SECURE DATA SHARING

Helios Data Inc.

SECURE DATA SHARING™

Helios Data's Secure Data Sharing™ solution is an enabler for companies to quickly proceed down the path to data monetization. Traditional companies, especially those rich in consumer data "reserves", such as telecommunication service providers, banks, retailers, transportation companies, and insurers are desperately seeking to embrace the new data-driven business paradigm, in order to monetize their own enormous, petabyte-magnitude, stores of consumer data.

GUIDING PRINCIPLES

Helios believes the following guiding principles are necessary for any secure data sharing initiative or solution. We strive to adhere to each of these from the beginnings of our solution concepts, through our development lifecycle, in our discussions with customers, and in the delivery of our products.

- Privacy by Design
- Assumption of Minimum Trust
- Minimum Loss
- Maximum Isolation
- Asymmetric Transparency.

With a privacy-by-design as a cornerstone, every step along the way should consider the impact of privacy for that specific step and the interactions of that step across the platform development. By having an assumption of minimum trust, one can be more assured that the sharing of knowledge and communications will have an appropriate mechanism to achieve the highest level of trust. As one develops a multi-party secure data sharing solution, one should always look to minimize the inadvertent loss of sensitive data representative of each party. We view the loss of data as not just the traditional leakage or hack of information, but also the unauthorized disclosure or transfer of one party's data or analytic product to their respective contractual partners. We recognize that to achieve the three prior elements requires the guiding principle of maximum isolation for the data elements, methodologies, and coordinated sharing of each party's IP ownership and operational trade secrets. Both parties recognize that their respective assets will temporarily co-mingle to create an entirely new clean data asset, which will bring new data monetization value to each of them. Finally, the combination of these cornerstones in our guiding principles can be summed up in the principle of Asymmetric Transparency. Asymmetric Transparency states that each of the two parties, individually have full visibility and transparency to their unique data assets, algorithms, and contracted outputs, while at the same time each party has an opaque view of the other party's data assets, algorithms and contracted outputs. This White Paper presents key aspects of our Secure Data Sharing solution, from which you will see the incorporation of our guiding principles.

Our Secure Data Sharing™ (SDS) solution provides companies the ability to securely share their data reserves with data analytics companies, so they can 'mine' their data reserves to unearth new business opportunities. An example of this scenario is a credit card company who wishes to utilize the data analytics of a marketing company. The credit card company, as the data provider, and marketing company, as the data processor, can

both be assured that their individual intellectual property assets will not be leaked or exposed during the SDS™ process to join the analytics with the data for the data mining. SDS™ provides a contract creation process whereby the two companies interactively identify the data sets, the individual data elements, and the specific runtime algorithm that will be applied to the data to deliver the desired results. The contract creation process provides a controlled sequence of negotiated steps beginning with output data product definition, algorithm development, testing, acceptance, and algorithm execution behavioral profile. The SDS™ solution minimizes risks of Personal Identifiable Information (PII) exposure and enforces appropriate algorithm behavior during production runtime execution. Secure Data Sharing™ is fully compliant with GDPR privacy and data ownership consent regulations.

A common misconception is that secure file sharing (SFS) is essentially the same as secure data sharing. Nothing could be further from the truth. Secure file sharing solutions have traditionally been used for secure file storage, secure file transmissions; and as such, require extensive integration across database management and user access control mechanisms. SFS solutions encrypt all the data as a single file entity, then decrypt the entire file prior to algorithm runtime. Even with the use of file level encryption, SFS solutions do not protect the data or algorithm from loss, misuse, repurposing, or redistribution since both are decrypted to plaintext for use in general purpose computing environments. More importantly, SFS has no awareness of an algorithm's behavior and therefore cannot protect the data from a misbehaving or errant algorithm. SFS solutions are wholly inadequate for achieving true secure data sharing and delivering on data monetization capabilities.

In contrast, Helios Secure Data Sharing™ is designed for combining advanced AI and data processing activities; so that, companies can expand the revenue potential of their data reserves through data sharing arrangements between companies. SDS™ assures that there will be no risk of their respective data or algorithm getting leaked, lost, misused, repurposed, or re-distributed beyond the specific usage agreed to by both parties.

PRIVACY-BY-DESIGN

At the core of our solution is the critical concept of "privacy-by-design." With Helios' data-centric mindset, we are approaching secure data sharing much differently than all the other solutions. Our SDS™ allows one party's data to securely flow into another party's domain for analytical, algorithm, execution in order to gain the benefits of combining the data and algorithm to create a new data product. SDS™ is involved in the actual processing of data and algorithm between multiple parties and not the typical approach of providing each party access to the other party's data or algorithm, respectively. Additionally, SDS™ securely shares ONLY the relevant data between the two parties, within a mutually approved framework, under a negotiated contract, and where the analytical results (outcomes) have been thoroughly vetted for the specific use without fear of data or algorithm loss, misuse, re-purpose, or re-distribution.

CONSTRAINED PROTECTION DOMAIN™(CPD™)

Helios implements a Zero-Trust model within our SDS™ solution. A foundation of our Zero-Trust security model is the Constrained Protection Domain™. The CPD™ represents our secure computational environment and utilizes a 'Check and Execute' process to provide a pristine environment for each data-algorithm execution activity.

The CPD™ is a totally self-contained execution environment that maintains the highest levels of secure computing regardless of the surrounding trust-worthiness of the underlying system and infrastructure. The CPD™, removes any residue of the preceding data and algorithm's presence prior to the ingestion of the next encrypted data block and encrypted algorithm. It is only after the CPD™, is deemed pristine, that the CPD™, decrypts the individual data block and algorithm and proceeds to data-algorithm execution. Our CPD™ is a closed system that does not allow any external entry and maintains a pristine executing and clean memory environment during processing.

As a part the contract creation algorithm evaluation process, the Helios platform executes an automatic pre-screening of the algorithm results, searching for any sensitive data, in this example PII data, based on the test data provided. This step verifies no PII exist in the test data output and certifies the algorithm performs as expected. Later during CPD™ production processing and the individual data block processing has completed; we again check the data output for any residual PII data. If no PII is found, the output data block is encrypted and transmitted to the appropriate party. If any PII data is found in the post-processing data block, the data block is held, the processing is halted, and the two parties are notified. At no point in the CPD™ processing is there more than a single data block exposed, and we perform a rigorous post-processing check to catch any adverse conditions.

Data encryption and use of secure sandbox processing are not new. However, the ability to contain a sensitive dataset as a correlation of similar data assets in a single Data Defined Network™ (DDN™) across multiple infrastructures with a zero-trust data management policy attached, while allowing Data Providers to maintain visibility and control of their data all the times, is a key differentiator of the Helios SDS™ solution. Had our SDS™ been deployed, the massive exposure and rogue misuse of PII data between Facebook and Cambridge Analytics would have be averted.

ALGORITHM BEHAVIOR ENFORCEMENT™

During the SDS™ contract negotiation stage, the algorithm goes through a testing and acceptance process between the Data Provider and the Data Processor (GDPR Data Controller). It is during the algorithm testing stage that SDS™ learns the algorithm's behavior and we are able to employ our Algorithm Behavior Enforcement ™(ABE™) capabilities.

Once the algorithm is thoroughly tested and vetted, the Data Provider, through the Helios platform, divides the production dataset into individual data blocks and encrypts each individual block with a separate encryption key, and then sends the encrypted dataset via secure a path to the Data Processor location, where the CPD™ resides. During CPD™ execution, the ABE™ function monitors the behavior of the algorithm; such that, any deviation from the expected behavior causes a halt to the computation on the specific data block and both parties are immediately notified.

HOMOMORPHIC ENCRYPTION

In the world of secure data sharing, a lot of research efforts are on providing a feasible solution based on homomorphic encryption. Homomorphic encryption allows computation on encrypted data. Thus, data can remain confidential while it is being processed, enabling useful tasks to be accomplished with data residing in untrusted environments. However, homomorphic encryption has major limitations. Homomorphic encryption often incurs large computational overhead, described as the ratio of computation time of the encrypted version versus computation time of the clear text version. It can only support a limited type and number of mathematical operations. When the complexity of an algorithm increases, homomorphic encryption may not be usable or effective anymore.

Helios creates a robust, yet secured, method targeted to data sharing use cases inside of big data environments, specifically Hadoop and spark. Our CPD™ provides a framework to evaluate the data processing algorithm, learn the normal behavior of algorithm operations, at the same time, securely protect the data and verified algorithm execution; while inside an untrusted environment. The CPD™ protects the data by preventing an unverified algorithm to access the data, and by enforcement of a verified algorithm, or any components inside the CPD™, that is acting abnormally. Fundamentally, Helios provides a cross infrastructure platform that allows customers to create a secured data delivery channel with a safe (trust) compute environment and a validated algorithm to support secure data sharing.

Compared to homomorphic encryption implementations, Helios' SDS™ solution does not have limitations on the algorithm complexity and under SDS™ the computation cost does not increase due to the level of algorithm complexity. Both homomorphic encryption and Helios SDS™ implementations provide for standard data protection; however, SDS™ uniquely protects the algorithm from being disclosed or used in unintended ways.

SECURE DATA SHARING & DATA RESIDENCY

The sharing of sensitive data across territorial jurisdictions has always been a heated topic, even more so with the emergence of global privacy and data residency regulations. The current market approach is a hodgepodge collection of network segmentation security techniques, database access controls, extensive user access controls (AD), and a cumbersome and complex set of multi- use encryption schemes. This approach has been marginally effective within an enterprise infrastructure and been increasingly problematic across infrastructures and geographically distributed platforms. We believe the extension of traditional network security or legacy database approaches to respond to data residency concerns may pass the initial data residency regulatory test, but they represent a shaky foundation from which to build on the future.

Our solution is based upon patent-pending Data Defined Network™ (DDN) and Data Micro-Segmentation™ (DMS) technologies that allow companies to determine in which territory to perform computational activities based on the regulatory restrictions, computational resources, or amount of target data residing in a specific territory. For example, if the data resides in Indonesia and the desired processing algorithm is in Singapore. Then our SDS™ customers could securely transfer the Indonesia data via SDS™ data encryption and execute their algorithms in Singapore without fear of PII data leak or unauthorized access or algorithm misuse while the data is being processed in Singapore. Our Constrained Protection Domain™ maintains privacy compliance with Indonesian data residency regulations that mandate no PII data reside outside of Indonesia.

MIT CONNECTION SCIENCE: TRUST DATA CONSORTIUM

Helios Data is a member of the MIT Trust Data Consortium. The consortium is representative of over 30 global companies, government entities, universities, individual researchers, and thought leaders. Consortium members collaborate on an exchange of ideas, technical frameworks, advanced working groups, business practice development, and pursue the global promotion of the core principles of the consortium to develop privacy-preserving identity systems and safe distributed computation, enabling an Internet of Trusted Data. Helios Secure Data Sharing solution achieves many of the objectives highlighted in the following consortium projects: Open Algorithm (OPAL), Digital Identity and Privacy, Personal Infomediary and Data Fiduciary, MIT Enigma, and OpenPDS.

MIT Trust Data Consortium https://www.trust.mit.edu/