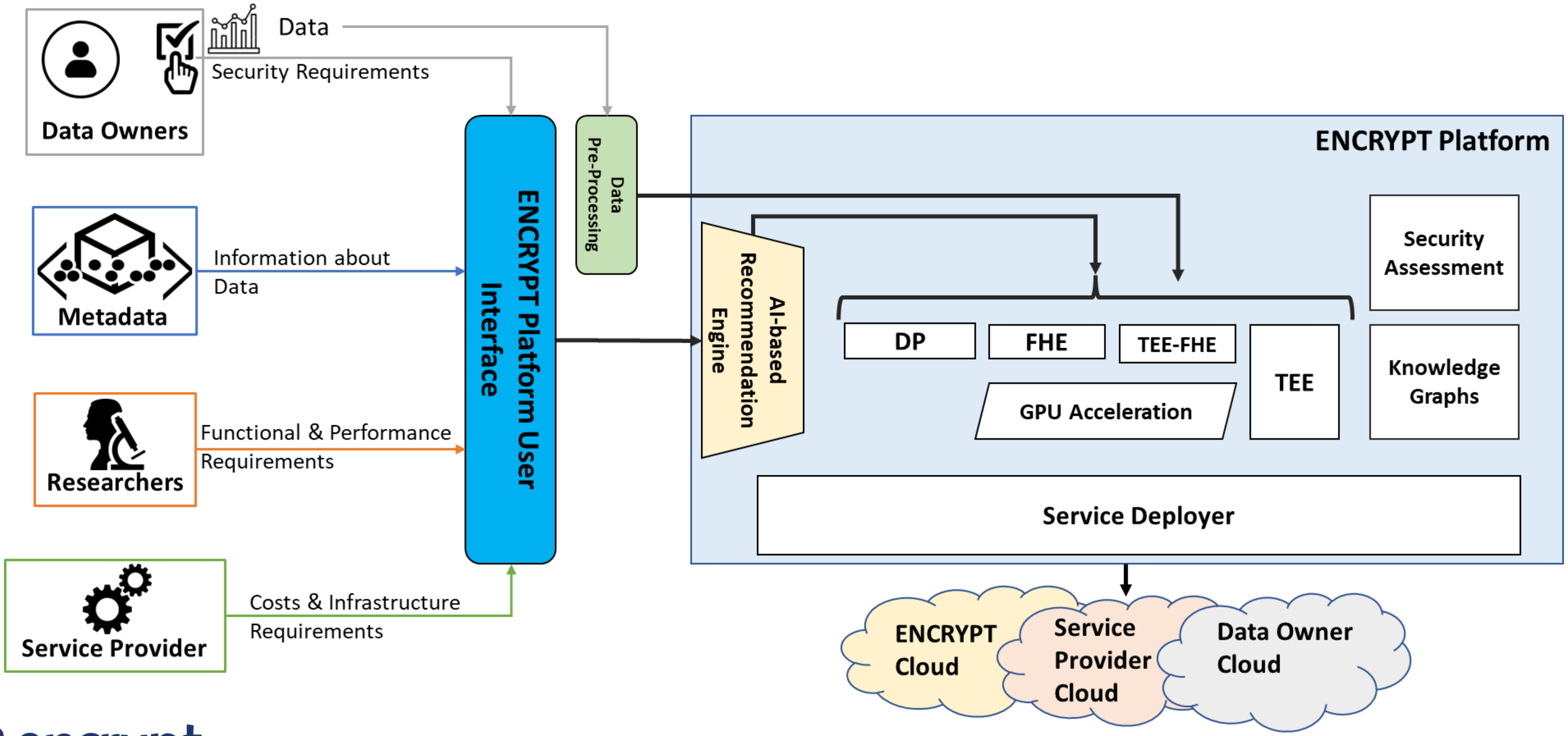
ENCRYPT Technologies

Supporting Technologies

Privacy Preserving Technologies

encrypt

ENCRYPT Technologies and ENCRYPT Platform



encrypt

ENCRYPT Supporting Technologies

- Pre-processing Tool
  - ✓Prepares data for secure processing
- Knowledge Graphs
  - ✓Enhances data interoperability and understanding
- Risk Assessment Tool
  - ✓Evaluates privacy and security risks
- AI Recommendation Engine
  - ✓Suggests optimal Privacy-Preserving Technologies
- User-Centric Design
  - ✓Interface designed for ease of use and accessibility across diverse set of users

ENCRYPT Pre-processing Tool

- Why Data Preprocessing Matters?
  - ✓ Preprocessing is the first step in privacy-preserving analysis
  - ✓ It ensures data is in the correct format for privacy tools and machine learning

- Objective of the Preprocessing Tool
  - ✓ Clean data to make it usable by privacy-preserving technologies
  - ✓ Prepare data for downstream tasks like Machine Learning

- Extracts Private Identifiable Information (PII)
  - ✓ Removes any information that can directly or indirectly identify a person.
  - ✓ Preprocessing aligns data with privacy and compliance goals making it safe for processing

ENCRYPT Pre-processing Tool

- Uses Regular Expressions, NEPRII Python Library, BERT-based NER

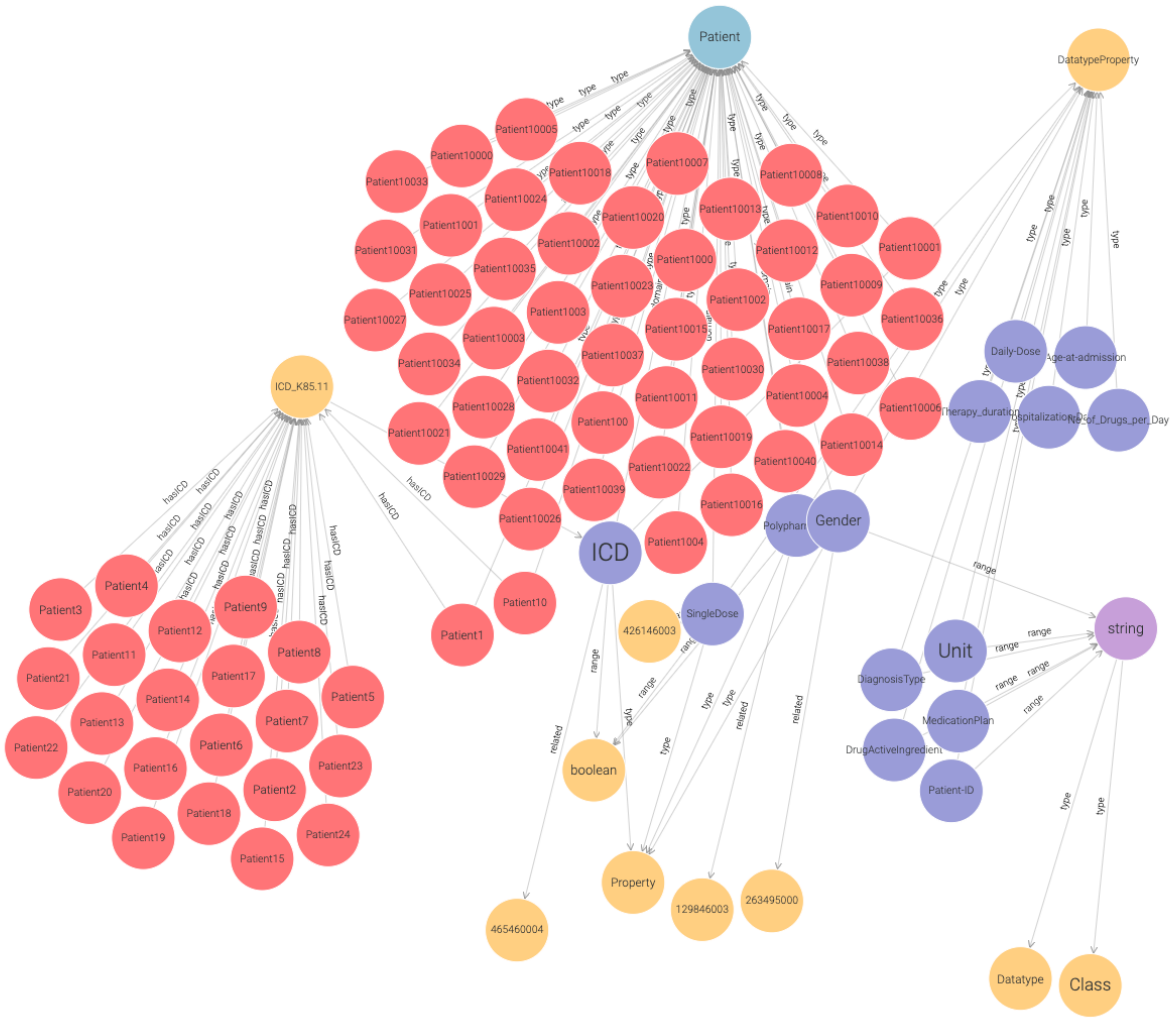
Step	Technique	Purpose
Remove NaNs	Clean dataset	Ensures data usability
Clean symbols (e.g., "-", "14/3")	Normalize data types	Reduces dimension explosion in FHE
Handle missing values	Data imputation or removal	Improves accuracy
Remove duplicates	Optimize size	Enhances performance
Convert categorical to numeric	One-hot encoding or label encoding	FHE compatibility
Feature selection	Remove correlated features	Minimize dimensions for FHE

ENCRYPT Knowledge Graphs

- Goal: Deliver user-friendly tools for privacy-preserving data integration
- Approach: Use Knowledge Graphs (KGs) for interoperability, extensibility and semantic data sharing
- Core Components: Standard ontologies, Mapping tools (LLM-enhanced), Reasoning layer
- Key stages:
  - ✓ Input Standardization → Extract schema
  - ✓ Ontology Mapping → Aligning with various frameworks
  - ✓ KG Refinement → Hierarchies + semantic relationships

ENCRYPT Knowledge Graphs

- In ENCRYPT, KGs linked with Data Privacy Vocabulary ontology
  - ✓ Describes and manages personal data processing activities
- KG is able to extracts & classify sensitive attributes
- Outcomes:
  - ✓ Seamless interoperability
  - ✓ Strong privacy compliance
  - ✓ Actionable insights across sectors



Knowledge graph representation of the MIRACUM Use Case, illustrating patient-medication relationships

ENCRYPT Risk Assessment Tool

- Why a Risk Assessment Tool is necessary
  - ✓ Privacy and cyber risks are intertwined in today’s digital systems
  - ✓ Users need structured support to identify and mitigate these risks
  - ✓ The tool enables this by offering a systematic, scenario-based approach

- Risk Assessment Tool Process
  - ✓ Define scenario & data assets
  - ✓ Assess cyber vulnerabilities & privacy threats
  - ✓ Generate mitigation-focused risk report

ENCRYPT Risk Assessment Tool

Empowering users to take a proactive stance on data protection.

MITRE ATT&CK

LINDDUN

Define the context
- Company domain
- Actors and roles

Identify business process
- PDP and processing role
- personal data type (KG)
- PPT in use

Assets and Vulnerability
- IT-OT assets mapping
- MITRE ATT&CK vulnerability assessment

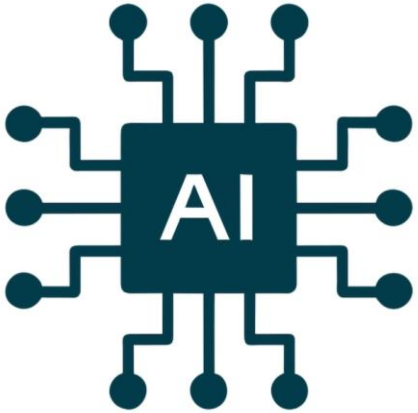
Privacy Threats
- Linking
- Identifying
- Non-Repudiation
- Detecting
- Data Disclousure
- Unawareness & Unintervenability
- Non-Compliance

Risk Evaluation
- Privacy and Cyber risks correlation
- Fill in the questionnaire based on ISO/IEC-27005:2022 to understand the maturity level of the PDP

- Users select system assets from a predefined taxonomy
- For each asset – Cyber and Privacy threats identified:
  - ✓ Cyber risks: evaluated using MITRE ATT&CK threat scenarios
  - ✓ Privacy threats: assessed via LINDDUN questionnaire
- Risk level depends on mitigation coverage and impact level

ENCRYPT AI Recommendation Engine

- Why a Recommendation Engine?
  - ✓Many privacy-preserving technologies (PPTs) exist in ENCRYPT
  - ✓Users struggle to choose the right one for their scenario
  - ✓**Solution:** AI-based Recommendation Engine to guide the selection



- How users interact with it
  - ✓Users rate on a 5 point scale a number of scenario traits:
    - Data sensitivity, Data size, Computational intensity, Computational constraints, Performance constraints
  - ✓Knowledge graph component also provides additional information

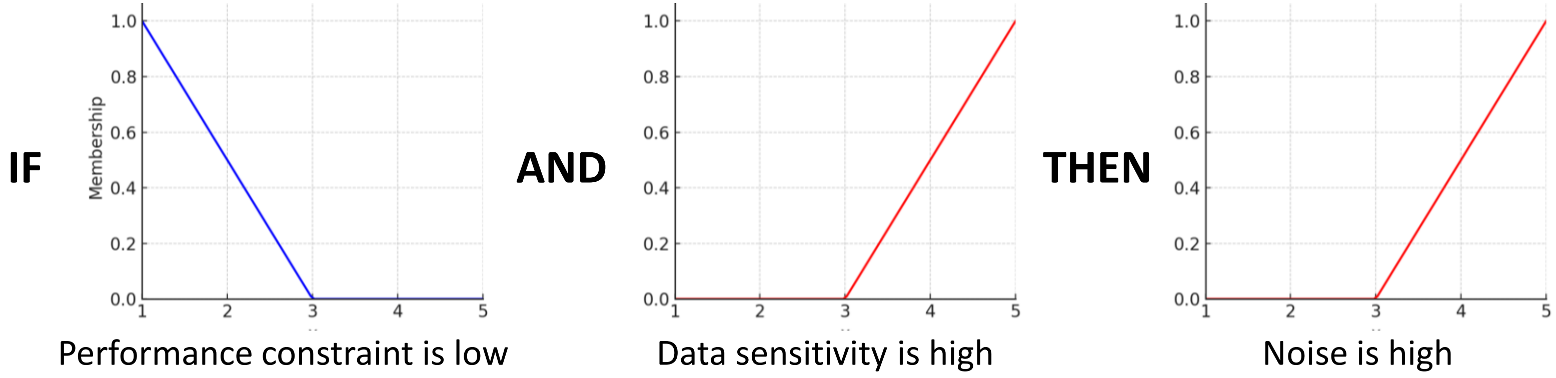
- **Output:** The most appropriate PPT to use and protect data

ENCRYPT AI Recommendation Engine

encrypt

ENCRYPT AI Recommendation Engine – How it works

- Fuzzy Logic is the used to select the most appropriate technology
  - ✓A way to encode human knowledge in natural language rules defined by experts
  - ✓Allows for ambiguous adjectives such as “good” or “low”
  - ✓Using the user inputs, each privacy preserving technology is given a score for how appropriate it is for the described scenario

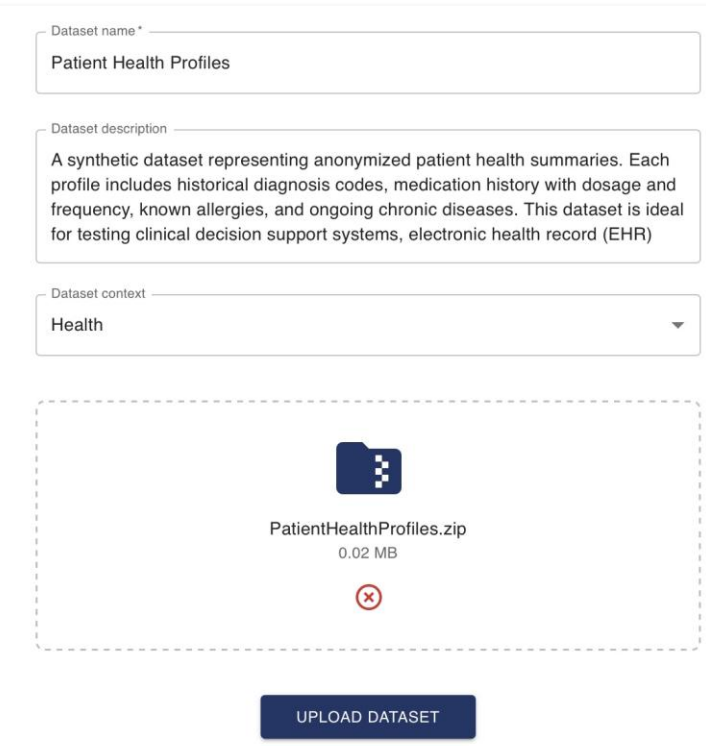


Example of fuzzy logic rule implementation

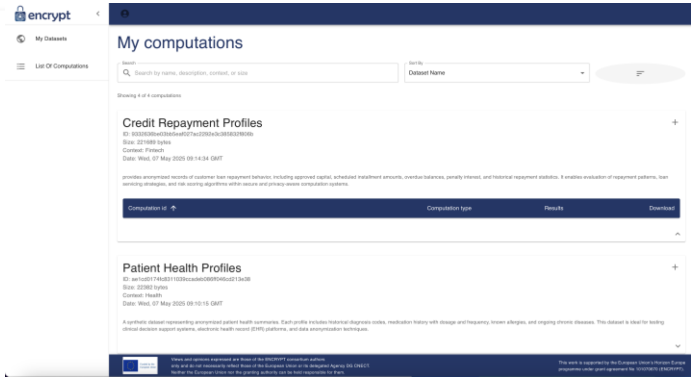
- The RE also **provides a justification** for the chosen solution
  - ✓Including an explanation of the technology itself
  - ✓Why it is most appropriate for the scenario

ENCRYPT User Interface

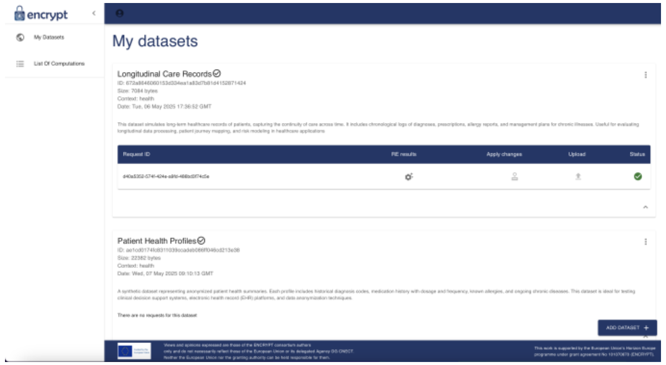
- ENCRYPT Platform includes an intuitive user interface
  - ✓ Allowing users to easily use features and ENCRYPT advanced technologies



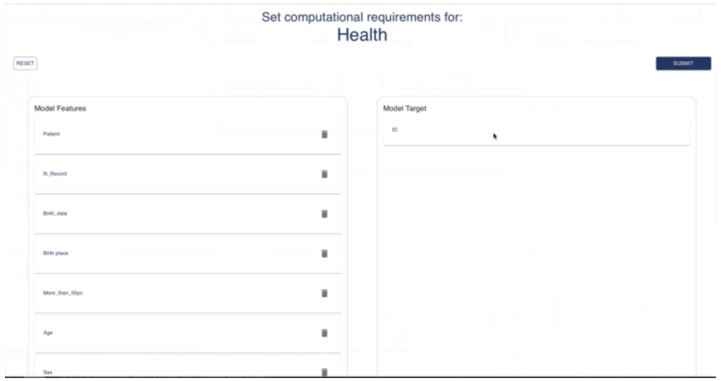
Upload Datasets



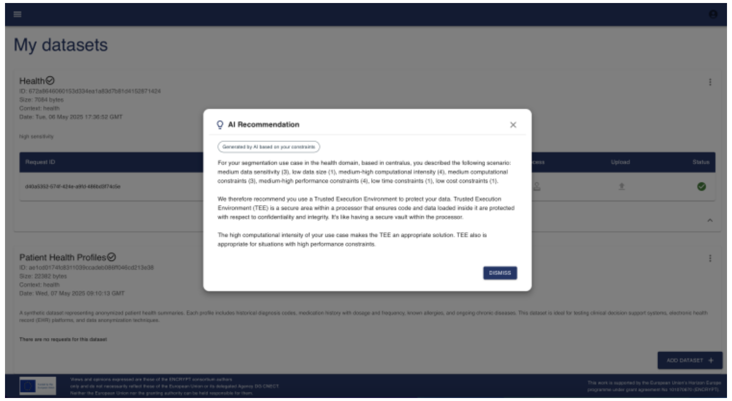
Select dataset for computation



Start data processing



Select features and targets



Get info about processing

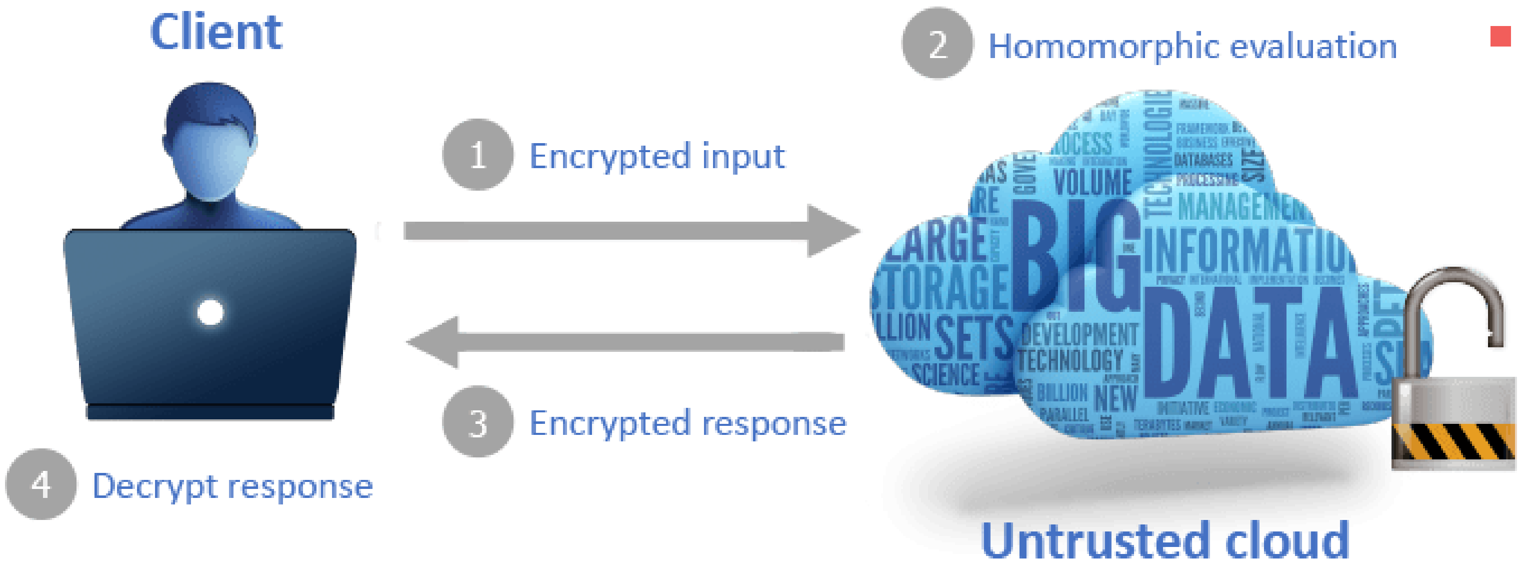
encrypt

Core Privacy-Preserving Technologies in ENCRYPT

- Fully Homomorphic Encryption (FHE)
  - ✓ Data Analysis on Encrypted Data
- Trusted Execution Environments (TEEs)
  - ✓ Ensuring Secure Data Processing with TEEs
- Differential Privacy (DP)
  - ✓ Sharing Data Without Sharing Secrets
- Hybrid Protection Services
  - ✓ Bringing together TEE and FHE
- Hardware Acceleration
  - ✓ Boosting Performance with Hardware Acceleration

Fully Homomorphic Encryption

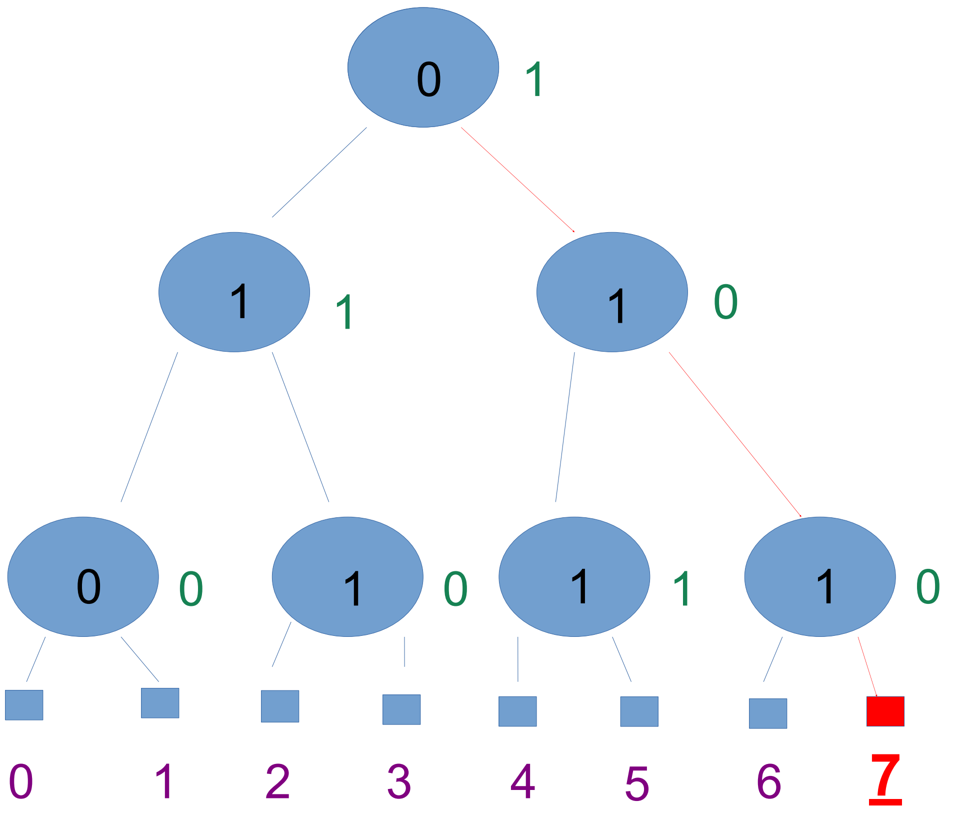
- What Is Fully Homomorphic Encryption (FHE)?
  - ✓FHE allows computations on encrypted data
  - ✓Keeps data secure during storage, sharing and processing
  - ✓Neither the data nor the result is exposed to the computing server



- Within ENCRYPT FHE has been used to:
  - ✓Finance: Evaluate decision trees on encrypted credit data
  - ✓Health: Query encrypted medical tables for flags
  - ✓Cybersecurity: Blindly search encrypted IP blacklists

Fully Homomorphic Encryption – Fintech Use Case

- FHE evaluation of binary decision trees from an already trained random forest model.
- Input:
  - ✓ An encrypted parameter per node (from a model owner)
  - ✓ An encrypted piece of data (from a user)
- Output:
  - ✓ The encrypted tag of some leaf



- Purpose: Know if a client/institution is eligible for a bank loan.
- Each bit comparison corresponds to a question
  - ✓ “Has the client a source of income greater than X euros per month?”
  - ✓ “Has the client any outstanding debt?”

Trusted Execution Environments

- What Is a Trusted Execution Environment?
  - ✓TEE→ Secure area in a processor that protects code and data
  - ✓Guarantees confidentiality and integrity, even if Operating System is compromised

- Key Features
  - ✓**Remote Attestation**: Verifies the integrity of TEE remotely
  - ✓**Confidentiality and Integrity**: Ensures secure data processing

- Technology Integration
  - ✓Integrated Gramine (a library OS) to run unmodified applications securely inside SGX enclaves
  - ✓Enabled container-based deployment using Gramine Shielded Containers, making secure execution accessible via Docker

Trusted Execution Environments within ENCRYPT

- Security Features
  - ✓Enforced strong isolation of data and code from the host system
  - ✓Enabled remote attestation to verify the integrity of the enclave before processing sensitive data
  - ✓Secured data at rest and in transit using sealed storage and enclave-terminated TLS connections

- Deployment & Usability
  - ✓Built an automated workflow for transforming standard container images into TEE-protected versions
  - ✓Offered a user-friendly interface for **selecting** between secure (TEE) or standard container deployment

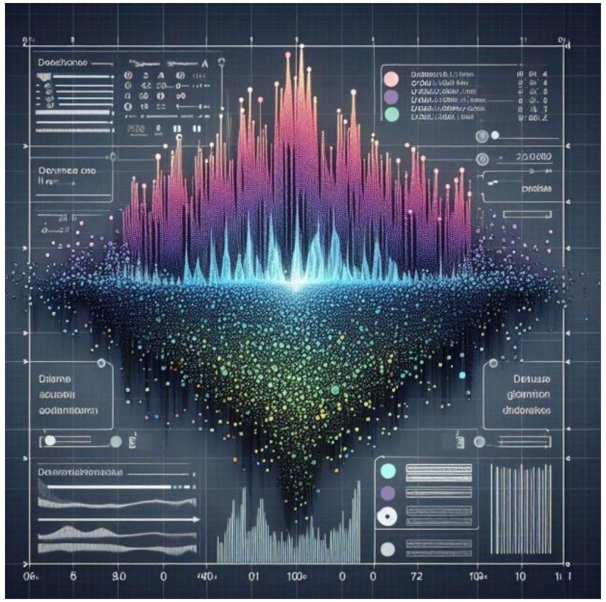
Differential Privacy

- Differential Privacy
  - Ensures data privacy by adding carefully calibrated noise to query responses
  - Prevents the identification of individual records (anonymity)
  - Still allows for accurate aggregate analysis
- Local Differential Privacy
  - Extends privacy guarantee by introducing noise at the source (individual data points)
  - Data contributors add noise (epsilon amount) to data before sharing it
  - Ensures privacy even when a central curator cannot be fully trusted
- Both prioritize privacy without sacrificing utility – useful for data analysis

Differential Privacy

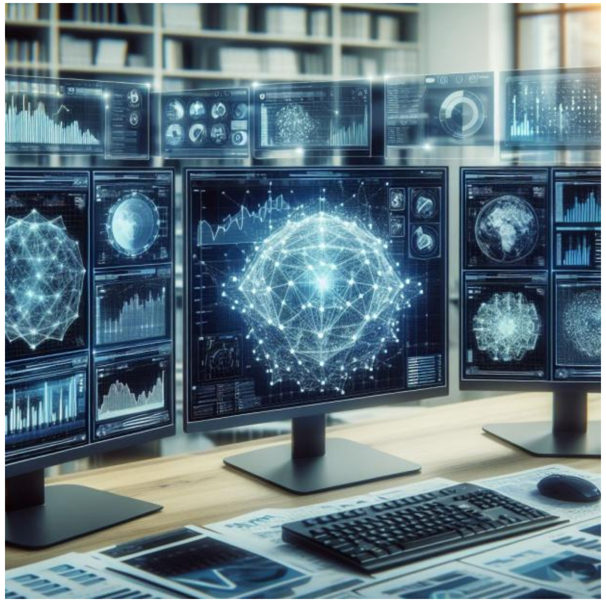
- Local Differential Privacy within ENCRYPT

DP suggested by Recommendation Engine with ε-value



Users will have to locally add noise to data **using the ENCYPT interface**

This noisy data will then be uploaded to the ENCRYPT platform



Machine Learning models can then be trained on the anonymized data

Differential Privacy model Training

- DP modules compares model training on a dataset
  - ✓ Comparison of model training with addition of noise (Differential Privacy case) and upon the original clean dataset
  - ✓ Models trained
    - Gaussian Naive Bayes Decision Tree Classifiers Random Forest Classifiers Neural Networks
  - ✓ Accuracy provided for both and in general, even though noise is added to a dataset, data analytics is still possible
- Most accurate models are serialized (pkl format) for future use
  - ✓ These models can be used for analytics on non-training datasets
- Code is general enough to work on any dataset
  - ✓ Just need to identify the features to use and target variable
  - ✓ The above is possible through the user interface

Hybrid Protection Services

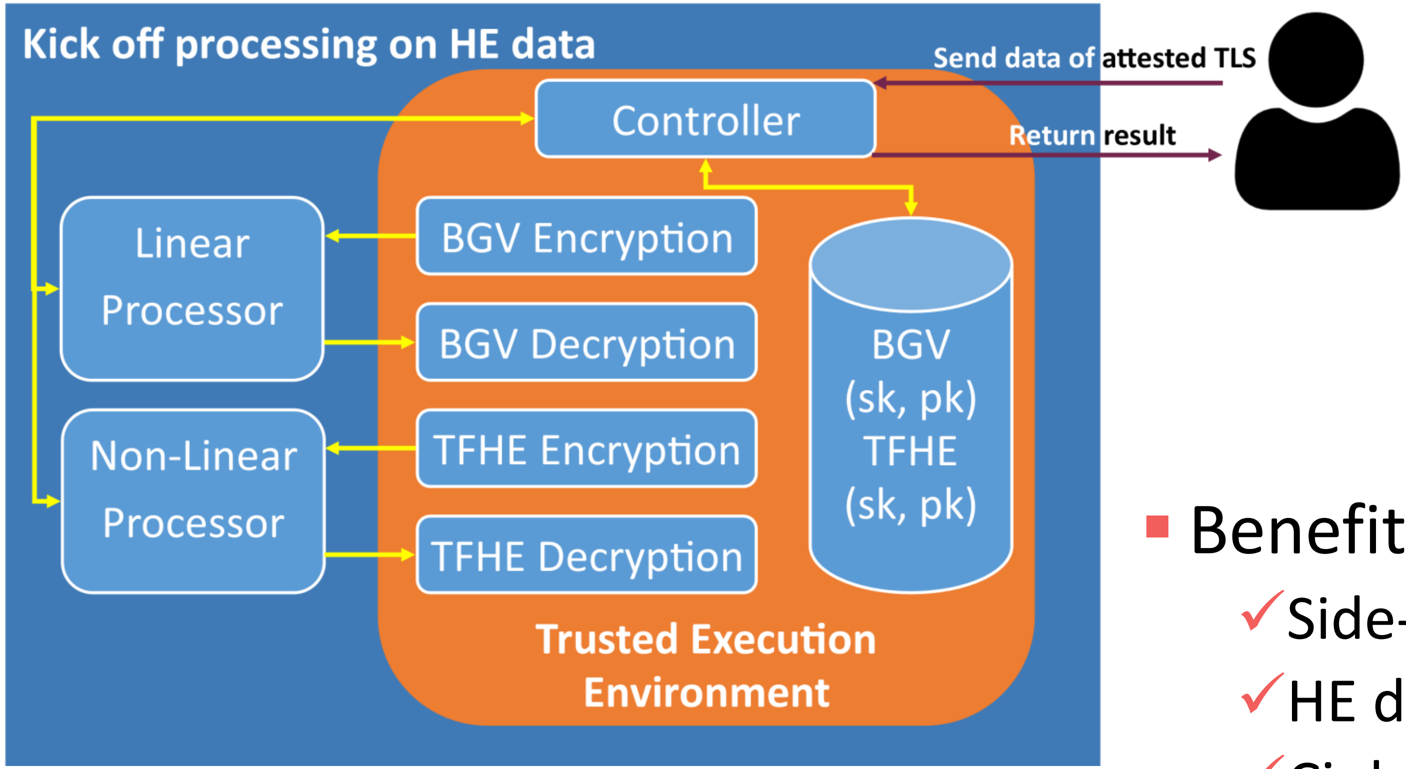
- Combining FHE and TEE for Practical, Secure Privacy-Preserving Computation

- Data in use remains vulnerable, even with encrypted-at-rest or -in-transit protections
- FHE allows computation on encrypted data, but suffers from:
  - ✓High computational overhead, Ciphertext Expansion, Unverifiable Conditionals
- Trusted Execution Environment executes fast, but:
  - ✓Is subject to side-channel attacks, Requires full trust in hardware vendor

- **Goal:** Combine FHE's privacy guarantees with TEE's performance in a hybrid architecture that reduces risks and improves efficiency

Hybrid Protection Services

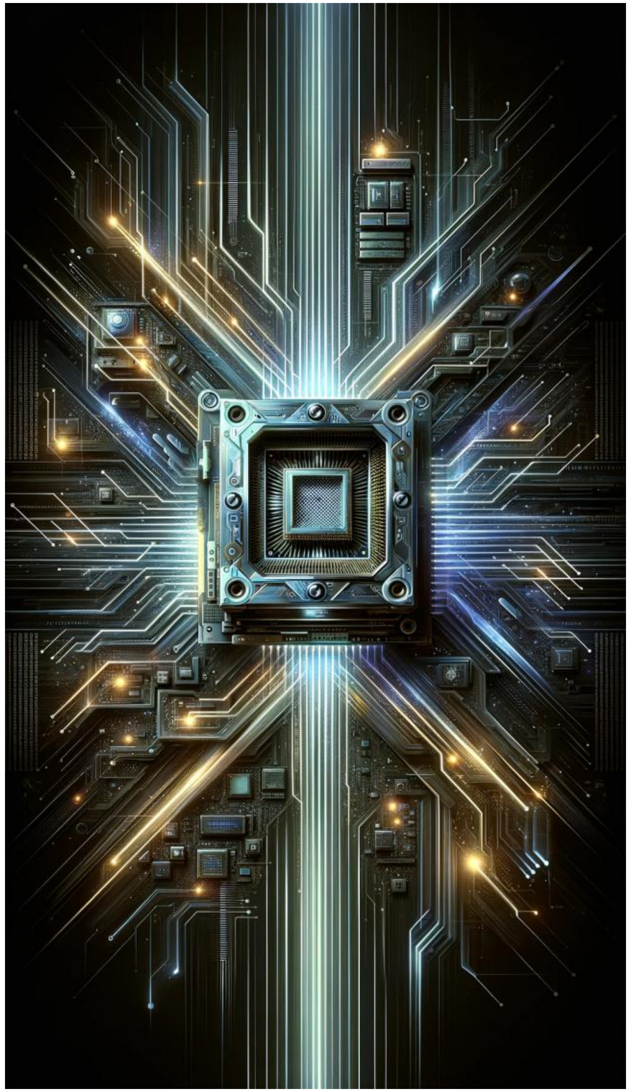
- **Solution**: SOTERIA - Hybrid Processing Framework for Secure, Practical FHE/TEE Integration



- Use of FHE for encrypted computation
- Use of TEE only for:
  - ✓ Conditional checks
  - ✓ Key generation/sealing
  - ✓ Crypto-scheme switching
- Benefits:
  - ✓ Side-channel window
  - ✓ HE decryption/encryption standardized
  - ✓ Ciphertext never revealed outside TEE

Hardware Acceleration

- Enhancing Computational Efficiency
  - ✓ Offloads intensive tasks to specialized hardware (eg, GPUs)
  - ✓ Reduces processing time for complex cryptographic operations
- Optimizing PPTs
  - ✓ Improves the performance of Privacy-Preserving Technologies
  - ✓ Makes advanced encryption methods more practical use
- Energy Efficiency
  - ✓ Decreases energy consumption during data processing
  - ✓ Supports sustainable and scalable privacy solutions
- Scalability
  - ✓ Enables processing of larger datasets without compromising speed
  - ✓ Critical for handling data in sectors like healthcare and finance
- Real-World Impact
  - ✓ Enhances the overall usability and adoption of the ENCRYPT
  - ✓ Facilitates faster and more secure data analysis

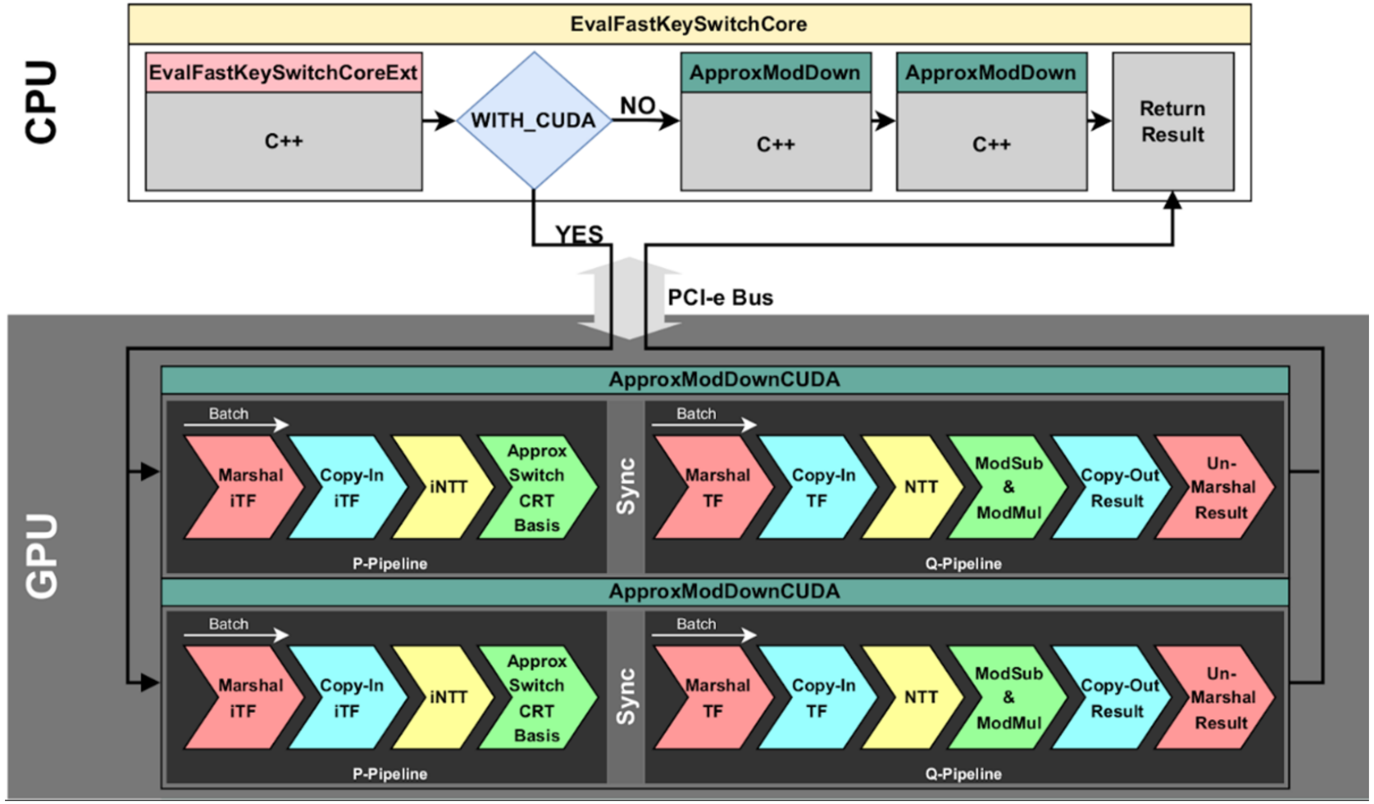


Hardware Acceleration

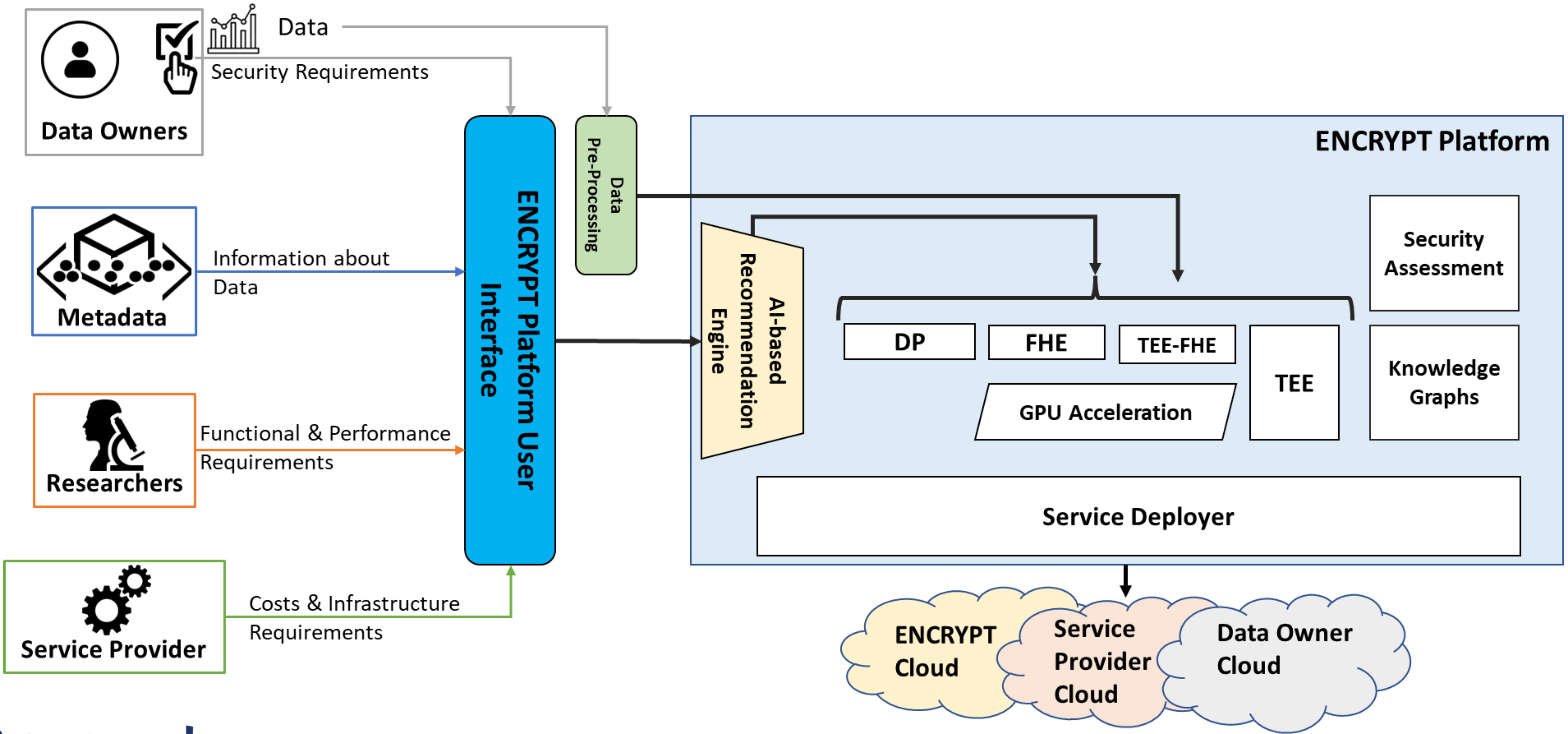
- Why GPU for FHE?
  - ✓ FHE is secure but computationally intensive
  - ✓ GPUs excel at parallelism → ideal for speeding up FHE
  - ✓ No existing GPU support for BGV in OpenFHE → we fill that gap
- What we Did?
  - ✓ Profiled OpenFHE → identified bottlenecks, Offloaded code to GPU
  - ✓ Enabled 128-bit integer support
  - ✓ Integrated seamlessly into OpenFHE
  - ✓ Reduced data transfer cost using batch processing, pipelining, CUDA streams

- Open Source Repository

Also funded by UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee for grant number 10039809.



ENCRYPT Technologies and ENCRYPT Platform



encrypt

Thank you!

Stay in touch

https://encrypt-project.eu/

encrypt-project

@encrypt_project