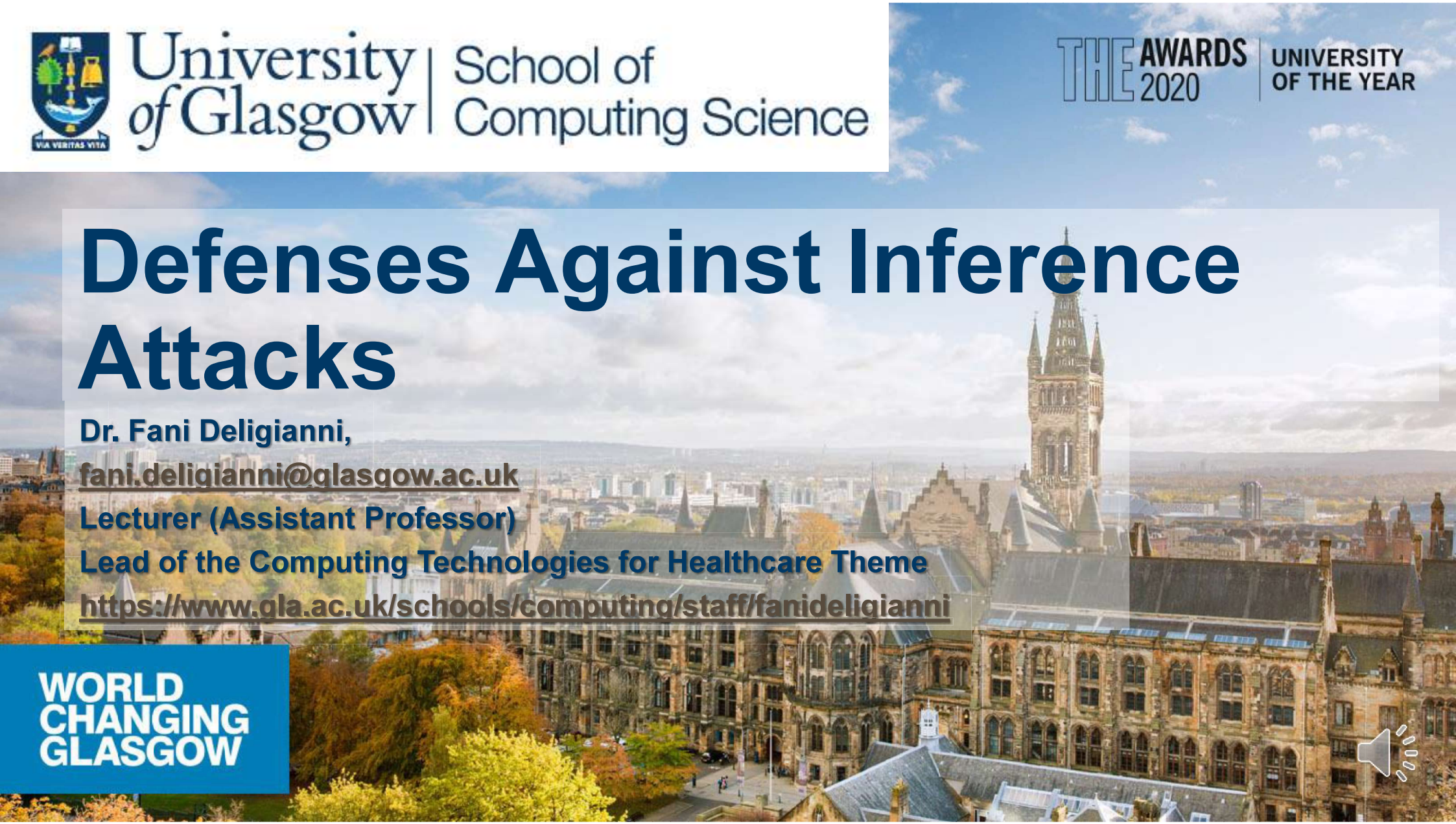
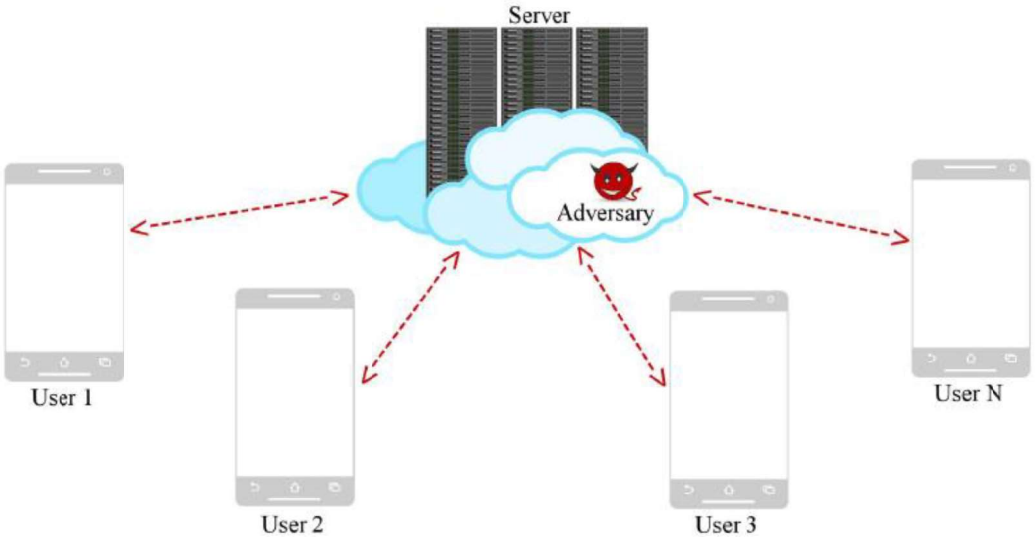


Defenses Against Inference Attacks

Dr. Fani Deligianni,
fani.deligianni@glasgow.ac.uk
Lecturer (Assistant Professor)
Lead of the Computing Technologies for Healthcare Theme
https://www.gla.ac.uk/schools/computing/staff/fanideligianni

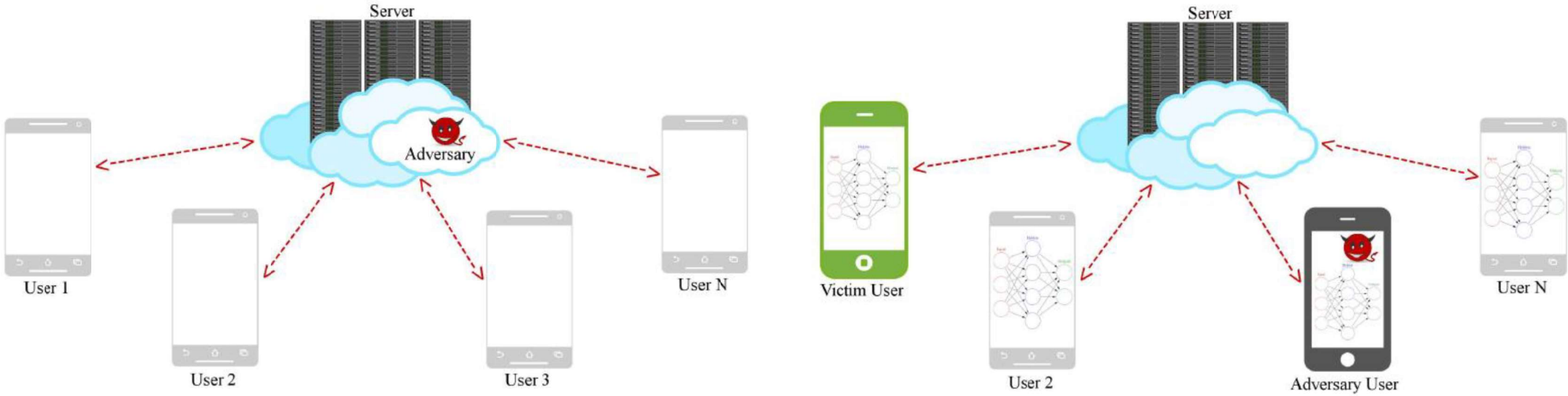
Centralised Learning



Centralised Learning

Hitaj et al. ‘Deep Models Under the GAN: Information Leakage from Collaborative Deep Learning’, ACM Conference on Computer and Communications Security, 2017.

Adversarial Attacks – Federated Learning



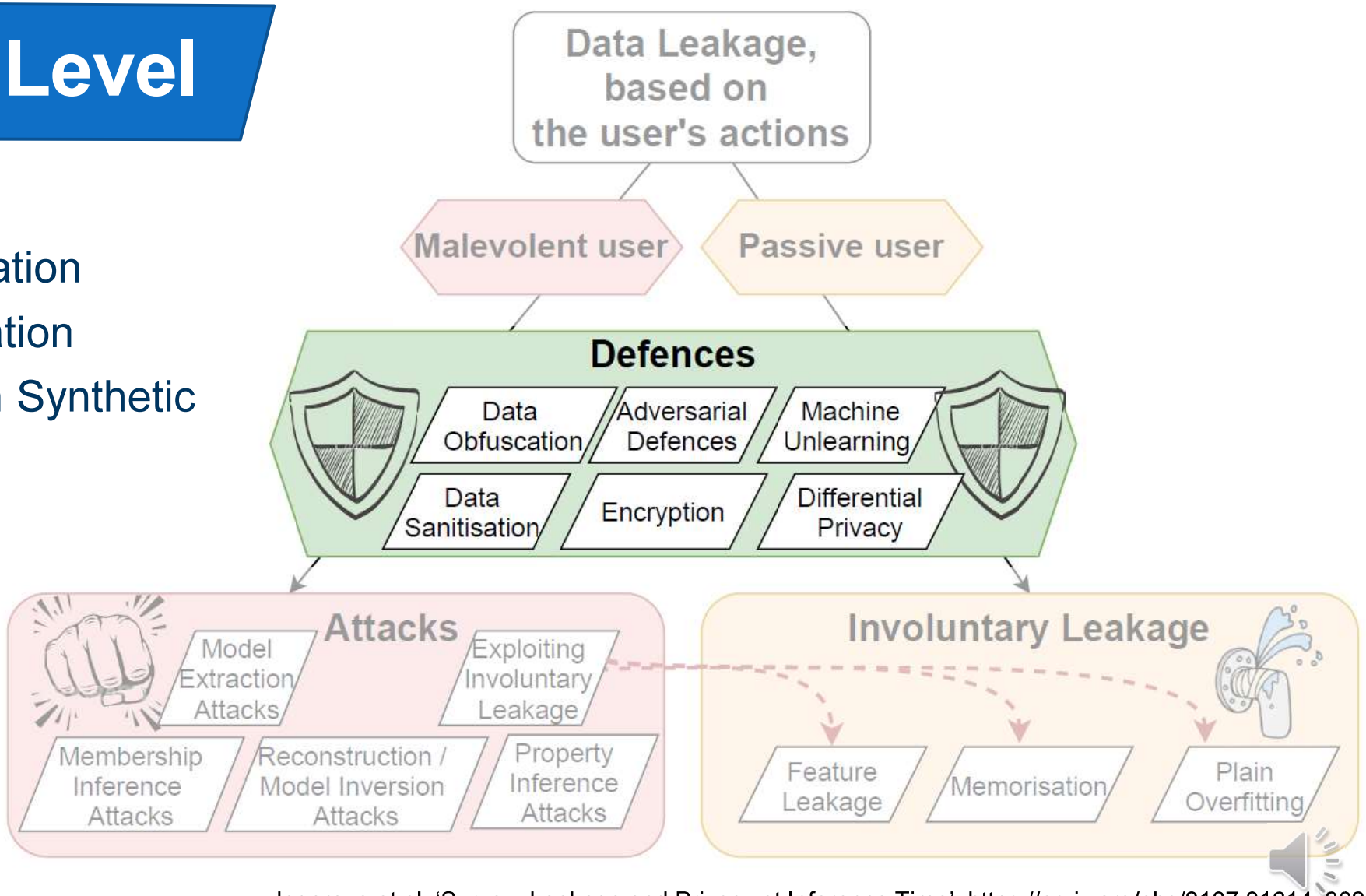
Centralised Learning

Federated Learning

Hitaj et al. 'Deep Models Under the GAN: Information Leakage from Collaborative Deep Learning', ACM Conference on Computer and Communications Security, 2017.

At Data Level

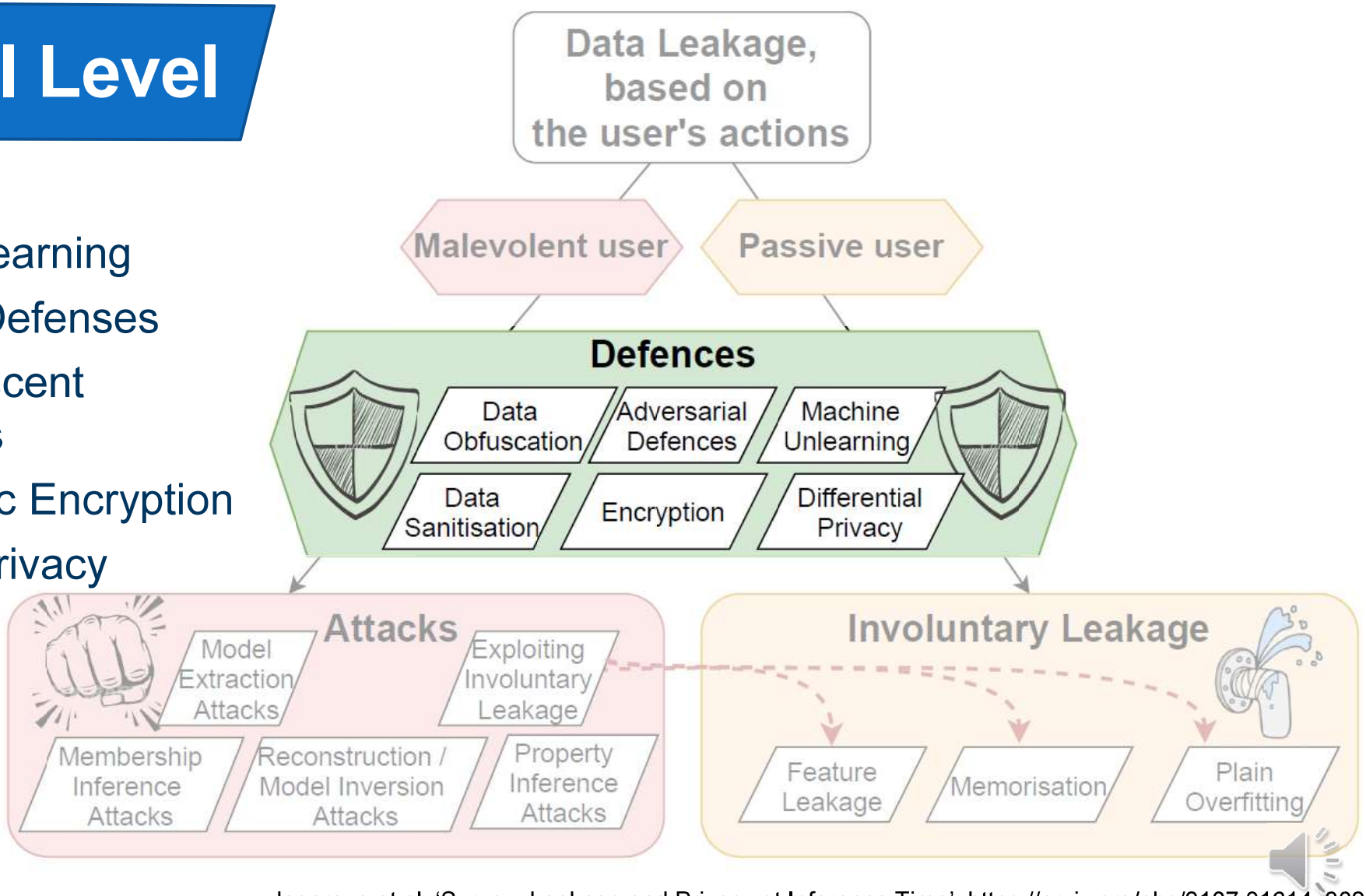
- Data Obfuscation
- Data Sanitisation
- Learning with Synthetic Data



Jegorova et al. 'Survey: Leakage and Privacy at Inference Time', https://arxiv.org/abs/2107.01614, 2021.

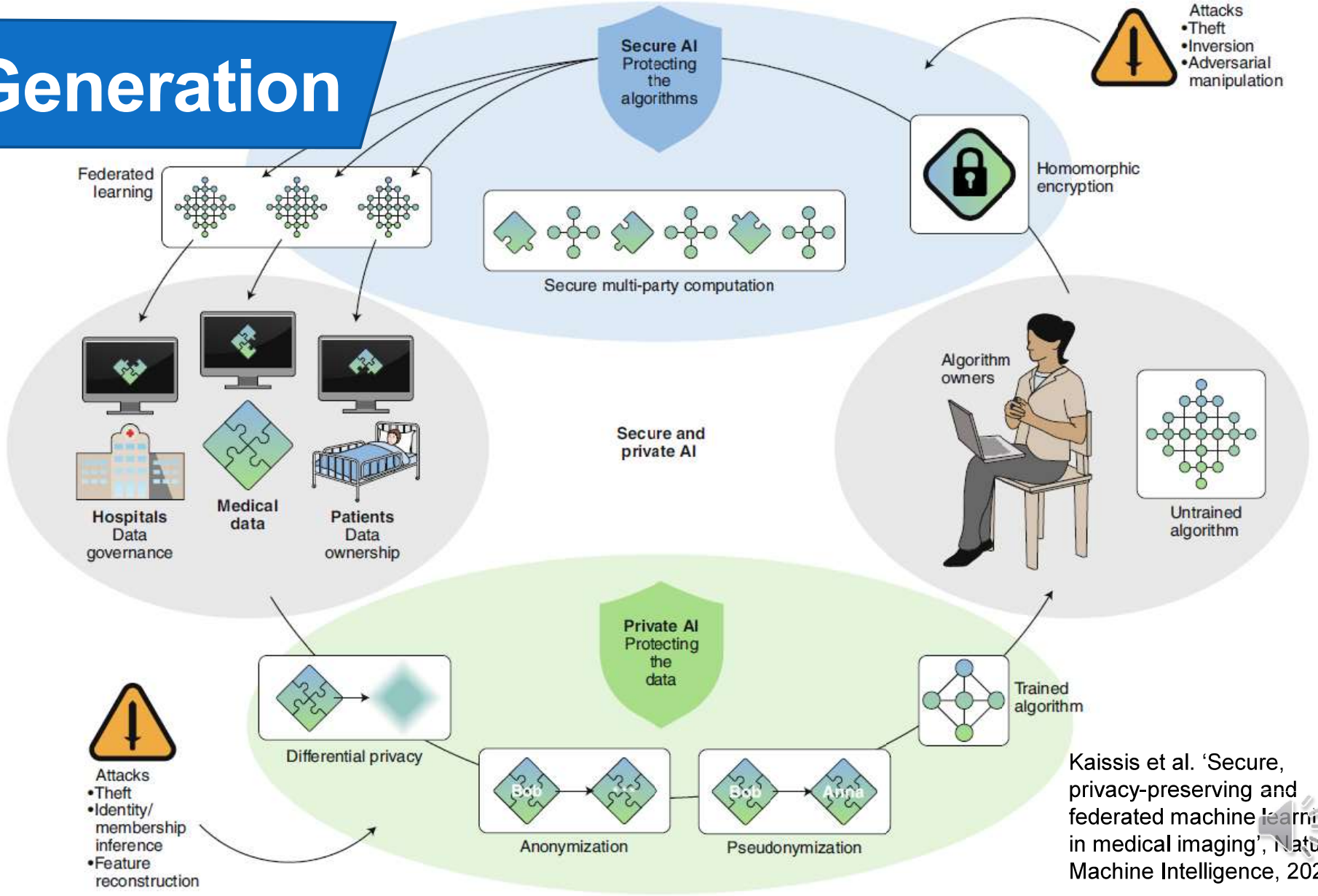
At Model Level

- Machine Unlearning
- Adversarial Defenses
- Gradient Descent Perturbations
- Homomorphic Encryption
- Differential Privacy



Jegorova et al. 'Survey: Leakage and Privacy at Inference Time', https://arxiv.org/abs/2107.01614, 2021.

Next Generation



Kaissis et al. 'Secure, privacy-preserving and federated machine learning in medical imaging', Nature Machine Intelligence, 2020.

Secure Hardware

- They have been used in secure processors
- The have been used in edge hardware and mobiles
- They can play a significant role in federated learning workflows

Open questions

- Does decentralized data storage and federated learning would enable privacy-preserving, cross-institutional research?
- Are encrypted deep learning approaches efficient enough?
- What is the optimum trade-off between accuracy, interpretability, fairness, bias and privacy?
- How to troubleshoot algorithms that are encrypted?

Summary

- Several mechanisms have been proposed to safeguard healthcare data with relation to inference attacks
- Some of these methods work on the data at the preprocessing level and some of them operate on the machine learning models.
- Federated learning have been also proposed as a way to minimize privacy risks while ownership of the data is also retained

References

- Kaissis et al. ‘Secure, privacy-preserving and federated machine learning in medical imaging’, Nature Machine Intelligence, 2020.
- Hitaj et al. ‘Deep Models Under the GAN: Information Leakage from Collaborative Deep Learning’, ACM CCS’17, 2017.
- Jegorova et al. ‘Survey: Leakage and Privacy at Inference Time’, https://arxiv.org/abs/2107.01614, 2021.