THE OPAQUE PLATFORM™

CONFIDENTIAL COMPUTING AND THE EXPLOSION OF CONFIDENTIAL DATA

Data teams across organizations globally are tasked with determining how to secure and process an increasing amount of confidential and sensitive data. It is estimated that today **confidential data that is locked down from access and use, is worth over $300 billion in value**. While the volume of confidential data grows exponentially, regulatory and compliance policies on accessing and using confidential data continue to get more stringent. For instance, violating regulations such as General Data Protection Regulation (GDPR) or the California Consumer Privacy Act (CCPA) or PIIA data privacy policies via a breach in accessor leaks may cost organizations as high as 4% of gross annual revenue.

Traditional approaches for protecting data and managing data privacy relied primarily on implementing strict access controls, access policies, and encrypting data at rest and in transit (i.e., while the data is in storage or traversing a network, respectively). But these approaches still leave data highly exposed and at risk when it is in use and undergoing processing by applications, (e.g., for analytics, and machine learning (ML) models). **Confidential Computing** – projected to be a $54B market by 2026 -- seeks to close this gap by protecting the data when in use using hardware hardware-based attested environments known as **Trusted Execution Environments** (TEE's). TEE's, also known as '**enclaves**' encrypt data while in memory and during computation, isolating data from access, exposure and threats. They are designed to specifically ensure that both code and data within enclaves is inaccessible to other users or processes that are collocated on the system.

However, **what's been lacking with TEE's are the software platforms that make it possible to run frictionless analytics and AI workloads within TEE's and further, enable secure, collaborative analytics and AI on confidential data**. This is needed as otherwise during analytical processing and AI/ML execution the data is decrypted in memory leaving it exposed to hackers and unauthorized access. Without the necessary AI and analytics software platforms, data scientists are left to struggle with finding secure ways to run AI and analytics on data encrypted in enclaves, share data between authorized users and teams and perform multi-party analytics on confidential data. This has significantly hampered organization's ability to effectively address **critical business cases** across industries -- such as fraud detection, identifying financial crime, healthcare analytics, marketing attribution and more in a timely manner.

THE CHALLENGE

Confidential Computing and trusted execution environments protect stored data and data in transit from internal and external threats, however maintaining security while data is in use has been a challenge.

A DEEPER LOOK AT THE CHALLENGES

There are numerous issues data and analytics teams face when analyzing data secured by TEE's. The biggest challenges span four areas which include secure analytics; governed data sharing; collaborative multi-party AI and analytics and scalable distributed data processing across teams and organizations; and ease-of-use and accessibility.

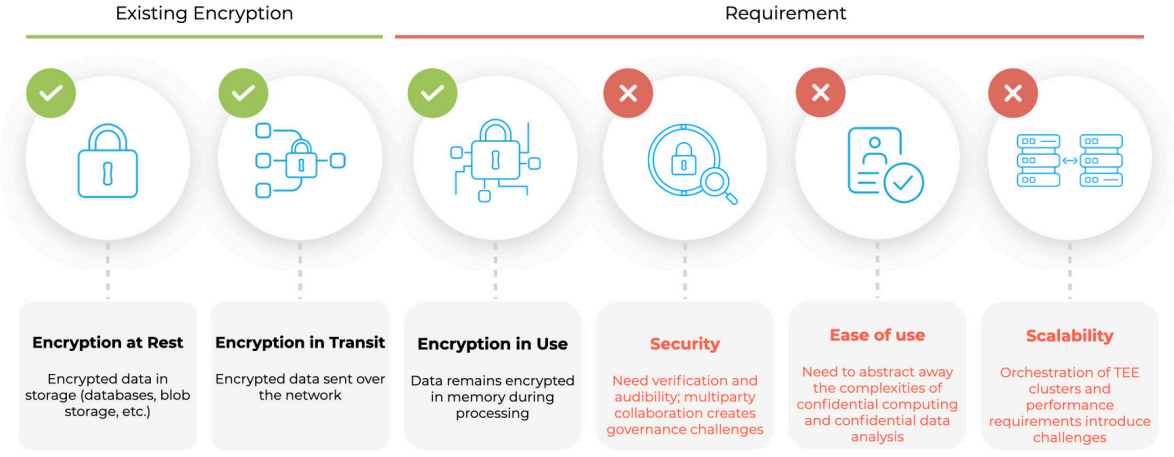
1. Security - In addition to processing data for the purpose of AI and analytics within TEE's, it is also necessary to ensure verifiability of the code and computation within the TEE's and enforce governance mechanisms for data sharing and multi-party analytics and machine learning. This has posed challenges for data scientists and analytics teams tasked with running analytics use cases.

2. Data Sharing and Data Privacy - Sharing data protected in enclave environments between multiple parties while adhering to data privacy policies is non-trivial and has been difficult for analytics teams. However, enabling secure, governed data sharing between multiple parties internal and external to the organization is necessary to unlock critical confidential computing business cases.

3. Performing Collaborative Analytics and Machine Learning on Encrypted Data - Performing analytics and machine learning directly on encrypted data has been highly difficult and collaborative analytics on encrypted data within TEE's has been impossible until now. Enclaves protect data at rest, in transit, and during processing but new analytics platforms are needed to leverage the core properties of enclaves and enable *collaborative analytics and machine learning at scale*.

4. Ease of Use - Enabling data scientists and analytics teams to leverage existing skillsets and work with familiar frameworks within these new platforms is critical

The Gap: Enabling Analytical Processing, Verifiability, and Multi-party Analytics and AI *on Encrypted Data in Enclaves*



to accelerating use cases. Further, for ease of use and frictionless adoption it is also necessary to Integrate and co-exist with an organization's existing analytics and AI workflows.

Confidential AI and Analytics Platforms have emerged to address these four major challenges and as a result of these innovations, organizations globally are able to now realize the full potential of confidential computing.

CONFIDENTIAL COMPUTING AND CONFIDENTIAL AI PLATFORMS

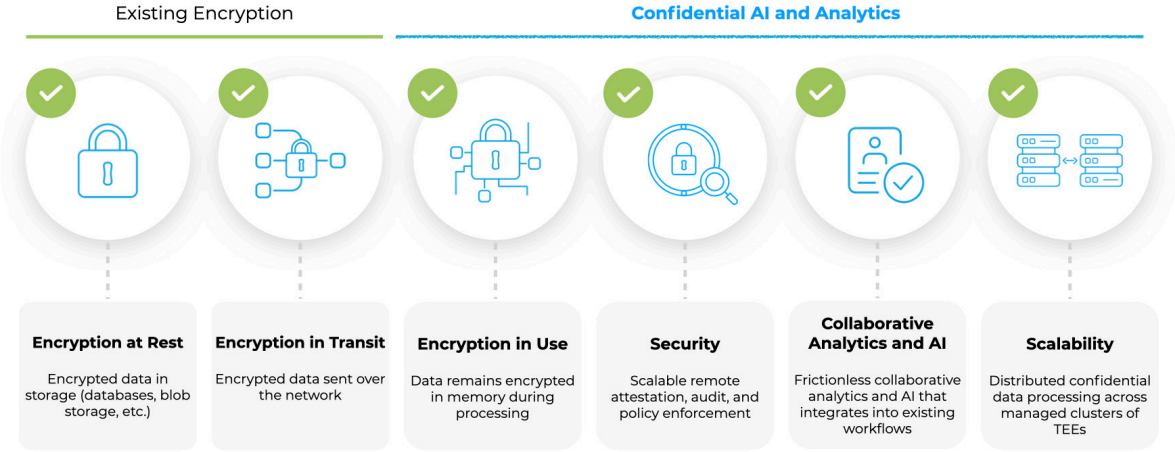
The optimal way to perform analytics and machine learning on sensitive data is to enable multi-party analytics on encrypted data so all confidential data is protected end-to-end; at rest, in transit, and encrypted during analytics and processing. The massive organizational **need to protect data throughout the lifecycle** has led to the emergence of new **Confidential AI and Analytics Platforms** that enable data analysts and machine learning practitioners to securely analyze data – without ever having to expose the data unencrypted during processing. The **Opaque Platform** is the leading Confidential AI Platforms in confidential computing that enables secure **multi-party collaborative analytics and AI on encrypted data.**

THE OPAQUE PLATFORM AND ADVANTAGES

Opaque's Confidential AI and Analytics Platform makes confidential data usable by enabling secure and scalable analytics and machine learning directly on encrypted data within enclaves.

The Opaque Platform: Confidential AI and Analytics

Protecting Data Throughout the Lifecycle and Enabling Analytics and AI on Encrypted Data within Enclave Environments



Organizations can use Opaque to encrypt their confidential data on-premises, accelerate the transition of sensitive workloads to enclaves in Confidential Computing cloud environments, and then analyze the encrypted data while ensuring it is never exposed unencrypted during the lifecycle of the computation. **Multiple data teams within or across organizations can collaborate** in Opaque's environment and **perform collaborative analytics or machine learning** on their collective data, while ensuring that each party is only privy to the data and insights that they are authorized to see. To prevent any malicious use of the data or access, Opaque's Platform ensures all data remains encrypted end-to-end throughout the analytics lifecycle.

Further, Opaque enables analytics on encrypted data within enclaves via a new innovative technology – pioneered by the Opaque founding team. The technology provides SQL-compatibility which means analytics teams everywhere, familiar with SQL and notebooks, can work with confidential data and accelerate analytical outcomes.

With Opaque's Confidential AI and Analytics Platform and unique collaboration capabilities, data owners retain full control over how their data is processed. By combining encrypted data that is distributed across silos and **performing collaborative analytics or training machine learning models on the joint datasets**, data teams within organizations and across organizations can obtain insights that would not otherwise be possible. Since the data remains protected even when in use, Opaque also makes it easier for data owners to comply with privacy laws and regulations.

The Opaque Platform is powered by cutting-edge technology based on years of academic research from UC Berkeley. It builds upon the open source MC2 Platform, which leverages a **novel combination of two key technologies—secure hardware enclaves and cryptographic fortification**. This novel combination of hardware and software ensures that Opaque's Platform is secure, fast, and scalable.

KEY CAPABILITIES AND BENEFITS

Hardware Enclaves with Cryptographic Fortification - In order to provide strong privacy guarantees for user data, Opaque leverages hardware enclaves. Any data or software placed within the enclave is encrypted and isolated from the rest of the system. No other process on the same processor—not even privileged software such as the OS or the hypervisor—can access the encrypted enclave memory. As a result, enclaves provide very strong protection to the data when it is being processed, protecting it from all other software on the system, hackers, as well as server administrators. The Opaque Platform utilizes enclaves for its advanced functionality and efficiency, but in addition fortifies them with cryptography to provide even stronger security guarantees. Opaque fortifies enclave execution in two ways: it ensures the integrity of distributed computation and uses data-oblivious techniques to strengthen enclaves against side-channel attacks. An attacker has to compromise both of these security layers to get access to sensitive data in Opaque; it does not suffice to subvert only one of these layers.

Data Protection Throughout the Lifecycle - Opaque protects all sensitive data (e.g., PII and SHI data) using advanced encryption as well as secure hardware enclave technology, throughout the lifecycle of the analytics computation—from data upload to applying the analytics models to obtaining the results. Opaque's technology ensures that the entire dataset is available for analysis and there is no loss in the quality of the analysis due to unavailability of data, while simultaneously remaining strongly protected throughout the entire data lifecycle.

The Opaque Platform protects the confidentiality of each customer's data and the integrity of computations from attackers at the cloud (e.g., cloud employees, hackers) and from other customer organizations. Opaque's approach to security is based on the "defense in depth" principle.

Secure Computation in the Cloud - Businesses today are increasingly migrating confidential data to the cloud and storing the data in cloud enclaves and Trusted Execution Environments. Opaque's Platform uniquely enables organizations to migrate their confidential data and confidential analytics to the cloud and apply advanced analytics models and machine learning directly to the encrypted data, keeping it protected throughout the lifecycle of the computation.

The Opaque Platform ensures that no data is ever exposed and data in use and in transit always remains encrypted. Additionally, Opaque's Platform also guarantees data privacy, trust, and compliance during data access, data processing and analytics.

Secure Multi-Party Collaboration and Data Sharing - Beyond deriving insights from their own data, data teams can benefit from combining their data together and analyzing it jointly to obtain mutually beneficial insights. These teams may belong to different departments/branches within the same organization, or different organizations that are unable to otherwise share their data with each other. For example, such data sharing can enable financial institutions to identify fraud and combat financial crime more effectively, or healthcare institutions to perform better medical studies. However, organizations are often unable or unwilling to share their data with each other owing to privacy concerns, regulatory hurdles, or business competition.

The Opaque Platform allows multiple data owners to pool their encrypted data together in the cloud, and jointly analyze the collective data. The solution ensures that the data of individual owners is never exposed to either the cloud environment or to other data owners. Data owners individually upload encrypted data to the cloud, and then jointly apply analytics models directly on the encrypted data. Each data owner retains full control over how their data is used.

Secure Collaborative Analytics - The Opaque Platform enables data scientists to perform deep analytics on encrypted data—for instance, using the popular Spark SQL framework. For ease of use, Opaque preserves the Scala, SQL, and Python APIs provided by Spark—data scientists who know SQL or have previously worked with Spark through Scala or PySpark, already know how to use Opaque. Data scientists can use Opaque to do much of what one can do with Spark SQL: run rich SQL-based analytics on the data; perform statistical analysis; or manipulate data using projec-

tions, filters, joins, sorts, and aggregations.

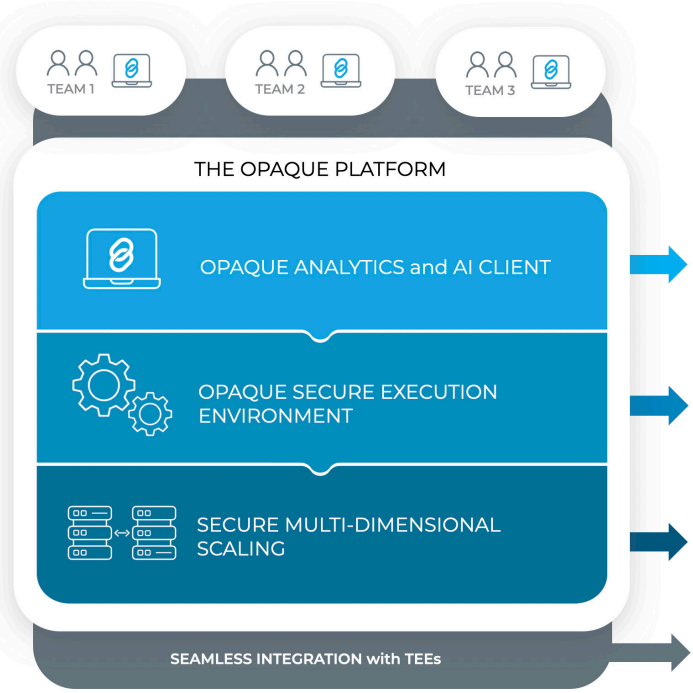
Secure Collaborative Machine Learning - Opaque's machine learning solution allows data scientists to train machine learning models efficiently and securely and/or serve predictions on confidential data. The Opaque Platform supports classical machine learning models such as linear and logistic regression, as well as advanced models such as gradient boosted decision trees (such as the popular XGBoost framework).

Scalability and Orchestration of Enclave Clusters - Orchestration of TEE clusters and meeting performance requirements for scalable data processing can be challenging. Opaque provides distributed confidential data processing across managed clusters of TEE's and automates orchestration and cluster management. Opaque also ensures secure inter-enclave communication. The platform additionally provides monitoring and simplifies management and across enclave clusters.

Compliance and Policy Enforcement for Data Sharing, Analytics, and Insights - Compliance with privacy laws and regulations is often accompanied by a decline in the quality of the analysis that can be performed on data. For instance, certain data fields (such as PII data) cannot be shared or revealed due to their sensitive nature, leading to lesser data available for analysis and poorer insights. While one could attempt to tackle this problem using techniques like anonymization, they are difficult to apply correctly, destroy useful information, and are also prone to re-identification. The Opaque Platform provides policy enforcement capabilities for inter- and intra-company analytics and AI, policy enforcement for multi-party data sharing and sharing of insights.

State-of-the-art Cloud Security - State-of-the-art cloud security consists of a concerted set of techniques including aspects such as employee training and verification, operational security (vulnerability management, malware prevention, monitoring, incident response), state-of-the-art access control (e.g., firewalls), physical security of data centers (e.g., cameras, alarms, authentication), encryption in transit and at rest, audits and third-party certifications, regulatory compliance and others. Opaque maintains the security of these advanced clouds and further provides the second layer of security based on hardware enclaves and cryptographic fortification, as described above. Opaque thus makes it exceptionally difficult for attackers to attempt to subvert both security layers of Opaque. As a result, the Opaque Platform provides a very high degree of security that is much stronger than the traditional security of even prominent clouds today.

Opaque Confidential AI and Analytics Platform



- Run analytics and ML on encrypted data
- Use familiar tools, Spark SQL and Notebooks
- Eliminate custom coded queries, analytics & ML workflows
- Set sharing and governance policies

- Executes confidential collaborate analytics and AI/ML
- Governs data sharing
- Delivers insights specific to each party

- Enables secure inter-enclave communication
- Automates orchestration
- Secures multi-cloud operations within and across clouds

- Seamlessly integrate with enclaves and confidential VM's

Opaque's technology is widely applicable and can help address a wide range of use cases across a diverse set of industries.
For more information on use cases visit: *https://opaque.co/solutions/*