6GSNS

AI/ML as a Key Enabler of 6G Networks

AI/ML Training Data sets & Security/Privacy

Marios Avgeris

Assistant Professor @ Multiscale Networked Systems (MNS),
University of Amsterdam

m.avgeris@uva.nl

DESIRE6G

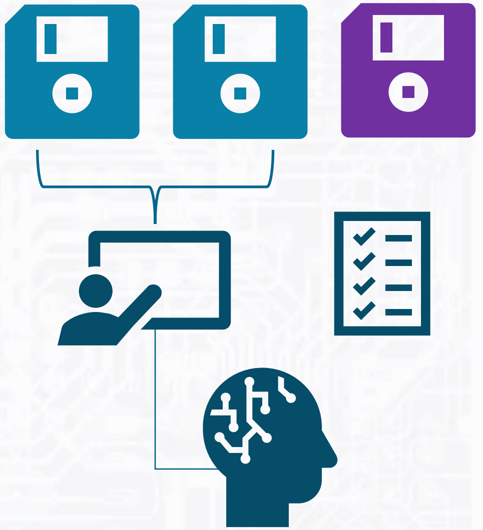
Universiteit van Amsterdam

March 18, 2025

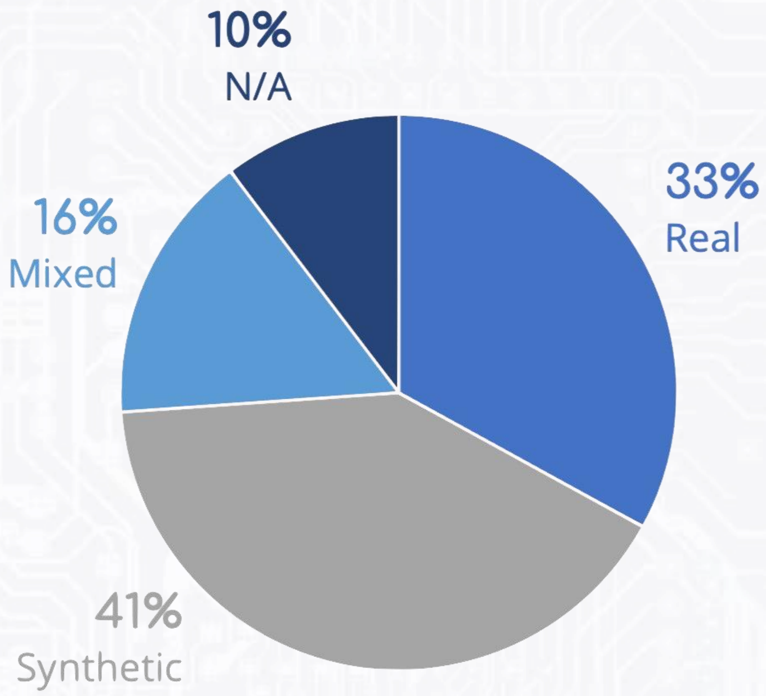
Training Datasets

Key Takeaways:

- Preference for synthetic and mixed datasets.
- Challenges in obtaining real-world datasets for AI/ML model training, particularly in networking-related AI



Source of Training Datasets



Training Datasets

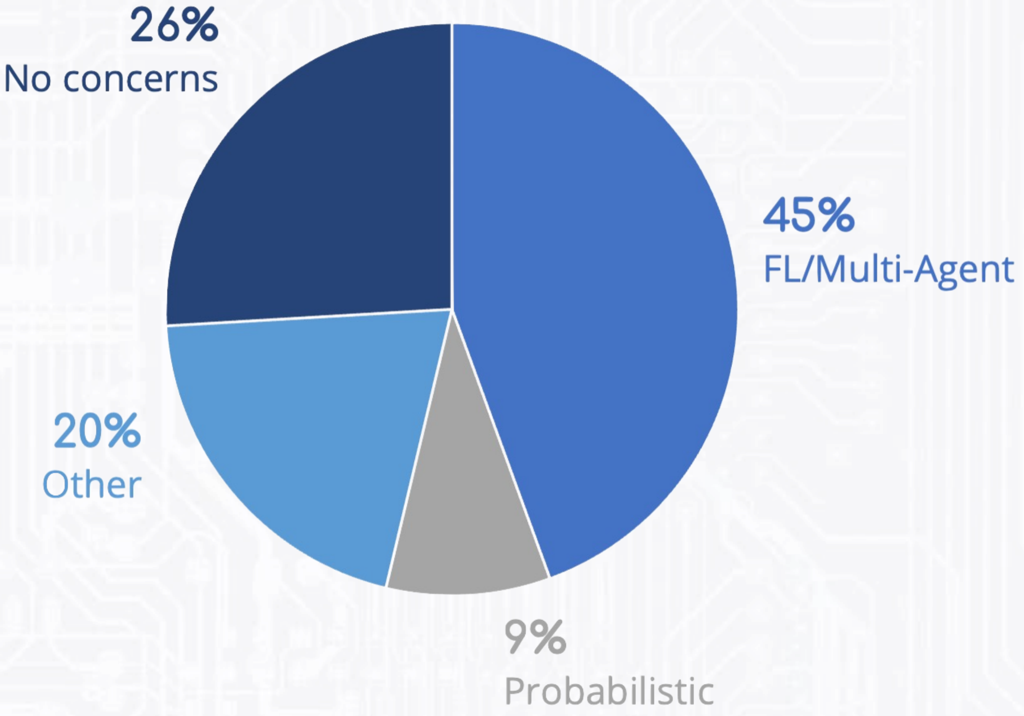
“Significant effort is underway to create an SNS JU database for sharing utilized AI datasets with the global research community”

Data Privacy & Security

Key Takeaways:

- Federated & Multi-Agent Learning gaining traction for decentralized security.
- Probabilistic techniques (e.g., Differential Privacy) used in some cases.
- Other security methods include cryptography, token-based approaches, and Explainable AI (XAI).
- Some projects report no significant privacy/security concerns due to:
  - Local training
  - Lack of sensitive data
- 73% of projects did not provide information on privacy/security strategies

Data Privacy and Security mechanisms


No concerns
FL/Multi-Agent
Other
Probabilistic