HARPOCRATES Technical Briefing Paper

Privacy-Preserving Analytics for Cybersecurity and Healthcare

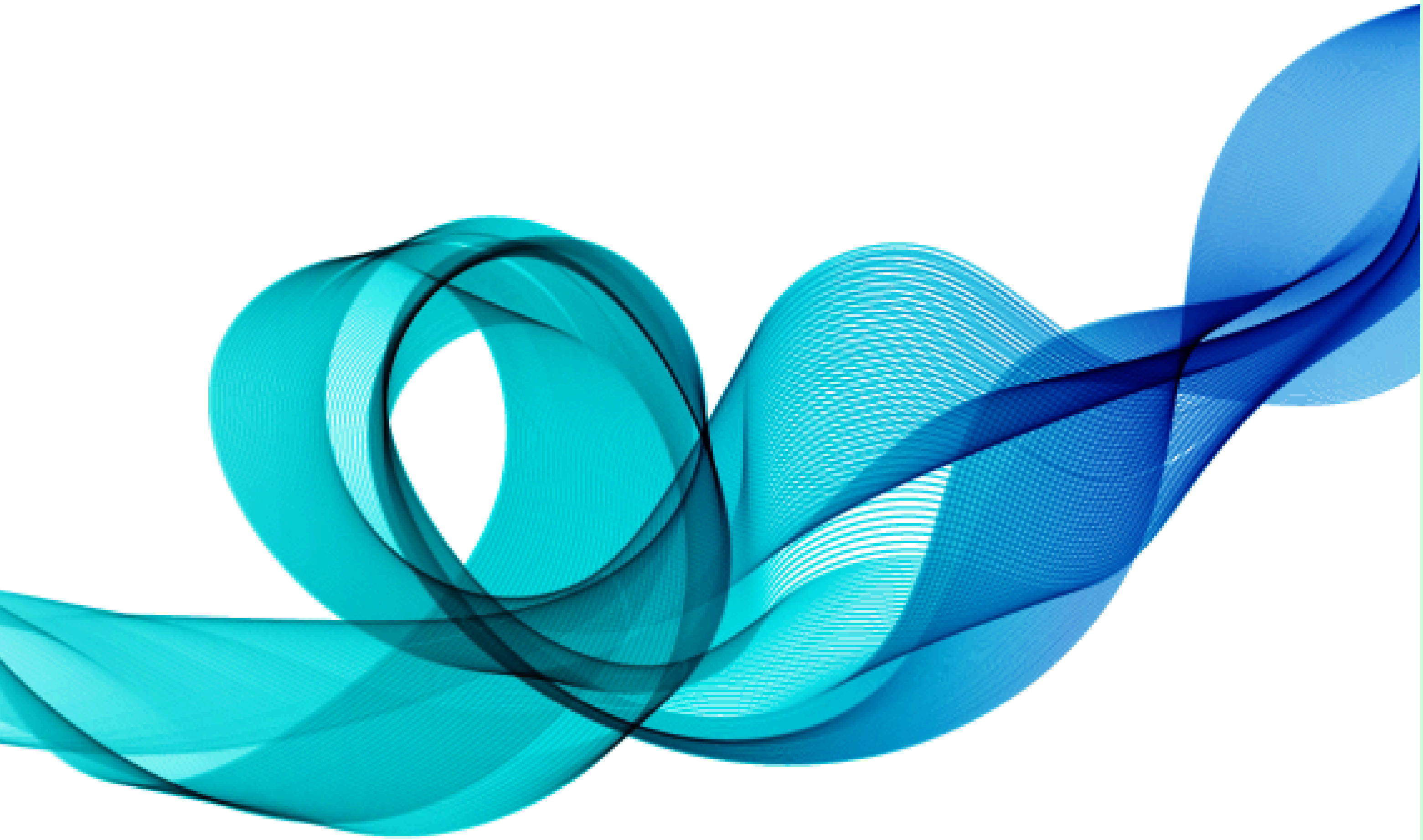


Table of Contents

Introduction

HARPOCRATES is a Horizon Europe research project involving 13 partners across 9 countries. The project designs and validates secure data processing tools that allow organizations to analyze and share sensitive information without revealing the raw data itself.

The project applies advanced cryptography and privacy-preserving machine learning across two main demonstrators:

- ✓ **Threat Intelligence Sharing for public administrations**
- ✓ **Sleep Medicine Analytics for clinical decision support**

Tools and methods are modular, GDPR-compliant, and extendable to domains like mobility, energy, and finance.

Project Objectives

The HARPOCRATES project is driven by the need to enable secure, privacy-preserving data analytics across sensitive sectors.

These objectives guide the technical and applied work of the project, ensuring that data confidentiality is maintained without compromising analytical capabilities.

- ★ Enable analysis of encrypted data
- ★ Combine cryptography with differential privacy
- ★ Train models collaboratively without data sharing
- ★ Improve system robustness against malicious input
- ★ Demonstrate use in cybersecurity and healthcare
- ★ Provide modular, GDPR-compliant tools
- ★ Support automated privacy assurance
- ★ Turn advanced research into operational solutions

Project Concept

The HARPOCRATES project is based on the idea that data privacy and utility can coexist. Organizations increasingly rely on sensitive data—such as health records or system logs—for analysis, but legal and ethical constraints limit data sharing.

HARPOCRATES addresses this by enabling analytics on encrypted data using advanced cryptographic methods like homomorphic and functional encryption, federated learning, and differential privacy.

Instead of a one-size-fits-all approach, the project uses a modular framework that supports different data types, environments, and threat scenarios. Components like encrypted ML models, anomaly detection tools, and GDPR modules are designed for interoperability and easy integration.

Core Technical Innovations

A range of cryptographic and privacy-enhancing techniques have been developed and integrated to support secure data analysis across different environments.

Homomorphic Encryption (HE): Processing data without decryption

Functional Encryption (FE): Controlled disclosure of computation results

Federated Learning (FL): Model training without centralizing data

Differential Privacy (DP): Statistical protection of individual data

Threat Detection Pipelines: Privacy-respecting anomaly detection

Encrypted ML Models: For sensitive signal classification

DEMONSTRATOR 1

Threat Intelligence Sharing

Secure cyber threat detection across public authorities

Objective: Allow local authorities to collaborate on detecting coordinated cyberattacks without exposing raw logs.

Data Source:

- Windows Sysmon logs
- Processes like powershell.exe cmd.exe explorer.exe

Workflow:

- Structured feature extraction from system logs
- Local model training per institution
- Anomaly detection per monitored process
- Encrypted alert exchange via Functional Encryption
- Cross-institutional correlation to detect multi-site campaigns

Results:

✓ Secure signal sharing

✓ GDPR-compliant threat intelligence

✓ Scalable to national CSIRT/CERT contexts

DEMONSTRATOR 2

Sleep Medicine Analytics

Encrypted clinical data analysis for sleep stage classification

Objective: Enable multi-institutional research on sleep data without compromising patient privacy.

Data Source:

- EEG-derived hypnograms
- Labeled stages Wake REM N1 N2 N3

Workflow:

- Local feature extraction from time-series data
- Encryption using HE and FE
- Training classifiers on encrypted data
- Inference over encryption
- Result sharing without disclosing input data

Results:

✓ Accuracy of up to 87.55%
✓ Preserves medical data confidentiality
✓ Supports clinical and research scalability

Tools and Components

To support the demonstrators, HARPOCRATES developed modular tools that combine cryptography, machine learning, and data protection. These are built for easy integration and GDPR compliance.

SPADE Engine: Modular system log anomaly detection

Encrypted ML Wrappers: Enable HE/FE training and inference

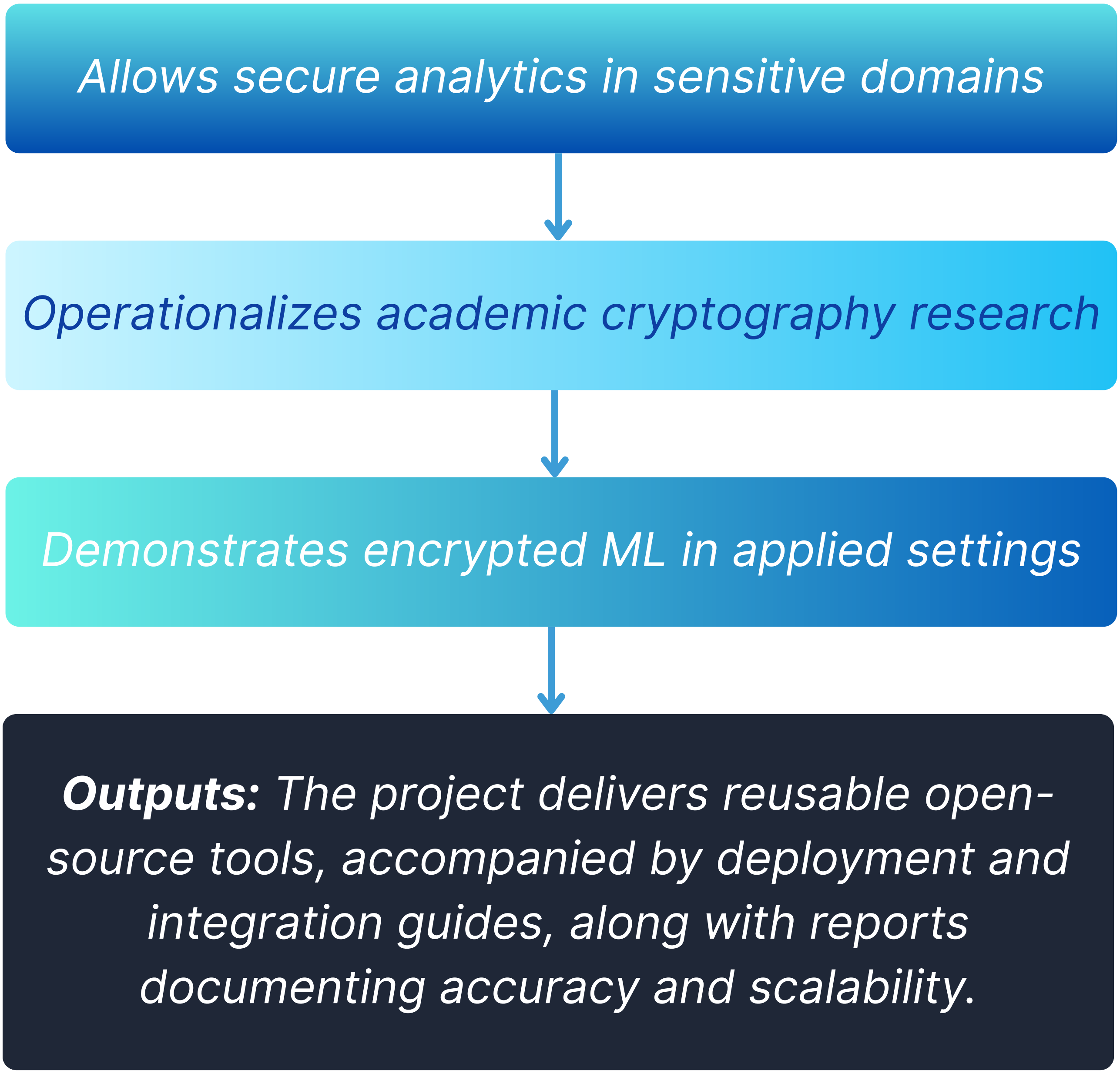
Privacy-Preserving Pipelines: For ML workflows

Testbeds and Evaluation Frameworks

GDPR Compliance Mapping Modules

Technical Impact

The technologies developed in HARPOCRATES are not only experimental but designed for real-world application. This section outlines how the project contributes to secure data processing in sensitive domains.



Allows secure analytics in sensitive domains

Operationalizes academic cryptography research

Demonstrates encrypted ML in applied settings

***Outputs:** The project delivers reusable open-source tools, accompanied by deployment and integration guides, along with reports documenting accuracy and scalability.*

Contact

Project Coordinator: Tampere University, Tampere, Finland

Technical Coordinator: University of Westminster, London, United Kingdom

Scan this QR Code to connect with HARPOCRATES

Co-funded by
the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101069535.

UK Research
and Innovation

UK participant in Horizon Europe Project HARPOCRATES is supported by UKRI grant number 10048312.