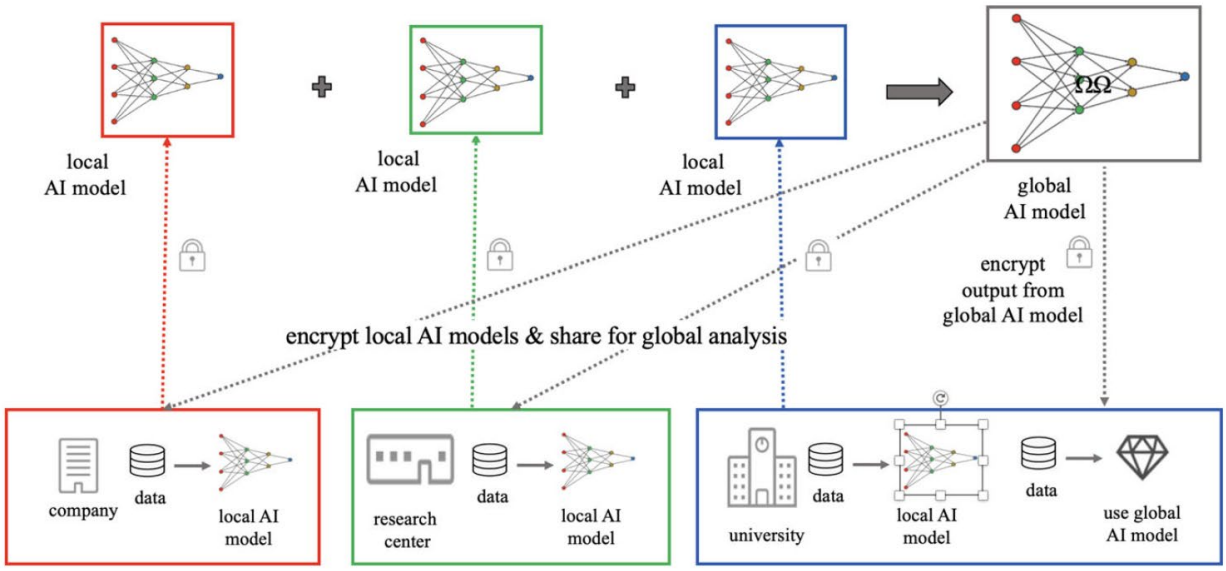
Federated Learning for Addressing Data Privacy and Scarcity Issues

Federated Learning is a relatively new type of multi-party computation, primarily developed to protect sensitive data. Federated Learning is a framework for collaboratively training an AI model in the field of medical health, finance, and public transport by including data differential privacy and homomorphic encryption, with multi-party computation. Federated Learning relies on edge computing and a distributed computing framework, in which processing of raw data takes place on local devices, close to where it was generated. Only results from the local processing of raw data using edge computing are sent to external data centers to train a global AI model collaboratively. A typical federated learning framework consists of training multiple local AI models and transmitting the encrypted weights of these local AI models to an external authority to compute the global AI model, as depicted below.



Cross-silo Federated Learning Framework

Previously, global AI model development involved multi-party computation and data communication, requiring encryption, transmission, and decryption of an entire massive set of raw data from each participant to an external authority, which raised issues of data security, privacy, and cost. Federated Learning addresses the data security and privacy issues by not requiring the transmission of raw data to an external authority. Construction of the global AI model using the Federated learning framework involves significantly less computational work than multi-party computation due to the encryption, transmission, and decryption work required only for the local weights, which typically differ by several orders of magnitude. Furthermore, training the global AI model using the federated learning framework requires less computational work, as federated learning averages the weights of the AI model. It does not necessitate training with a massive volume of data. Frequent updating of the global AI model is also possible with federated learning since data encryption, transmission, and decryption costs are comparatively very low, and computation of the global AI model is also very low. Furthermore, federated learning also addresses the issue of insufficient volume of high-quality data required by each participant to train an AI model. The underlying idea of federated learning is that learning from the collective knowledge and information in all of the participants' datasets is likely to lead to a more accurate and robust global AI model than from any single local model.