swiss made software

TUNE INSIGHT

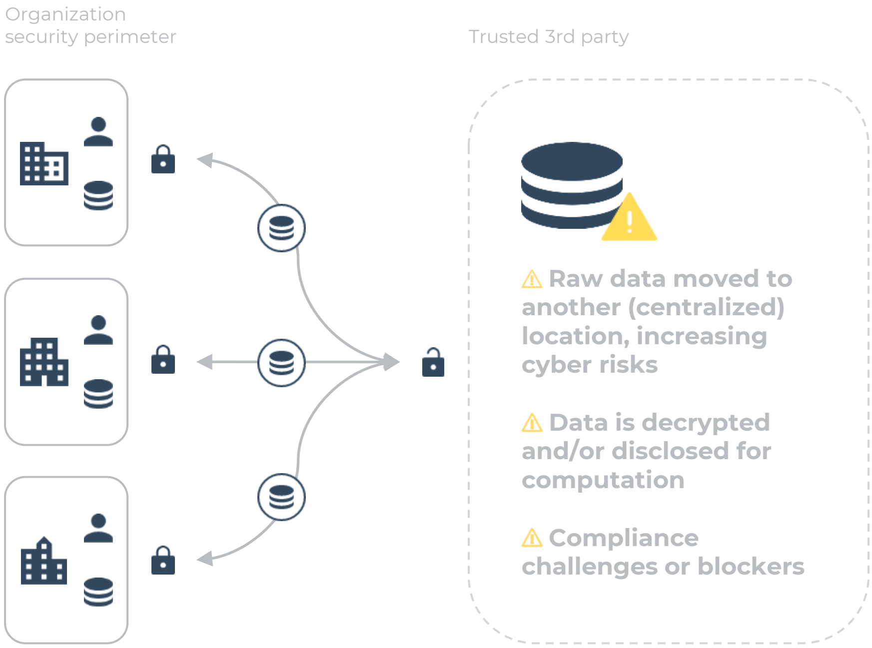
Federated Encrypted Computing: Share Insights, protect the data

Basel, May 16 2024

Juan R. Troncoso, co-founder & CEO

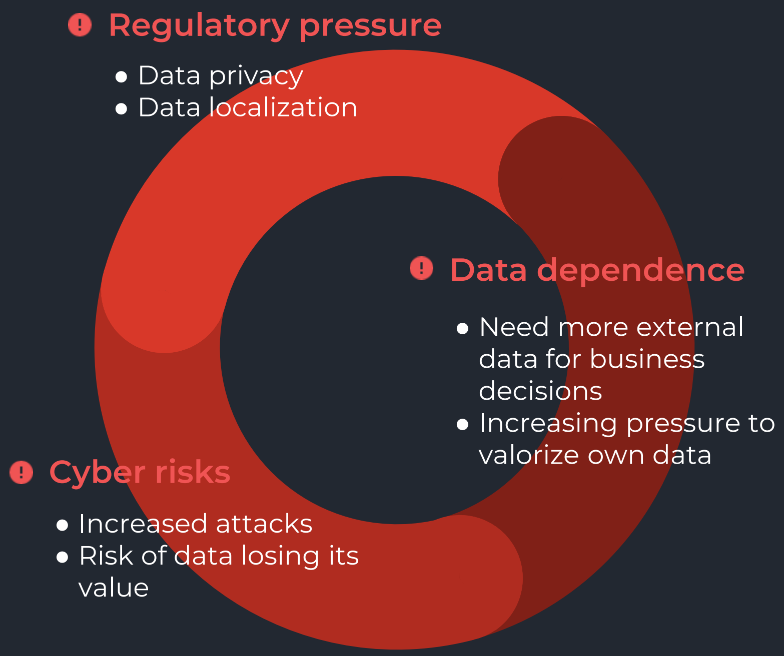
juan@tuneinsight.com

Regulatory pressure and cyber risks prevent companies to access external data they need & to valorize data they own



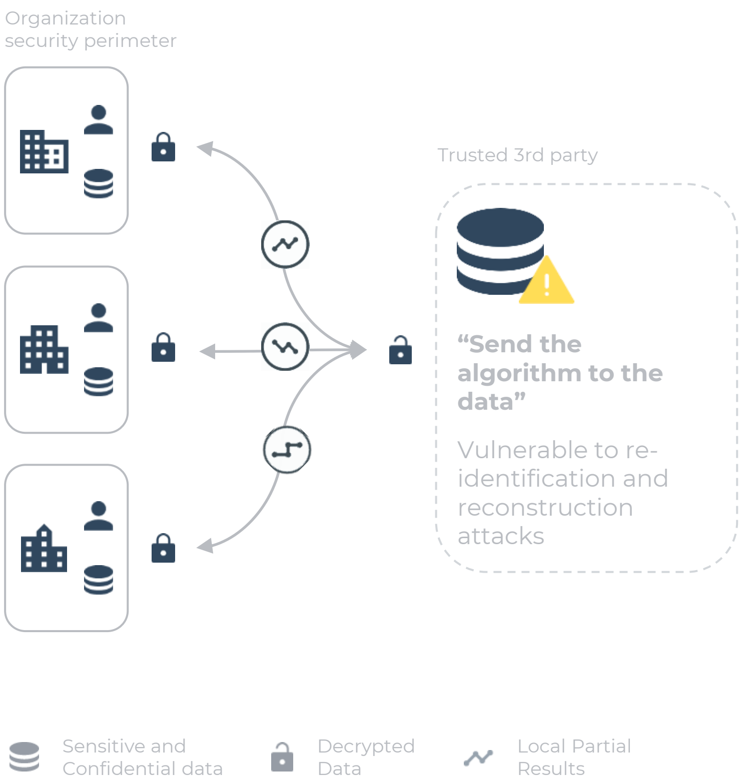
Sensitive and Confidential data

Decrypted Data



Vicious data circle

“Vanilla” Federated Learning



TUNE INSIGHT

- **Requires trust on the aggregation server**
- **Vulnerable to re-identification and reconstruction attacks**
  - B. Hitaj, G. Ateniese, and F. Perez-Cruz. Deep models under the GAN: Information leakage from collaborative deep learning. In ACM CCS, 2017.
  - Z. Wang, M. Song, Z. Zhang, Y. Song, Q. Wang, and H. Qi. Beyond inferring class representatives: User-level privacy leakage from federated learning. In IEEE INFOCOM, 2019.
  - L. Zhu, Z. Liu, and S. Han. Deep leakage from gradients. In NIPS. 2019.
  - L. Melis, C. Song, E. De Cristofaro, and V. Shmatikov. Exploiting unintended feature leakage in collaborative learning. In IEEE S&P, 2019.
  - M. Nasr, R. Shokri, and A. Houmansadr. Comprehensive privacy analysis of deep learning: Passive and active white-box inference attacks against centralized and federated learning. In IEEE S&P, 2019.

Federated Health Data Space: Technologies

TUNE INSIGHT

Homomorphic Encryption
encrypted computation
→ Raw data is never shared, only encrypted data is used for computations.

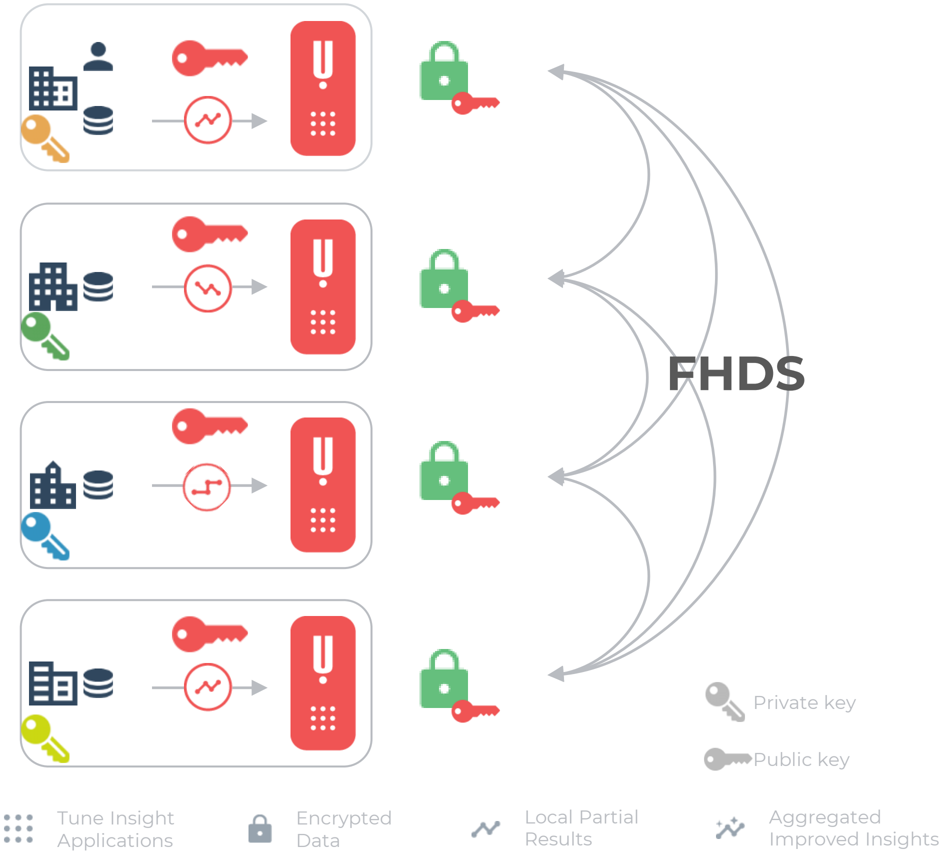
Secure Multiparty Computation
collective key
→ Decryption of the results can only be run collectively.

FL/FA Aggregation
of individual data
→ Only encrypted aggregates are shared.

Differential Privacy
And disclosure prevention controls
→ Minimize re-identification risk

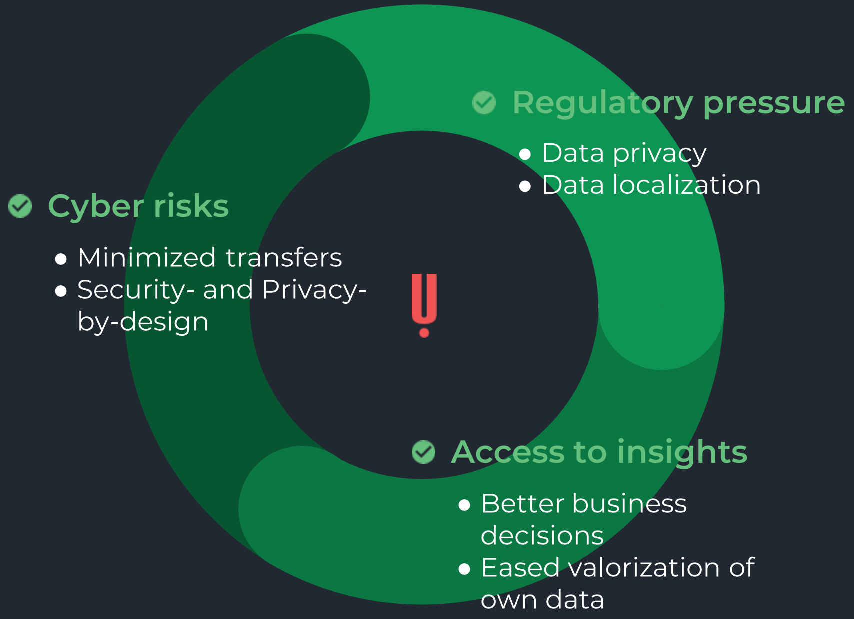
Synthetic Data
Unlinkability
→ Limited leakage for exploratory workflows

- ✓ **Embedded policy enforcement**
- ✓ **Raw data does not move**
- ✓ **Encrypted computation**
- ✓ **Only aggregated data is decrypted**

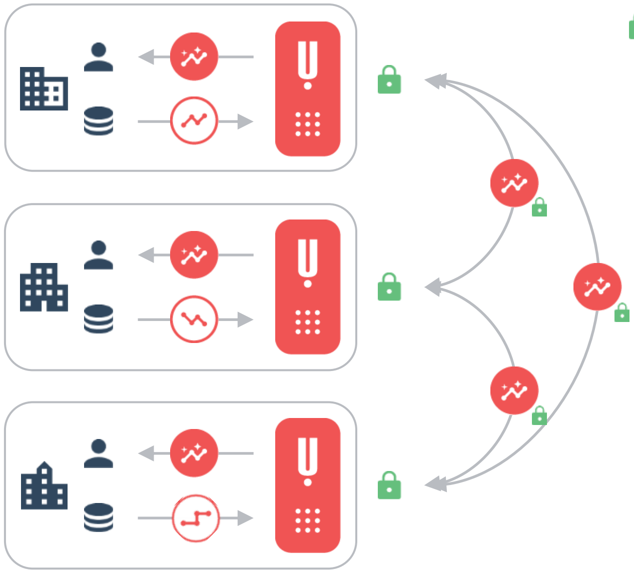


tuneinsight.com

This combination of technologies minimizes risks and streamlines compliance



Virtuous insights circle



Combination of Federated Learning and Privacy Enhancing Techniques

Based on 10+ years of world-class EPFL research, published in Nature Comms

Raw data does not move

Controlled computation over encrypted data

Tune Insight Applications · Encrypted Data · Local Partial Results · Aggregated Collective Insights

TUNE INSIGHT

Platform used across verticals

Hospitals & Pharma

Collective survival analysis in oncology

Lab reference data

Train image classifiers in dermatology

CHUV

Universitätsspital Basel

USZ

INSELSPITAL
UNIVERSITÄTSSPITAL BERN
HÔPITAL UNIVERSITAIRE DE BERNE

HUG Hôpitaux Universitaires Genève

+ others

Insurance & Re-Insurance

Train collective risk models

Cross-vertical collaboration (Value-Based Healthcare)

groupemutuel

Future of Health Grant

+ others

Cyber Security

Cross-organization alert enrichment

Collective threat intelligence models

Private search of IoCs/alerts

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

armasuisse
Science and Technology

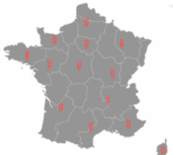
+ others

Financial Services

Collaborative analytics

Sensitive data pooling, AML-CFT

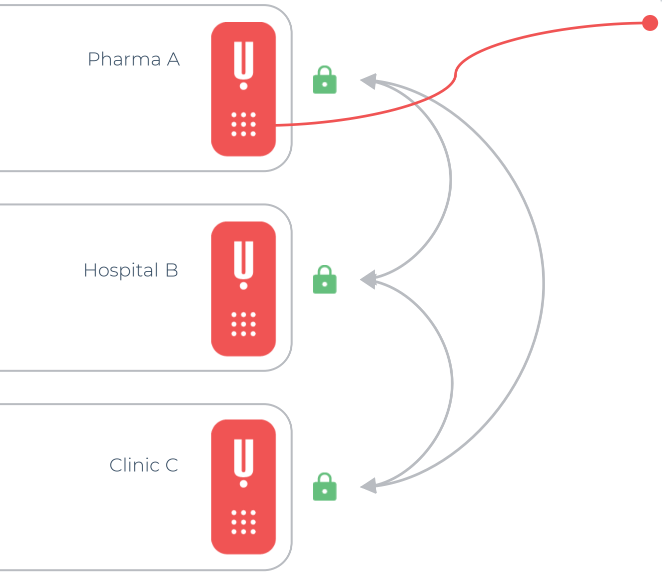
ACPR
BANQUE DE FRANCE



Participated in the Tech Sprint organized by ACPR on Confidential Data Pooling for AML-CFT in 2022

Confidential Collaborative Analytics, Machine Learning and AI

Pharmas, hospitals and insurers can collaborate without transferring or disclosing patient-level data

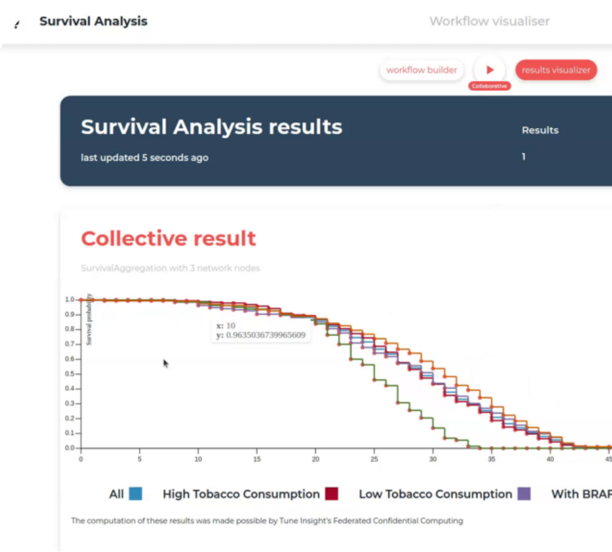


Example Healthcare

Relying only on their own data, hospitals and clinics lack representative datasets to provide personalized care

With Tune Insight, they can collaborate with others to recommend precision treatments without moving or disclosing any raw patient data, and include private players in the collaboration

Developed frontend and backend integrations



Python SDK
Use case: developer-friendly experience for data analysts and integrations

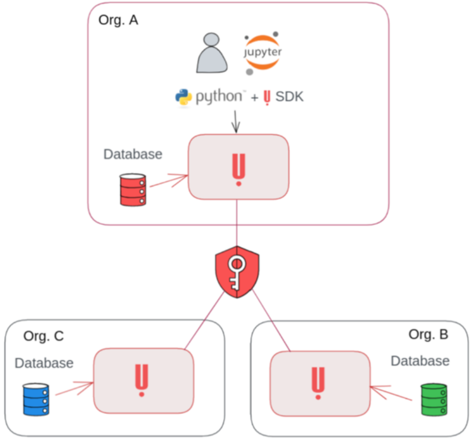
Tune Insight Documentation TI4Health tuneinsight.com

- Getting Started
- Generic Architecture
- Integrations
- TI4Health
- Python SDK

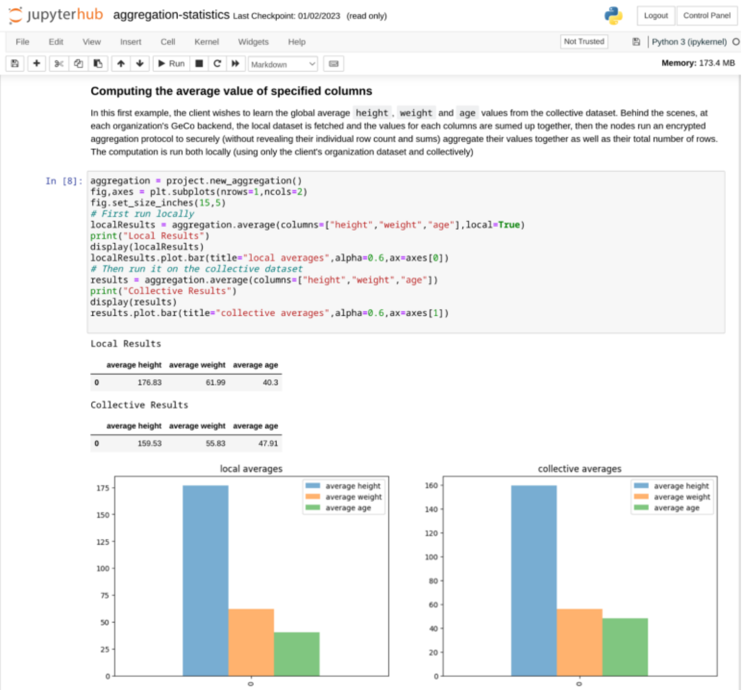
Python SDK

Python SDK

The Diapason Python SDK is a client library for the Tune Insight API. It allows users or programs to interact with t[…] using Python. For example, computation workflows can be created and executed using Juptyer notebooks or oth[…] Python programs.



jupyter python



Example: Cross-jurisdiction Federated Learning for Dermatology

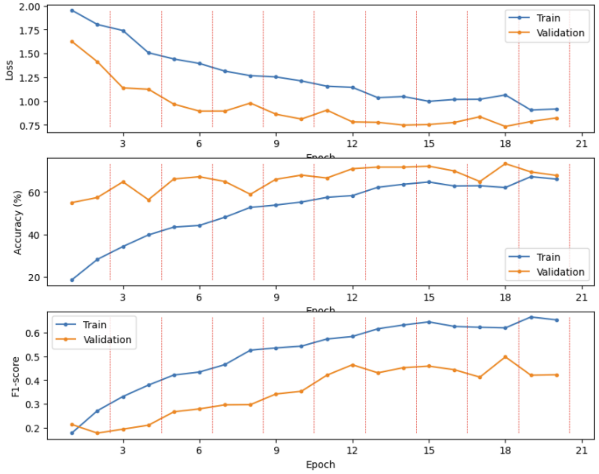
Final aggregated model path

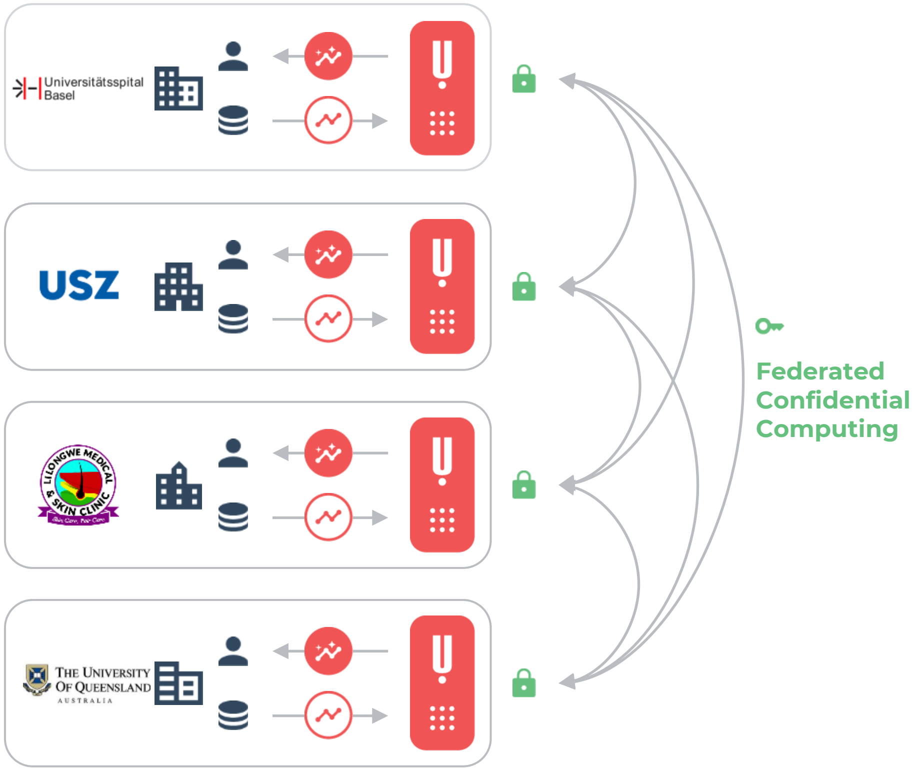
In [11]:
```
result_path = ml_result.result_path
result_path
```

Out[11]: 's3://ti-bucket0/ef63cc9a-2d62-4f3d-9b7c-5220f3c3be08_mlshare_public'

In [12]:
```
history = ml_result.history
```

In [13]:
```
hybrid_fl.plot(history)
```





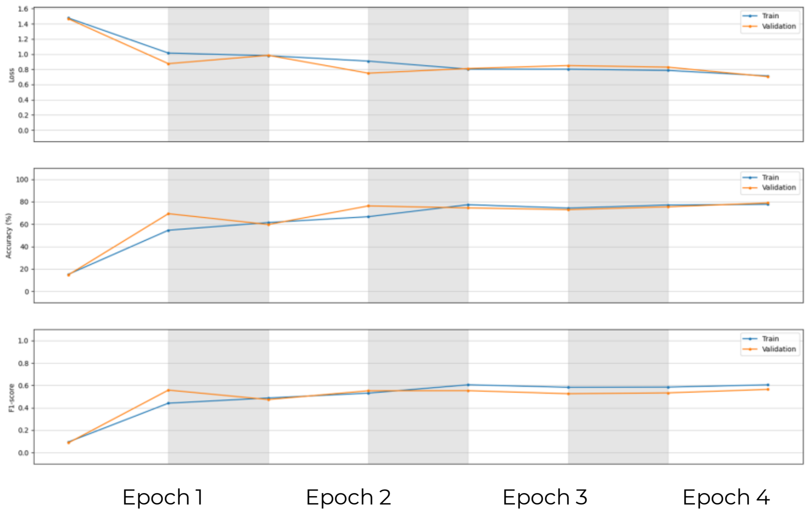
Tune Insight Applications

Encrypted Data

Local Partial Results

Aggregated Improved Insights

Example: Secure Federated Training of Deep Neural Networks on Dermatology Images



4 epochs	**Local training baseline**	**Secure federated training**
Nodes	1 node with 10909 samples	3 nodes (~3635 samples each)
Training accuracy	72.16%	77.65%
Training F1-score	0.279431	0.604438
Validation accuracy	72.13%	78.88%
Validation F1-score	0.279364	0.564171
Privacy params	N/A	ε = 1.0, δ = 0.0001
Time overhead	0	~10% (w.r.t. vanilla FL)

Dataset: Fitzpatrick17k, ~30k images (https://github.com/mattgroh/fitzpatrick17k)
Model:
Type: ViT with 4-layers embedding
Size: 5,528,259 parameters, 44.3MB

100 seconds/epoch on a g4dn.2xlarge AWS EC2 instance with a Nvidia T4 GPU (16GB memory)

Legal analysis - GDPR Compliance

*“**Technical solutions such as multiparty homomorphic encryption (MHE)** that combine these three technical measures while still allowing for the possibility to query and analyse encrypted data without decrypting it have significant potential to **provide effective security measures that facilitate cross-borders transfers of personal data in high-risk settings**.”*
Compagnucci et al., Supplementary Measures and Appropriate Safeguards for International Transfers of Personal Data after Schrems II (February 23, 2022). https://ssrn.com/abstract=4042000

Contact us for a full analysis of the platform benefits and risk minimization, addressing the relevant GDPR recitals.

Article 25 Data protection by design and by default	Article 32 Security of processing	Article 33 Breach notification to supervisory authority
Article 34 Breach communication to the data subject	Article 35 Data protection impact assessment	Article 46 Transfers subject to appropriate safeguards

Data Protection Impact Assessment (DPIA) for multisite medical data analysis (Example, June 2021)

TUNE INSIGHT

Centralized approach with standard pseudonymization

Threat	Threat likelihood	Threat impact	Risk	Risk level
Unlawful access to the system	Unlikely	High	Loss of data confidentiality	Moderate
Malicious use of the system	Possible	High	Loss of data confidentiality	High
Loss of data	Unlikely	Minor	Loss of data integrity, data unavailability	Minor
Data leak of host/cloud	Possible	High	Loss of data confidentiality	High
Collusion of host/cloud	Possible	High	Loss of data confidentiality	High
Corrupted or malicious host/cloud	Possible	High	Data unavailability, loss of data integrity, loss of data confidentiality, loss of data correctness	High
Unavailability of host/cloud	Possible	Minor	Data unavailability, loss of data correctness	Moderate
Re-identification/attribute inference	Possible	High	Loss of data confidentiality	High

Federated approach enhanced with TUNE INSIGHT

Threat	Measure introduced TUNE INSIGHT	Threat likelihood	Threat Impact	Risk	Risk level
Unlawful access to the system	1	Unlikely	Minor	Loss of data confidentiality	Low
Malicious use of the system	1, 2, 4, 10	Possible	Minor	Loss of data confidentiality	Low
Loss of data	3, 5	Unlikely	Minor	Loss of data integrity, data unavailability	Low
Data leak	4, 5, 8, 9, 10	Unlikely	Minor	Loss of data confidentiality	Low
Collusion between nodes	4, 9	Unlikely	Moderate	Loss of data confidentiality	Moderate
Corrupted or malicious nodes	2, 5, 6, 7, 8, 9	Unlikely	Moderate	Data unavailability, loss of data integrity, loss of data confidentiality, loss of data correctness	Moderate
Unavailability of of nodes	6, 7	Possible	Minor	Data unavailability, loss of data correctness	Moderate
Re-identification or attribute inference	1, 2, 4, 9, 10	Unlikely	Minor	Loss of data confidentiality	Low

Contact

swiss made software

Juan R. Troncoso, co-founder & CEO

juan@tuneinsight.com

https://tuneinsight.com

https://linkedin.com/company/tuneinsight

Federated Encrypted Computing: Share Insights, protect the data

Basel, May 16 2024