National Edge AI Hub

Artificial Intelligence Theme

Data Quality and Model Quality Challenges

The vision of WS4 is to establish research directions for **developing fundamental concepts** and techniques that can **guard the data and AI algorithm learning quality** against cyber-disturbances impacting the EC architectures

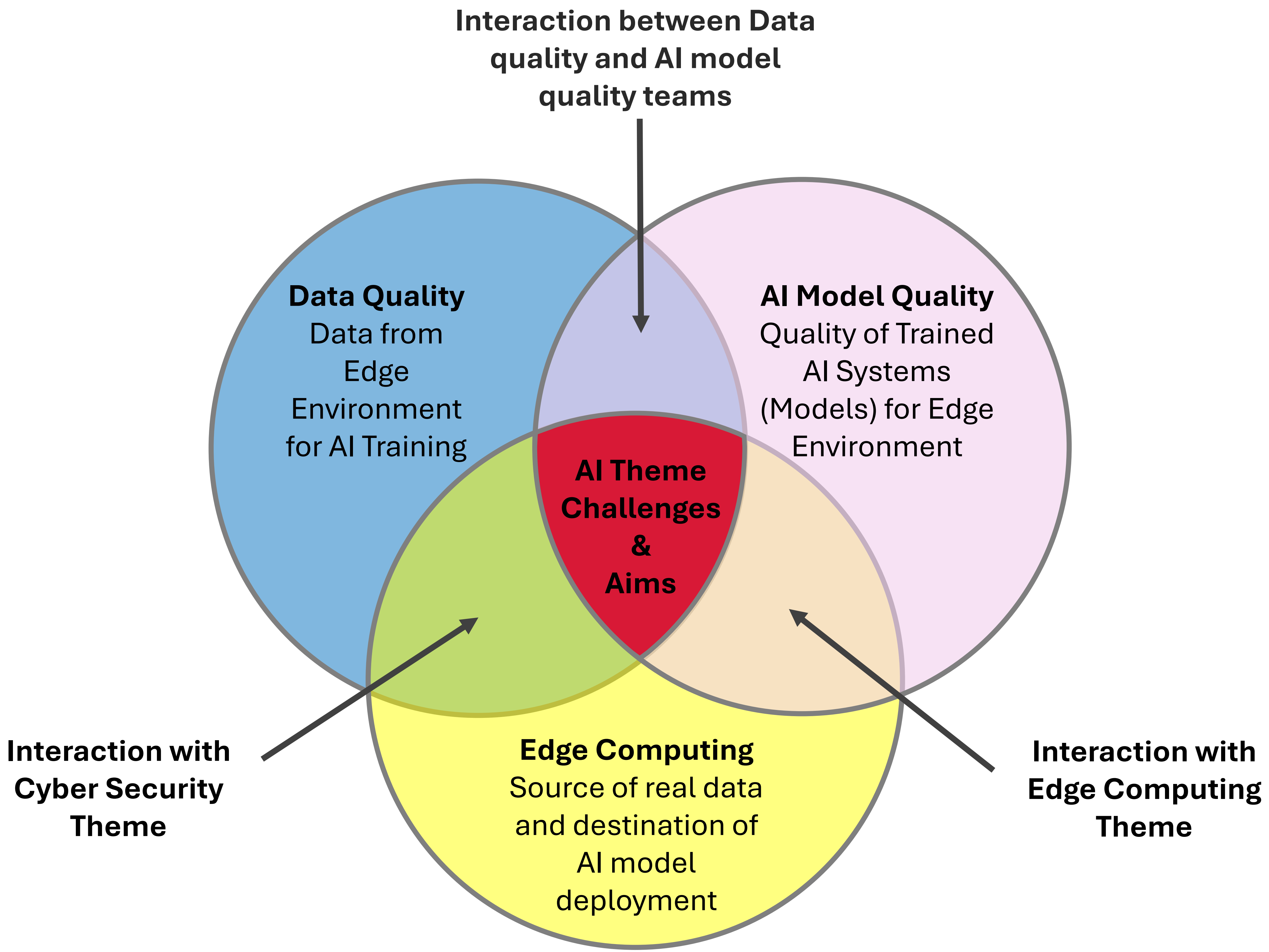
Team: Newcastle, Durham, Hull, Swansea, and QUB

by Dr Varun Ojha

3 May 2024

@

National Edge AI Hub Lunch Day, Newcastle University, UK


Interaction between Data quality and AI model quality teams
Data Quality
Data from Edge Environment for AI Training
AI Model Quality
Quality of Trained AI Systems (Models) for Edge Environment
AI Theme Challenges & Aims
Edge Computing
Source of real data and destination of AI model deployment
Interaction with Cyber Security Theme
Interaction with Edge Computing Theme

AI Theme Challenges / Research Aims

- **Monitoring of Data/Model Quality**

  How to monitor cyber-disturbances impact AoD, AI algorithms learning quality and the overall application resilience?

- **Recovery of Data/Model Quality**

  How to recover/ensure data and AI model quality that are impacted by cyber-disturbances and ensure suitability for AI model deployment on devices at Tiers 1, 2 of EC architectures ?

- **Assurance of Continuity of Data Quality and Model Quality**

  How to AI algorithms continually adapt to EC environments where unknown cyber-disturbances that were not presented in the original training dataset?

Potential Research Problems

- **Monitoring**
  - **RP1.** Investigates, characterise, and develop ontologies of data challenges and models challenges for EC environment.
  - **RP2.** Data and model quality assurance to data quality challenges, faults, missing data, hardware failure, sensor degradation; diverse data source; sensor/data heterogeneity.
- **Recovery**
  - **RP3.** Investigates and develop data and model quality certification/robustness to various challenges such as data distribution shift, impurities, adversarial attacks, hardware resources limitations, etc.
  - **RP4.** Investigates the model quality certification/robustness to cyber disturbances, cyber-attacks, on federated/distributed EC environment.
- **Assurance**
  - **RP5.** Data/Model quality verification/assurance. This will aim to identify quality issues with AI models implementation on edge and offer mitigation strategies to resolve the challenges.

Our Smart City Testbench

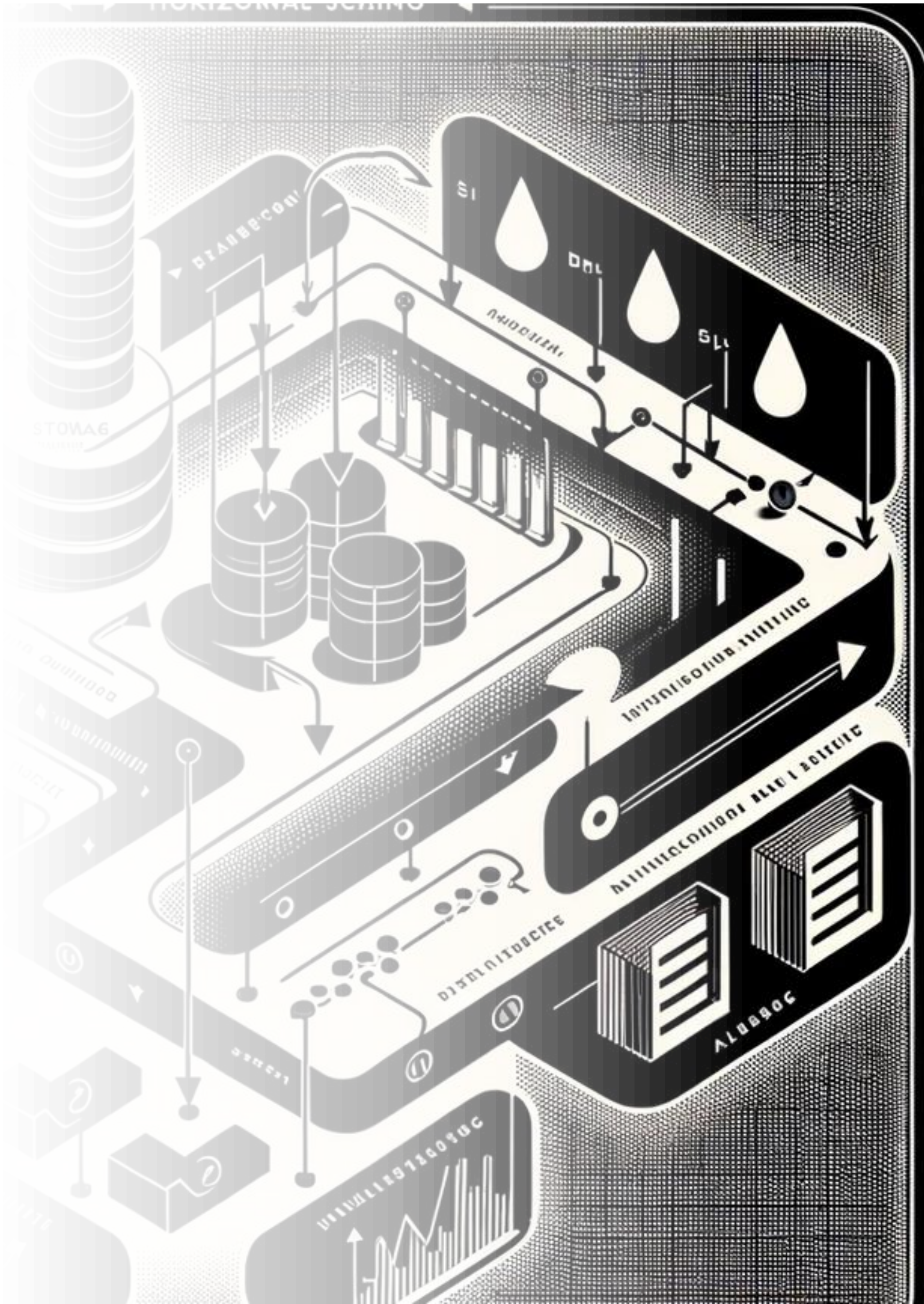
Newcastle University's Urban Observatory Sensors

- Air Quality: pollutants, particulates
- Weather: Temperature, precipitation, wind speed, humidity
- Traffic: Vehicle counts, speeds, classification
- Footfall: Pedestrian counts, movement patterns, poses
- Water Quality: pH levels, contaminants, biological indicators
- Sensor, image and Video (CCTV) feeds

Source: Phil et al (Newcastle)

Our Experience with Data Quality Challenges

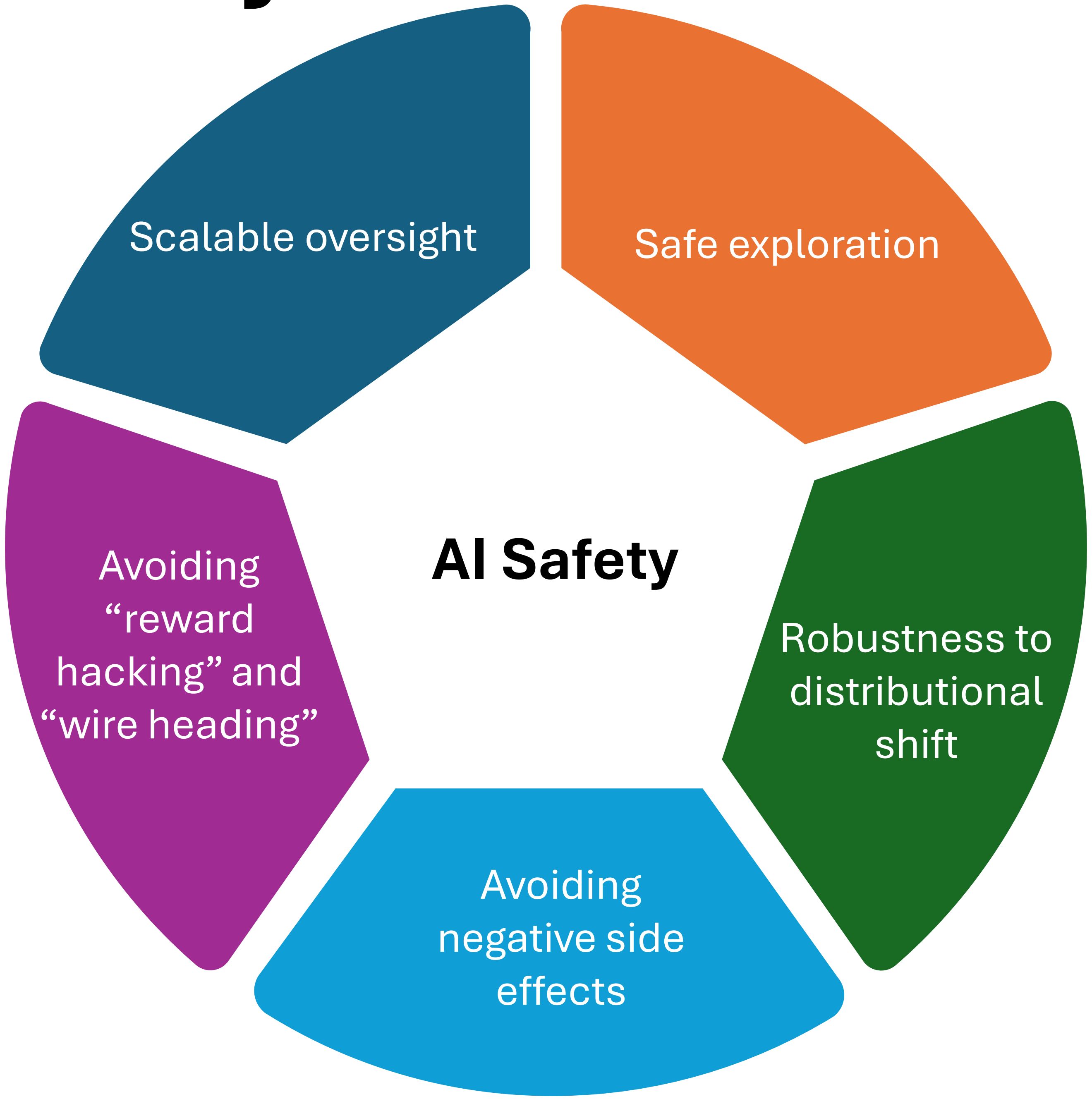
- **Data quality**
  - degradation of sensors over time
  - Anomalous values, random spikes, exogenous anthropogenic or environmental issues
  - Data out of range, Out distribution, uncertainity
- **Data stream issues**
  - Data Retrieval: Source API failure
  - Comms issues: Comms to source API failure, network failure, network overload
  - System Throughput: Queues building up, hardware issues
  - Asynchronicity issue with external APIs
- **Cyber security**
  - Adversarial attacks
  - Denial of Services, spoofing
- **Failure**
  - Hardware failure at sensor



Source: Phil et al (Newcastle)

Dimensions of EdgeAI Safety

Our focus has been on: Robustness to Distributional Shift using SafeML: Issues related to changes in the AI's operational environment that differ from its training environment, which can lead to unexpected or harmful behaviour.



Source: Thakker et al (Hull)

SafeML & Data Certification – Example Solution

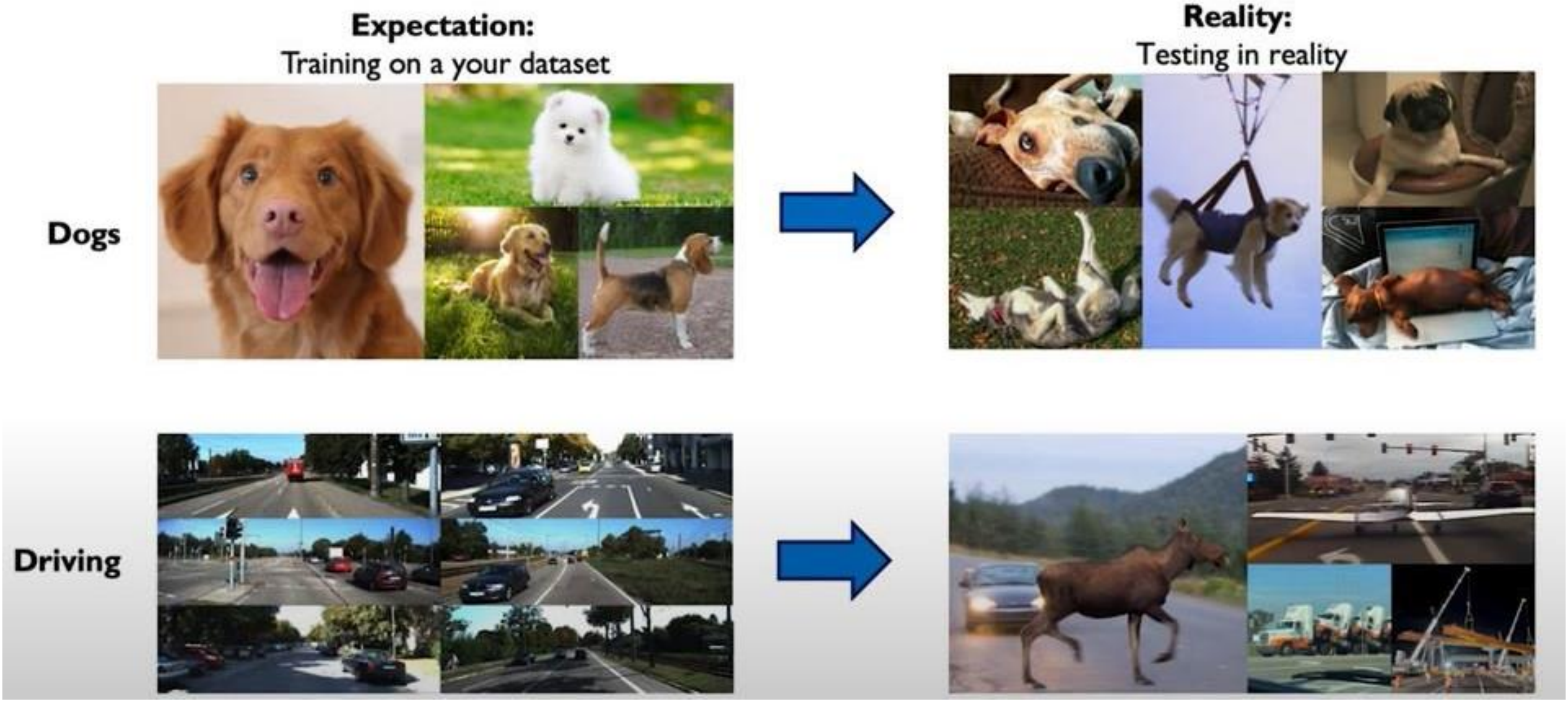
Trusted dataset for AI model training

Existing Solutions:

- SafeML – is a framework for safety monitoring of ML models at run time focusing on distribution (see example)
- D-ACE – framework for certifying training datasets using a number of characteristics

How we will extend for EdgeAI

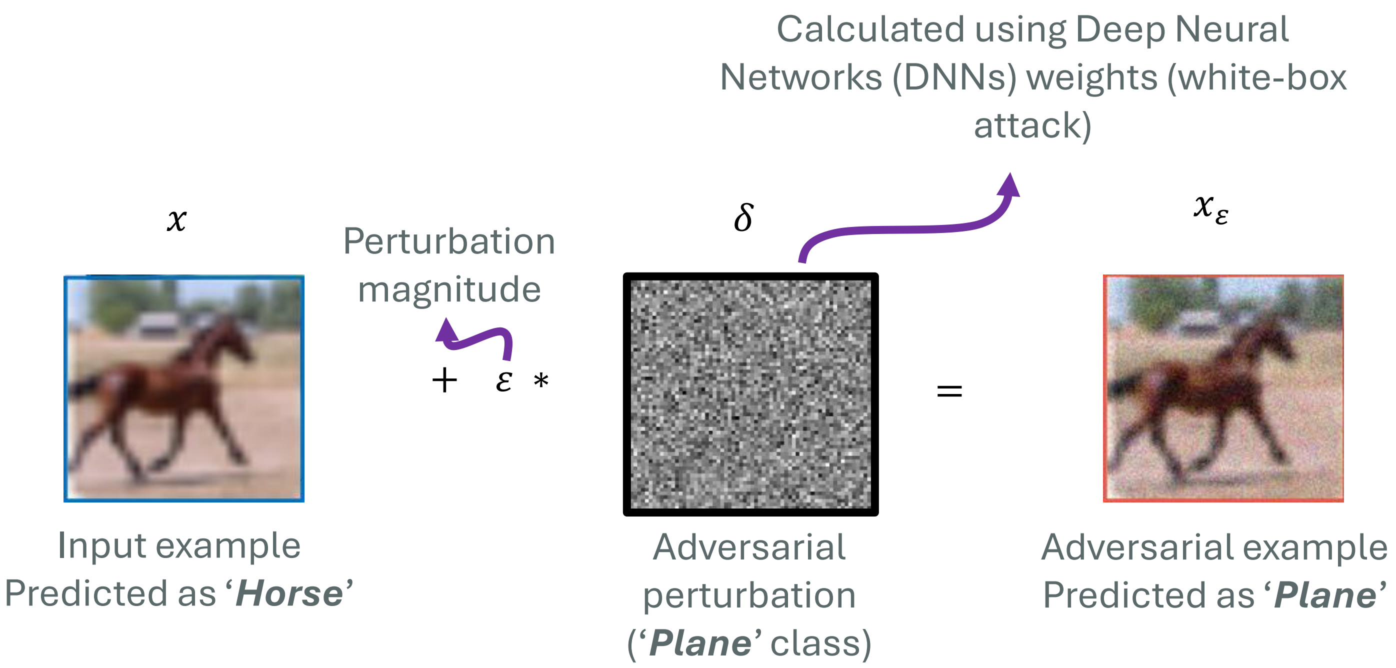
- Ensuring Safety of Federated Learning algorithms in EdgeAI architecture
- Extend D-ACE for certifying datasets in federated Edge AI architecture

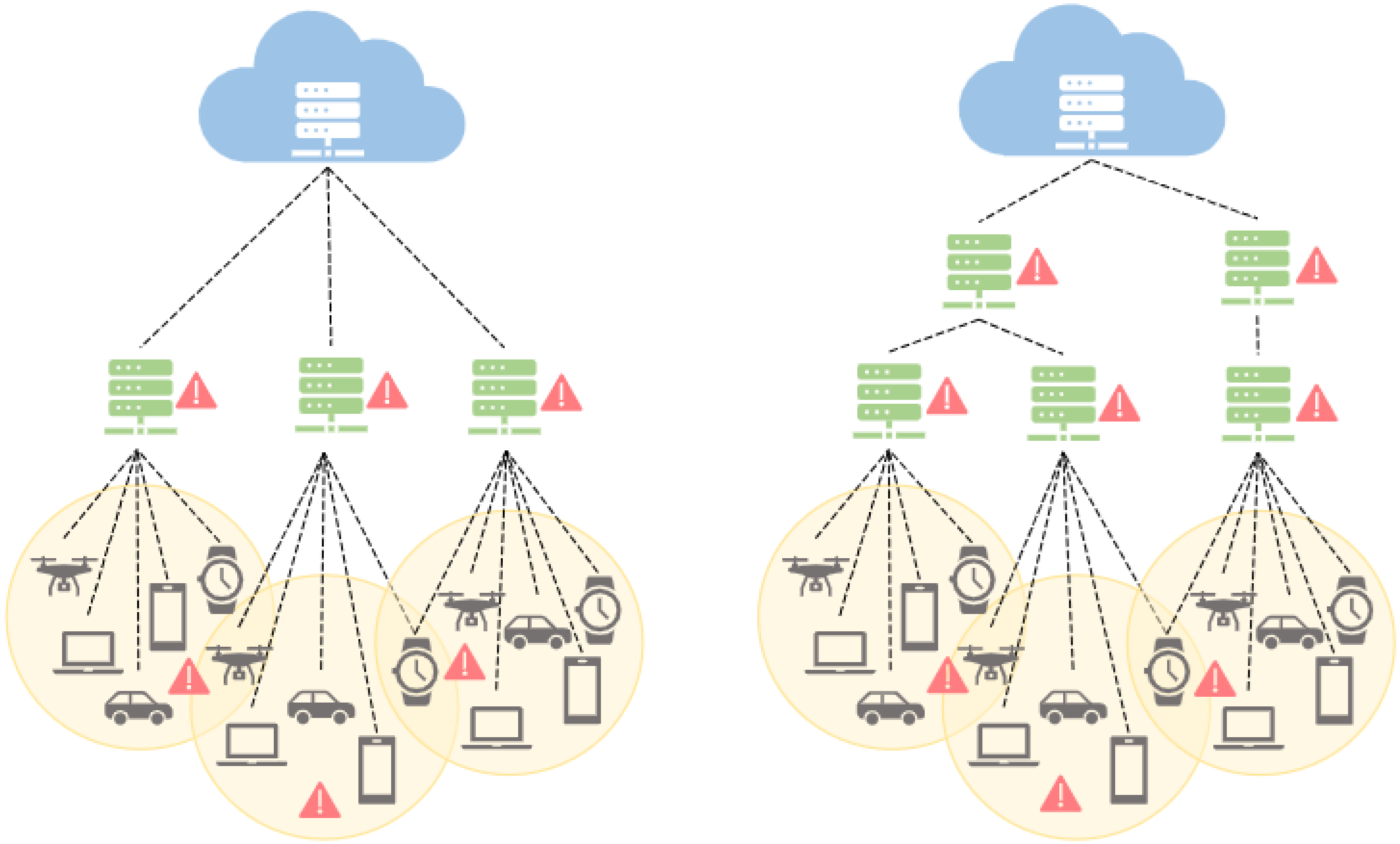


Source: Thakker et al (Hull)

Model Certification – Example Solution

Models Adversarial Attacks Mitigation on Federated Edge Environment





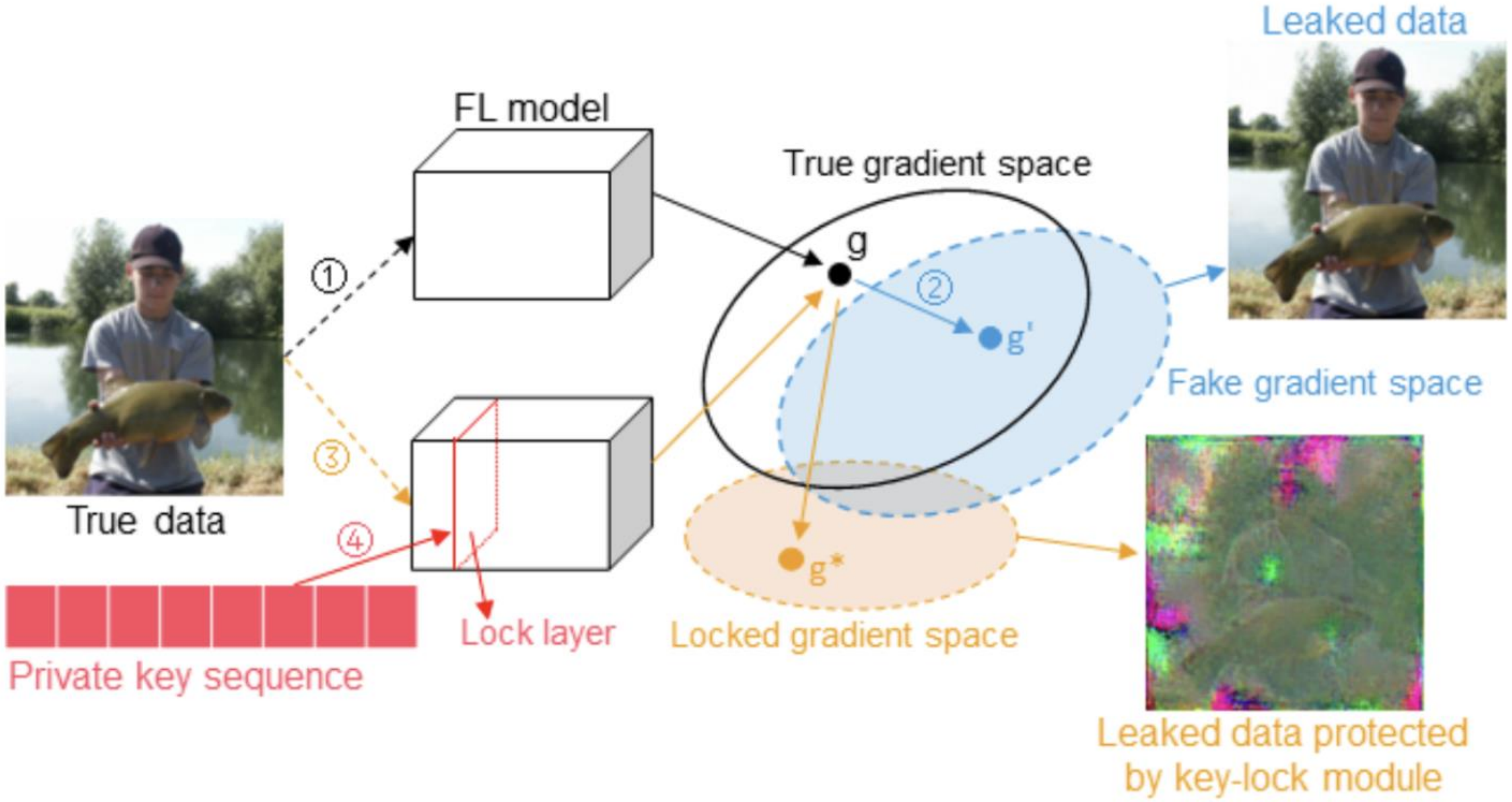
Attack of Federated and Distributed Edge Environment

The general premise of a robustness analysis is to subject DNNs to the '**worst case**' conditions and evaluate the ***ability for a DNN to remain invariant*** under such settings.

Source: Ojha et al (Newcastle)

Gradient Leakage and Protection for Federated Learning

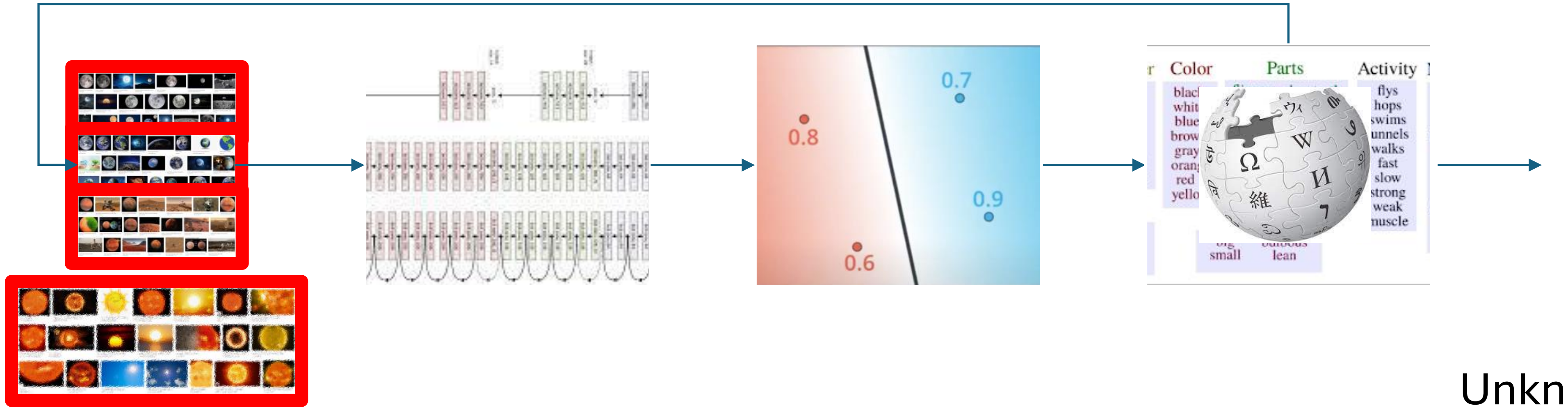
- The root cause of gradient leakage has been explored anc proved mathematically.
- Based on the above findings, to block private information leakage during the propagatior of the gradient.
- Negligible impact on model performance.
- No need to balance defense level and model performance.



Source: Xie et al (Hull)

Data and Model Quality Transparency

Transparent Zero-Shot Knowledge Transfer



- Earth
- Moon
- Mars
- Sun (✓)

Unknown New Tasks

Knowledge Generated Data

Knowledge Representation
- Ontological System
- LLM-driven

Large Neural Networks
- Black-box
- Data Driven

Transparent Inference
- Machine learning
- Statistical models
- Heuristic methods
- Symbolic approaches

Source: Yang long (Durham)

We are Team AI

Rajiv Ranjan (Newcastle)
Internet of Things
raj.ranjan@newcastle.ac.uk

Phil James (Newcastle)
Urban Data
philip.james@ncl.ac.uk

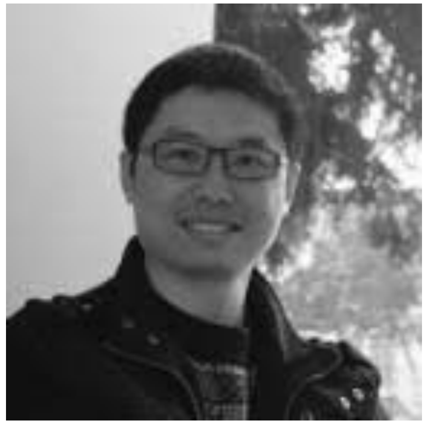
Varun Ojha (Newcastle)
Artificial Intelligence
varun.ojha@newcastle.ac.uk

Mutaz Barika (Newcastle)
Big Data
mutaz.barika@newcastle.ac.uk

Dhaval Thakker (Hull)
AI and Internet of Things
d.thakker@hull.ac.uk

Yang Long (Durham)
Computer Vision
yang.long@durham.ac.uk

Xianghua Xie (Swansea)
Machine Learning
x.xie@swansea.ac.uk



Zheng Li (Queens Belfast)
Software Engineering
zheng.li@qub.ac.uk