encrypt

A scalable and practical privacy-preserving framework

ENCRYPT White Paper

The ENCRYPT Project

Building Trustworthy and Scalable Privacy Preserving Protection for the Digital Era

The ENCRYPT project addresses the growing challenge of safeguarding sensitive data in an era of accelerated digital transformation, where cross-border collaboration and regulatory compliance are paramount. By integrating advanced Privacy-Preserving Technologies - including Fully Homomorphic Encryption, Trusted Execution Environments, Differential Privacy and Hybrid Protection Services - ENCRYPT delivers a robust platform for secure data processing in federated environments. The platform combines cryptographic and hardware-based protections with AI-driven orchestration, hardware acceleration, risk assessment tools and knowledge graphs, ensuring both technical scalability and user-centric accessibility.

Validated across healthcare, finance and cybersecurity use cases, ENCRYPT demonstrates the feasibility of privacy-aware analytics in highly sensitive domains. Its modular and extensible architecture supports statistical analysis, secure enclaves for confidential computing, differential privacy-enhanced model training and GPU-accelerated homomorphic operations. These capabilities enable GDPR-compliant data sharing, predictive modeling and collaborative intelligence without compromising confidentiality.

Through iterative development and real-world validation, ENCRYPT illustrates that hybrid, multi-layered privacy-preserving solutions can meet domain-specific requirements while remaining adaptable to future threats and evolving regulations. The project provides a comprehensive, regulation-compliant framework for confidential computing, highlighting that privacy and innovation can coexist to foster trust, collaboration, and sustainable progress across industries.

Contents

Funded by the European Union

encrypt

A scalable and practical privacy-preserving framework

ENCRYPT White Paper

ENCRYPT Project

The accelerating digital transformation of society has triggered a surge in data generation, unlocking new possibilities for innovation, research and service delivery. However, this rapid progress brings heightened concerns over the privacy and security of sensitive data. For organizations managing distributed and cross-border datasets, compliance with stringent regulations like the GDPR remains a significant challenge. This is particularly true in sectors such as healthcare, finance and cybersecurity, where trust, legal adherence and responsible data use are critical. The ENCRYPT project was launched in response to the growing tension between the potential of data-driven innovation and the increasing demand for privacy-preserving, regulation-compliant solutions.

To address these challenges, ENCRYPT developed a robust, flexible platform designed for secure data analysis in federated environments. It integrates cutting-edge privacy-preserving technologies, including Fully Homomorphic Encryption (FHE), Trusted Execution Environments (TEE) and Differential Privacy (DP), supported by a suite of advanced tools - such as intelligent data preprocessing, knowledge graphs, privacy risk assessment frameworks, AI-powered recommendation systems and a dual-mode (GUI/API) user interface. The platform is now in a mature stage, offering seamless, privacy-aware computations suitable for real-world deployment.

ENCRYPT has been validated across three high-impact domains. In healthcare, it supports secure data sharing between clinicians during multidisciplinary consultations, protecting patient privacy across institutions. In cybersecurity, the platform enables organizations to collaborate on cyber threat intelligence without exposing sensitive internal data, using privacy-preserving techniques to safely correlate threat indicators. In finance, ENCRYPT facilitates privacy-compliant debt collection optimization by enabling machine learning on sensitive customer data without compromising confidentiality or regulatory obligations.

These use cases underscore ENCRYPT's flexibility in meeting complex, domain-specific requirements. The final platform introduces several key enhancements, including accelerated Homomorphic Encryption for Fintech applications, combined use of FHE and TEE in both Health and Finance domains and a more sophisticated decision-support system powered by a risk assessment tool and a recommendation engine.

The final prototype has been thoroughly tested using both internal and external datasets, confirming its robustness, efficiency and adaptability.

In conclusion, ENCRYPT offers a robust and modular solution for privacy-preserving data processing that is ready for deployment across diverse sectors. By combining secure computation frameworks, performance-enhancing technologies and user-focused tools, the platform is well-positioned for widespread adoption beyond its initial project boundaries.

ENCRYPT Platform

The ENCRYPT platform is the central technological outcome of the project, designed to operationalize advanced privacy-preserving technologies (PPTs) in a scalable, user-oriented and regulation-compliant ecosystem. Building upon the foundations introduced in the first phase of the project, the second iteration of the platform extends its architecture with enhanced modularity, orchestration mechanisms and integration workflows that ensure seamless deployment of PPTs across diverse industrial domains.

Funded by the European Union

At its core, the platform is designed to support the secure processing, analysis and sharing of sensitive data under GDPR and broader EU data protection regulations. It achieves this by combining state-of-the-art cryptographic techniques, trusted hardware components and privacy-by-design methodologies within a unified architecture. The platform's overarching goal is to lower the barrier to adoption of complex PPTs, enabling data controllers and processors to derive value from sensitive datasets while maintaining strict guarantees on confidentiality, integrity and accountability.

Integrated Privacy-Preserving Technologies

The ENCRYPT platform integrates a selected suite of PPTs, each addressing different aspects of the data protection challenge:

- **Fully Homomorphic Encryption (FHE)**: Enables computation on encrypted data without decryption, ensuring confidentiality even during active processing. This is particularly relevant for sensitive domains such as medical research and financial analytics.
- **Trusted Execution Environments (TEEs)**: Provide secure enclaves for computation, isolating data and operations from potential system-level threats. TEEs complement cryptographic approaches by delivering performance-efficient security guarantees.
- **Differential Privacy (DP)**: Ensures that statistical insights can be extracted from datasets while preventing the re-identification of individuals, thus safeguarding privacy in large-scale data analytics and AI training.
- **Hybrid Protection Services**: An advanced orchestration layer that allows combined deployment of PPTs (e.g., coupling FHE with TEE-based execution) to create multi-layered protections adapted to specific risk profiles and operational needs.

AI-Powered Orchestration and Recommendation

A key innovation of the platform is its AI-driven recommendation system, which guides users in selecting the optimal PPTs for their specific use case. By analyzing parameters such as dataset size, sensitivity levels, latency constraints and compliance requirements, the system provides tailored user guidance. This functionality is particularly valuable for organizations with limited cryptographic expertise, democratizing access to advanced data protection.

The recommendation system is part of the ENCRYPT platform's dynamic orchestration, where recommended PPTs are not only suggested but also automatically deployed and managed within the platform's workflow. This reduces configuration overhead and minimizes misconfigurations that could weaken privacy guarantees.

Hardware Acceleration and Performance Scaling

Given the computational intensity of PPTs, the ENCRYPT platform integrates hardware acceleration mechanisms to optimize performance. This includes GPU-based acceleration, which reduces latency and increases throughput for cryptographic operations. As a result, the platform supports larger datasets and more complex computations without compromising usability, making advanced privacy-preserving data analytics feasible for real-world, large-scale applications.

Cross-Sector Applicability

Although initially driven by project use cases in healthcare, finance and cybersecurity, the platform's modular and extensible design ensures applicability to a wide range of sectors, from government services to other industry applications. Its flexible architecture supports both research-focused scenarios and production-grade deployments, enabling organizations of varying maturity levels to adopt advanced privacy protection.

Funded by the European Union

Privacy Preserving Technologies

The ENCRYPT project addresses the increasingly complex challenge of preserving privacy in the digital age where data breaches and privacy concerns are present and rising. As part of its mission, the project is dedicated to developing a scalable, practical and adaptable privacy-preserving framework that enables GDPR-compliant processing of sensitive data which are stored in federated cross-border data spaces. ENCRYPT utilises state-of-the-art privacy-preserving technologies (PPTs), including Fully Homomorphic Encryption (FHE), Trusted Execution Environments (TEE), Differential Privacy (DP) and Hybrid Protection Services, which are also supported by hardware acceleration techniques developed within ENCRYPT activities. Together, these technologies represent ENCRYPT's approach to securing data privacy while allowing for the use of data for computational data analytics to be carried out.

By bringing these technologies together, ENCRYPT aims to bring to real-world practical applications, the theoretical potential of PPTs. The project's implementation of FHE, TEE, DP and hardware acceleration solutions is key to its commitment to advancing the field of data privacy, offering robust protection against unauthorized access and ensuring the integrity and confidentiality of sensitive information. In this way, ENCRYPT will be able to play a role in the setting of new standards for data security and privacy.

Fully Homomorphic Encryption

FHE is an advanced privacy-preserving technology achieving robust data security while maintaining functionality. This cryptographic paradigm enables the execution of complex mathematical operations on encrypted data, which allows any data processing in theory and more kinds of data processing in practice. As data are processed in ciphertext form, confidentiality of sensitive information is retained. This makes FHE one of the key technologies of the ENCRYPT project.

ENCRYPT's use of FHE is not just an academic exercise but a concerted effort to translate cutting-edge cryptographic research into practical applications - especially in private-sensitive industry sectors such as healthcare and finance.

FHE involves complex mathematical operations that can be computationally intensive. ENCRYPT's approach to embedding FHE focuses on optimizing these operations to make them feasible for practical applications. This involves addressing challenges related to noise management in FHE schemes - a fundamental aspect that affects the accuracy of encrypted computations.

Application to ENCRYPT and related work

Formally speaking, homomorphic encryption (HE) is a relatively recent cryptographic technique that allows a remote server to perform blind computation on encrypted data without information leak on clear data, even on the result of its own computation (the output being encrypted). The purpose is to ensure data confidentiality not only during storage and sharing as usual, but also during computation.

Three analytic tools have been developed for the ENCRYPT platform (one per internal use-case), each implementing a specific HE evaluation algorithm for the needs of the associated use-case. Thus, homomorphic encryption is used for blind statistical analysis on an encrypted numerical table in the context of the health use-case, for blind search in an encrypted database in the context of the CTI use-case and for blind evaluation of binary decision trees in the context of the FinTech use-case. Client-side tools have also been provided for each of these use-cases, mainly for key generation, encryption and decryption. Here, we provide a description of the FinTech application in more detail.

FinTech use-case

For the FinTech use-case, the analytic tool based on HE evaluates (in a blind way) a complete binary decision tree with one bit (as a model parameter) to be compared with one bit (as specific client data input) at each node, with all these bits being encrypted. The purpose of this evaluation is to know if a client/institution is eligible for a bank loan, with each bit comparison corresponding to a question such as "has the client a source of income greater than X euros per month?" or "has the client any outstanding debt?". The tree is assumed to be from an already trained random forest, so that it can be evaluated independently from the other ones.

Funded by the European Union

A homomorphic encryption is described as “fully” when it allows blind evaluation of additions and multiplications of some ring structure. FHE allows blind decision tree evaluation by evaluating all the decisions as Boolean polynomials specific per internal node. Then for each path (from the root to a leaf), the product of the decisions at all the nodes of the path is evaluated. The final output is evaluated by adding all these partial results.

In the context of the FinTech use-case, this was implemented using the mainstream FHE library OpenFHE, which is currently chosen by the research community. In this context, we use the OpenFHE implementation of the FHE cryptosystem BGV for compliance with specifications from the hardware acceleration tool.

Other use-cases

Whereas the implementation for the health use-case is also based on the OpenFHE implementation of BGV, the HE-based tool for the CTI use-case is implemented using a CEA internal library that implements the Paillier cryptosystem with an extra layer (also internally developed) that allows one level of homomorphic multiplication. Indeed, the blind search algorithm that we implemented is based on a Private Information Retrieval (PIR) protocol that needs exactly one level of multiplication. PIR is a cryptographic protocol that allows a user to retrieve a specific piece of data from a database without the database server knowing which piece of data was requested.

The Paillier cryptosystem has a higher security level than the usual HE schemes such as BGV and allows easier compliance with standards since it is based on more classical mathematics, similar to traditional cryptographic schemes like RSA.

Results of applying homomorphic encryption to the FinTech use-case

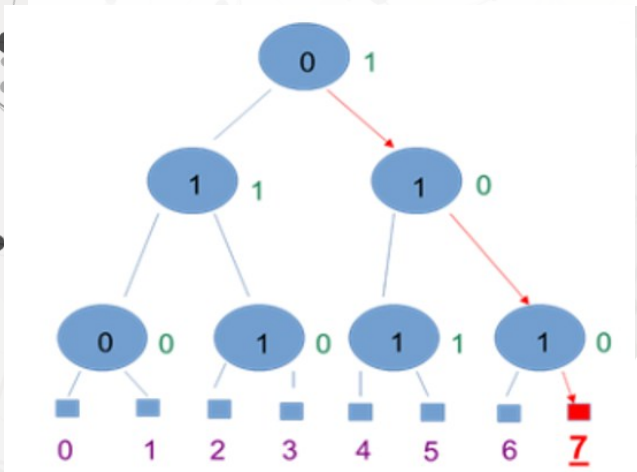
For the FinTech use-case, one analytic tool has been integrated in the main platform of the project to perform blind computation on encrypted data and two client-side tools have been provided in addition, which are an encryptor and a decryptor. These client-side tools have been dockerized for the purpose of the integration. The encryptor generates keys and encrypts clear input data and the decryptor decrypts results from blind analysis.

These tools are implemented in C++ with OpenFHEv1.1.2. The chosen cryptosystem is BGV with multiplicative depth 8, which allows to run tests on trees of depth 3 (that is, with 3 stages of internal nodes). The three dockers (encryptor, analytic tool and decryptor) have the same environment for the compatibility of the OpenFHE objects (cryptographic context, keys, ciphertexts) when turned into serialized txt files and reversely. Such serializations are produced using the native OpenFHE serialization function Serial::SerializeToFile parametrized with SerType::Binary.

Funded by the European Union

Algorithm

Now, we focus on how the blind decision tree evaluation has been implemented. Each non-leaf node of a binary decision tree is tagged with a fixed binary parameter, which is compared to an input binary value. The tree is browsed from the root to a leaf in a way such that at each node, the browser turns left if the parameter is equal to the input value and right otherwise. Then, the output is the tag of the final leaf.



In the picture on the left, the browser turns right at the first step because the fixed parameter associated to the root (1) is different from the input value associated to the root (0). Then the browser turns right again because at the right son node of the root, the parameter (0) is different from the input value (1). Then the browser turns right again because the parameter (0) and the value (1) are still different at the last step and goes to the leaf 7, so that the output result is 7.

Since the building blocks of FHE evaluation are additions and multiplications in discrete rings, this algorithm must be turned into a polynomial form.

To achieve this, each path from the root to a leave is browsed with computing d_i := x_i + t_i + dir_i ("+" being a XOR) at each node, where x_i (in black in Figure 2) is the input value at the i-th node of the path, t_i (in green in Figure 2) is the parameter associated with the i-th node and dir_i is the direction of the path at the i-th node (0 for right, 1 for left). The only path that leads to the correct leaf gives d_i = 1 for all i (in our example this path is [right, right, right], that is, [dir_0, dir_1, dir_2] = [0, 0, 0] for this path). For any other path, there is some i with d_i = 0. Hence, the product p_j of all the d_i associated to the path leading to a leaf j is 1 for the good path (j = 7) and 0 for all other paths. Hence, the sum of all the jp_j gives the expected output result (in our example, this sum is 0x0 + 0x1 + 0x2 + … + 0x6 + 1x7 = 7).

Workflow

The clear input data for the encryptor are two txt files (one for the client data, one for the model parameters), each one containing bits organized through the same complete binary tree structure, depicted with one step per line and with spaces as separators between 2 nodes of the same step (see an example of such encoding below).

1
0 1
1 0 1 1

The encryptor outputs:

- An OpenFHE "cryptontext" object (which encodes the public parameters of the cryptosystem),
- A public encryption key,
- A secret decryption key,
- An evaluation key (to perform the homomorphic multiplications),
- One HE-encrypted model parameter bit per internal node of the tree,
- One HE-encrypted client data bit per internal node of the tree.

Funded by the European Union

The inputs of the analytic tool are exactly the outputs of the encryptor, except the secret decryption key that must be sent to the decryptor only. The output of the analytic tool is an encrypted integer that encodes the tag of the selected output leaf of the tree. Each of these data (cryptographic context, key, encrypted bit) is encoded as a txt file, as well as the output ciphertext. Finally, the inputs of the decryptor are the cryptographic context, the secret decryption key and the ciphertext outputted by the analytic tool. The decryptor outputs the tag of the selected leaf in clear, also as a txt file.

This workflow is depicted in the following diagram. Note that the HE workflow for the two other use-cases have the same basic structure with one first client round for key generation and encryption, one server round for analysis and one final client round for decryption.

Performances

The full workflow has been successfully tested on a laptop running Ubuntu 22.04.5 with a 13th Gen Intel(R) Core(TM) i7 -13700H processor (20 cores), with trees of depth 3 in input, filled with uniformly random synthetic parameters and uniformly random synthetic client inputs.

The costliest part of the workflow is the homomorphic analysis itself, which takes almost 35 seconds in the average case with low variability. The encryptor is run in almost 1 second and the decryptor is run in less than 1 second. In comparison, the homomorphic analysis for the health use-case is more performant, since it also takes less than 1 second in an analogue framework.

The following table gives the approximative memory size of the data generated and shared by the three tools. The cryptocontext file does not figure in this table since its size is negligible in comparison with all other generated data (less than 4 kB). The heaviest data transfer is the one from the encryptor to the analytic tool (after encryption and before analysis). The total size of the data transferred in this context is almost 100 MB. The other data transfer steps are far less heavy.

Data	Public key	Evaluation key	Private key	Encrypted parameter bit	Encrypted client bit	Result from blind analysis
Approximative	6.5 MB	20 MB	2.5 MB	5 MB	5 MB	2 MB

Work has taken place in parallel on an optimized version of the algorithm for blind tree evaluation. However, this version is not integrated in the ENCRYPT platform for the moment. According to the first tests, the optimized version allows homomorphic evaluation of trees of depth up to 7 in few seconds, with a small integer comparison at each node, these small integer values being encoded on 4 bits.

Trusted Execution Environment

Trusted Execution Environments (TEEs) are an important aspect of ENCRYPT's multi-technology privacy-preserving framework. A TEE is a secure enclave within the processor, isolated from the main operating system and untrusted applications. This isolation provides a protected execution context that guarantees the confidentiality and integrity of both code and data, even in adversarial settings where the broader system might be compromised. Within ENCRYPT, TEEs can be used to enable confidential computing, ensuring that sensitive data remains protected throughout its processing lifecycle.

Funded by the European Union

Role of TEE in the ENCRYPT Platform

While encryption ensures protection of data at rest and in transit, conventional paradigms expose data during processing. TEEs directly address this gap by allowing computations to be executed inside a secure enclave, where neither the host system nor potential attackers can interfere. This significantly reduces the attack surface, as illustrated by the shrinkage in exploitable vectors once enclaves are introduced.

In ENCRYPT, TEEs are deployed as execution backends for sensitive operations that require both high performance and strong security guarantees. They are particularly effective for tasks where homomorphic encryption alone would be too computationally expensive, or where real-time responsiveness is required.

Integration with Other Privacy-Preserving Technologies

The ENCRYPT architecture integrates TEEs in a hybrid protection model that combines the strengths of multiple PPTs. For example:

- **TEE + FHE Hybrid Workflows**: While FHE enables computation over encrypted data, it is computationally demanding. TEEs can offload parts of the computation, allowing efficient scheme-switching and practical deployments without compromising security.

This layered integration ensures robustness against diverse threat vectors, aligning with GDPR's principles of "data protection by design and by default."

Design and Implementation in ENCRYPT

The ENCRYPT project has adopted a TEE design approach based on Gramine, an open-source framework that facilitates running applications inside Intel SGX enclaves. This implementation replaces earlier prototypes, providing stronger guarantees and greater configurability.

Key functionalities developed within the ENCRYPT context include:

- **Secure Execution of Privacy-Preserving Algorithms**: Machine learning models, data queries and cryptographic functions are executed inside enclaves to ensure that sensitive intermediate values are never exposed.
- **Remote Attestation**: TEEs support remote verification mechanisms, allowing users in federated environments to verify that computations are executed within genuine and untampered enclaves before sharing data. This builds trust in cross-organizational collaborations, such as healthcare research networks or cyber threat intelligence exchanges.

Differential Privacy

Differential Privacy is a mathematical framework that offers strong guarantees on the privacy of individual data entries when data is shared. Within ENCRYPT, Differential Privacy serves as a privacy preserving technology which helps ensure that sensitive information can be processed, analyzed and shared without risking the privacy of individual-specific details, fully aligning with GDPR and ethical guidelines.

ENCRYPT Implementation and Optimizations

Differential Privacy works by introducing noise to datasets, while still allowing for analytics upon these datasets to take place. A practical challenge is balancing privacy guarantees with data utility, which considers how much noise to add to a dataset to still maintain its privacy while still allowing for the data to be used for accurate model predictions.

Funded by the European Union

Key innovations and implementation details include:

- **Data Preprocessing & Noise Addition**: Sensitive datasets were normalized using standard scaling and Laplacian noise was added to these dataset.
- **Experiments with Multiple Classifiers**: Extensive experimentation was performed on Random Forest Classifiers, Decision Trees, Gaussian Naive Bayes, Logistic Regression, Neural Networks and Linear Regression. Both standard (non-private) and differentially private versions of these classifiers were implemented for experimentation and comparison purposes.
- **Differentially Private Model Training**: A hyperparameter tuning pipeline was developed for Differentially Private versions of classifiers and grid search with k-fold cross-validation was used to select optimal hyperparameters under noisy data constraints.
- **Adaptive Optimization Loop**: In some scenarios, models were trained to iterate with different noise levels and configurations, stopping when a predefined accuracy threshold or maximum number of iterations was reached. The most accurate model achieved was then serialized for future deployment.
- **Integration into ENCRYPT Architecture**: The Differential Privacy modules were integrated into ENCRYPT's Platform and closely coupled to ENCRYPT's AI-powered recommendation system, where use-case requirements allow for the right amount of noise to be added to a dataset.
- **Use-Case Driven Validation**: The Differential Privacy component was successfully validated in both Fintech and Health sector use cases. For Fintech, models predicted loan default risks, while for Health, blind statistics over sensitive medical attributes were computed.

Trade-offs and Research Directions

ENCRYPT's extensive experimentation with Differential Privacy highlighted the following key trade-offs which need to be considered:

- **Accuracy vs. Privacy**: As privacy guarantees increase when more noise is added to dataset, model accuracies decrease. Strategies such as hyperparameter optimization and feature scaling helped partially mitigate these effects.
- **Training Stability**: Noisification of datasets can destabilize the convergence of certain classifiers (e.g., Neural Networks). Hence, certain models need to run for longer training durations.
- **Scalability**: Noise addition and privacy-preserving model training can be computationally expensive, especially for ensemble methods like Random Forests. Performance optimizations through parallelized grid search are one feature which can be used to combat this.

Hardware Acceleration

To address the computational demands of PPTs such as FHE, the ENCRYPT project integrates hardware acceleration as a fundamental enabler for practical adoption. While FHE offers strong privacy guarantees, its high computational cost often limits usability. By leveraging Graphics Processing Units (GPUs), we can significantly reduce processing time and improve energy efficiency - making advanced cryptographic techniques more accessible and scalable.

Funded by the European Union

Objectives and Scope

The goal of the Acceleration Service in ENCRYPT is to enhance the performance of computationally intensive privacy-preserving operations through GPU acceleration. Our focus is on the Brakerski-Gentry-Vaikuntanathan (BGV) encryption scheme, which is widely used in FHE applications. Currently, support for GPU acceleration in BGV is limited, which presents a valuable opportunity for innovation.

Our solution builds on OpenFHE, a leading open-source FHE library. We extend its capabilities by offloading selected high-cost operations to GPUs. This is done by first identifying performance bottlenecks, then re-implementing those parts using GPU-compatible code and finally integrating the accelerated components back into the library for seamless use.

Profiling and Bottleneck Identification

To determine where acceleration would have the most impact, we first analyzed an application that combines common FHE operations such as addition, multiplication and rotation. This profiling revealed that homomorphic multiplication is the most time-consuming operation during typical application workflows.

We then examined a more focused use case with repeated multiplications to understand the underlying performance costs. This deeper analysis showed that a specific internal function - ApproxModDown - accounts for a significant portion of the overall runtime. As a result, we selected this function as our primary target for acceleration.

GPU Acceleration Strategy

To accelerate the ApproxModDown function, we restructured it into smaller, independent steps and implemented these on the GPU. This allowed us to take advantage of GPU parallelism while maintaining compatibility with the existing OpenFHE library. We also introduced a mechanism to switch between CPU and GPU execution paths, enabling flexibility depending on system capabilities.

To make the most of the GPU's performance potential, we optimized data processing and scheduling so that computation and memory transfers can occur in parallel. This helps reduce idle time and ensures efficient use of resources across different stages of the operation.

Integration with OpenFHE

Our GPU acceleration is fully integrated into OpenFHE's execution flow. When enabled, the library automatically offloads selected computations to the GPU without requiring changes from the user. We also parallelized certain parts of the flow to further improve throughput, especially in scenarios involving large inputs or repeated operations.

Performance Analysis

We benchmarked our GPU-accelerated implementation against the standard CPU-only version of OpenFHE. At low workloads, the overhead of transferring data to and from the GPU offsets the benefits of acceleration. However, as the workload increases, the GPU implementation becomes significantly more efficient.

At higher operation counts, we observed noticeable speedups both in specific components and in overall execution time. For example, one critical operation showed more than a fourfold speedup, while overall application runtime improved by up to 8%. Importantly, these gains were achieved by accelerating only a small part of the full computation flow, suggesting even greater potential as more components are optimized.

Funded by the European Union

Key Takeaways and the Road Ahead

This work highlights how targeted GPU acceleration can make privacy-preserving technologies, such as FHE, significantly more practical and efficient. Our approach centered on optimizing a critical bottleneck - the ApproxModDown function, while ensuring compatibility with the widely used OpenFHE library.

Key Innovations include:

- **Selective Acceleration**: Focused GPU offloading of the most expensive FHE operation without overhauling the entire software stack.
- **Seamless Integration**: Designed to work natively within OpenFHE, allowing users to benefit from acceleration without changing their code.
- **Modular and Scalable Design**: The pipeline-based implementation supports parallelism and lays the groundwork for broader acceleration.

Future Research Directions include:

- **Expanding Acceleration Coverage**: Offloading additional FHE operations to further enhance overall performance.
- **Exploring New Hardware Platforms**: Extending support to other accelerators such as FPGAs or custom hardware.
- **Applying to Real-World Use Cases**: Integrating GPU-accelerated FHE into practical domains like secure data analytics, privacy-preserving machine learning and federated systems.

These directions will not only extend the technical impact of this work but also contribute to the broader adoption of efficient, privacy-respecting technologies across industries.

Privacy Supporting Tools

Pre-processing Tool

The ENCRYPT platform integrates a specialized pre-processing module designed to enhance its privacy-preserving functionality. This module utilizes datasets for secure computation within the ENCRYPT framework, ensuring robust detection and management of Personally Identifiable Information (PII) in compliance with stringent privacy and data protection standards. Its primary objective is to optimize data utility while minimizing privacy risks.

The pre-processing module performs an extensive range of data cleaning and transformation operations to prepare inputs for downstream PPTs, including FHE and DP. Key operations encompass missing value imputation, deduplication of records and the application of sophisticated feature selection algorithms. By employing feature selection techniques, the module enhances computational efficiency, facilitating scalable processing of large-scale datasets while preserving processing performance and privacy assurances.

Central to the module's functionality is the automated identification and extraction of PII, achieved through advanced Machine Learning (ML) models and Natural Language Processing (NLP) techniques, particularly Named Entity Recognition (NER). This process isolates sensitive data elements, generating PII metadata that drives the platform's Recommendation System. The system subsequently proposes the optimal PPT - such as DP, FHE, or Trusted Execution Environments, for secure data handling. This automated workflow significantly reduces the risk of privacy violations, bolstering the security and regulatory compliance of the ENCRYPT data processing pipeline, including adherence to GDPR and other pertinent data protection frameworks.

Funded by the European Union

Recommendation System

The ENCRYPT Recommendation Engine is a decision-support component integrated within the ENCRYPT platform, designed to help users identify the most appropriate Privacy-Preserving Technologies (PPTs) for their data processing scenarios. Unlike traditional matching tools, this engine uses a fuzzy logic-based framework that enables natural language-style rules to inform numerical decision-making. This architecture provides both flexibility and transparency - the system is lightweight, deterministic and inherently explainable, with justifications derived from the rule base that mirror the engine's internal logic.

The engine evaluates each scenario using a multidimensional input set. Key criteria include data sensitivity, data size, data type and several operational constraints - namely computational, cost, time and performance constraints, as well as computational intensity and location of computation. These factors determine the appropriateness of various PPTs. For instance, the availability of GPU acceleration at the specified location informs whether FHE is a viable option given time constraints. The same is true of the availability of TEE.

The knowledge base behind the recommendation engine is modular and easily extensible. While it does not dynamically incorporate new research or regulations, it can be updated with new technologies or revised rule sets without modifying the engine's underlying structure. This flexibility was critical during the project's development, where updates were introduced to handle additional use case-specific inputs - most notably cost constraints and location of computation, to improve the precision of recommendations, especially in scenarios requiring acceleration technologies provided by the University of Manchester.

In terms of outputs, the engine provides natural language justifications that summarize why a particular technology is appropriate. To strengthen trust and interpretability, ENCRYPT also experimented with large language models (LLMs) to augment and refine these explanations. Moreover, justifications now include contextual information from the ENCRYPT risk assessment tool. While the risk scores themselves do not directly alter the recommendation, they are referenced in the output to prompt users to reflect on the broader threat landscape, especially when asset risk scores exceed critical thresholds.

The system was tested across four representative use cases from different domains:

- In the Fintech use case, the engine recommended varying technologies (e.g., FHE, TEE, DP, FHE+ACC) depending on inputs related to performance and cost constraints,
- For healthcare, recommendations typically included FHE or a TEE+FHE combination, accounting for high data sensitivity and regulatory expectations,
- In the Cyber Threat Intelligence (CTI) use case, TEE+FHE was consistently recommended due to stringent performance and confidentiality demands,
- Finally, in the MIRACUM use case (clinical research network), DP or FHE were selected depending on the type and size of data and constraints provided.

The component was significantly hardened during development: improvements in error handling and missing data detection were introduced to ensure robustness, especially as the engine now processes a broader and more nuanced range of inputs. Despite this increased complexity, system performance remained unaffected due to the engine's computationally lightweight nature.

In summary, the ENCRYPT Recommendation Engine is not only technically mature but also aligned with the practical needs of real-world data custodians. It exemplifies how explainable, rule-based AI can support secure and privacy-aware decision-making in complex data environments.

Recommendation system

ENCRYPT incorporates Knowledge Graphs as core components of its data processing and enrichment infrastructure. These graphs serve not only as repositories of interconnected information, but also as dynamic models that facilitate understanding, accessibility and semantic interoperability across diverse data domains. By establishing relationships between entities and concepts, the platform enables the construction of a semantically rich, federated web of knowledge that supports both detailed analysis and privacy-aware automation.

Funded by the European Union

The system ingests heterogeneous and sensitive data from multiple use cases, including financial data, health records, cybersecurity event logs and research datasets. These inputs are provided in formats such as CSV, JSON and TXT and are semantically aligned using domain-specific ontologies. Financial datasets are mapped against standards such as FIBO, health data utilise SNOMED and FHIR, while cyber threat intelligence leverages structured frameworks like MITRE ATT&CK. This modular approach ensures accurate representation of entities, traceability of relationships and coherent integration across heterogeneous data sources. All transformation processes are aligned with GDPR principles and further support the discovery of personal and sensitive information that may not have been explicitly annotated in upstream systems. The framework's adherence to W3C standards such as RDF and OWL ensures long-term semantic consistency and interoperability.

The enriched Knowledge Graphs provide a semantic backbone for ENCRYPT's AI-powered services, enhancing their interpretability, accuracy and contextual relevance. By embedding domain knowledge and aligning with privacy constraints, they enable downstream functionalities such as intelligent recommendations, risk detection and decision support to operate over a structured and meaningful data space. A dedicated reasoning layer further enables the extraction of actionable insights and reinforces the interpretability of decision-support mechanisms. Moreover, the interconnected nature of the graphs fosters a collaborative ecosystem in which information can be securely shared and semantically reused across actors, tools and sectors - amplifying both the practical utility and governance of data within the ENCRYPT framework.

Risk Assessment Tool

The purpose of the ENCRYPT's Risk Assessment Tool is to define and implement a methodology that permits in the identification, evaluation and mitigation of privacy risks associated with business processes.

In this first release, the Risk Assessment Tool performs analysis on three different domains, which are Healthcare, Fintech and Cyber Threats Intelligence (CTI).

To enhance the effectiveness of the analysis, domain-specific asset taxonomies have been defined in collaboration with subject matter experts. These taxonomies provide a structured understanding of the typical assets within each domain and enable the identification of potential threats (LINNDUN-based) and vulnerabilities to which these assets may be exposed (MITRE ATT&CK-based).

The approach to carrying out the risk evaluation is illustrated in the figure below.

First, the user must identify their business processes (BPs) and for each of them, they must identify the personal data processes (PDPs). For every PDP, the user chooses the assets involved; after that, they can perform an Impact Analysis.

In the Impact Analysis phase, the user indicates the privacy threats that can affect the assets. The threats from which the user can choose are those made available by LINDDUN.

Funded by the European Union

For every selected threat, the user must indicate their impact in terms of high, medium and low. At the end of this phase, the system has computed the likelihood of an impact.
The next phase is performing a Vulnerability Assessment for each asset, starting from the known vulnerability, the user can indicate new vulnerabilities and select the mitigation for them. The vulnerability and mitigation used are all provided by the MITRE ATT&CK. At the end of this phase, based on the selected vulnerability and mitigation, each asset must have a vulnerability score.

In the last phase, the user can review the results of the previous phase and check the risk score of each asset. The risk score is given by the product of the likelihood of impact and the vulnerability score.
To check the maturity level of a PDP, in the last phase, the user should fill out a questionnaire based on the ISO/IEC 27005:2022. Every question will be answered with an answer that has a score from 0 (very negative answer) to 5 (very positive answer); the sum of the scores indicates the maturity level.

Level	Description
1 – Initial	In this stage, the organization is in the early phases of event-based risk management. There is limited awareness and the processes are ad-hoc or non-existent. Significant improvements are needed in all areas.
2 – Managed	The organization has started to establish a basic event-based risk management process. There is growing awareness and some processes are in place. However, significant enhancements are required for consistency and effectiveness.
3 – Defined	At this stage, the organization has a defined event-based risk management process. It follows standardized methodologies and maintains an updated list of potential events. However, there's room for improvement in terms of alignment and documentation.
4 – Quantitatively Managed	The organization has achieved a mature state of event-based risk management. Processes are well-documented, risks are quantitatively assessed and measures are aligned with assessed risk levels. Continuous improvement and integration are emphasized.
5 - Optimizing	This is the highest level of maturity. The organization actively promotes a culture of event-based risk awareness and continuous improvement. Lessons learned are integrated and there is strong alignment with the ISMS.

Funded by the European Union

At the end of the whole process, the user can send the assessment to the ENCRYPT's Kafka broker, thus making the compiled data available.
The sharing of assessment data also allows integration with the recommendation engine. In this way, the user can clone the assessment and recompile it according to the recommendations provided by the recommendation engine.
The Risk Assessment Tool has been developed with a strong focus on usability, offering an intuitive interface that facilitates seamless navigation throughout the assessment process. Its design ensures accessibility for a broad spectrum of stakeholders, including data protection officers, IT security professionals, project managers, junior researchers and policy makers. By lowering barriers to use, the tool enables active engagement in the management of privacy and security aspects related to data processing activities. Furthermore, the tool is built to accommodate asset taxonomies from additional domains, allowing for future extensibility and adaptability to evolving sector-specific needs.

Use Cases

The ENCRYPT project presents three use cases which illustrate the practical application and impact of these PPTs across the sectors of Fintech, Health and Cyber Threats. Each use case has been designed to address the unique challenges and requirements of its respective domain, showcasing the adaptability and effectiveness of the ENCRYPT platform in safeguarding sensitive information while facilitating essential data processing activities.

Fintech

The Fintech use case demonstrates how the ENCRYPT platform applies privacy-preserving technologies to real-world challenges in the financial sector, particularly in secure data sharing, regulatory compliance and risk management. The project has achieved notable progress in enhancing technical scalability and enabling practical, real-world integration.
A key development in this use case was the successful integration of HE, allowing sensitive financial data - such as transaction histories, demographic information and behavioral scores - to be processed without exposing the underlying content. In addition to HE, the Fintech use case explored the use of DP to protect customer data during analytics. Following a Local Differential Privacy (LDP) model, noise is added directly on the user's device before data is uploaded to the ENCRYPT platform. This ensures that sensitive financial information remains private from the outset. The DP functionality was embedded into the ENCRYPT user interface, giving institutions a seamless way to anonymize datasets while retaining analytical value - a critical step for GDPR compliance and responsible data usage.
The use case also validated ENCRYPT's modular design through integration with the Financial Industry Business Ontology (FIBO), enhancing data interoperability and enabling automated risk assessment through semantic annotations. The creation of a knowledge graph supported compliance and informed decision-making via dynamic querying and reasoning.
Another key milestone was deploying the Acceleration Service on Microsoft Azure, allowing scalable, cloud-based GPU support for FHE tasks. The containerized environment, also available on DockerHub, ensures flexibility across on-premise and cloud infrastructures.
In testing with over 3,000 anonymized EPIBANK records, the Recommendation Engine was able to suggest suitable privacy-preserving techniques - such as FHE, TEE, or DP, based on constraints defined by data owners and researchers.
The Fintech use case confirms ENCRYPT's ability to meet evolving industry needs. With upcoming validations on external datasets, it offers a scalable model for privacy-preserving analytics in areas like consumer lending and risk profiling. Through secure and compliant data handling, ENCRYPT positions privacy not as a barrier, but as a strategic advantage.

Funded by the European Union

Health

The Health Use Case within ENCRYPT exemplifies the application of privacy-preserving technologies (PPTs) in one of the most sensitive and heavily regulated sectors: healthcare. Protecting patient information while enabling meaningful analysis and collaboration is a core requirement of modern healthcare systems. ENCRYPT activities demonstrate how advanced cryptographic and hardware-based protections can make this possible, ensuring compliance with GDPR and other privacy regulations while fostering innovation in medical research and clinical practice.

Addressing Challenges in Healthcare Data

Healthcare data presents unique challenges: it is highly sensitive, often distributed across multiple institutions and must be processed in ways that respect strict regulatory frameworks. Traditional anonymization techniques are increasingly insufficient in preventing re-identification risks. The ENCRYPT platform addresses these challenges by providing end-to-end confidentiality of health data - during storage, transfer and computation, using a combination of FHE, TEE and DP.

Privacy-Preserving Analytics for Clinical Research

In the Health Use Case, new tools were developed to enable blind statistical analysis over encrypted medical datasets. This allows healthcare professionals and researchers to compute critical indicators without exposing individual-level data. By leveraging homomorphic encryption, statistical queries can be run directly on encrypted tables, while DP mechanisms add controlled noise to further mitigate privacy risks.

Cyber Threats

The ENCRYPT Cyber Threats use case highlights the important role PPTs can have in strengthening cybersecurity measures against increasingly sophisticated cyber threats. The use case also demonstrates how the ENCRYPT platform employs advanced cryptographic and privacy-enabling technologies to enable secure and private sharing of Cyber Threat Intelligence (CTI) across organizations. Specifically, for this use case, FHE and TEE have been used. These technologies enhance collective defense mechanisms while ensuring that sensitive information remains protected.

The challenge of this use case is the balance between the need for open sharing of threat intelligence and the necessity to protect the confidentiality of the shared data. Cybersecurity teams require access to timely and relevant threat data to effectively prevent, detect and respond to cyber-attacks. However, the sharing of this data often involves sensitive information that could reveal vulnerabilities, proprietary security measures, or even personal data, thereby posing privacy risks. The ENCRYPT project addresses this challenge using PPTs. These technologies ensure that data can be analyzed, correlated and exchanged without exposing the actual content, thus maintaining the privacy of the data subjects and the security of the organizations involved.

Initially, data pre-processing is used in the Cyber Threats use case. The Cybersecurity domain data holders use the tool on their premises to clear their data (e.g. remove duplicate entries) and identify PIIs in the dataset. The Cyber Threats use case also uses the ENCRYPT's Recommendation System (RE) and the Risk Assessment tool, which is tailored for the cybersecurity domain. The RE assists organizations in selecting the most suitable PPTs based on the specific context of shared threat intelligence, considering factors such as the sensitivity of the data and the required level of data utility. In parallel, the Risk Assessment tool enables organizations to identify and evaluate the potential privacy and security risks associated with sharing CTI, providing actionable insights for risk mitigation.

After that, the organization's datasets are sent to the TEE for secure extraction of CTI. Inside the TEE, functions that extract STIX-compliant CTI are running along with functions that extract patterns from the dataset, for example, how many times a malicious IP has been identified in the dataset.

Funded by the European Union

The use case also highlights the importance of collaboration and interoperability in combating cyber- threats. By facilitating secure and privacy-preserving sharing of CTI, the ENCRYPT platform enhances the collective cybersecurity posture of participating organizations. It enables a more coordinated response to cyber threats, reducing duplication of efforts and accelerating the countering of threats and dissemination of critical threat intelligence. Specifically, in the use case, a correlation procedure is performed inside TEE between the organization datasets that share the same characteristics. For example, if two datasets from different organisations contain IP-related attacks, then correlation is happening between the datasets, so some general statistics are shared between the organizations. In this way, the organisations can have a broader view of the threats that target their infrastructure.

Conclusion

The ENCRYPT project has delivered a comprehensive platform for privacy-preserving data processing, demonstrating how diverse Privacy-Preserving Technologies - including Fully Homomorphic Encryption, Trusted Execution Environments, Differential Privacy and Hybrid Protection Services, can be effectively integrated into a unified framework. The platform was successfully aligned with GDPR principles, ensuring that sensitive data can be processed, shared and analyzed without compromising confidentiality or integrity.

Overall Assessment

Across its use cases in healthcare, finance and cybersecurity, ENCRYPT has shown that advanced PPTs can move from theoretical constructs to practical, deployable solutions. The iterative development process revealed that:

- **Integration is key**: No single technology can cover all privacy and performance needs; hybrid approaches combining FHE, TEE proved useful and essential.
- **Trust mechanisms are critical**: Features such as remote attestation in TEEs were indispensable for establishing trust across federated data collaborations.
- **Efficiency matters**: Hardware acceleration and scheme-switching approaches reduced the performance overheads typically associated with PPTs, making them viable for larger datasets and real-world scenarios.
- **User-centricity increases adoption**: The inclusion of an AI-driven recommendation system and risk assessment tools lowered the barrier for organizations to adopt complex PPTs, enabling informed decision-making even without deep cryptographic expertise.

Lessons Learnt

During the design and deployment of the platform, several key lessons emerged:

- **Balance between security and usability** – Strong privacy guarantees must be matched with practical runtime performance and ease of integration into existing infrastructures.
- **Adaptability across domains** – While healthcare, finance and cybersecurity shaped the platform's requirements, the modular architecture demonstrated potential applicability across many other sectors.
- **Importance of interoperability** – The ability to combine different PPTs and tailor them to specific workflows was one of the most impactful outcomes of the project.
- **Continuous evolution of threats** – The platform's hybrid and layered approach will be crucial for adapting to emerging risks and regulatory demands in the future.

Project Closure and Outlook

The ENCRYPT project not only delivered a functional platform but also provided a framework for the broader adoption of privacy-preserving technologies in Europe and beyond. It demonstrated that federated data spaces can operate securely, fostering trust, collaboration and innovation without sacrificing data protection.

ENCRYPT leaves behind a scalable, adaptable and regulation-compliant framework that serves both as a technological foundation and as a proof-of-concept for future initiatives in confidential computing. The project's outcomes underline that privacy and innovation need not be in conflict: with the right architectures and tools, sensitive data can be a driver of progress while remaining fully protected.

Funded by the European Union

encrypt

A scalable and practical privacy-preserving framework

ENCRYPT project in a nutshell

Project Title	A scalable and practical privacy-preserving framework
Acronym	ENCRYPT
GA No	101070670
Start	01 July 2022
End	30 June 2025
Budget	4.392.540 €
EU Funding	4.392.540 €
Call	HORIZON-CL3-2021-CS-01
Funding	RIA - Research and Innovation action
Topic	HORIZON-CL3-2021-CS-01-04

Consortium

EXUS · ENGINEERING THE DIGITAL TRANSFORMATION COMPANY · CERTH CENTRE FOR RESEARCH & TECHNOLOGY HELLAS · EIGHTBELLS Independent Research & Consultancy · cea

TrustUP A PARTHENOPE UNIVERSITY SPIN-OFF · ARISTOTLE UNIVERSITY OF THESSALONIKI · DBC EUROPE S.A. · TILBURG UNIVERSITY Understanding Society · UNIVERSITÀ DEGLI STUDI DI NAPOLI FEDERICO II

BANK OF EPIRUS COOPERATIVE YOUR LOCAL BANK · GOETHE UNIVERSITÄT FRANKFURT AM MAIN · JG|U UNIVERSITÄTSmedizin. MAINZ · MANCHESTER 1824 The University of Manchester

Stay in touch!

https://encrypt-project.eu · encrypt-project · @encrypt_project

encrypt

A scalable and practical privacy-preserving framework

Funded by the European Union