Thesis proposal

Topic:	**Privacy-preserving federated learning**
Supervisor:	Axel Wisiorek, Haotian Ye
Examiner:	Hinrich Schütze
Level:	MSc
Summary:	Many recent applications of LLMs require large amount of data to be collected and stored centrally, causing concerns related to the privacy and security of the data. In contrast, federated learning allows users to collaboratively train a shared model while keeping the data locally, thus removing the necessity to share sensitive information with a centralized server. However, concerns have been raised that naively applying federated learning may not guarantee data privacy from a honest-but-curious central server, which may recover some information that identifies a specific individual through, e.g., information leakage and membership inference attacks. In this project, we would like to explore privacy-preserving techniques like differential privacy and discuss their limitations (e.g. the privacy-utility tradeoff factor).
Requirements:	good programming skills, ability to use different Transformer models; knowledge in a federated learning framework is a plus.

References:

- Cynthia Dwork et al. (2006). "Calibrating noise to sensitivity in private data analysis". In: *Theory of Cryptography: Third Theory of Cryptography Conference, TCC 2006, New York, NY, USA, March 4-7, 2006. Proceedings 3*. Springer, pp. 265–284
- Ligeng Zhu, Zhijian Liu, and Song Han (2019). "Deep leakage from gradients". In: *Advances in neural information processing systems* 32