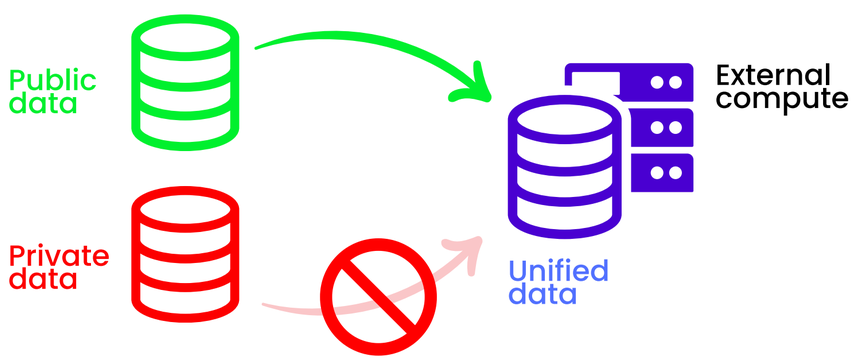
CAN AI STILL LEARN WHEN DATA CANNOT BE SHARED?

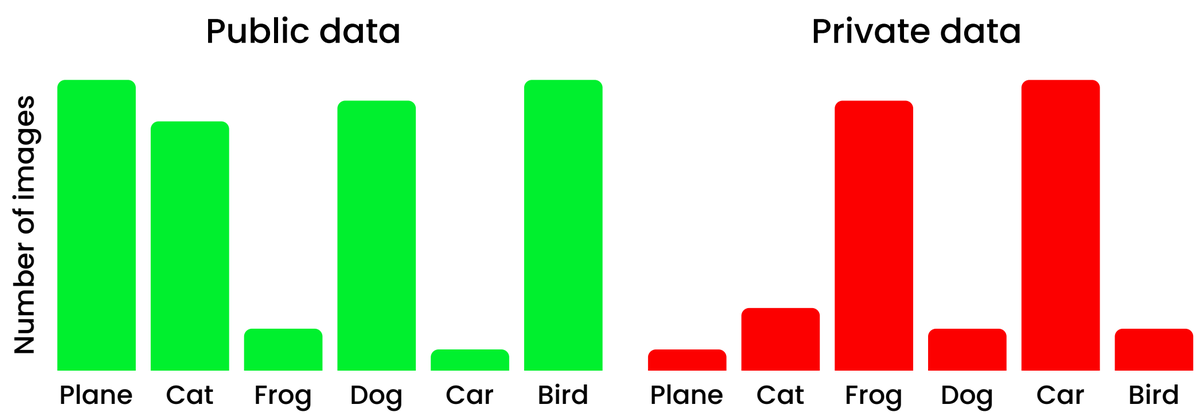
Improving Model Performance While Keeping Sensitive Data Private

Strong AI models rely on large and diverse datasets. And conventional training requires these data to be **brought together** in a single training environment.

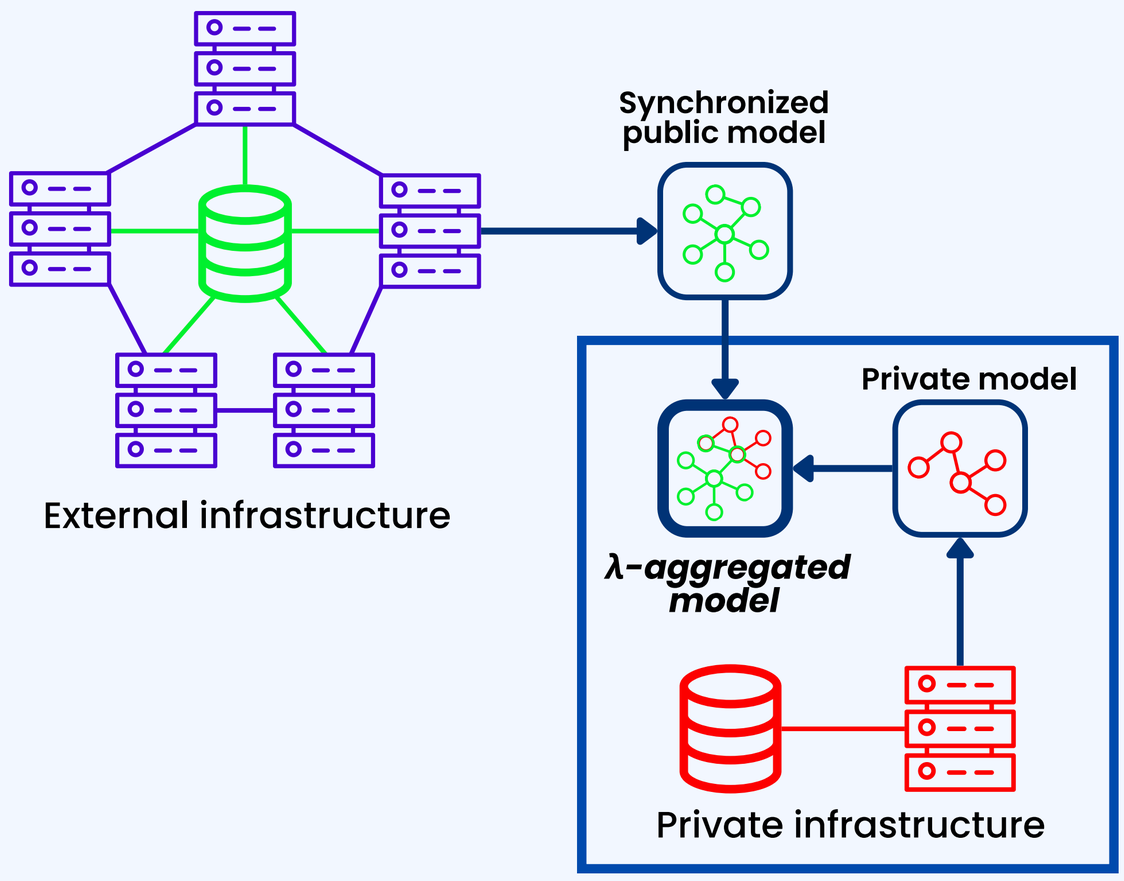


When local compute is insufficient, training relies on external infrastructure. Public data can be processed there, but **sensitive private data must stay local** [1], preventing the two data sources from being used jointly.

Public and private data are often **complementary**; using either alone leaves blind spots.



Existing approaches still trade privacy for performance or exchange updates that may leak sensitive information [2,3].
How can we learn from both without exposing private information?



Our answer? Move *knowledge*, not data.

Two models train in parallel: one on public data, the other on private data. **The public model is periodically transferred to the private infrastructure**. The private model is then updated using a weighted average of their parameters, balancing private and public knowledge:

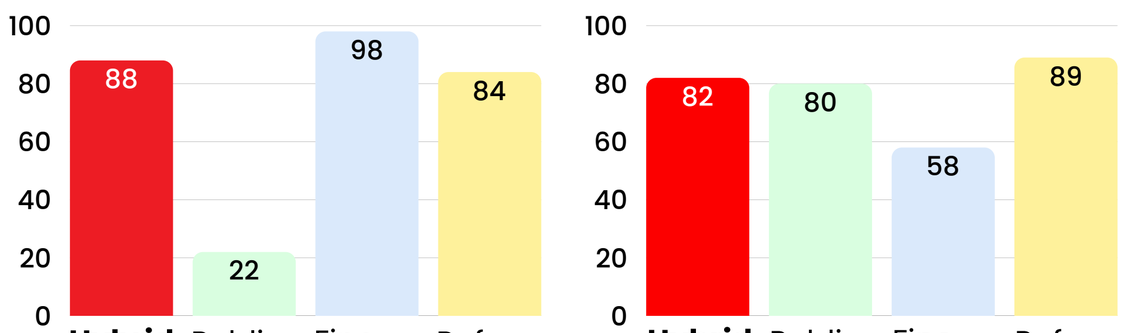
$$\theta_{\text{new}} = \lambda\,\theta_{\text{private}} + (1-\lambda)\,\theta_{\text{public}}$$

θ: model parameters · λ: weight given to the private model

No private data or model updates are sent back.

Hybrid learning dramatically improves private accuracy

22% → 88%
vs. public-data-only training



Fine-tuning shown here is another common alternative: a model trained on public data is then retrained on private data. It achieves the highest private accuracy but hurts overall performance.
Hybrid stays strong on both, nearly matching a reference model trained jointly on all data.

Where this matters

The architecture helps AI adapt to local needs while preserving overall performance — all without sharing private data.

Healthcare
Local patient populations

Finance
Local fraud patterns

Industry
Site-specific operations

Tested on CIFAR-10 and CIFAR-100 with LeNet and ResNet-20. Research conducted at SaCS, EPFL.
References, full experimental details & acknowledgements via **QR code**.

Dimitri URBAIN
dimitri.urbain@epfl.ch

SaCS EPFL LAIDLAW SCHOLARS