Vendor Profile

Anonos: Embedding Privacy and Trust Into Data Analytics Through Pseudonymisation

Ralf Helkenberg

IDC OPINION

The digital transformation revolution is well underway. By leveraging data and technologies, organizations across all industry sectors are undergoing significant transformation in their business models. With greater demand for information, they are producing ever-more amounts of data. This data is being produced, processed, and shared in more places. Big Data adoption is set to rapidly increase beyond 2020 as more organizations use powerful analytics, increasingly infused with artificial intelligence (AI)/machine learning (ML) to achieve faster innovation, enhanced business performance, and competitive advantage. Cloud's promise of agility, scale, and flexibility is accelerating adoption. Organizations are embracing edge computing and hybrid multicloud architectures, shifting their compute strategies from centralized on-premises databases to distributed data infrastructure models. Distributed processing of large data sets and the richness of the data, much of it sensitive information, make Big Data highly open to data leakage and breaches. Meanwhile, privacy regulation is tightening, adding further complexities to turning data into business insight. In a GDPR and CCPA world, negligence of data privacy protections will not be tolerated and will result in higher fines. In this new reality, data privacy and security must follow the data.

Extract Value From Data in a Secure, Ethical Way

The sharing of data for the purposes of data analysis and research has many benefits. At the same time, concerns and controversies about data ownership and data privacy elicit significant debate. So how do organizations utilize data in a way that protects individual privacy but still ensures that the data is of sufficient granularity that analytics will be useful and meaningful?

IDC believes instilling trust in the use of data is a precondition for fully realizing the gains of data analytics. Encouraged through data protection regulation, encryption has become the default modus operandi for many to securing personal data. Yet it is primarily a security measure for making data unintelligible against unauthorized users. In environments where data is constantly moving between different parties and combined with other data, encryption – though providing effective data protection – is an inhibitor to creating valuable business insights. As a result, advanced analytics and data science projects that require fast access to data are slowing down or coming to a halt.

De-identification (functional separation) through pseudonymisation and anonymisation are important enablers of data analysis without requiring a compromise in data privacy and security. Although they may appear similar at first, they perform different functions in data protection law, such as the GDPR. The difference between anonymisation and pseudonymisation rests on whether the data can be re-identified. Data that has been irreversibly anonymised ceases to be personal data and does not require compliance with data protection law.

August 2020, IDC #EUR146734820

However, uncertainties exist as to whether such procedures can provide a sufficient degree of anonymity. Studies have shown that even within independent anonymised datasets, identifying individuals is not that difficult. Researchers were able to develop a machine learning model capable of correctly identifying 99.98% of Americans in any anonymised dataset using just 15 characteristics. A different MIT study of anonymised credit card data found that users could be identified 90% of the time using just four relatively vague points of information. This means knowing whether anonymisation has been achieved is rarely a black-and-white proposition and a challenging assessment to make. A further downside to anonymisation is that it results in a decrease in data utility. To preserve levels of utility, traditional anonymisation techniques restrict data processing to enclaves or silos to mitigate the risks of reidentification.

Pseudonymisation as a Way Forward

Pseudonymisation protects sensitive data by replacing one or more identifiers (direct or indirect) with pseudonyms or codes, which are kept separately and protected by technical and organizational measures. More importantly, the process can be reversed if the re-identification of data is required. While not a new privacy-preserving technique, pseudonymisation has been newly redefined and gained special prominence within the GDPR in which the benefits of proper pseudonymisation in protecting sensitive data are referenced 15 times. The European Union Agency for Cybersecurity (ENISA) provides further endorsement, where in its publication "Recommendations on shaping technology according to GDPR Provisions" highlights the following benefits from GDPR-compliant pseudonymisation:

- Pseudonymisation serves as a vehicle to "relax certain" data controller obligations, including:
  - Lawful repurposing (further processing) in compliance with purpose limitation principles
  - Archiving of data for statistical processing, public interest, scientific, or historical research
  - Reduced notification obligations in the event of a data breach
- Pseudonymisation supports a more favorable (broader) interpretation of data minimization.
- Pseudonymisation goes beyond protecting "real-world personal identities" by protecting indirect identifiers.
- Pseudonymisation provides for un-linkability between data and personal identity, furthering the fundamental data protection principles of necessity and data minimisation.
- Pseudonymisation decouples privacy and accuracy, enabling data protection by design and by default while enabling data about individuals to remain more accurate.

While pseudonymisation has many benefits, using it effectively requires significant expertise. The same ENISA report recognizes that effective pseudonymisation is highly context-dependent and requires a high level of competence to prevent compromise while maintaining data utility.

Data protection solutions that automate and simplify functional separation implementation have an important role to play in helping organizations realize their data strategies in a privacy-compliant manner. Technology vendors in this market space are few. One such company is Anonos, which have been at the forefront of shaping the pseudonymisation market with its BigPrivacy platform. The platform enables the sharing, collaboration, and analytics of personal data while enforcing security and data protection policies. It does this by employing state-of-the-art data de-identification technologies, controlled re-identification and re-linking of data, and a data-centric approach to security.

Its patented dynamic pseudonymisation technology differs from the more traditional approaches to data protection that use anonymisation, tokenization, static pseudonymisation, and generalization, and do not protect personal data from unauthorized re-identification when data sets are combined and used for multiple use. The BigPrivacy platform stands out for the range of pseudonymisation capabilities it can offer and its ability to meet specific technical requirements for achieving GDPR-compliant pseudonymisation. The flexibility in control settings that enable the relinking of de-identifiers back to individuals to support lawful business purposes is a plus.

A secure, scalable, and versatile platform, IDC believes BigPrivacy is a significant step up from traditional centralized data protection technologies, and it is well placed to resolve organizations' need for faster data insights while controlling data use risk across multiple environments and data-sharing partners.

IN THIS VENDOR PROFILE

This IDC Vendor Profile looks at Anonos, a technology provider of privacy-preserving data solutions. The profile focuses on the company's patented BigPrivacy platform, which pseudonymises personal data, thereby enabling organizations to undertake complex and sophisticated data analytics in a privacy- and security-compliant manner.

SITUATION OVERVIEW

Many organizations are trying to obtain more value from their data to improve their products and services. For example, more chief data officers and data analytical roles are being created to drive such data-enabled transitions. However, data privacy has become a flashpoint in the drive to achieve digital transformation. Concerns over potential privacy violations and the prioritization of locking data down through security measures has mistakenly led many organizations to forego the benefits of data insight. It need not be an either/or choice, since dynamic pseudonymisation organizations can achieve both.

Company Overview

Successful business partners for 20 years, Gary LaFever and Ted Myerson have a track record of developing privacy-preserving technology that help organizations turn regulation into a competitive advantage. Anonos was founded in 2012, on the premise that a whole new approach to data control, stewardship, and protection was necessary to unlock the maximum value of data and to turn it into a business asset without violating privacy, security, or regulatory restrictions. The outcome is the BigPrivacy pseudonymisation platform. The platform capitalized on growing market demand for privacy-compliance solutions, particularly the GDPR, and found adoption across financial, healthcare, telecom, and other data-intensive industries that rely on consumer data insights.

In 2019, the company secured a $12 million growth investment led by private equity firm Edison Partners.

Technology Proposition: BigPrivacy

Anonos' BigPrivacy platform supports a risk-based approach to data protection. This is accomplished by empowering data scientists and privacy engineers to set privacy controls at the data element level by applying a combination of traditional anonymisation techniques, GDPR-compliant pseudonymisation, and its own patented reidentification risk management technology – Controlled Linkable Data. Anonos enhanced the platform's pseudonymisation capabilities by leveraging the 50 technology recommendations by ENISA, and it can meet the specific technical requirements for achieving GDPR-compliant pseudonymisation.

The platform is comprises the following key components.

Variant Twins

The core of BigPrivacy's capabilities is centered around Variant Twins (i.e., the final pseudonymised dataset). The patented system leverages dynamic pseudonymisation to replace personal identifiers, such as a person's name and date of birth, with unique de-identifiers that prevent attribution of the data to a specific person without permission. Privacy control settings can be fined-tuned to provide the type and level of identifiability needed for each authorized use case. Because all Variant Twins are derived from the original source data, rather than permanently altered, organizations suffer no degradation in data value or accuracy.

GDPR-Compliant Pseudonymisation

Anonos' patented controlled relinkable dynamic de-identifiers are an advancement on the traditional pseudonymisation techniques of applying the same static tokens to direct identifiers across datasets. While useful in centralized environments, this traditional approach provides limited protection against unauthorized re-identification of individuals through data linkage and inference attacks. BigPrivacy uniquely combats the relinkage of data (Mosaic Effect) by using dynamic de-identifiers to introduce uncertainty (entropy) at the data element level for both direct and indirect identifiers. The product supports a range of pseudonymisation policies including the ENISA-defined fully randomized and deterministic pseudonymisation, and three additional intermediate-level policies: field, table, and document deterministic pseudonymisation.

Data Use Risk Management

K-anonymity sets out to address the risk of re-identification of anonymised data through linkage to other datasets – the higher the K value, the higher the degree of anonymity. The BigPrivacy platform incorporates a Data Use Risk Management module that leverages the k-anonymity concept. The pseudonymised dataset (Variant Twins) is filtered for reidentification risk to suppress records that do not meet the required k-anonymity threshold.

Lawful Insights API

Maximum data value often comes from combining and sharing data sets across multiple environments and with different partners. Anonos' Lawful Insights API enables lawful and multiparty processing both inside and outside of an organization's environment. Utilizing the same techniques and technology as BigPrivacy, it leverages endpoint interfaces to reduce data transfer friction and accelerate the process of safely and securely sending and receiving data for analytical processing. This is helpful when organizations experience trouble getting access to third-party data to augment the value of their data assets, or when third parties express concern about potential liability.

Company Strategy

Anonos' strategy aligns with addressing key privacy challenges in the sharing, collaboration, and analytics of personal data in a compliant manner to the GDPR. The main use cases are the following.

Compliant AI/ML Data Use

Data is a key driver for many of the new emerging technology innovations such as artificial intelligence and machine learning. Though AI/ML offer enormous business and innovation opportunities, they also pose privacy risks and regulatory challenges. The GDPR requires processing of personal data be carried out for specific purposes, no more personal data than is necessary to achieve those purposes is processed, and that personal data is only processed for as long as necessary to achieve those purposes. Tensions arise between these data privacy principles and AI, since the development of an AI system can often result in data being used for unexpected purposes, and often requires vast amounts of data to be inputted into the system for it to meaningfully detect patterns and trends. It also makes relying on consent as a lawful basis for many kinds of sophisticated data analysis impossible. BigPrivacy enables organizations to overcome the limitations of consent by using GDPR-compliant pseudonymisation to enable legitimate interests as a lawful basis for processing data.

IoT Data Protection

The number of devices connected to the Internet (i.e., Internet of Things or IoT) continues to grow exponentially. IDC forecast there will be 41.6 billion connected IoT devices generating 79.4ZB of data in 2025. Many of these devices exist within the automotive, healthcare, and consumer goods fields. Privacy and security though are big issues; through the data risk management controls of BigPrivacy, privacy-preserving data collection and management can be enabled across distributed IoT environments.

Compliant International Data Transfers

The GDPR sets out the legal mechanisms for the transfer of personal data outside the EU. The July 16, 2020, Schrems II decision of the European Court of Justice (CJEU) invalidated the EU-US Privacy Shield, a mechanism for transferring personal data from the EU to the U.S. At the same time, the CJEU reaffirmed the validity of standard contractual clauses, but added the caveat that they are only valid if they contain effective safeguards to ensure compliance with the protections provided by EU law. Anonos' BigPrivacy software is well placed to satisfy the Schrems II requirements for appropriate safeguards by creating pseudonymised versions of personal data (Variant Twins). Variant Twins ensure that desired processing results are achievable without providing third parties, including country authorities, the ability to re-identify individuals.

Data Protection by Design and by Default

The GDPR introduces the concept of "data protection by design and by default" into formal legislation, whereby organizations must integrate data protection into their processing activities and business practices from the design stage and throughout the life cycle. This includes adopting appropriate technical and organizational measures in implementing the data protection principles effectively. Pseudonymisation is one of several measures that organizations are urged to adopt to transition to the data protection by design and by default posture. The risk management controls in BigPrivacy help support data use minimization by enforcing selective access to data and ensuring employees only have access to the data required for them to do their jobs.

Data Processing in the Cloud

With the growing trend of moving workloads to the cloud, organizations can take advantage of cloud services without fear of breach of data privacy laws, as BigPrivacy pseudonymises the data at the point of ingestion. The solution enables organizations to create a global data lake in the cloud and also meet data sovereignty requirements.

FUTURE OUTLOOK

Concerns over data privacy and security have never been stronger, with the GDPR significantly influencing the way personal data is processed and protected by organizations. IDC believes many organizations have not yet fully recognized the benefits of GDPR-compliant pseudonymisation in deriving value from data while remaining compliant with data regulations and business rules. Many privacy professionals are not yet fully attuned to its potential, with their focus having concentrated in the past few years on ensuring regulatory compliance through policies and procedures. But a rapid shift is underway, with value-creation from data becoming a primary concern for organizations. In this new environment, privacy professionals are realizing they cannot just be compliance gatekeepers — they need to step up as business enablers to support organizations implement their data-driven business models. This means giving analytics and data science teams dynamic and frictionless access to datasets while enabling privacy-preserving measures to work in the background.

IDC thinks COVID-19 will be an inflection point for accelerated adoption of functional separation control techniques. Clinical and technological research projects that have arisen to mitigate the spread of COVID-19 have, in many cases, necessitated inter-organizational and cross-border data collaboration. Pseudonymisation has proved instrumental in providing a legally compliant approach to link and share sensitive datasets and provide access to secure analytical environments for researchers.

As the adoption of digital and data-driven business models accelerates and the need to use and share data in an ethical and trustworthy manner becomes a prerequisite, we believe Anonos is well-positioned with its state-of the-art pseudonymisation capabilities and granular protection control settings to capitalize on demand for privacy-preserving data analytics.

ESSENTIAL GUIDANCE

Advice for Anonos

Anonos has set itself up as an innovator and technology leader within the de-identification market space, and it is leading the charge to evangelize the benefits of pseudonymisation.

Though these concepts are not new, there remains much misunderstanding around the terminology and use of anonymisation and pseudonymisation. A legally and technically complex subject matter, Anonos needs to continue to push awareness around deployment best practices. Establishing a center of excellence in this field might further its cause.

It has rightly recognized the positioning of BigPrivacy as a legal compliance solution doesn't necessarily resonate with data audiences, but it is going in the right direction with its pitch that data privacy and data utility need not be an either-or choice and that both are possible with its leading-edge technology.

To prove its capabilities and credentials with new audiences, it needs to showcase through high-profile use cases how enterprises across different sectors have used its BigPrivacy platform to turn sensitive data into compliant business assets.

The technology also has a role to play in the hybrid-multicloud era. As distributed processing of Big Data and analytics migrates to cloud, cloud service providers are seeking to integrate and upgrade their privacy and security capabilities to secure this business, much of which is encryption focused. For Anonos, there is an opportunity to address this gap and enhance cloud providers' data application value proposition with its state-of-the-art pseudonymisation technology.

LEARN MORE

Related Research

- *IDC Market Analysis Perspective: European Data Privacy 2019* (IDC #EUR145752519, January 2020)
- *Anonos' SaveYourData - a EuroPrivacy Certified Solution "Deep Freezes" Enterprises' Existing Personal Data Sets as They Plan Analytics Strategies* (IDC #EMEA44411718, November 2018)
- *Europe's Political Leaders Put Ethics at the Heart of AI Strategy* (IDC #EMEA43324118, April 2018)
- *Anonos: Helping Businesses Become Data-Driven Without Compromising GDPR Compliance Obligations* (IDC #EMEA43641318, March 2018)

About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications and consumer technology markets. IDC helps IT professionals, business executives, and the investment community make fact-based decisions on technology purchases and business strategy. More than 1,100 IDC analysts provide global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries worldwide. For 50 years, IDC has provided strategic insights to help our clients achieve their key business objectives. IDC is a subsidiary of IDG, the world's leading technology media, research, and events company.

IDC U.K.

IDC UK
5th Floor, Ealing Cross,
85 Uxbridge Road
London
W5 5TH, United Kingdom
44.208.987.7100
Twitter: @IDC
idc-community.com
www.idc.com