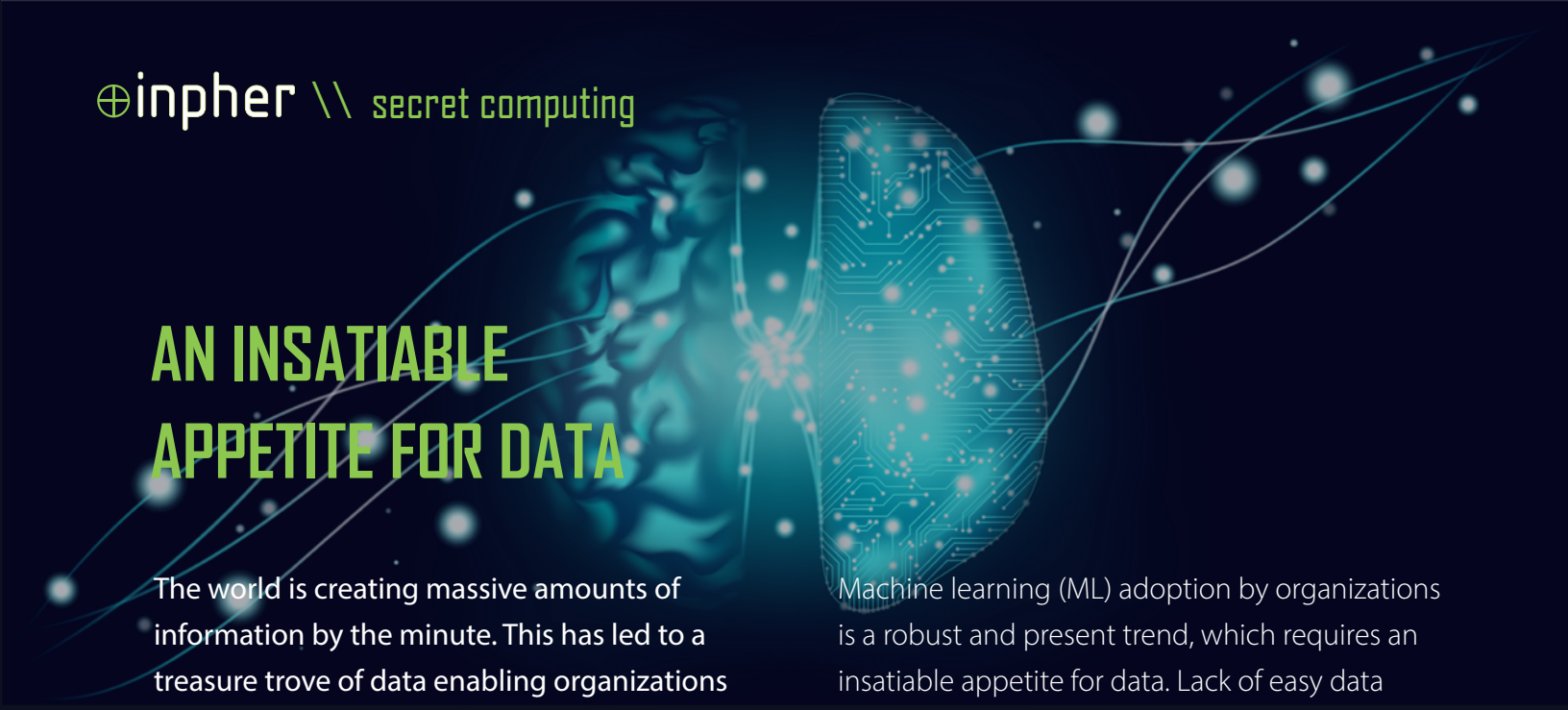
WHITE PAPER

inpher

Privacy & Precision: Data Analysis & Machine Learning with Secret Computing®



AN INSATIABLE APPETITE FOR DATA

The world is creating massive amounts of information by the minute. This has led to a treasure trove of data enabling organizations to derive significant business value from hidden insights or even launch a new business model by monetizing their data.

As part of this effort, businesses are exploring various ways of data sharing to support advanced analytics and machine learning both within and across the organizations. However, sufficient and compliant access to private data inputs are holding back acceleration. Organizational compliance privacy zones, compliance mandates like GDPR, and other data-sharing roadblocks make it near impossible for data scientists to analyze all of the data they want to make better, more precise models. These data privacy challenges get in the way of true data usability.

Current privacy policies also make it difficult for data scientists to work on the data, making privacy and ease of use two different, but equally difficult problems to contend with for organizations who want to share data.

Machine learning (ML) adoption by organizations is a robust and present trend, which requires an insatiable appetite for data. Lack of easy data access and current anonymization schemes hampers organizations' abilities to do so.

Without data usability, business leaders miss key insights and opportunities to innovate by foregoing true access to all relevant data. Data scientists, for example, may not be able to easily build an accurate drug discovery models and develop cures for diseases due to the confidentiality of patient data. Retail chains may not have a true understanding of current consumer demands and trends if buyer data is confidential and inaccessible.

In this white paper, we explore the challenges of data sharing and usability, the technology that can surmount these issues, and current and future applications with **Secret Computing®**, a compute engine built by Inpher for privacy-preserving analytics and AI applications.

TODAY'S DATA SHARING CHALLENGES FROM A DATA SCIENCE PERSPECTIVE

In business today, there are many stakeholders who want to be able to share private data for analytics, ML model training, and inference.

An organization may want to share data with a third party to take advantage of capabilities not available inside the data owner's organization. Or they may want to access complementary data for enhanced analytics and insights. But there are many roadblocks to why data cannot be shared and transferred with another party; there might be compliance issues and legal requirements, or there may be trust concerns.

One such challenge to data sharing is the abundance of privacy and data regulations that must be complied with, including the European Union's General Data Protection Regulation (GDPR), the California Consumer Privacy Act (CCPA), or the Health Insurance Portability and Accountability Act (HIPAA) to name a few. All have specific requirements that must be strictly adhered to or else organizations risk serious penalties, including large fines.

For example, in 2019 Facebook was given a **$5 billion fine** by the Federal Trade Commission (FTC) for allegedly violating privacy practices and mishandling user data in connection with the Cambridge Analytica incident (but more on that later).

Regulations like HIPPA, for example, make it challenging for healthcare entities to share private medical records with researchers due to its privacy rules. GDPR has strict regulations pertaining to the cross-border transfer of personal data from an EU country to a non-EU country.

Both of these requirements make certain types of data-sharing impossible, or so expensive, complex, and time-consuming that the business case for doing so is weakened.

Another challenge when it comes to data sharing is protecting an organization's intellectual property. IP can include many sensitive elements that cannot be exposed for competitive reasons, including trade secrets, patents, processes, employee information, or plans for product launches. Companies also run the risk of any exposing competitive knowledge that could be misused by third parties when they share data with outside parties.

In addition, organizations brush up against ethical considerations when it comes to data sharing because, in many instances, the data owner, or customer, has not consented to the data being shared or used in certain circumstances.

The case of Cambridge Analytica is an example of this. In that incident, the personal data of 50 million Facebook users was sold to the company for reasons other than what users agreed to when they gave up their data.

The information was not used for the academic purposes, which is what some users had consented to, but to craft detailed psychographic profiles to target audiences with digital advertisements, mostly political in nature.

THE TECHNOLOGY OF SECRET COMPUTING

How can an organization enjoy both data usability and data privacy without violating rules, risking private data or violating customer trust?

The answer to secure, compliant data sharing may seem simple: encrypt it. But it is not that easy. Typically, encryption techniques have taken one of two forms: **encryption at-rest** and **encryption in-transit**.

Encryption at-rest means encrypting data that does not move. Encryption in-transit means encrypting data that moves through a network. But the two forms of encryption methods do not encrypt data while it is being processed, therefore data is still vulnerable if it is not encrypted during processing.

Addressing the challenge of data encryption in use is where Secret Computing comes in. There is a momentum building for Secret Computing – which allows data scientists to easily, conveniently, compliantly, securely, and privately compute on distributed data without ever exposing or moving it.

Privacy Preserving Techniques and Methodologies

There are several privacy-preserving techniques and methodologies that have the potential to change the dynamics of data-sharing by eliminating or reducing the risks of data sharing.
They include:

+ Differential Privacy
+ Federated Analysis
+ Zero Knowledge Proof
+ Homomorphic Encryption
+ Secure Multiparty Computation (SMPC)

Of these techniques, only SMPC with FHE, the foundation of Secret Computing, can keep data encrypted while it is being processed by leveraging these two complementary encryption in-use techniques.

Secret Computing technology allows for encryption in-use with two complementary encryption techniques: **Secure multiparty computation (MPC/SMPC) and fully homomorphic encryption (FHE).**

SECURE MULTIPARTY COMPUTATION (MPC/SMPC)

Secure multiparty computation (MPC/SMPC) is a cryptographic protocol that distributes a computation across multiple parties where no individual party can see the other parties' data.

Secure multiparty computation protocols can enable data scientists and analysts to compliantly, securely, and privately compute on distributed data without ever exposing or moving it.

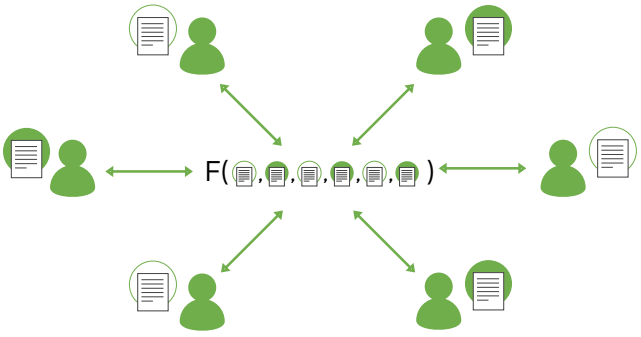


Figure 1. [SMC relies on "secret sharing", where sensitive data from each contributor is distributed across every other contributor as encrypted "shares." These shares, if intercepted by a malicious third party or misused by an individual contributor, would be worthless, since they are decipherable only when combined with the information distributed across many other parties. Through analysis and calculations, sensitive data is not shared between parties, but the correct end result can still be derived.]

FULLY HOMOMORPHIC ENCRYPTION (FHE)

Fully homomorphic encryption (FHE) is an encryption scheme that enables analytical functions to be run directly on encrypted data while yielding the same encrypted results as if the functions were run on plaintext.

Research firm **Gartner says**, "homomorphic encryption enables businesses to share data without compromising privacy."

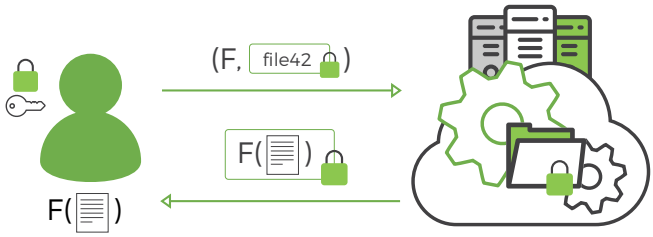


Figure 2. [A customer homomorphically encrypts its records and sends them to the researcher's cloud computing environment. Because the data is encrypted, it is fully protected and private in the cloud. Next, the researcher runs its analytical functions on the homomorphically-encrypted data in the cloud, manipulating the data while it remains encrypted. Last, the researcher downloads the encrypted output, and decrypts the result to reveal the plaintext answer. Notice that the sensitive record data is encrypted end-to-end, and is only decrypted when revealing the final answer behind organizational firewalls.]

With these two capabilities, cryptographic Secret Computing technology allows for secure, privacy-preserving analytics and machine learning that can transform a business data strategy to one that will unleash the insights that true data usability can deliver.

SECRET COMPUTING'S POTENTIAL FOR NOW - AND THE FUTURE

Secret Computing is not a futuristic concept.
It is available for use now and it enables data scientists to unlock sensitive data for their machine learning and analytical models while meeting their organization's privacy, security, and compliance requirements.

Secret Computing is currently allowing leading global financial services, technology, and manufacturing companies to enjoy true data usability benefits now, including:

+ Secure sharing of data-driven insights
+ Privacy-compliant machine learning models built on distributed data sources
+ Migration to zero-knowledge cloud computing
+ Monetization of insights without giving away the data
+ Secure data sharing in inter-organizational partnerships

Regulators See Promise for Secret Computing

Regulators are always seeking solutions to problems that have historically challenged highly regulated industries.

For example, according to a **2018 report by Refinitiv**, financial institutions lose an estimated $1.45 trillion annually from financial crimes while spending $1.28 trillion to prevent them in the first place. Regulators see great promise for Secret Computing to address these issues.

In a **recent speech**, the Head of Financial Conduct Authority (FCA)'s Financial Crime Department, Rob Gruppetta, asserted that "data analytics and machine learning are widely seen as the approaches with the greatest potential to improve current practices."

To understand the current-day applications of Secret Computing, we can examine the following case studies:

1. Secret Computing enables better fraud modeling in financial services

Fraud is always an issue for financial institutions, and a large multinational financial services firm wanted to improve their fraud detection and risk models by accessing more data.

But fraud data sharing initiatives are difficult to implement for several reasons, including jurisdictional restrictions, data confidentiality concerns, and trade secret sensitivities. Secret computing enables the financial services firm to train their fraud and risk models on distributed data from participating partner banks while maintaining all parties' security and privacy requirements.

Double blind machine learning for fraud detection at BNY has been colloquially known as "Sinking Pirates," according to BNY's Data Science Lead, Hays W. "Skip" McCormick.

> Now that cloud computing has provided the elastic scale required to use Data Science as our next strategic tool for fighting fraud, the next big hurdle is assembling large-enough, current-enough training data sets. Investing in obtaining more training data is a lot more effective than work to improve models. Yet, we will continue do both because the cost of financial fraud robs us all of resources better used elsewhere while funding criminal activities that threaten civilization.
>
> The 'bad guys' have all the same technologies we do. But the one thing they cannot obtain is the scale of training data we can through collaborative sharing, such as through what Inpher offers with Secret Computing. Training models to prevent and counter fraud using data that we cannot or would prefer not to otherwise share offers an excellent advantage for beating the criminals and keeping them beaten.
>
> – **Hays W. "Skip" McCormick**, BNY's Data Science Lead

2. Secret Computing leads to better detection of heart disease in healthcare

In this next scenario, a large healthcare provider seeks to facilitate early prognosis of cardiovascular diseases through the use of private data. But privacy laws such as HIPAA and confidentiality concerns prevent contributors from sharing data outside of organizational firewalls.

With Secret Computing technology, researchers can privately compute across organizational data sources in order to increase both sample size and patient attributes, leading to improved model performance and disease prognosis.

Inpher is also enabling privacy-preserving ML in **Moore4Medical**, an EU project pioneered by Philips and top European healthcare providers and hospitals.

With Secret Computing, Philips and M4M consortia will collaborate with partner hospitals to build a distributed ML platform across their infrastructure, hospitals and sleep monitoring devices to enable early detection of heart diseases.

3. Secret Computing allows for more accurate forecasting of employee attrition

Employee attrition and recruiting is always a costly investment for organizations. That's why best-in-class corporations strive to retain productive and valuable employees. Proactively identifying at-risk and high performers enables them to decide how to invest in opportunities to keep them with the company.

But important predictive information indicative of attrition is highly sensitive and distributed departmentally across the business, and often exists in siloed systems.

Secret computing enables the company in this scenario to privately compute across private HR data repositories. Departmental and attribute level privacy is preserved, and this allows for beneficial, compliant analytics to be run as required. The company can now better predict employee attrition by training an attrition and retention model across siloed human resources data sources.

4. Secret Computing leads to predictive maintenance for distributed fleets

A large defense contractor needed to monitor the performance of its fleets in order to improve usage analysis. But, due to the sensitive nature of the data, was unable to obtain certain critical information to aid the analysis. Overall maintenance performance suffered due to the lack of key information.

Secret Computing enabled the defense contractor to privately compute distributed data across fleets in order to improve maintenance models. The result is improved uptime, reduced operational costs, and more operational value for fleet managers.

SECRET COMPUTING SOLVES THE MOST COMMON DATA SHARING BARRIERS

Restrictions on data transfers are no longer an issue with Secret Computing technology because it enables knowledge sharing without the need for personal data transfers. There is no disclosure or sharing of personal data in the computing process. Data residency and data sovereignty are maintained. Cybersecurity concerns are also addressed because the technology protects against data breaches Secret Computing offers a quantifiable surety against unauthorized data access and actually prevents the harm of a breach instead of simply remedying them after they occur.

New paths to collaboration are realized through Secret Computing as multimodal and interoperable AI training allows secure collaboration for greater coordination and scalability. Multiple parties can access data while also ensured of data privacy, utility, and accuracy.

Find out how Secret Computing drives business value through more accurate analytics and ML models.

Access a deeper and broader pool of sensitive information distributed across restricted data silos with XOR Secret Computing Engine.

Visit Inpher today and learn more.

inpher


\\ WHAT IF
YOU COULD ACCESS MORE DATA?
WHAT IF
YOU COULD MAKE BETTER MODELS?
WHAT IF
YOU COULD DO IT SECURELY & PRIVATELY?
IT'S NOT MAGIC
IT'S INPHER SECRET COMPUTING
inpher
Secret Computing©
inpher.io
info@inpher.io
@inpher_io
inpher-inc
36 West 25th St., Suite 300
New York, NY 10010
EPFL Innovation Park Bâtiment A
1015 Lausanne, Switzerland