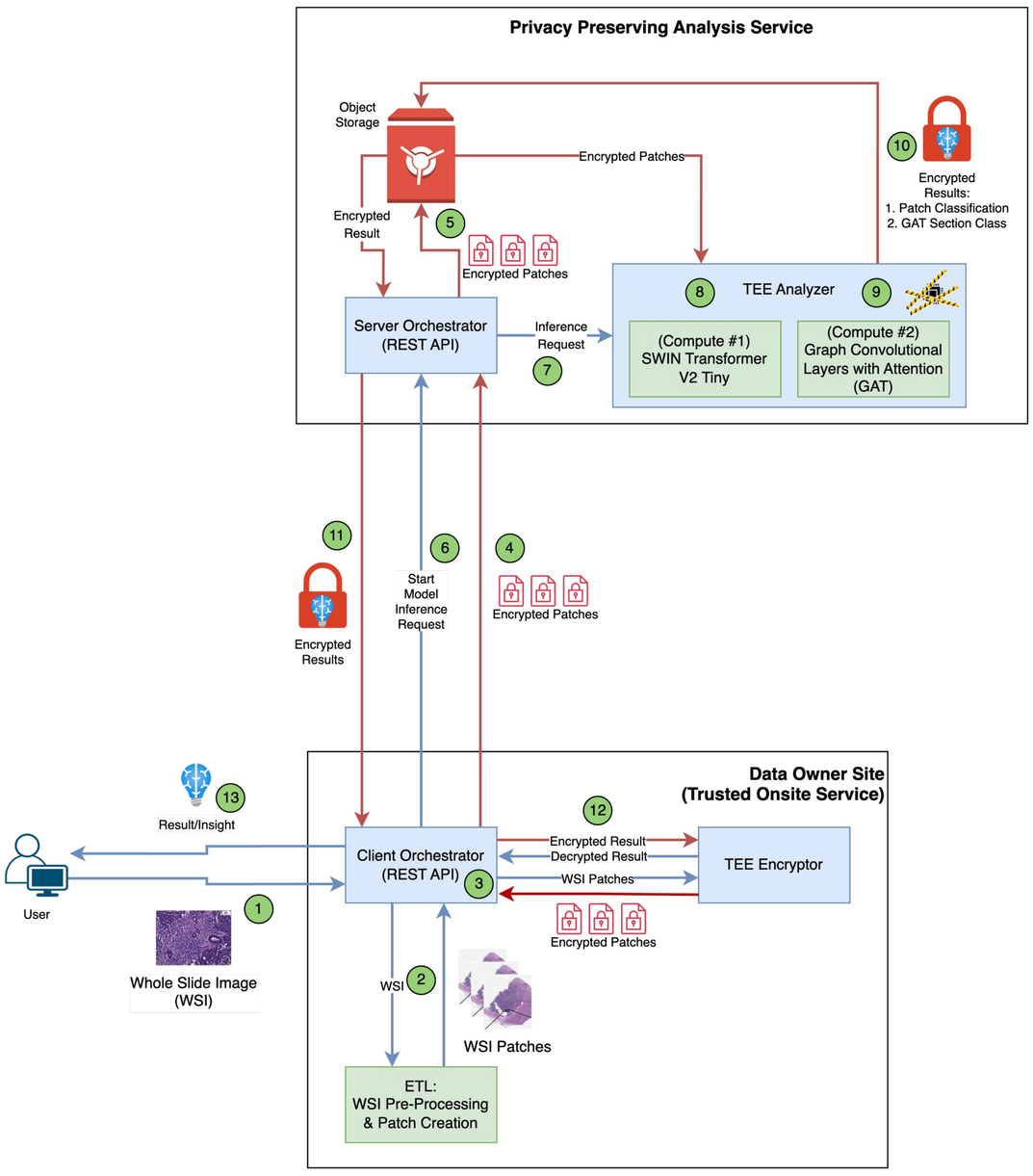
PRIVACY ENHANCING AI-BASED WHOLE SLIDE IMAGE ANALYSIS

Rodel Felipe Miguel[1], Benjamin Hong Meng Tan[1], Jin Chao[1], Deng Xiaoxia[1], Chan Fook Mun[1], Khin Mi Mi Aung[1], Khor Li Yan[2], Heng Seow Ye[2], Dr. Cheng Chee Leong[2]

*[1] Institute for Infocomm Research (I²R), Agency for Science, Technology and Research (A*STAR), Singapore*
[2] Singapore General Hospital Pte. Ltd.



Abstract

Digital pathology begins when a thin section of human tissue is mounted on a glass slide and scanned into a **Whole-Slide Image (WSI)** – typically 80,000 x 60,000 pixels and a size of 500MB to 5GB even in compressed form. WSIs drive state-of-the-art computational pathology, but hospitals typically restrict their analysis to isolated, air-gapped workstations because these gigapixel slides contain highly sensitive patient data. On such systems the workflow for a single case is onerous:

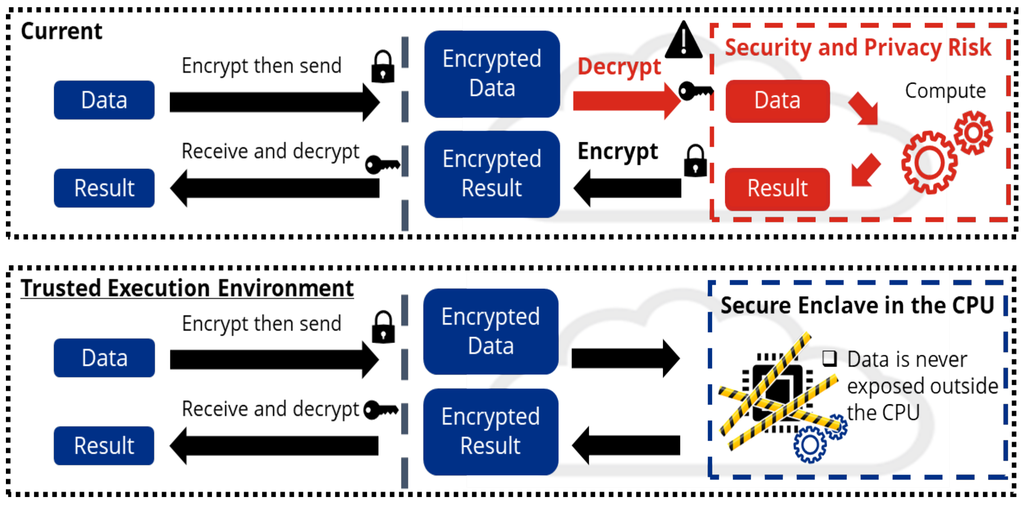
i. Technicians copy the multi-gigabyte WSI to a removable medium and walk it to the secure workstation
ii. The slide is partitioned into patches (≈7 min)
iii. Deep-learning inference (≈20 min)

With sequential processing and manual hand-offs, throughput stalls well **below the 50 cases per day target** for routine diagnostics.

We present a privacy-preserving, cloud-enabled pipeline that removes the physical-transfer bottleneck while maintaining strict confidentiality guarantees. The solution hinges on hardware-based trusted execution environments (TEEs): **TEE Encryptor (on-premises) and TEE Analyzer (cloud)**.

Because computation now runs on elastic cloud hardware, multiple TEE Analyzer instances can be launched in parallel. A deployment with ten enclaves cuts effective turnaround to minutes per case and comfortably exceeds the 50-case-per-day target, all without exposing WSIs or predictions in plaintext to the cloud operator. This solution offers order-of-magnitude improvements in digital-pathology throughput while preserving patient privacy.

Technology Features



- **Trusted Execution Environment (TEE)** *[AMD SeV-SNP/Intel TDX]*: A hardware-isolated, memory-encrypted VM in the cloud. Only code loaded with the correct cryptographic "measurement" can run; the cloud provider cannot peek inside.

- **TEE Encryptor (Inside the Hospital Firewall):** After verifying a remote enclave's attestation, it establishes an ephemeral (temporary) session key, preprocesses each WSI locally, slices it into patches, encrypts the patches with the session key, and transmits them over TLS.

- **TEE Analyzer (In the Cloud):** Hosted in an AMD SEV or Intel TDX enclave, it decrypts patches only inside protected memory, executes the two-stage deep-learning cascade, re-encrypts results with the same session key, and stores all artefacts in an object store accessible solely within the private analysis service. Encrypted results return to the hospital and are decrypted by the TEE Encryptor.

- **WSI Inference:** A 2–stage model inference that runs inside the TEE Analyzer. This 2-level inference comprises of patch-level classification (SWIN-v2 Transformer) and WSI-level aggregation (Graph Attention Network). The inference mimics a pathologist's holistic view while keep computations private.

Setup	Step 1 - Data Preparation	Step 2 – WSI Inference
1x Encryptor VM: Ubuntu 22.04 LTS, 64 GB RAM **Analyzer Host**: 2x AMD EPYC 9554 processor, 256 GB Memory, Ubuntu 25.04 **5x Analyzer VM:** Ubuntu 25.04, 32 GB RAM, 16 CPU threads **WSI:** 5x 1.5GB TIFF file	**Location:** Encryptor VM **Process:** Convert WSI into patches, encrypt, and sent to Analyzer for inference **WSI Conversion Time:** 4 min 21 sec **Encrypt and Submit Time:** 23 sec **Memory:** 11 GB **CPU Threads:** 10	**Location:** Analyzer VMs **Process:** 5x 2-stage inference tasks in 5 AMD SeV-SNP TEE VMs **Total Runtime of 5 Tasks:** 17m 57s **Average Runtime of 5 Tasks:** 16m 50s **Memory per VM:** 32 GB **CPU Threads per VM:** 16 **Cost per Inference in GCP:** $0.214

Architecture and Workflow

Components:

1. **Data Owner Site**
   - **WSI Pre-processing & Patch Creation**
   - **TEE Encryptor** (responsible for encrypting patches and decrypting final results)
   - **Client Orchestrator** (handles API calls to the server side, coordinates tasks)

2. **Privacy Preserving Analysis Service**
   - **Server Orchestrator** (manages job scheduling, routes data)
   - **Secure Object Storage** (stores encrypted patches/results, accessible only within this service)
   - **TEE Analyzer** (runs on AMD SEV/Intel TDX nodes, handles patch-level CNN and GAT classification in encrypted form)

3. **Attestation Service** (AMD or Intel Service)
   - Verifies the integrity/measurement of the TEE Analyzer
   - Facilitates session key establishment

Data Flow Summary:

1. WSI is **pre-processed** on-prem (removing artifacts, creating patches).
2. **Patches** are **encrypted** with a session key, established after attestation.
3. **Encrypted patches** are stored in **secure object storage** in the Privacy Preserving Analysis Service.
4. The **TEE Analyzer** fetches and **decrypts** patches in-enclave, runs inference.
5. **Results** are encrypted again, placed back in **object storage**.
6. The **Data Owner** retrieves **encrypted results** and uses the TEE Encryptor on-prem to **decrypt** them.

Acknowledgements

- This research is supported by the **National Research Foundation, Singapore** and **Infocomm Media Development Authority** under its **Trust Tech Funding Initiative (DTC-RGC-01)**. Any opinions, findings, and conclusions or recommendations ex-pressed in this material are those of the author(s) and do not reflect the views of National Research Foundation, Singapore and Infocomm Media Development Authority.
- This work was carried out in close partnership with **Singapore General Hospital Pte. Ltd.**, whose collaboration and clinical insights were essential to the success of this research.
- This research/project is supported by the **National Research Foundation, Singapore** under its **AI Singapore 100 Experiments Programme (AISG Award No: AISG2-100E-2023-108)**. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not reflect the views of National Research Foundation, Singapore.

www.a-star.edu.sg/i2r

facebook.com/i2r.research/

institute-for-infocomm-research

For more information or collaboration opportunities, please contact:

Rodel Felipe Miguel
Principal Research Engineer
Cybersecurity

Vernetia Loo
Industry Development Manager
A*STAR I²R