SNIA CSTI | CLOUD STORAGE TECHNOLOGIES

The Rise of Confidential AI

Live Webinar

July 26, 2023

10:00 am PT / 1:00 pm ET

Today's Presenters

Parviz Peiravi
Global CTO, Financial Services
Industry Solutions
Intel

Dr Richard Searle
Vice President of
Confidential Computing
Fortanix

Erin Farr
Storage CTO Office, IBM
Vice Chair SNIA Cloud Storage
Technologies Initiative

The SNIA Community

200 Corporations, universities, startups, and individuals	**2,500** Active contributing members	**50,000** Worldwide IT end users and professionals

SNIA CSTI | CLOUD STORAGE TECHNOLOGIES

What We Do

Educate vendors and users on cloud storage, data services and orchestration

Support & promote business models and architectures: OpenStack, Software Defined Storage, Kubernetes, Object Storage

Understand Hyperscaler requirements Incorporate them into standards and programs

Collaborate with other industry associations

SNIA Legal Notice

The material contained in this presentation is copyrighted by SNIA unless otherwise noted.

Member companies and individual members may use this material in presentations and literature under the following conditions:

- Any slide or slides used must be reproduced in their entirety without modification
- SNIA must be acknowledged as the source of any material used in the body of any document containing material from these presentations.

This presentation is a project of SNIA.

Neither the author nor the presenter is an attorney and nothing in this presentation is intended to be, or should be construed as legal advice or an opinion of counsel. If you need legal advice or a legal opinion please contact your attorney.

The information presented herein represents the author's personal opinion and current understanding of the relevant issues involved. The author, the presenter, and the SNIA do not assume any responsibility or liability for damages arising out of any reliance on or use of this information.

NO WARRANTIES, EXPRESS OR IMPLIED. USE AT YOUR OWN RISK.

Today's Agenda

- Security Is a Business Imperative
- Protect Data and Code In-Use
- The Need for Confidential AI
- Technology and Use Cases
- Real World Case Studies
- Q & A

Security Has Never Been More Important

^50%
Increase in 2021 cyberattacks YoY [1]

Regulatory fines

$458.9B
Projected global cybersecurity spending by 2025 [2]

Brand and reputational damage

$4.35M
Average cost of a single data breach in 2021 [3]

Fraud, loss of sensitive data or IP

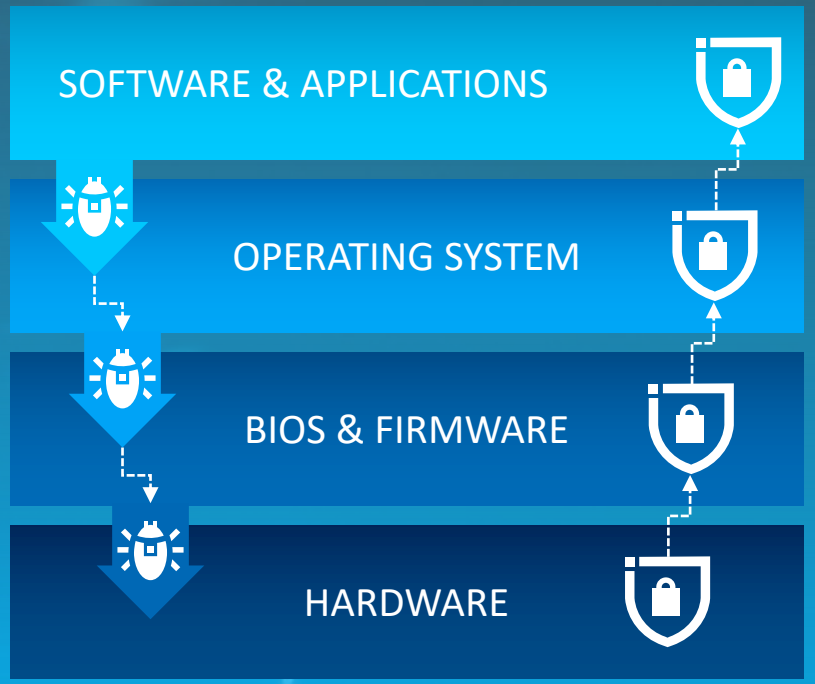
277 days
Average time to detect and contain a data breach [3]

93%
Networks estimated to be vulnerable to cyberattacks [4]

Business downtime & recovery

2 seconds
Pace of new ransomware attacks by 2031 [1]

Threats are growing and risks have never been higher



Attacks are going lower in the stack

1. https://spanning.com/blog/cyberattacks-2021-phishing-ransomware-data-breach-statistics/
2. https://cybersecurityventures.com/cybersecurity-spending-2021-2025/
3. https://www.ibm.com/reports/data-breach
4. https://www.forbes.com/sites/chuckbrooks/2022/06/03/alarming-cyber-statistics-for-mid-year-2022-that-you-need-to-know/?sh=47668cf17864

Security Is a Business Imperative

Customers Are Demanding It

Your customers trust you to protect their data and privacy.

How do you meet the demands?

Governments Are Regulating It

Rules mandating how data is stored and shared are increasing.

How do you maintain compliance?

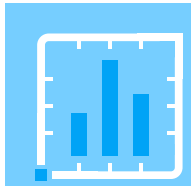
Cloud Leaders are Driving It

Trust is an area of differentiation used to drive competitive advantage.

How do you stack up?

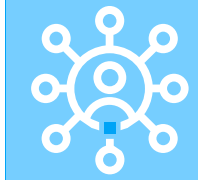
Data Protection Is Driving Business Opportunity

Corporations of all sizes recognize the accelerated pace of global business and have modernized their technology infrastructure to provide deeper business insights and to bring new products to market faster.



A Fundamental Requirement

- World data will grow 61% year-over-year by 2025[1]
- By 2020, about 1.7 MB/sec new data by every human on the planet[2]



Workload Placement Decision Factors

- Regulation and Compliance
- Trust
- Cost and Complexity
- Zero Trust Security Strategy

Trust

Enable business to innovate with velocity while remaining safe and trusting the CSP. Enable availability of backup and recovery resources. Impact is visible in industry studies that cite data security concerns as a barrier to adoption of cloud computing (business model barrier). Impacts enterprises and service providers.

Cost and Complexity

Infrastructure is getting more complex and becoming more expensive. If infrastructure is not current, your hardware is not as secure. You also hear from enterprises that legacy infrastructure is a barrier to innovation and makes it harder to enable security tools, features, etc. At the infrastructure level, you're not as protected.

Regulation and Compliance

Adding regulatory and compliance requirements for data protection, GDPR, and ensuring compliance with current regulations related to data security and privacy, data sovereignty, and transparency and control over operations.

1 Reference: https://www.networkworld.com/article/3325397/idc-expect-175-zettabytes-of-data-worldwide-by-2025.html
2Reference: https://quantium.com/wp-content/uploads/2016/08/BFM_Quantium.pdf

Protect Data and Code In-Use

Data at Rest

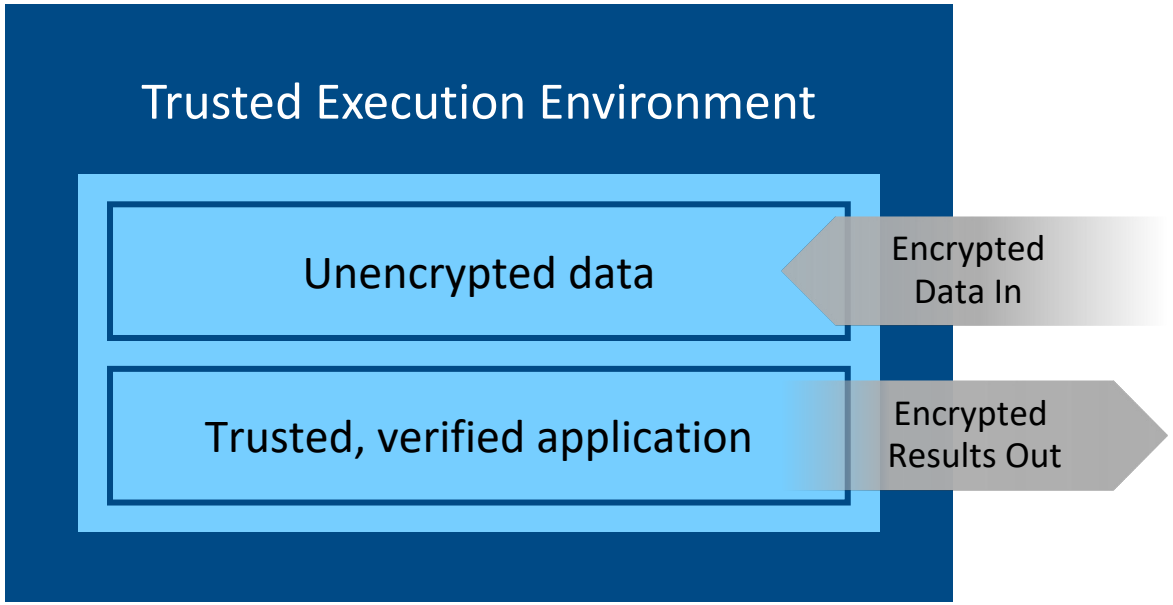
Storage Encryption

Data in Transit

Network Encryption

Data in Use

Confidential Computing



Isolation

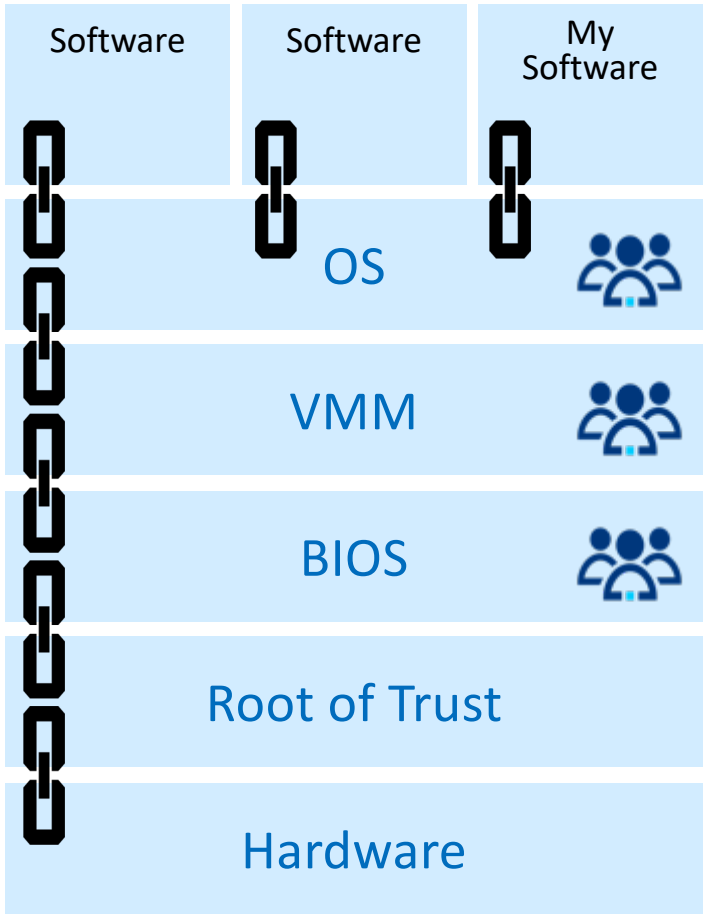
Designed so only authorized application software inside the TEE can access confidential data

Attestation

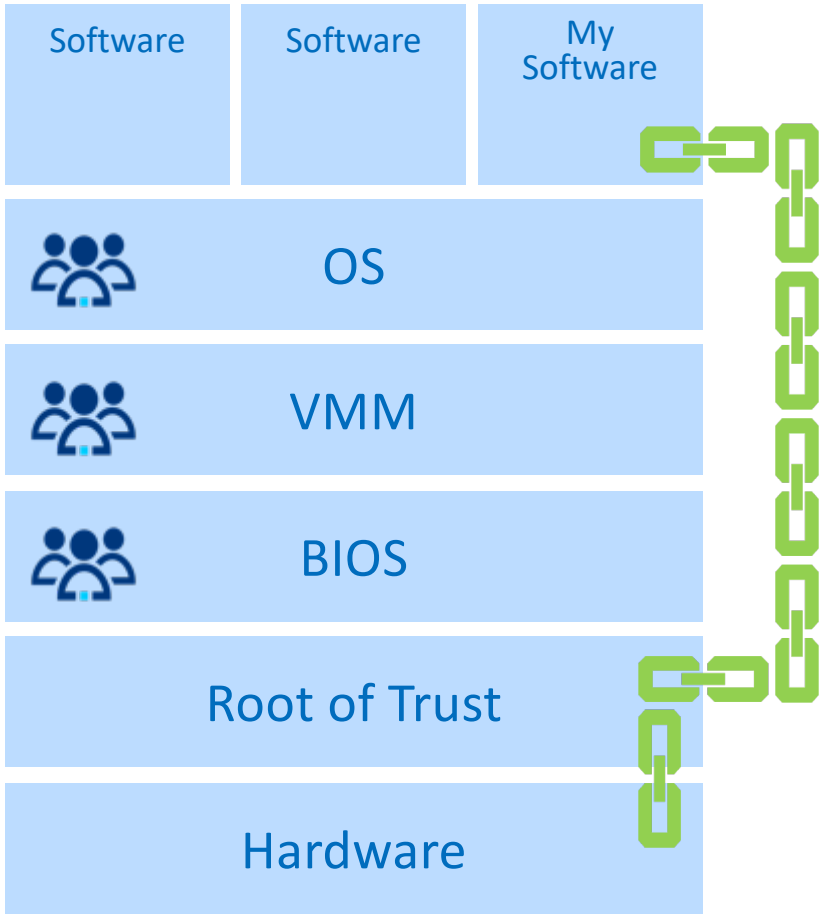
Confirmation that HW & SW configuration is valid & exactly as expected

Rethink the Way Trust Works

Traditional Chain of Trust



Confidential Computing Chain of Trust



When is this **Useful?**

Multi-party sharing scenarios

- Data usage without exposing the underlying data to any of the other parties

Moving sensitive workloads to managed infrastructure (e.g., cloud)

- Attestable Integrity and Confidentiality of the data and code
- Removing the Infrastructure owner (e.g., Cloud Service Provider) from the trusted computing base
- Hardware enforced isolation from other tenants

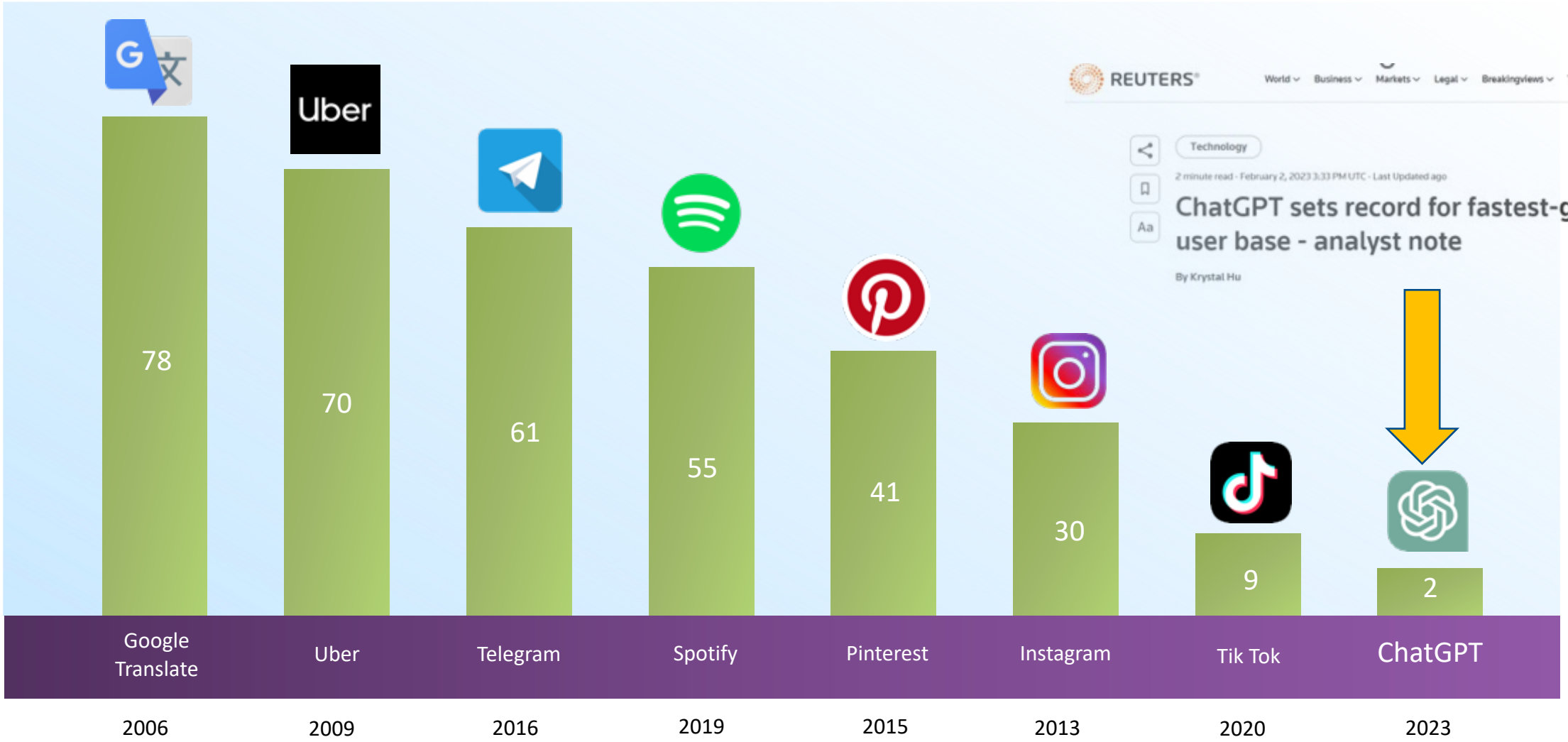
Workload with high privacy, security, and regulatory requirements

- Ability to meet privacy and security regulations in new ways

Explosion of AI use cases across industries and the need for Confidential AI solutions

“Cambrian Explosion” of AI

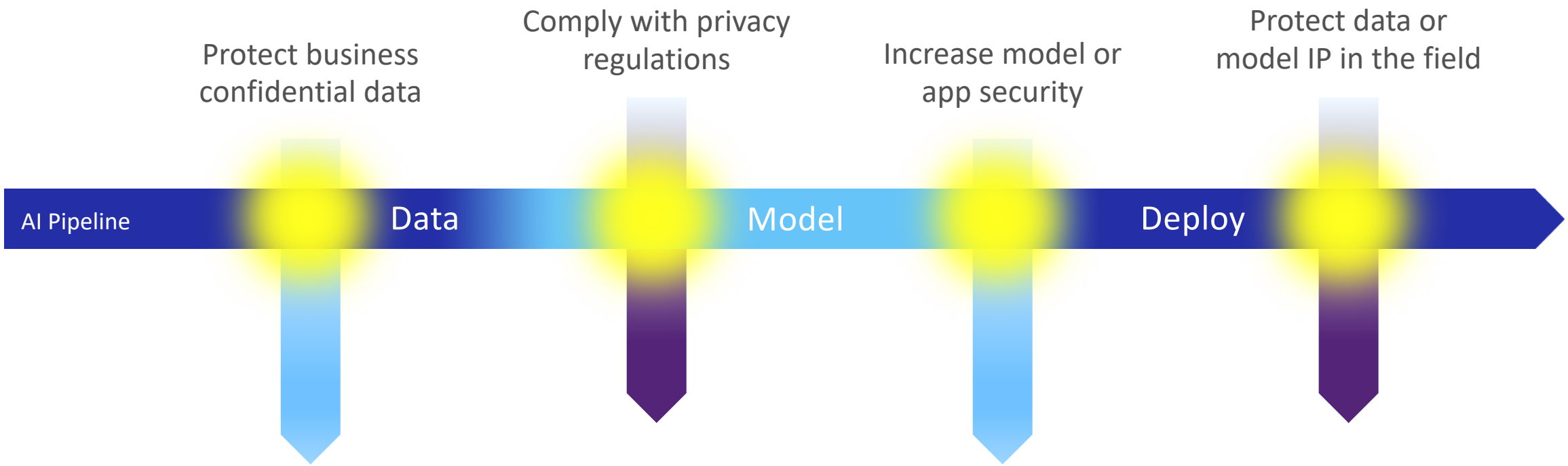
Months to get to 100 million global Monthly Active Users



Source: UBS/Yshoo Finance @EconomyAPP APPEconomy INSIGHTS

SNIA CSTI | CLOUD STORAGE TECHNOLOGIES

Confidential AI:
Strategy for elevated privacy, security and compliance



Confidential AI in Practice

Confidential AI

Securing AI/ML workflows with Secure Enclaves

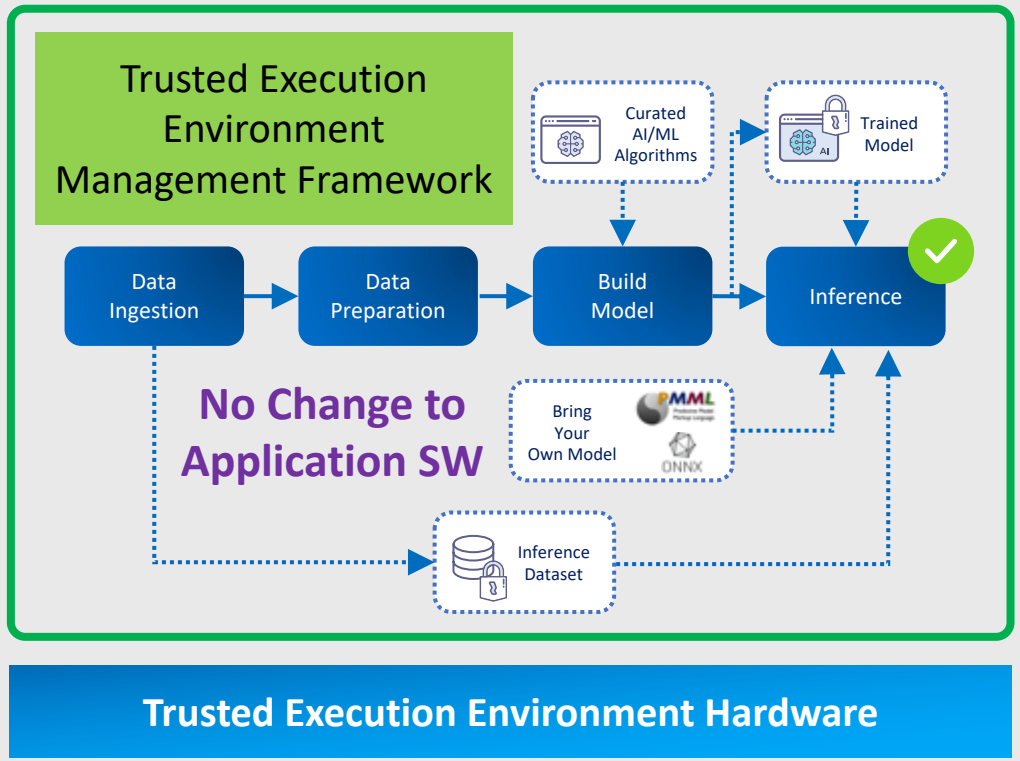
- No-hassle SaaS deployment of managed infrastructure
- Easy dataset connectors with AWS S3 accounts or local data upload
- Protect intellectual property in AI/ML models using secure enclaves
- Meet privacy requirements with auditable logs for proof of execution
- Broad AI/ML framework support:

PyTorch TensorFlow scikit learn

- Bring Your Own Model support using:

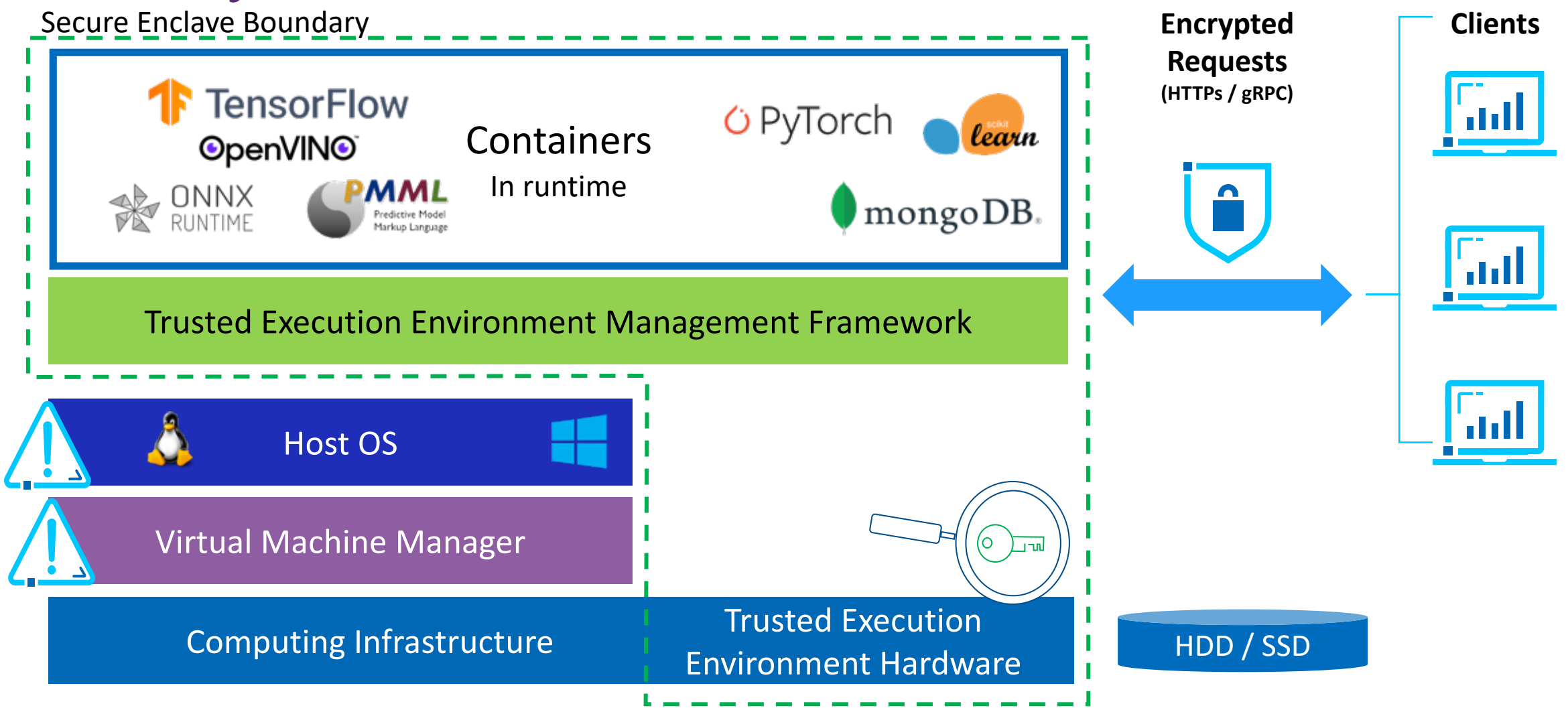
ONNX PMML Predictive Model Markup Language

Secure Enclave - Zero Trust Computing



"What is Confidential Computing and Why Should I Care?" SNIA CSTI, May '21
"Confidential Computing Protecting Data in Use," SNIA CSTI Webinar, June '21
"How to Easily Deploy Confidential Computing," SNIA CSTI webinar, July '21

Security Solution for Confidential AI



[Example Trusted Execution Environment Cloud Solution Delivery]
[Example Trusted Execution Environment Cloud Architecture Characteristics]
[Example Trusted Execution Environment Management Framework]
[Example Trusted Execution Environment Infrastructure]

Cross Industry Use Cases and Case Studies

Confidential Computing Use Cases

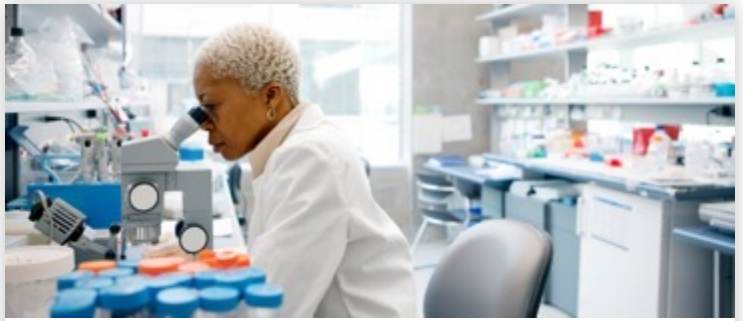
Business transformation, not just risk mitigation

Government

- Digital identity
- Critical infrastructure
- Anti-corruption
- Cyber-crime prevention
- Judicial proceedings and case management
- Deployed and disconnected operations
- Safeguarding / vulnerable population protection (including child exploitation, human trafficking, etc.)

Financial Services

- Anti-money laundering
- Digital currencies
- Secure Payment Processing including Credit Card and Bank Transactions
- Fraud prevention
- Credit risk assessment and qualification from combined bank records
- Capital Markets e.g.: Securing Quantitative Hedge Funds code and models
- Proprietary analytics / algorithms

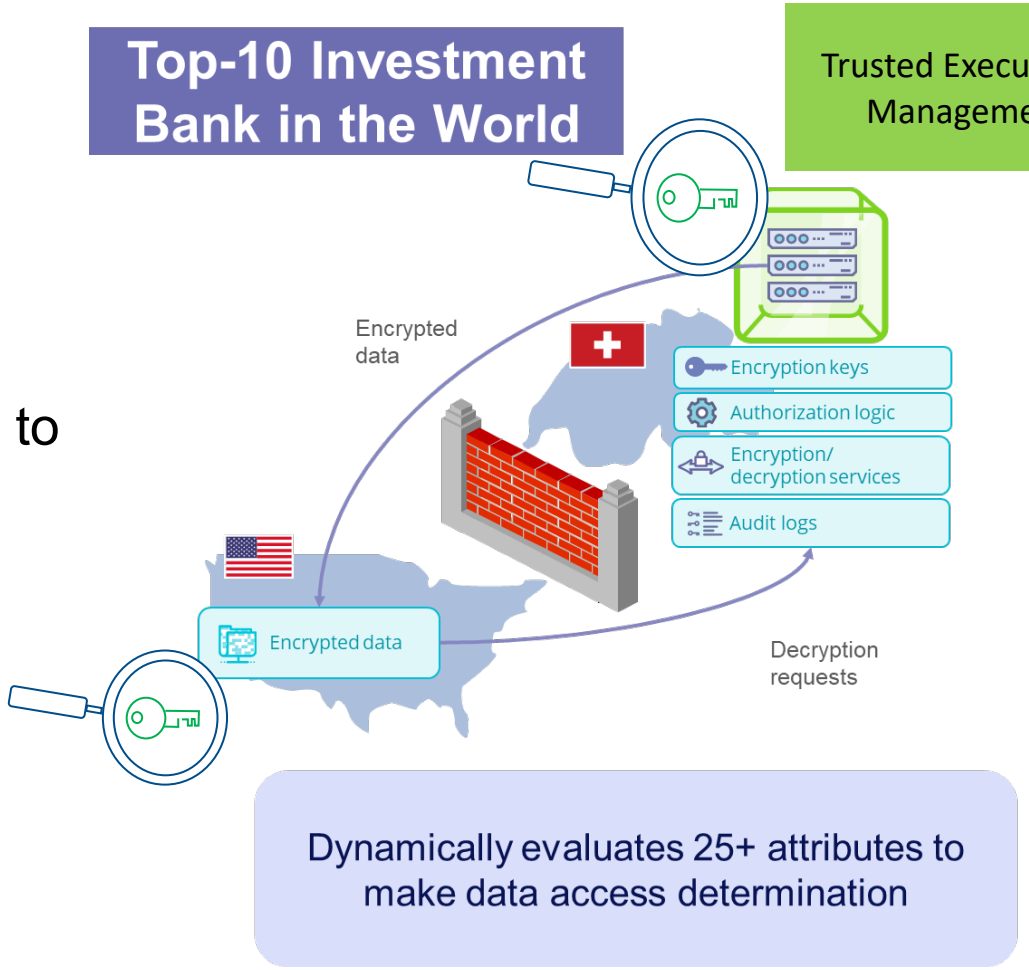


Healthcare

- Disease diagnostic
- Insurance fraud prevention
- Drug development
- Contact tracing
- Records and evidence management
- Insurance fraud, waste, and abuse prevention

Protecting Swiss Banking Data

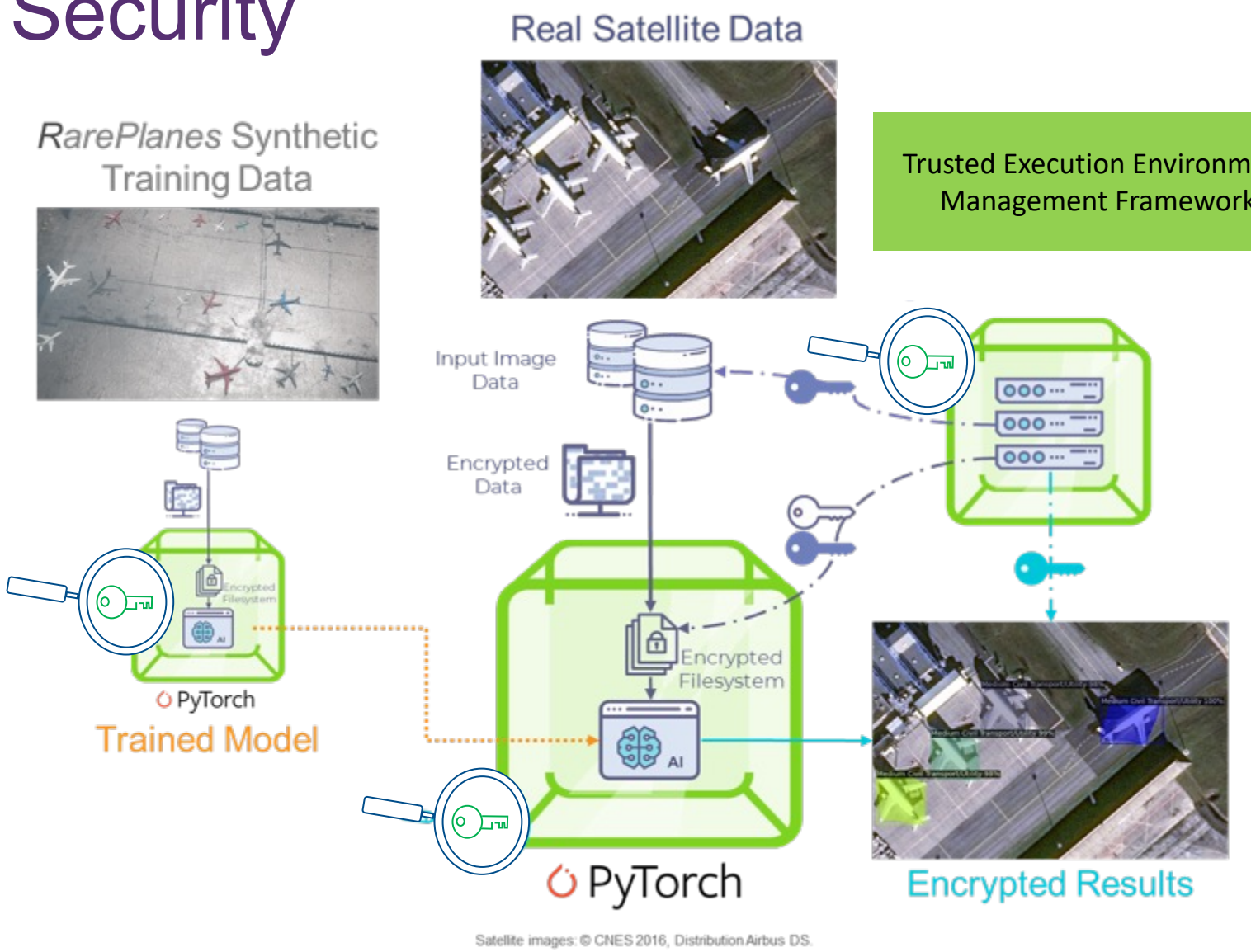
- Swiss regulations (Art.47, Banking Act) prohibit disclosure of private customer information to a third party.
- Protect cryptographic keys and policy evaluations for data access control.
- Segregation of data and keys is enforced to ensure compliance with data protection legislation and organizational policies covering use of Personally Identifiable Information (PII).



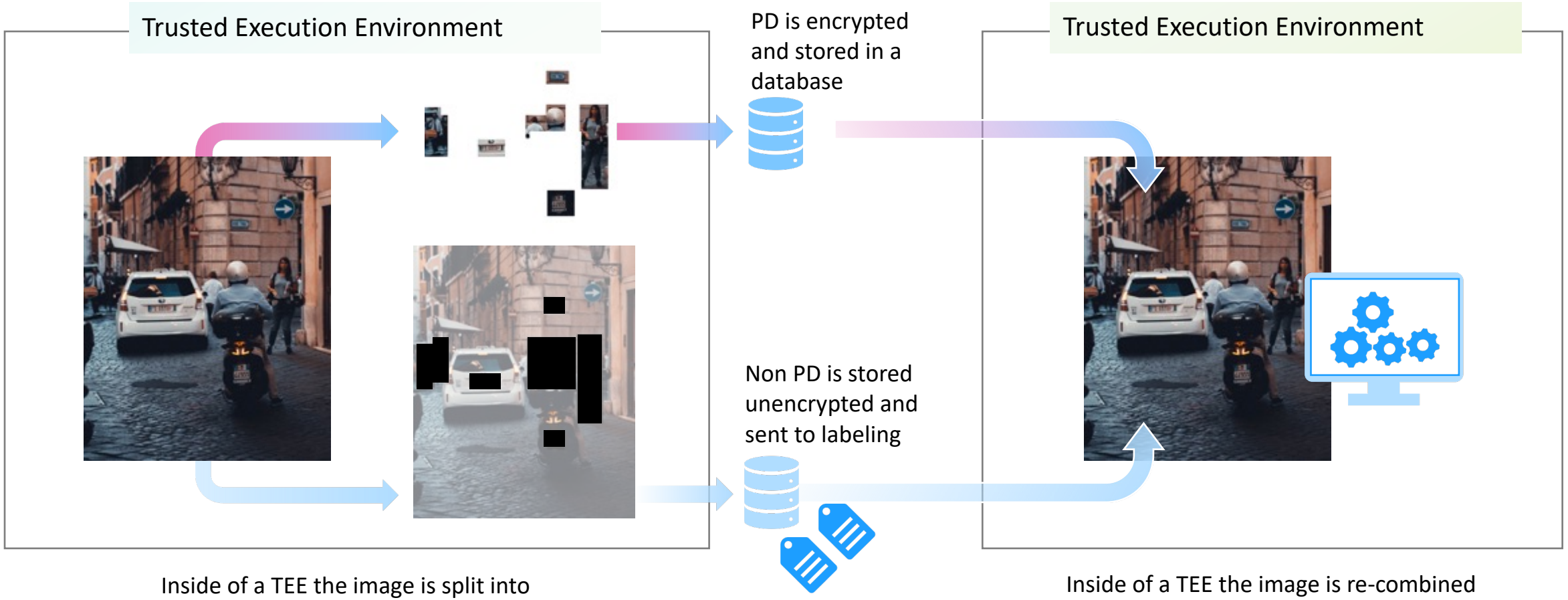
Protecting National Security

- Secure object detection and classification:
  - Training
  - Transfer Learning
  - Inference
  - Adversarial AI Defense
- Comprehensive AI/ML framework support

PyTorch TensorFlow scikit-learn ONNX PMML Predictive Model Markup Language

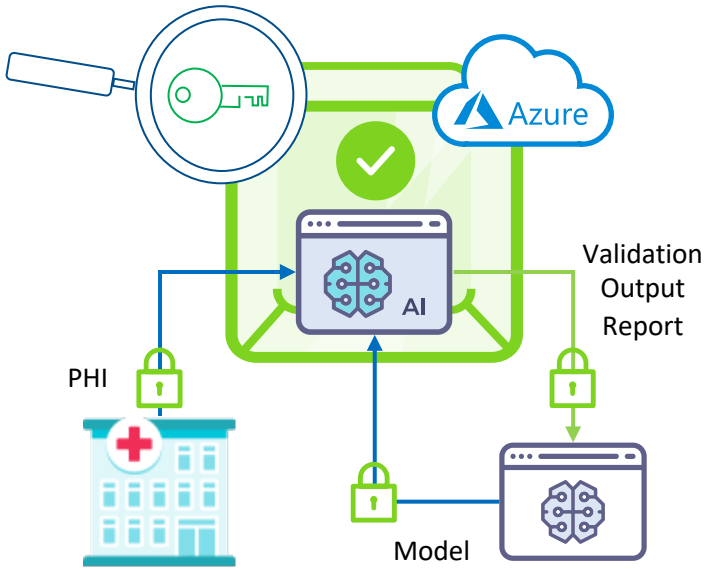


Data Storage and Protection



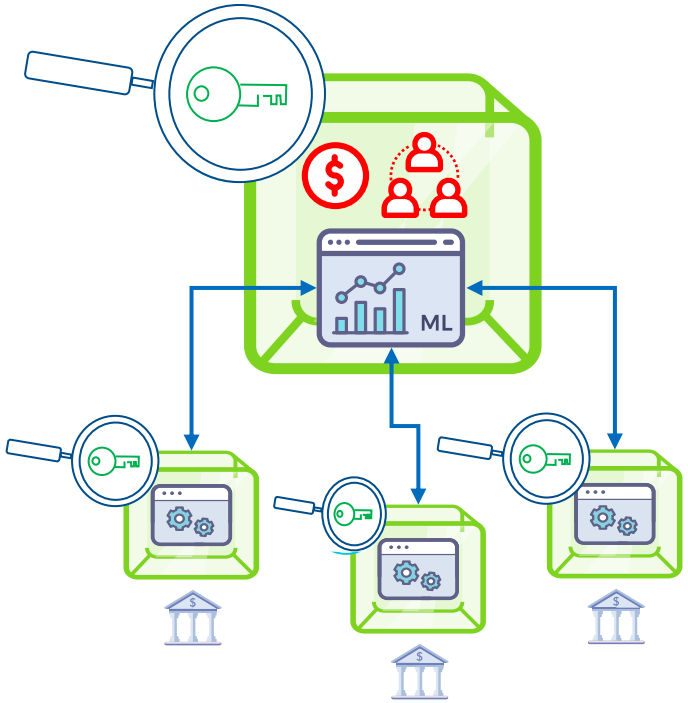
Inside of a TEE the image is split into Personal Data and non-PD.

Inside of a TEE the image is re-combined and used to train a model, on un-modified, original data.



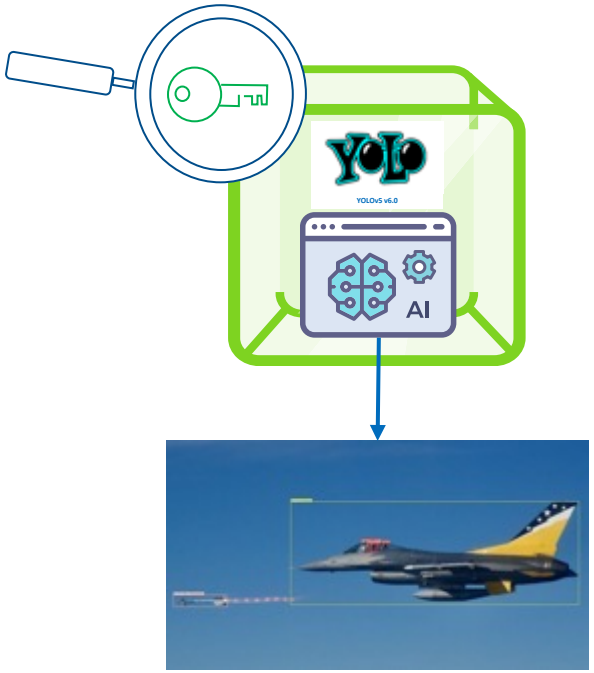
Clinical AI Validation

- HIPAA-compliant validation of AI models for disease diagnosis and drug development
- PHI data protection at rest, in transit, and in-use
- Secured IP to protect clinical AI developers' investment in new models



Financial Crime Prevention[1]

- Federated Machine Learning (FML) at scale for Anti-Money Laundering (AML) and financial crime detection
- Secured IP, with parameters served via secrets injection at runtime
- Workflow orchestration over distributed infrastructure



Secure Object Detection & Classification[2]

- Secure training, inference, and transfer learning for object detection and classification with YOLOv5 v6.0
- Defense against adversarial machine learning with secure enclave attestation

1 Reference: https://ieeexplore.ieee.org/document/10021108
2 Reference: https://www.spiedigitallibrary.org/conference-proceedings-of-spie/12113/121130C/Establishing-security-and-trust-for-object-detection-and-classification-with/10.1117/12.2618303.short

Q&A

Additional Resources

- Confidential AI based on Intel Security Solution, here
- Fortanix Confidential AI, here
- Accelerated AI Inference with Confidential Computing, here

Thanks for Viewing this Webinar

- Please rate this presentation and provide us with feedback
- This webinar and a copy of the slides are available at the SNIA Educational Library https://www.snia.org/educational-library
- A Q&A from this webcast will be posted to the SNIA Cloud blog: www.sniacloud.com/
- Follow us on Twitter @SNIACloud

Thank You!

