1. **Training Data Privacy**: The guarantee that a malicious actor will not be able to reverse-engineer the training data.
2. **Input Privacy**: The guarantee that a user’s input data cannot be observed by other parties, including the model creator.
3. **Output Privacy**: The guarantee that the output of a model is not visible by anyone except for the user whose data is being inferred upon.
4. **Model Privacy**: The guarantee that the model cannot be stolen by a malicious party.

<https://www.usenix.org/system/files/sec19-carlini.pdf>

„The Secret Sharer: Evaluating and Testing Unintended Memorization in Neural Networks“

2019

Kaip neuroniniai tinklai prisimena treniruojamus duomenis. Kaip to išvengti.

<https://ieeexplore.ieee.org/abstract/document/8857960>

Privacy-Preserving Artificial Intelligence: Application to Precision Medicine

2019

SCALABLE PRIVATE LEARNING WITH PATE

(Private Aggregation of Teacher Ensembles)

<https://arxiv.org/abs/1802.08908>

Homomorphic Encryption

*4 generacijos ir jų skirtumai, patobulėjimas. Panaudojamumas. Realizavimas (4-os generacijos).*

*3rd gen:*

*https://tfhe.github.io/tfhe/*

Secure Multiparty Computation (MPC)

Federated Learning

*Modelio kūrimas kompiuteriams neapsikeičiant mokymo duomenimis.*