  - Privacy in transmission of data
  - Techniques: federated learning, differential privacy, data anonymization

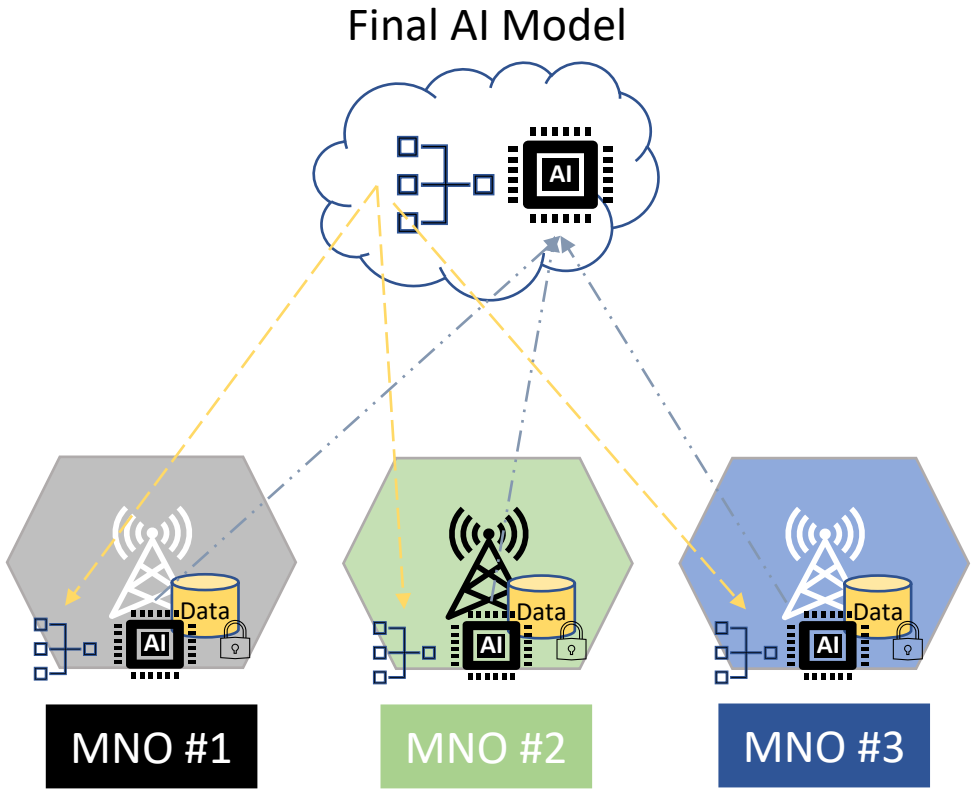
Privacy methods

Differential Privacy





Federated Learning

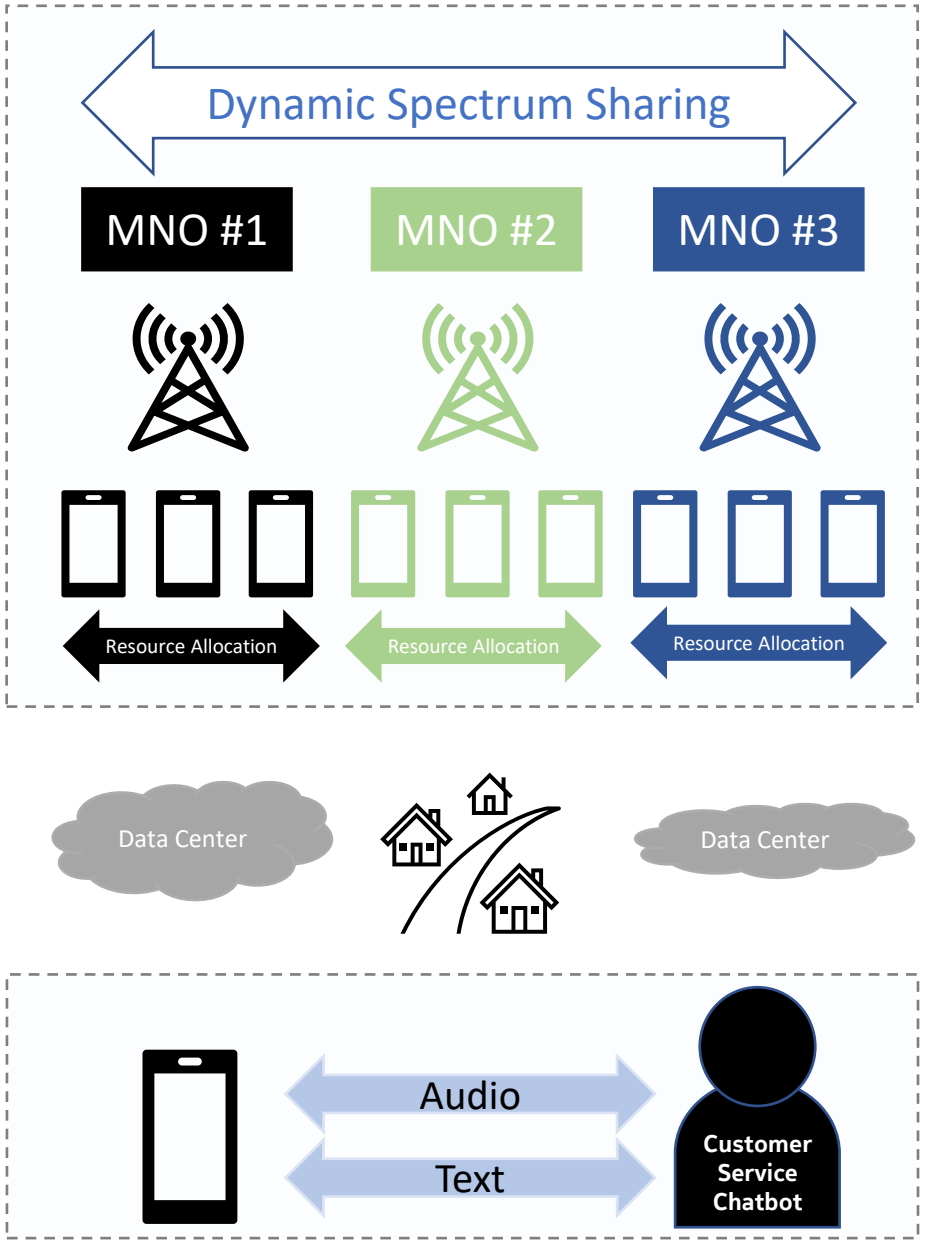


Fairness

AI systems should be designed and used to maximize fairness, non-discrimination, accessibility and universal design that promotes inclusivity, by, e.g., correcting data biases and algorithmic biases.

In Telecom, this includes:

- AI used in network resource and RF allocation must be fair wrt users – while ensuring prioritizations per govt regulations and business models
- AI used in dynamic spectrum sharing must be fair
- AI for network planning and sustainability must not bias against geographical regions or neighborhoods
- Different communications modes must be provided for users with different needs, i.e. vulnerabilities or disabilities, including with customer support chatbots



Accountability

- AI systems should be developed and deployed by consulting, involving, collaborating, and empowering all the stakeholders and, crucially, making accountability possible (e.g., the impacts of AI systems are appropriately distributed, the appropriate stakeholders are consulted, long-term effects are planned for, fundamental rights are met, and adequate remedies are provided).

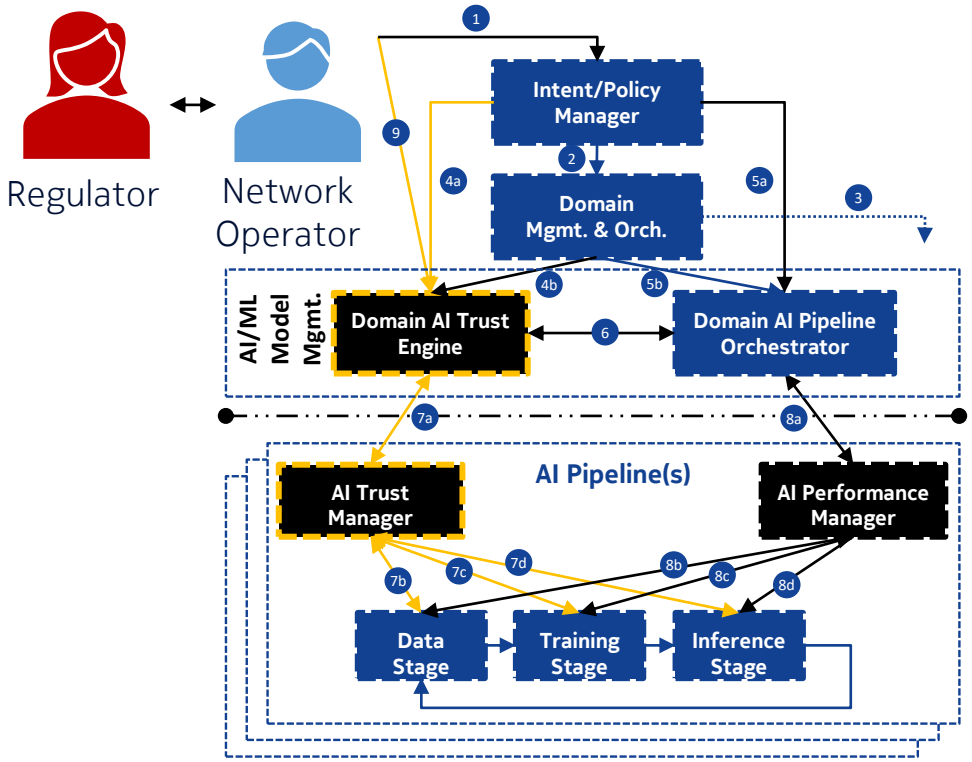
- Stakeholders accountable for results of AI to achieve Responsible AI goals in Telecom
  - Network operator
  - Equipment vendor
  - AI solution vendor if different from equipment vendor or network operator

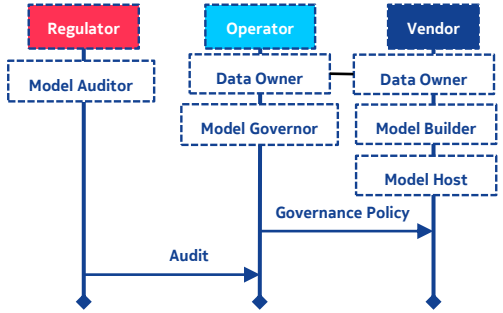
Telecom Operator Perspectives

- Verizon & Vodafone
  - ✓ Published commitment to Trustworthy AI/ML
  - ✓ Released several blog posts in the last two years
  - ✓ Many job advertisements on Ethical AI across all hierarchies in the last two years
- Deutsche Telecom
  - ✓ Defines own governance policy with nine self-binding guidelines for the development of their AI-based products and services in the future
- ETNO and GSMA
  - ✓ Endorses the TAI guidelines proposed by the EU Commission's AI High Level Expert Group (HLEG)

- Orange
  - ✓ Advocates Trustworthy AI/ML by design
  - ✓ One of 52 members in EU Commission's High Level Expert Group (HLEG) to produce guidelines for AI Regulatory Act
  - ✓ Creates an internal Data and AI Ethics Council for internal governance
  - ✓ Released several blog posts in the last two years
- Telenor
  - ✓ Endorses EU recommendations on Trustworthy AI/ML
  - ✓ One of the 52 members of HLEG
  - ✓ Representatives in AI/ML conference panels to advertise commitment to Trustworthy AI/ML

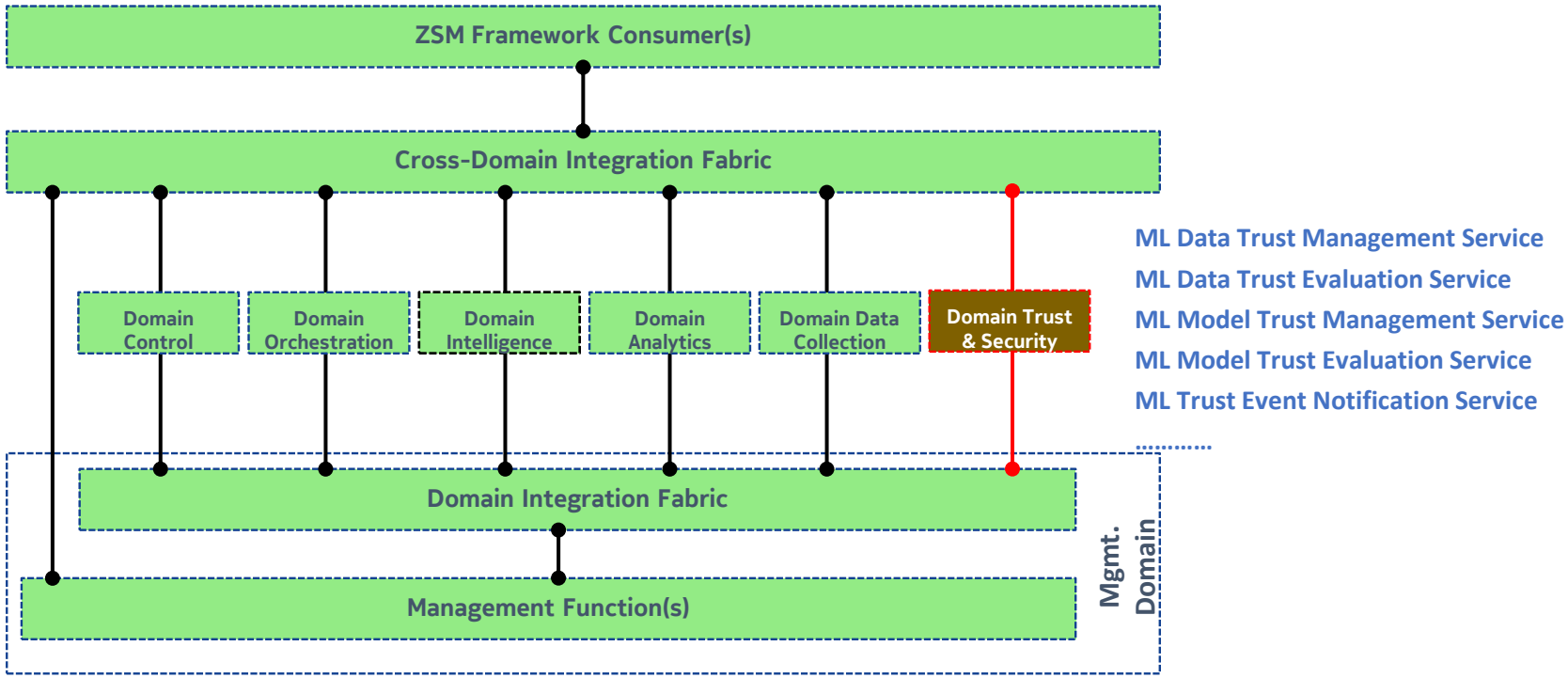
3GPP standards – proposed Trustworthy AI/ML framework





I/f #	Description	Example
1	Customer intent for network service	Trust intent
2	Network QoS & network QoT requirements	SLA / RLA (risk level agreement)
3	Network service control	Instantiation of a service
4a, 4b	AI QoT requirements	Requirements related to the network service (e.g., fairness)
5a, 5b	AI QoS requirements	Requirements related to the network service (e.g., accuracy)
6	AI QoT vs. QoT info exchange	Get AI QoT, AI QoS metrics to control tradeoffs
7a	AI pipeline trust control	Set TAI level (e.g. fairness level)
7b	Data trust control	Get data TAI metrics (e.g., report pre-processing fairness metric)
7c	Training trust control	Get training TAI metrics (e.g., report in-processing fairness metric)
7d	Inference trust control	Get inference TAI metrics (e.g., report post-processing fairness metric)
8a	AI pipeline performance control	Set accuracy level
8b	Data performance control	Get training data metrics
8c	Training performance control	Get trained model accuracy
8d	Inference performance control	Get prediction confidence
9	Human operator interface	Get human-readable explanations

ETSI standards – Trustworthy AI/ML related management services in ZSM012



* Similar Trustworthy AI/ML related management services are foreseen in 3GPP SA5 (Telecom Management) as well.

ZSM: Zero touch network and Service Management

Conclusions

- Responsible AI is essential to the adoption and successful usage of AI
- Toolkits are emerging to help address some aspects of Responsible AI
- Research is needed to reach Responsible AI goals for Telecom, e.g.:
  - Make AI transparent and explainable
  - Reduce carbon footprint of AI itself
  - Use AI to reduce carbon footprint of Telecom systems
  - Make AI secure and reliable
  - Ensure data privacy
  - Ensure fairness
  - Assignment of accountability
- Certification for each aspect of Responsible AI can incentivize acceleration to find solutions.

NOKIA

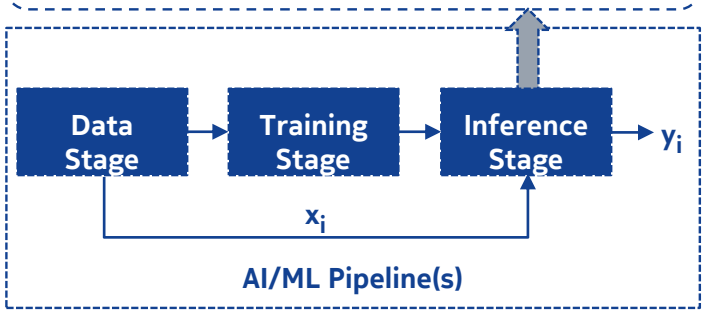
NOKIA

Many of the concerns introduced in general Responsible AI goals have actually been addressed and continue to be addressed in classical Telecom design, regulations, etc. , e.g. for safety, security, privacy, use of scarce resources (e.g. radio frequencies), functional content (standards), net neutrality, universal access, competitive markets…

As Telecom adopts AI, we need to ensure that the new AI solutions do not violate the regulatory and operator specific requirements that are also relevant to Responsible AI, but rather help to maintain and improve them, while also serving the business objectives as well.

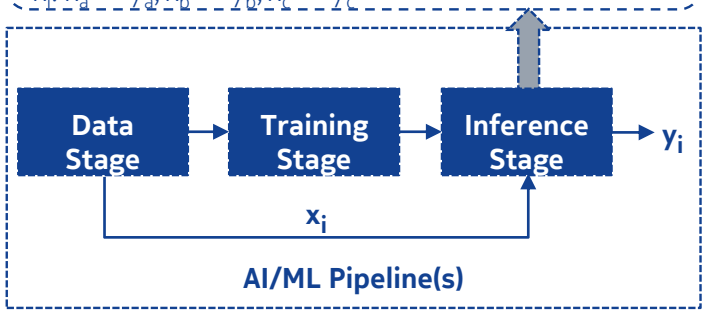
The techniques for Explainable AI/ML

Local Explanation: "The output for x_i is y_i because $x^2 > \beta$"



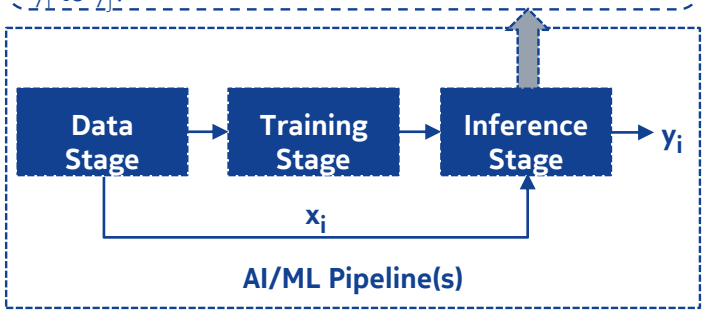
Text-based Explanations

Local Explanation: "Example outcomes from the model for other input samples that are similar to x_i: x_a -> y_a; x_b -> y_b; x_c -> y_c"



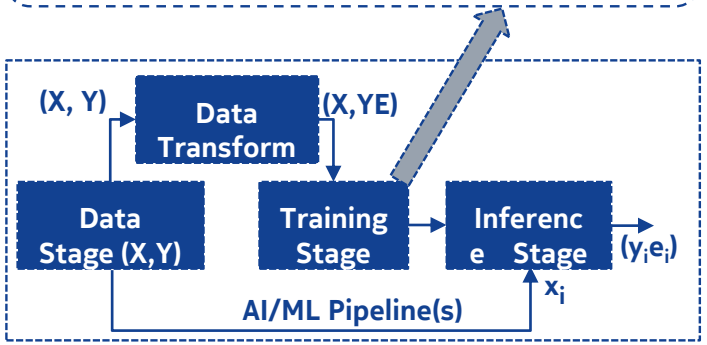
Examples-based Explanations

Local Explanation: "Which features in x_i should be modified by how much to switch the output from y_i to y_j?"

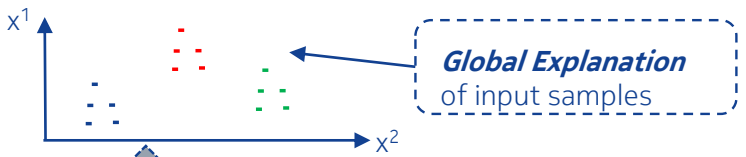


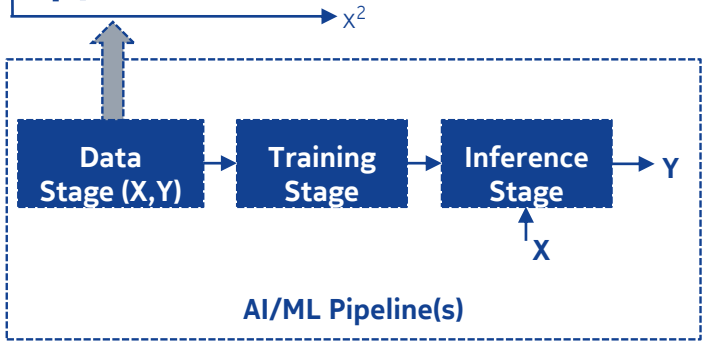
Contrastive Explanations

Local Explanation: "The model is trained together with labelled explanations and predicts both y_i and e_i for an instance x_i"

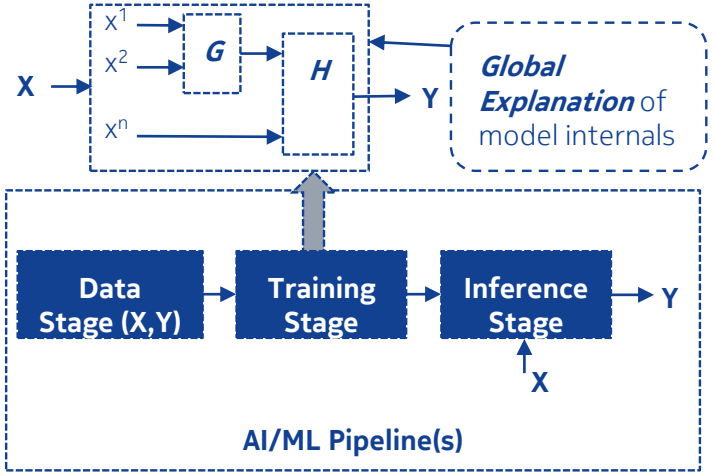


Joint Prediction & Explanation





Visualization-based Explanations



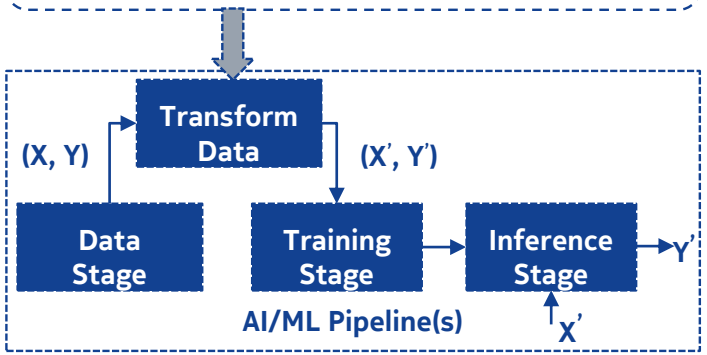
Model Simplification

REF-4: 'https://aix360.mybluemix.net/', Linux Foundation Artificial Intelligence.

NOKIA

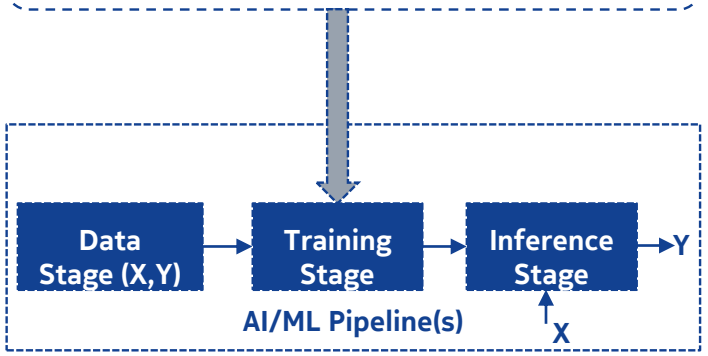
The techniques for Fair AI/ML

"***Modifies feature values and/or labels*** to mitigate unfairness towards certain groups"



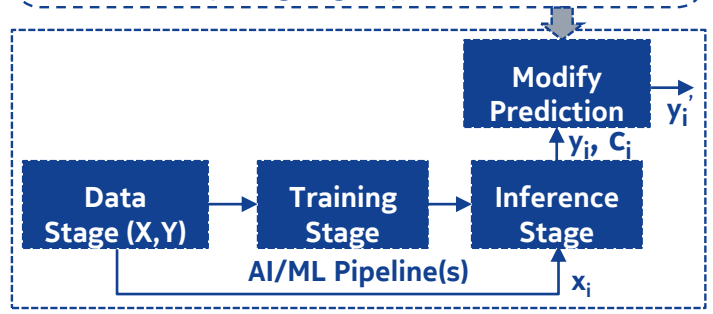
Pre-Processing Fairness – Type 1

"***Adds discrimination-aware regularization term*** to mitigate biased outcomes"



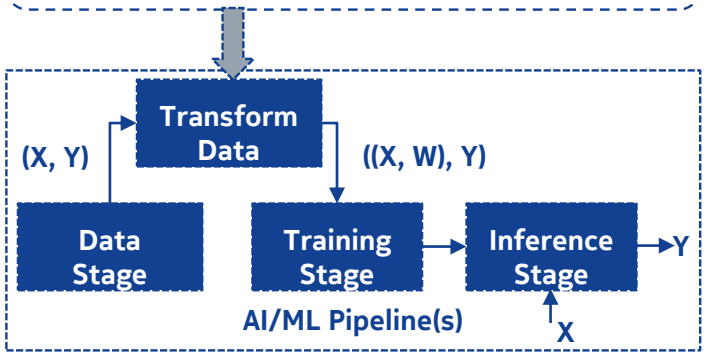
In-Processing Fairness – Type 1

"***Changes predictions from a classifier based on confidence values (e.g., c_i)*** to ensure favorable outcomes to unpriviliged groups and unfavorable outcomes to priviliged groups"



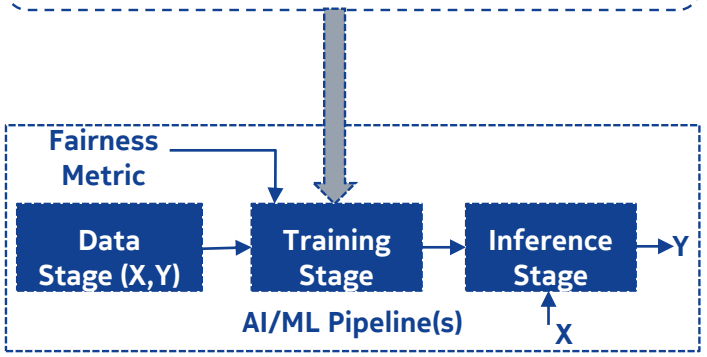
Post-Processing Fairness – Type 1

"***Adds weights to the training data samples*** to mitigate unfairness towards certain groups"



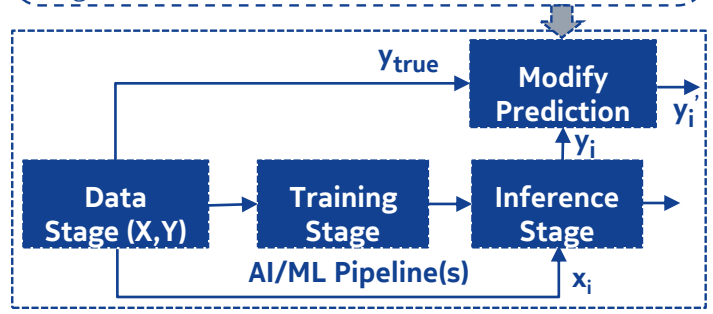
Pre-Processing Fairness – Type 2

"The model learns a fair classifier by ***optimizing the fairness metric provided in the input***"



In-Processing Fairness – Type 2

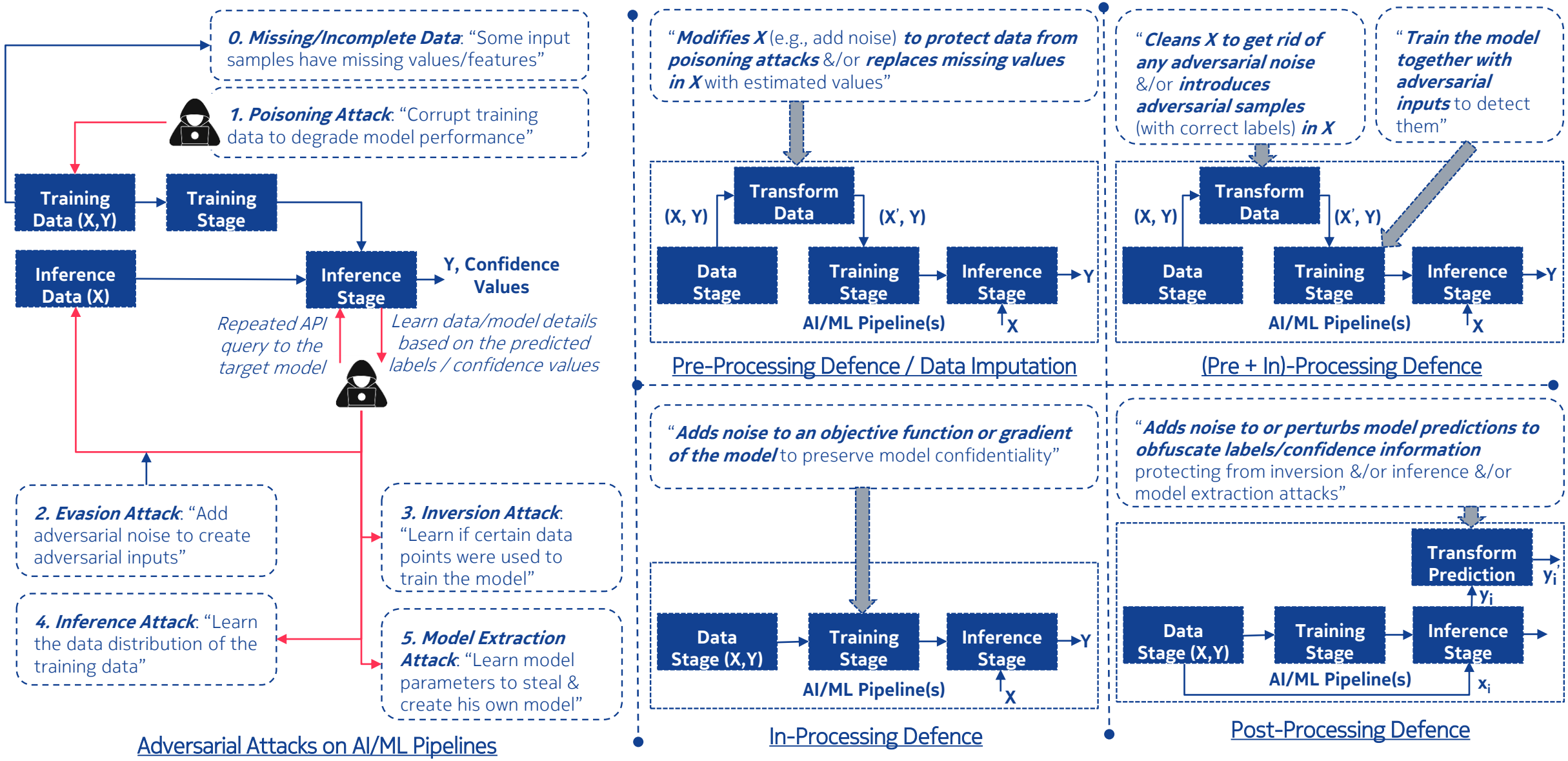
"***Modifies predicted labels using an optimization scheme*** to ensure both privileged & unprivileged groups have same false-positive and/ false-negative rates"



Post-Processing Fairness – Type 2

REF-5: 'https://aif360.mybluemix.net/', Linux Foundation Artificial Intelligence.

The techniques for Robust AI/ML



REF-6: 'https://github.com/Trusted-AI/adversarial-robustness-toolbox', Linux Foundation Artificial Intelligence.