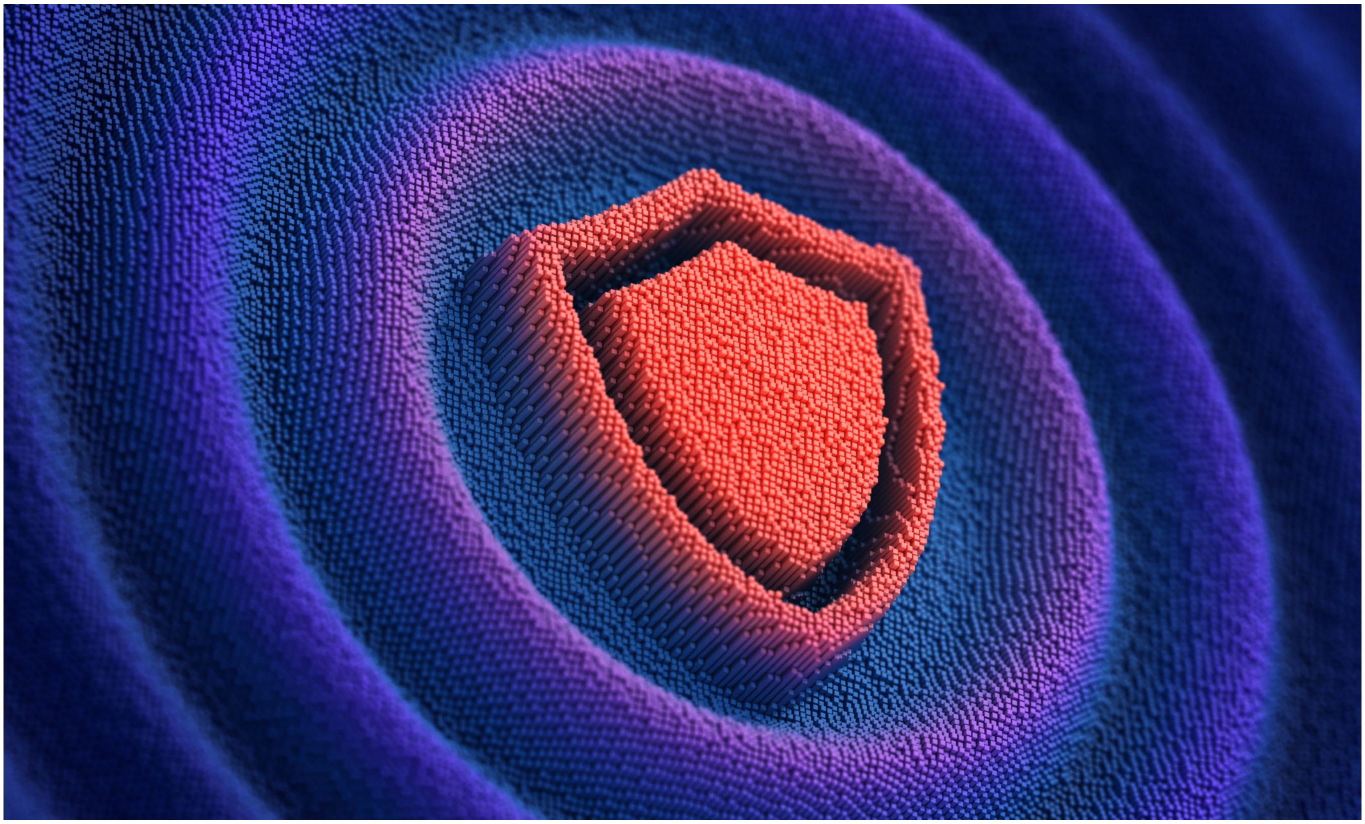
Duality uses Google Cloud's Confidential Computing to enable privacy preserving federated learning

July 2025



Duality

Founded by world-renowned cryptographers and expert data scientists, Duality empowers organizations to securely collaborate on sensitive data with their business ecosystem: customers, suppliers, and partners, or within the same organization in regulated industries.

By operationalizing Privacy Enhancing Technologies (PETs), Duality enables secure analysis and AI on data they can't analyze due to privacy, confidentiality, security, or operational challenges — while complying with data privacy regulations and protecting valuable IP.

Leading industry and government organizations partner with Duality to maximize the value of their data, including DARPA, Intel, Scotiabank, Oracle, IBM, World Economic Forum (WEF), and more.

Duality for secure data collaboration

Data collaboration is critical for AI innovation, but maintaining privacy and security is a persistent challenge, especially in regulated industries like healthcare and finance. Duality Technologies addresses this by integrating Google Cloud's [Confidential Space](), NVIDIA FLARE, and in-house privacy-enhancing capabilities into a unified, easy-to-deploy federated learning platform.

Duality simplifies deployment across multiple organizations with an intuitive installation process and seamless infrastructure integration.

The platform facilitates secure data ingestion and preprocessing, ensuring proper alignment of disparate datasets. It also provides robust project and participant management, defining roles and responsibilities for seamless execution. Additionally, Duality enforces governance and privacy policies, including differential privacy mechanisms, to protect sensitive insights. Finally, automated encryption and attestation ensures all computations are conducted securely within Trusted Execution Environments (TEEs), such as Google Cloud's Confidential Space, eliminating architectural complexities.

Drawbacks of standard federated learning and analytics

While federated learning (FL) keeps raw data decentralized, it does not inherently secure the computed intermediate results. The locally trained model weights, when aggregated, are often exposed in plaintext. This can lead to unintended information leakage, as adversaries can statistically infer underlying data patterns.

Recent academic research has demonstrated that local model updates can be analyzed to reconstruct sensitive training data, presenting a significant privacy risk. Studies have shown that image reconstruction techniques can be used to recover original training data (NeurIPS 2023). Furthermore, research highlights how an aggregator can infer local training data from FL participants (arXiv 2021). Additionally, statistical results have been used to re-identify data, demonstrating potential vulnerabilities (Latanya Sweeney 1997).

Google Cloud's Confidential Space for model aggregation in federated learning

To address these vulnerabilities, Duality has integrated Google Cloud's Confidential Space into its platform. Secure aggregation ensures that AI model updates remain encrypted throughout the federated learning process, preventing unauthorized access or inference attacks.

With Google's Confidential Space, all computations occur within a TEE that is instantiated from the "confidential-space" image provided by GCP. This hardened image does not allow any access unless explicitly enabled, ensuring a high-security execution environment (GCP Confidential Space).

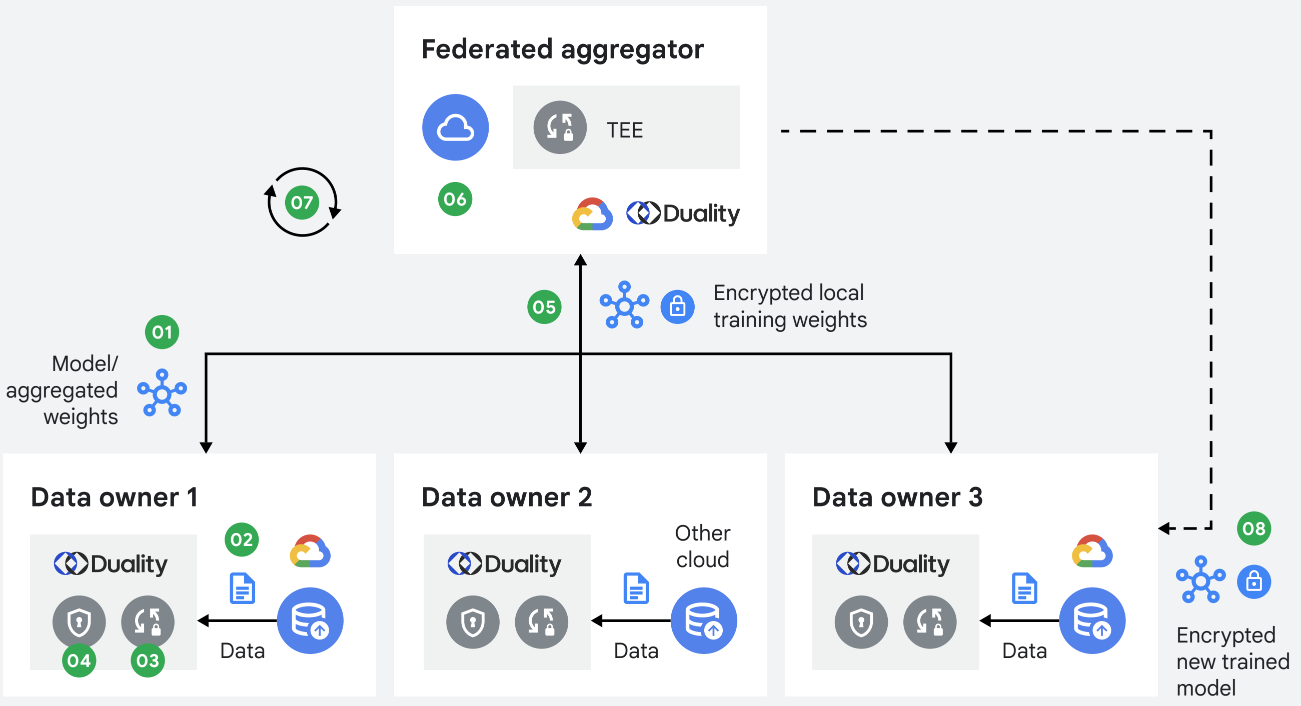
Protect from any potential adversary

With Google Cloud's Confidential Space, all computations occur within a TEE, meaning that:

- Even Duality, as the platform provider, or the account owner, or Alphabet as the Google Cloud owner, cannot access the raw model updates.
- No external party can tamper with or extract sensitive information during the aggregation process.
- Institutions can collaborate with full confidence that their proprietary data and model insights remain protected.

This integration creates a truly secure federated learning ecosystem that combines decentralized training with end-to-end encryption.

Secure federated learning with FLARE and Confidential Space computation flow



Duality's secure federated learning flow ensures privacy and efficiency across multiple organizations by following these key steps:

01 Federated aggregator distributes the data preprocessing and training parameters to participants.

02 Data is digested and preprocessed.

03 Preprocessed data is used for training.

04 Local intermediate results are encrypted.

05 Encrypted local training weights are sent to the federated aggregator for aggregation.

06 Federated aggregator decrypts inside the TEE and aggregates the intermediate results to get the new model weights.

07 The process repeats until model convergence for training. For analytics, this step is not required as it focuses on aggregated insights rather than iterative model refinement.

08 New trained model is sent to any selected party.

Use case:

Oncology research at Dana-Farber Cancer Institute

Cancer research relies on large-scale digital pathology data from multiple institutions. However, regulatory and privacy concerns make data sharing a slow and complex process. Pathology images, classified as Protected Health Information (PHI), cannot be freely exchanged between organizations, creating barriers to collaboration.

To overcome this, Dana-Farber Cancer Institute partnered with Duality Technologies to implement a secure federated learning framework. Using Duality's platform:

+

Hospitals could collaboratively train AI models on digital pathology images without transferring patient data.

+

Model training occurred locally, with only encrypted weights transmitted to an aggregation server.

+

Google Cloud Confidential Space ensured model aggregation was conducted securely, preventing data exposure.

Key features, technical implementation, and results

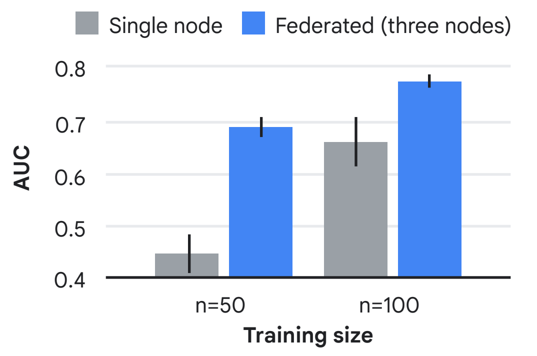
The federated learning model was implemented with two participating collaborators, each contributing 100 whole slide images which are high-resolution digital scans of entire pathology slides. The images, each approximately 1GB in size, are segmented into smaller images, each being passed through a pre-trained Convolutional Neural Network (CNN) for feature extraction.

The resulting vectors are then used to train a state-of-the-art cancer classification model named CLAM. CLAM, using an architecture of self-attention and SVM, is trained to provide a slide-level diagnosis of the entire initial image. The model was tested and compared in two different scenarios:

Federated training:
Data processed at all locations locally, with training results aggregated centrally.

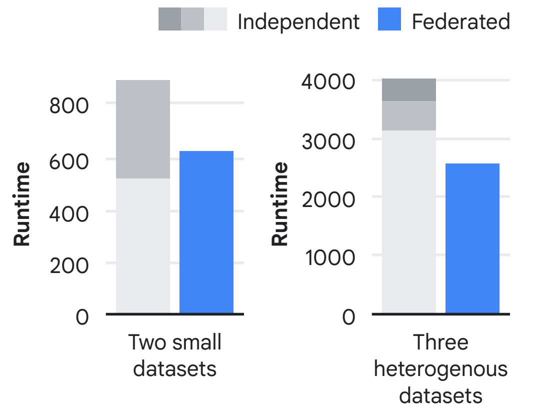
Single site training:
Train on a single site data.

Federated learning achieves much higher accuracy

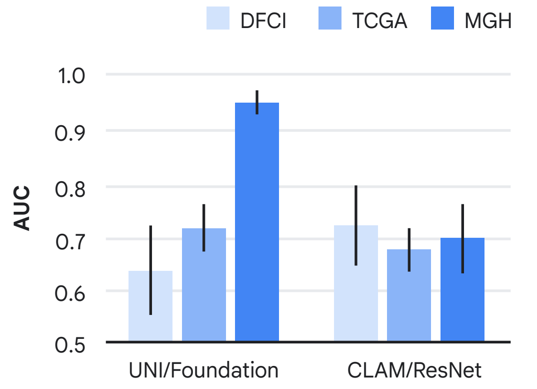


*Comparable to accuracy in the literature

Federated learning is more efficient



Results: Substantial heterogeneity of data across institutions



Interpretation of model performance can be highly contigent on the dataset. Federated learning appears to be robust to heterogeneity (but more testing needed).

The results demonstrated that secure federated learning enabled organizations to train a high quality model, specifically comparable to centralized AI training, while ensuring privacy compliance. Additionally, it surpassed the performance of single-site training while allowing data to remain in place, eliminating the need to transfer large-scale datasets, highlighting its advantage in collaborative model development.

Digital pathology models can be effectively trained in a federated environment.

Federated framework provides increased accuracy and efficiency.

Google Cloud’s Confidential Space is essential to guarantee the security of local results.

Cross-institutional analyses are critical and can substantially change the research conclusions.

The Duality platform provides custom loss functions, network architecture, and evaluations.

Google Cloud

Learn more at dualitytech.com