



Verifiable Compute to Trust AI in Regulated Workflows

PHUSE US Connect 2025



Agenda

1. **What** is Trusted AI Infrastructure?
2. **Evolving Trust** from Confidentiality to Verifiability
3. **Real World** Verifiable Compute for Life Sciences

I. Trusted AI Infrastructure

Calls to use trusted execution environments (TEEs) for AI have been growing since May of 2024:

- OpenAI
- Anthropic
- Google DeepMind
- RAND Corporation

All have endorsed using **SLSA Security Level 3** guarantees as the most advanced methods to protect AI models and systems.

SLSA Security Levels 		
Level	Description	Example
1	Documentation of the build process	Unsigned provenance
2	Tamper resistance of the build service	Hosted source/build, signed provenance
3	Extra resistance to specific threats	Security controls on host, non-falsifiable provenance
4	Highest levels of confidence and trust	Two-party review + hermetic builds

[SLSA Security Levels](#)

2021: Securing Patient Records in Healthcare - Intel

Solution Brief
Intel® Software Guard Extensions (Intel® SGX)

intel.

Maximum Security at the Processor Level: Intel® SGX Protects Electronic Patient Record

The Electronic Patient Record (ePA)

The electronic patient record (ePA) can store diagnoses, treatment data, medication plans, allergies and other health data. It is also possible to save individual health information, such as a diary of blood glucose levels, to it. Use of the ePA is voluntary and free of charge and is available uniformly across all statutory health insurance funds across sectors and cases. From 2022, additional information such as vaccination card, maternity log, the yellow booklet for the obligatory children's U-examinations and the dental bonus booklet can be stored on it.

Maximum security at the processor level

It takes innovative solutions to protect the security of highly sensitive data. Intel Software Guard Extensions (Intel SGX) fulfills extremely high security requirements while being completely transparent to the user. The AOK-specific implementation of the electronic patient record (ePA) in Germany is a prime example of the advantages of this approach.

In recent years, the requirements for protecting personal data have been strengthened, especially by the General Data Protection Regulation (GDPR) and the Patient Data Protection Act (PDSG). This is especially true for health-related information because this data is particularly sensitive – the collection, storage and processing are subject to the highest level of security precautions. Such special requirements can be implemented with reasonable effort in controlled environments like a doctor's practice or within a hospital. But it becomes a lot more difficult when insured persons and other service providers are involved. It was precisely this challenge that AOK – Die Gesundheitskasse, a statutory health insurance company, faced in the implementation of the electronic patient record (ePA). In addition to a strong focus on ease of use, the project participants particularly focused on data security. The ePA was approved by gematik in version 1.0 at the end of 2020 and has been in operation since January 1, 2021.

The concept of ePA

Germany: AOK is a statutory health insurance company with around 27 million insured individuals.

2024: Confidential AI Computing - Apple

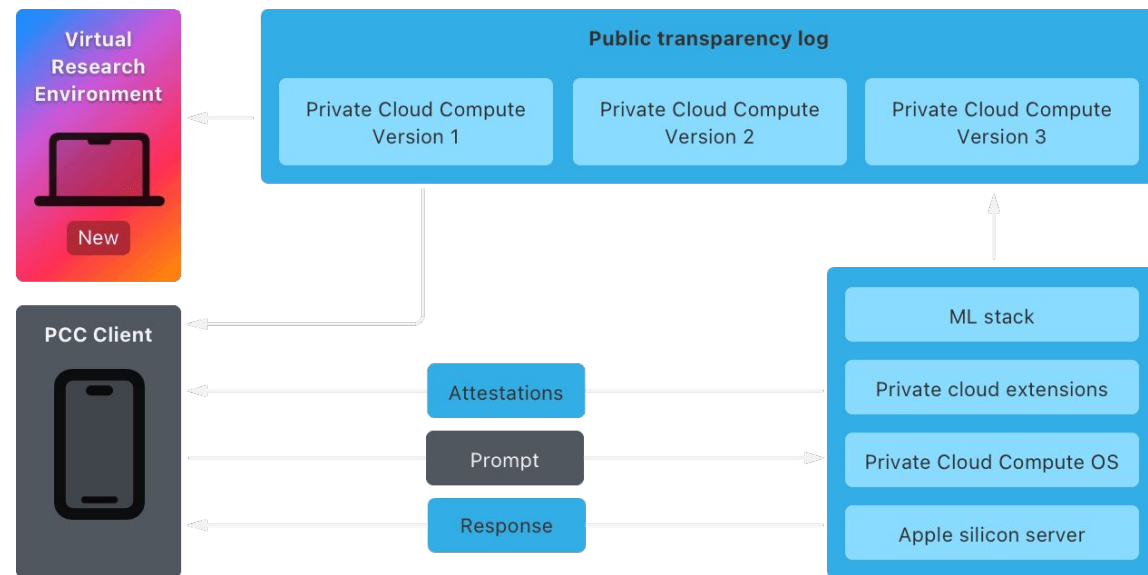
In Q2 2024, Apple announced that hardware-based AI compute as a key part of their phone and server side solutions.



<https://security.apple.com/blog/private-cloud-compute/>

Verifiable Transparency

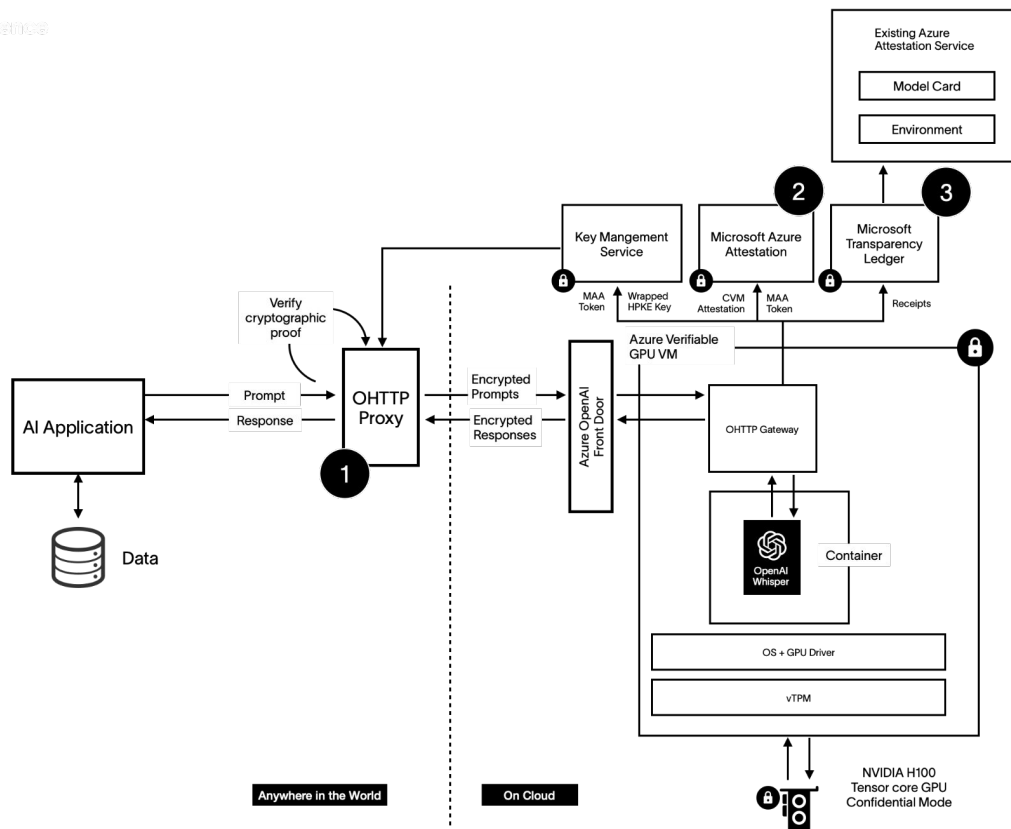
"Security researchers must be able to verify the security and privacy guarantees of Private Cloud Compute, and they must be able to verify that the software that's running in the PCC production environment is the same as the software they inspected when verifying the guarantees."



2024: Confidential AI Inference - Microsoft Azure

The 3 elements of the Azure reference implementation:

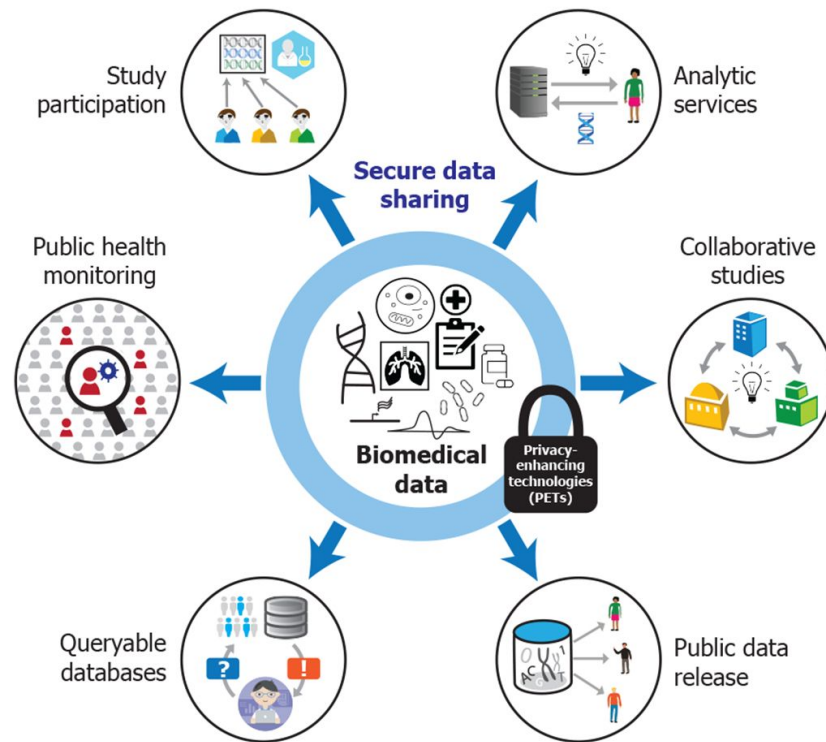
- 1 Oblivious HTTP (OHTTP) is used to encrypt the prompt from the client to the TEE through a protected Azure OpenAI front door.
- 2 Microsoft Azure Attestation service brokers a key release policy.
- 3 Microsoft Transparency Ledger registers the TCB and model card.



Use cases in Life Sciences are maturing.

“In the biomedical domain, TEEs have **gained significant traction** due to their ability to securely outsource the analysis of biomedical data and to facilitate the development and deployment of health artificial intelligence (AI) tools on a large scale”

Cho H, Froelicher D, Dokmai N, Nandi A, Sadhuka S, Hong MM, Berger B. Privacy-Enhancing Technologies in Biomedical Data Science. Annu Rev Biomed Data Sci. 2024 Aug;7(1):317-343. doi: 10.1146/annurev-biodatasci-120423-120107. PMID: 39178425; PMCID: PMC11346580.



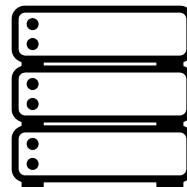
SUMMARY

Conventional Attestations for Confidential Compute



Proof of Confidentiality:

Verifiers can ensure that any confidential information involved in the computation remains confidential during computation as expected.



Proof of Environment:

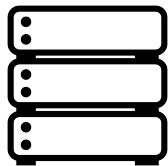
Verifiers can see that the computation was executed in their expected environment.

II. From Confidentiality to Verifiability

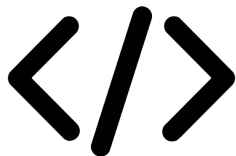
From the same root of trust:



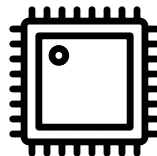
**Proof of
Confidentiality**



**Proof of
Environment**



Proof of Correctness:
Verifiers can check that the inputs (including the algorithm and data) and outputs are as expected, and executed as expected.



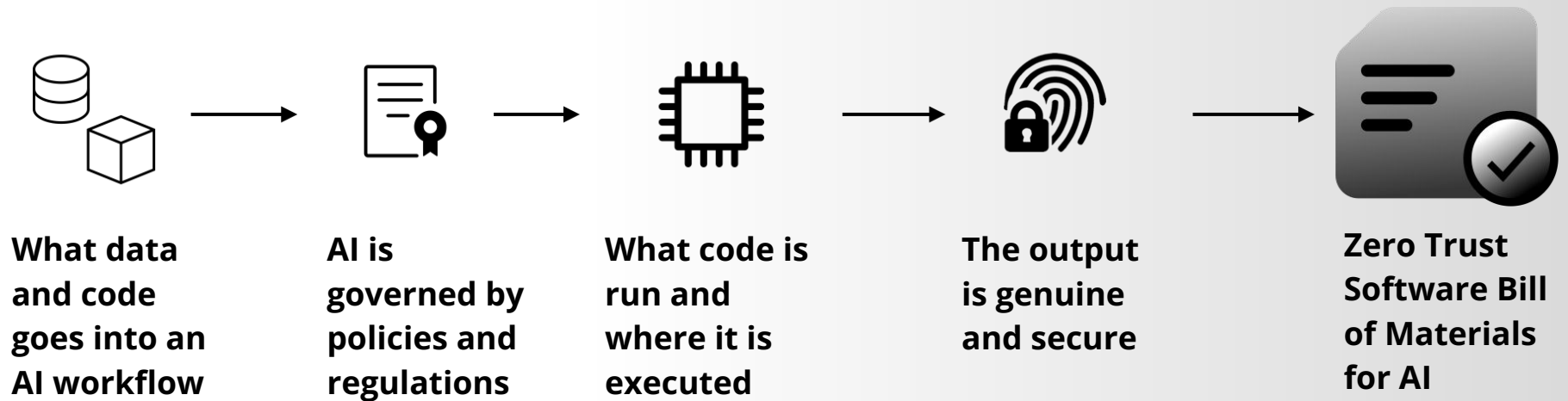
Proof of Computation:
Verifiers can confirm that the amount of work expended in the computation is as expected, provided that metrics are collected, signed, and emitted from the VM.



Proof of Governance:
Verifiers check for compliance with business or regulatory policies program execution at runtime. If these controls are not met, application logic may halt and/or provide an instant audit

NEW: Verifiable Compute

Gives you independent verification of:



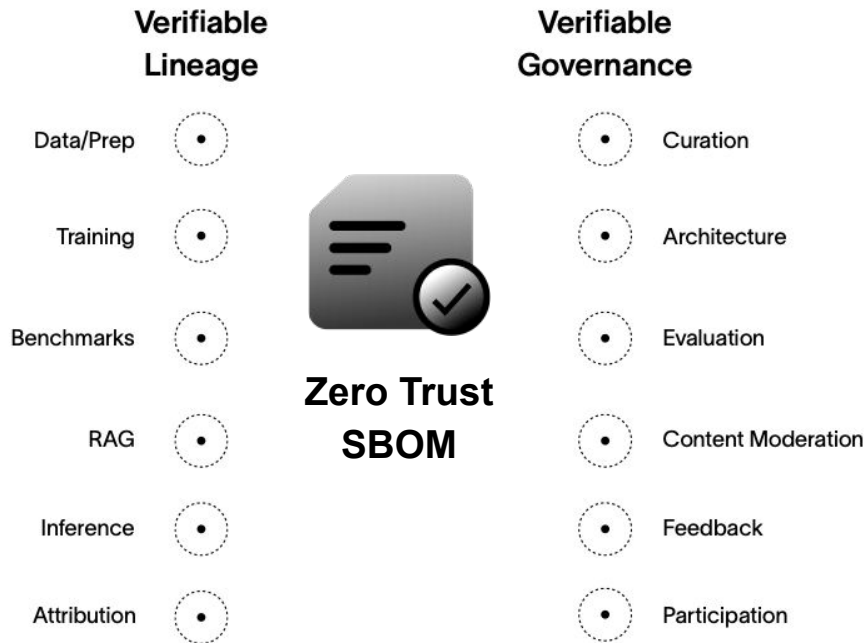
These attestations are self-contained, non-interactive proofs that are also composable across sessions. The artifact's certificate may contain attestations from different processes in very different computing environments and attested to by different organizations, thereby creating a true picture of its lineage, all verifiable without a centralized registry that keeps track of the end-to-end supply chain.

NEW: Evolving Trust for the AI Era

Traditionally, the cornerstone of responsible AI is reproducibility. We trust that AI models are genuine and performant through written documentation and benchmarks that can be validated by third parties. In many cases, however, reproducibility for AI training or inference is infeasible as it may be:

- **Prohibitively expensive** as workloads can cost millions to re-run
- **Impractical** given AI systems are often too brittle or complex to reanimate.
- **Impossible** if AI results are non-deterministic.
- **Ill-advised** if full transparency could lead to economic impact or serious harm.

NEW: End-to-End Trust for AI



- **Persistent:** available instantly in a browser or can be audited decades later.
- **Auditable:** a tamperproof record of compute that can be independently verified.
- **Composable:** action remediation and even compensation for stakeholders.

III. Verifiable Compute for Life Sciences

Traditionally, the cornerstone of *Clinical Trials* programming is reproducibility... We trust that clinical trials programs are genuine and performant through written documentation, thorough validation processes (such as double programming) and benchmarks that are often audited. In many cases, however, reproducibility for Clinical Programming is infeasible as it may be:

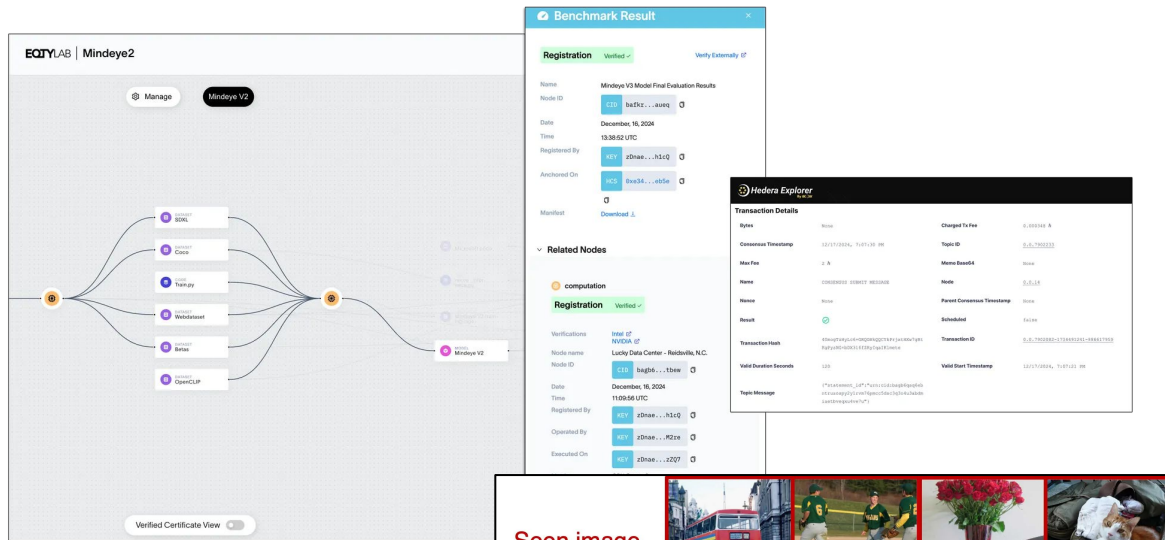
- **Prohibitively expensive** to procure an exact same system to replicate upon.
- **Impractical** given systems are often too complex or old to reanimate.
- **Impossible** if data are unreadable, or more unlikely, lost.

Clinical trials data are also incredibly sensitive, and caution should be taken to ensure the security and privacy of such data.

Advancing Brain-Computer Interfaces: The MindEye Project

MindEye2, a collaboration between MedARC and Princeton University, advances AI brain-computer interfaces and underscores the need for new trust standards as these technologies move into sensitive, large-scale medical use.

MindEye demonstrates remarkable capabilities in reconstructing visual perception from fMRI brain data. As this research proves the viability of processing real individuals' brain data, establishing rigorous frameworks for trust and verification becomes crucial for transitioning from laboratory research to real-world applications.

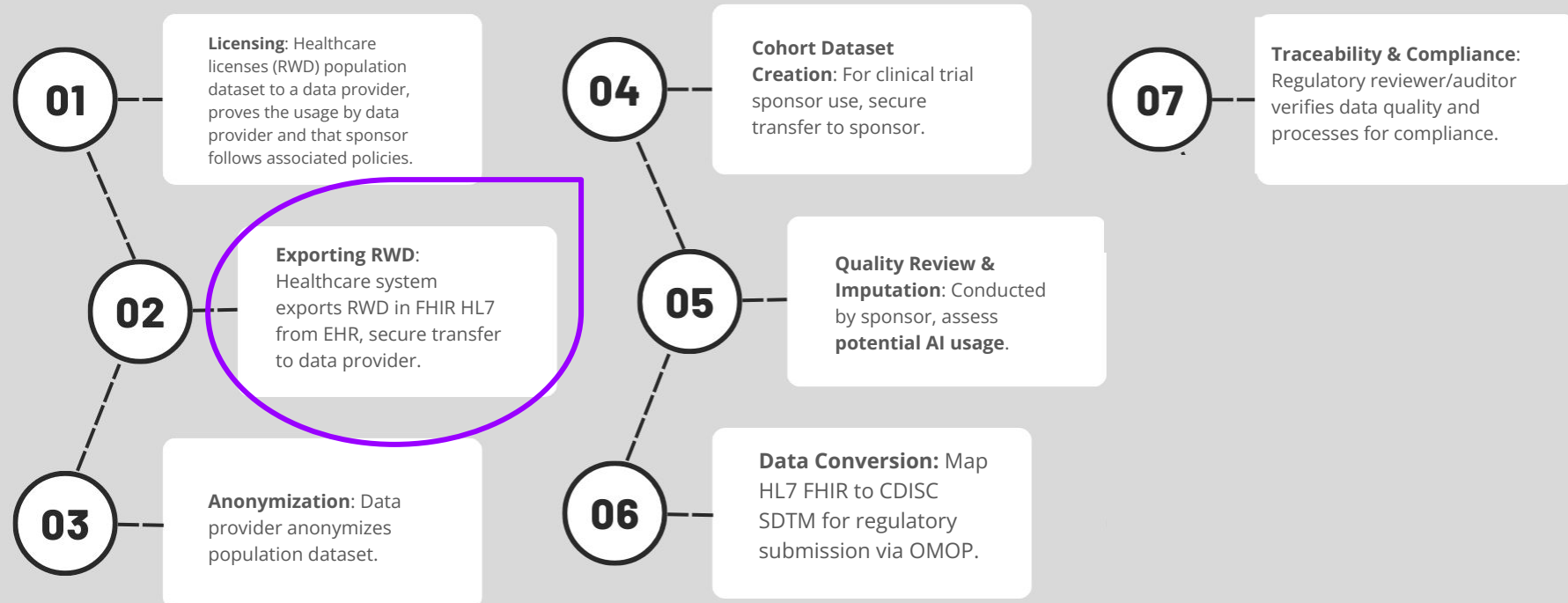


MindEye2: Shared-Subject Models Enable fMRI-To-Image With 1 Hour of Data (arXiv:2403.11207)



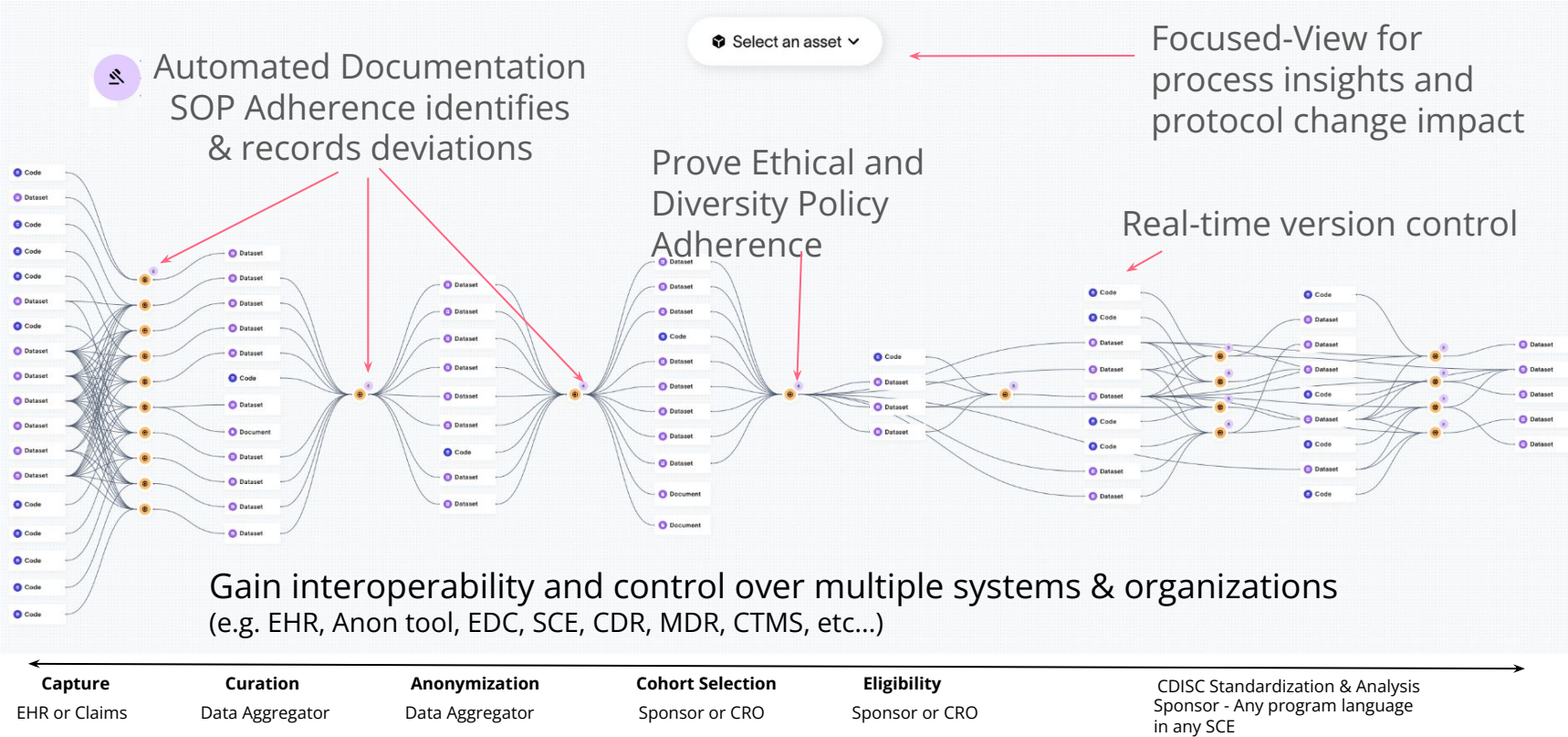
CRYPTOGRAPHIC HASHING

Real-World Evidence



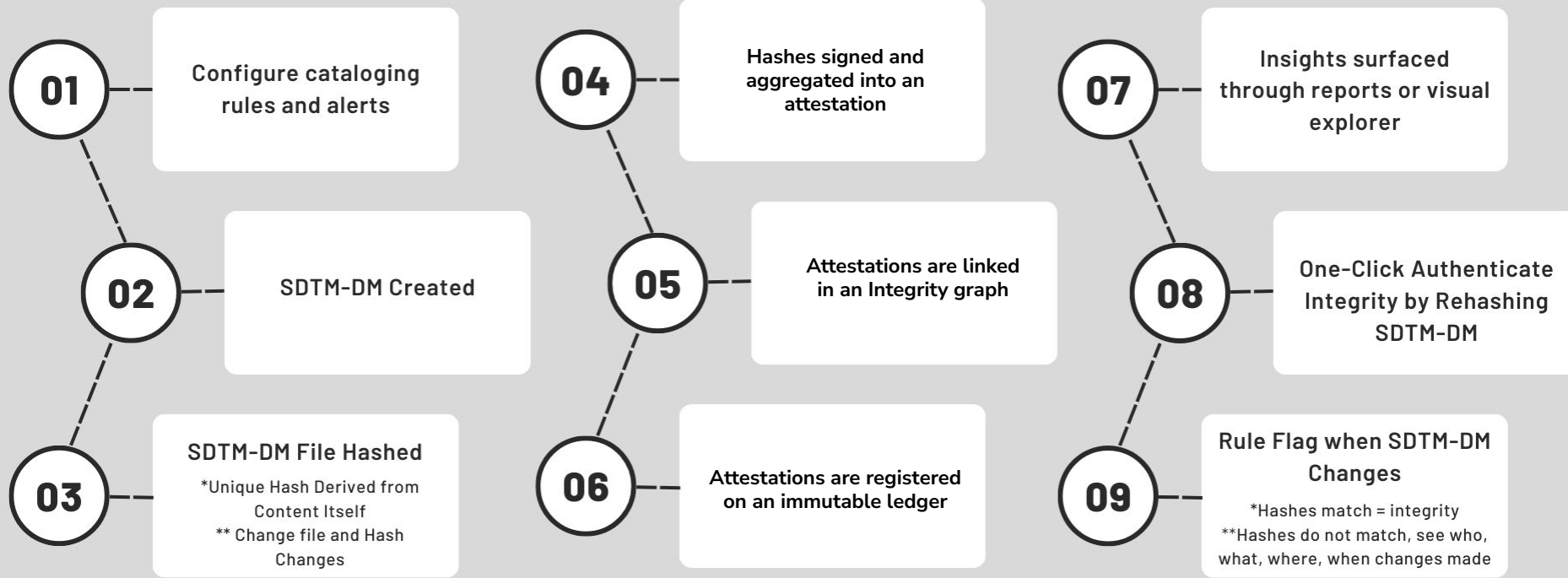
Real-World Evidence: Reliable Data and Controlled Processes

Audit entire data journey across workflow without access to third-party systems



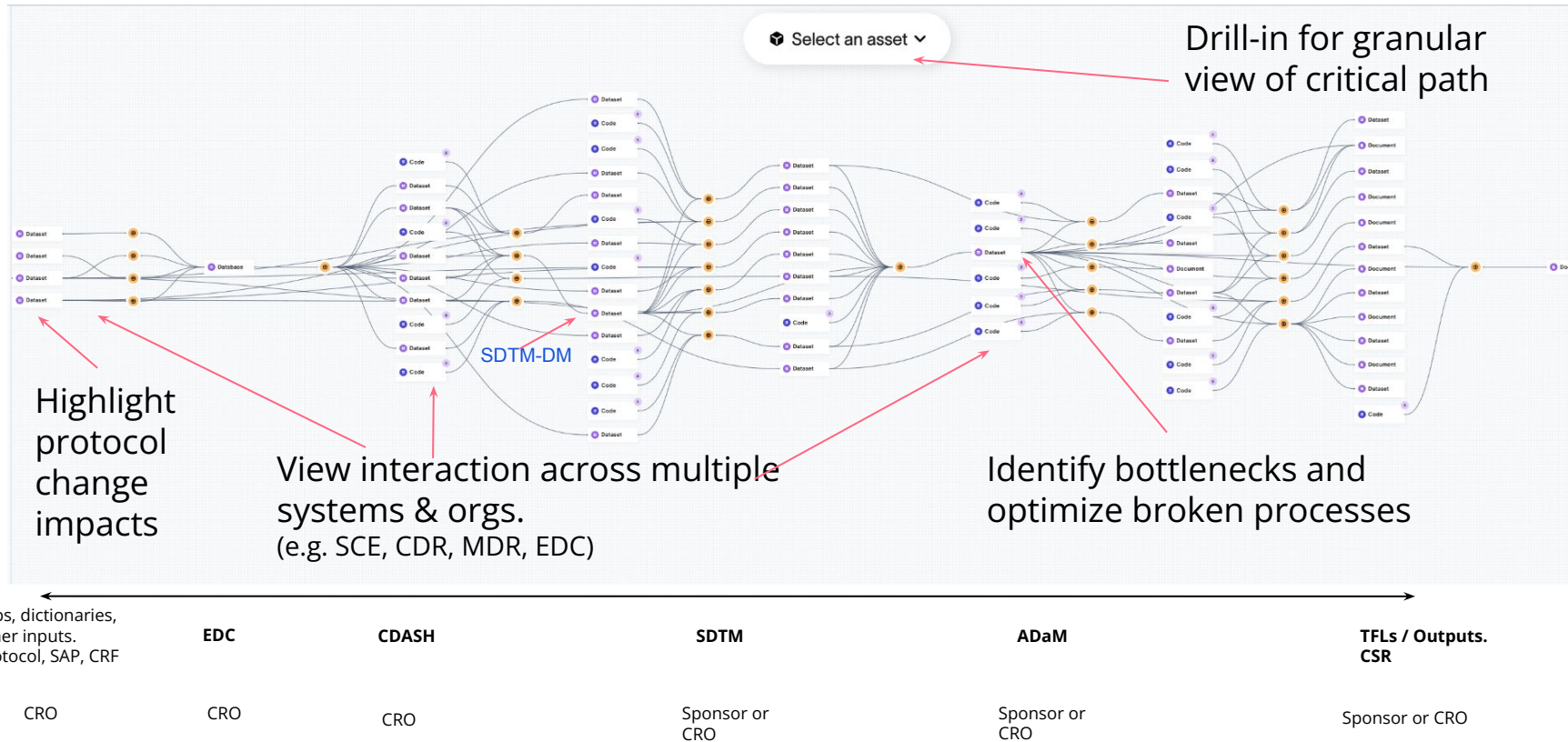
CRYPTOGRAPHIC HASHING

A Clinical Data Perspective



Clinical Workflow - Operational Efficiency

Unify governance across the workflow to embed compliance and quality.





To Learn More go to:

<https://www.eqtylab.io/life-sciences>

**Verifiable Compute
Whitepaper available at:**

<https://vcomp.eqtylab.io/>