MARKET PERSPECTIVE

Anonos' SaveYourData — a EuroPrivacy Certified Solution — "Deep Freezes" Enterprises' Existing Personal Data Sets as They Plan Analytics Strategies

Archana Venkatraman Martin Whitworth

EXECUTIVE SNAPSHOT

FIGURE 1

Executive Snapshot: Anonos Develops SaveYourData Software — a EuroPrivacy-Certified, GDPR-Compliant Solution — to Deep-Freeze Enterprises' Existing Personal Data

Anonos launched its EuroPrivacy-certified SaveYourData to meet the legal and technical requirements for GDPR pseudonymisation. It offers a pseudonymised "data safe" for enterprises' accumulated data so they have a GDPR-compliant legal basis to save existing data and plan analytics strategies. SaveYourData is available as an Anonos-Hitachi Vantara joint solution and comes when enterprises are deleting, bulk encrypting, or getting data reconsented, making analytics and GDPR compliance as an either-or choice.

Key Takeaways

- Certified by EuroPrivacy for GDPR compliance, Anonos' SaveYourData is a dynamic pseudonymisation technology that differs from traditional approaches to data security to enable legitimate interest processing. It offers a privacy-compliant option to deep-freeze personal data accumulated by enterprises while they outline their solutions to address analytics processing issues.
- Organizations are under pressure to become data-informed and insights–driven to gain a competitive advantage and accelerate digital transformation, making analytics, AI, and ML cornerstone initiatives.
- A long-term sustainable analytics strategy ensures that personal data is always appropriately protected, and that privacy and security by design is adopted from the outset to remain GDPR-compliant.

Recommended Actions

- Anonos should leverage its alliance with Hitachi Vantara and its network of global systems integrators as well as collaborate with other key storage providers with large installed bases to provide enterprises with an alternative to deleting or encrypting personal data for compliance.
- It needs to increase awareness of the ability of pseudonymisation to maximize data value and demonstrate how its patented dynamic pseudonymisation technology differs from anonymization, tokenization, static pseudonymisation, and generalization approaches to data protection.
- Due to the uncertain, time-sensitive nature of the one-off right to transform illegal data under the GDPR into a new legal format, organizations should evaluate their options and take appropriate action.

Source: IDC, 2018

October 2018, IDC #EMEA44411718

NEW MARKET DEVELOPMENTS AND DYNAMICS

In the digital era, companies are technology or software businesses first. However, it is also true that companies are becoming information businesses first, given the value and importance of being insights-driven. In fact, IDC research finds that 43% of European organizations are focusing their digital transformation efforts on data capitalization and data monetization to create new data-driven revenue streams.

But to capitalize on data, organizations need to prioritize analytics, data management, artificial intelligence (AI)/machine learning (ML), privacy and data security, compliance, and information governance as cornerstone initiatives.

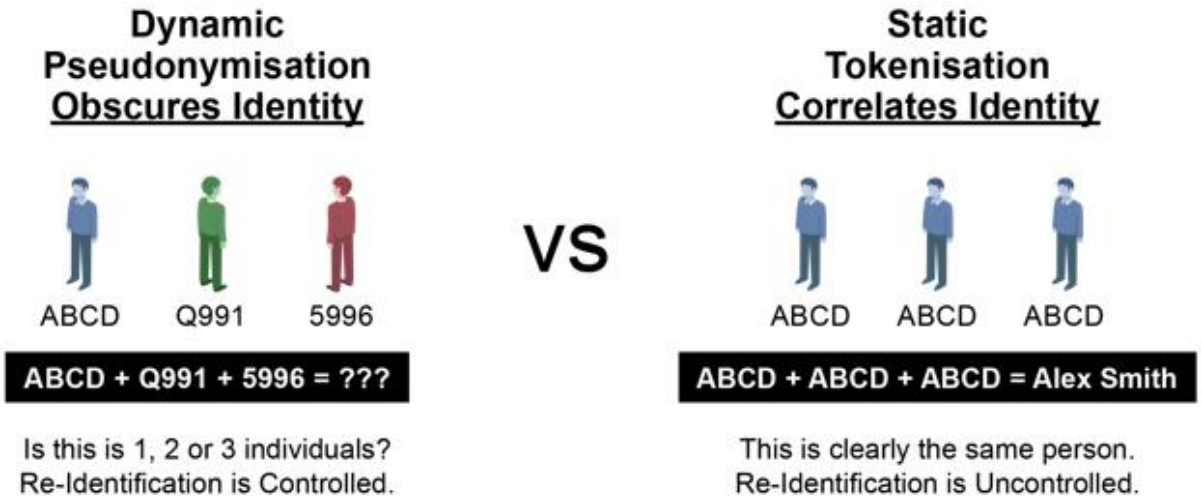
There can hardly be an enterprise boardroom or strategy committee discussion that is not talking about analytics and machine learning or artificial intelligence. These Innovation Accelerators unleash insights to drive better and smarter products, engaging customer experiences, personalized services, brand loyalty, improved collaboration, efficiency, and even newer revenue streams. Organizations then benefit from more meaningful business insights, improved employee engagement, new opportunities, or monitoring and modeling behavior – this creates competitive advantage in the digital era. These use cases require access to large amounts of data, much of which is personal data.

The introduction of new privacy laws, such as the EU GDPR, have become game-changing regulations and brought consideration of storage, use, and protection of personal data front and center. These regulations aim to put privacy and personal data under the control of consumers, making businesses reassess how they can comply with the new requirements while continuing business initiatives. This focus requires organizations to deliberate on and adopt a privacy-by-design approach, which necessitates understanding various approaches to personal data protection available, including:

- **Encryption.** This is typically used to protect the confidentiality of data by making all data unreadable without access to a secure key. There are developments in encryption techniques that may allow the processing of encrypted data in the future (e.g., homomorphic encryption).
- **Tokenisation.** This was made popular by card payment schemes (PCI-DSS) and usually used to replace a sensitive data value (e.g., card number) with a static, non-sensitive value or token. The static token is a reference, stored separately, that maps back to the original sensitive data value and allows the original data to be recovered.
- **Anonymization.** A much misunderstood term, this refers to irreversibly replacing a sensitive data value (usually a personal identifier) with data that cannot be used to re-identify the original (so anonymization is one-way).
- **Pseudonymisation.** It was introduced by the GDPR as a new technical approach to enable greater use of data in a privacy-respectful manner that is NOT the same as anonymization. In simple terms, pseudonymisation means replacing sensitive personal identifiers (usually personal details) with another unique identifier, typically generated through an algorithm. More importantly, the process can be reversed if the re-identification of data is required. Pseudonymization is different from static tokenization in that with the latter, any value is always tokenized into the same token and thus can be re-identified using the mosaic effect. On the other hand, pseudonymization tokens are dynamic, and the algorithm produces different results every time.

FIGURE 2

Dynamic Pseudonymisation vs. Static Tokenisation



Source: Anonos, 2018

These are necessarily brief descriptions, and a full explanation requires a separate discussion and is outside of the scope of this paper. But as can be seen, there are many options to ponder, and each use case requires careful consideration to enable selection of the most appropriate solution. While considering these important design decisions and before implementing appropriate controls, ensuring that you have a valid legal basis to store and use data is essential. In this paper, we refer to pseudonymising data and storing the results in a restricted, access-controlled environment only accessible by authorized persons as the first step in converting to legitimate interest processing as a deep-freeze state.

GDPR is not prescriptive regarding the technologies required to enable compliance. However, it recommends the implementation of encryption and pseudonymisation as approaches to protect sensitive data and manage data subject risk.

> "In order to maintain security and to prevent processing in infringement of this Regulation, the controller or processor should evaluate the risks inherent in the processing and implement measures to mitigate those risks..." (GDPR Recital 83).

GDPR Article 32 further states:

> "Taking into account the state of the art, the costs of implementation and the nature, scope, context and purposes of processing as well as the risk of varying likelihood and severity for the rights and freedoms of natural persons, the controller and the processor shall implement appropriate technical and organizational measures to ensure a level of security appropriate to the risk, including inter alia as appropriate:
>
> (a) the pseudonymisation and encryption of personal data;"

GDPR mentions encryption as an approach to mitigate risks of data processing because it renders personal data unintelligible to unauthorized individuals.

Encryption offers great opportunities to effectively protect personal data, but IDC believes encryption is not a panacea for good data management because it can impair business functionality and analytics strategies. Encrypted data may not always be usable for expected use cases such as search, sort, or analytics without application changes — unless decryption is also baked in to the process, rendering the data unprotected.

Organizations certainly need to comply with data protection regulations, but more importantly, they need to function as businesses and progress their data-driven initiatives to gain a competitive advantage. Using encryption can get in the way of this — in fact, organizations use copies of personal data for a variety of purposes (such as analytics), and testing and blanket encrypting these versions of personal data could render these purposes impossible (by preventing processing and analysis of encrypted data). Besides, encryption can also be complex — key management is non-trivial, and lost keys could result in catastrophic data loss. Format-preserving encryption is good at preserving business functionality, but it is not commonly used.

Alternatively, the use of pseudonymisation can help to reduce risks to data subjects while helping data controllers and processors meet their compliance obligations by minimizing exposure of personal data and opportunities to identify data subjects.

GDPR defines pseudonymisation as:

> "the processing of personal data in such a way that the data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organizational measures to ensure that the personal data are not attributed to an identified or identifiable natural person."

The pseudonymisation approach involves replacing personal identifiers within data (such as name, date of birth, email addresses, address) with alternate data that cannot be used to identify the data subject. But unlike anonymization, it is possible to tie the assigned reference value back to original data subjects if the data controller has the access to corresponding personal information that is kept separately.

At the heart of pseudonymisation are three key principles:

- The mechanism replaces all direct and indirect personal identifiers with different tokens.
- The personally identifiable information necessary to relink tokens back to individuals is stored in a separate datastore, and technical measures are implemented to protect the data and render it unattributable without authorized linking for a specific use case.
- Pseudonymised data is also within the scope of GDPR and should be considered as personal data.

These three principles distinguish pseudonymised data from anonymized (encrypted) data where relinking and re-identifying data subjects linked to pseudonymised data is possible under controlled conditions, making it valuable for analytics and insights-driven business decisions.

However, organizations need to consider data subjects' re-identification risks involved in pseudonymisation and should take all measures to protect pseudonymised data and implement the necessary measures (such as controlled access) to prevent unauthorized relinking of pseudonymised datasets.

Processing of personal data under GDPR is legal only if one of the following legal bases apply (Article 6, GDPR):

- **Consent.** Personal data is obtained by specific consent of the individual concerned; the data processing must be limited to the explicit and unambiguous purpose for which it was consented and collected for.
- **Contract.** The data is necessary for the performance of a contractual obligation between the company/organization and the individual data subject.
- **Data controller legal obligation.** Processing is necessary to meet a legal obligation of the data controller under EU or member state national legislation.
- **Data subject vital interest.** Processing is necessary to protect the vital interests of an individual.
- **Public interest.** Processing is necessary for the performance of a task carried out in the public interest under EU or member state national legislation.
- **Legitimate interest.** For the organization's legitimate interests while using technical and organizational safeguards such as pseudonymisation to avoid negatively impacting the fundamental rights and freedoms of the person whose data is processed. This assessment depends on the individual circumstances of each case in question.

The European Data Protection Board (EDPB) endorsed the Article 29 Working Party requirements for consent under Regulation 2016/679 (WP259 rev.01) that specify the requirements for and limitations of using consent as a legal basis for processing EU personal data under the GDPR. Consent Requirements acknowledge that the GDPR changed the definition of consent and that all data – collected before and after the GDPR – that fail to meet new strict GDPR consent requirements for specificity and unambiguity are no longer legal to "process" – which term under GDPR Article 4(2) includes mere storage of data. The GDPR has no "grandfather provision" or exemptions allowing for continued use of data collected using (now) illegal non-compliant consent. Storing or processing this data exposes organizations to regulator injunctions blocking access and use of data in addition to significant penalties under the GDPR. It is unclear how long organizations will have to exercise their one-off opportunity to transition data to support an alternate (non-consent) legal basis as outlined in the Consent Requirements. Data protection authorities (DPAs) are looking for proof that organizations have taken good-faith steps to comply with the GDPR.

Due to the uncertain time-sensitive nature of the one-off right to transform data that is otherwise illegal under the GDPR into a new legal format under Consent Requirements, organizations should evaluate their options immediately and take appropriate action.

As a result of these requirements, many organizations have simply resorted to either

- Adopting blanket data encryption that renders meaningful analytics impossible and does not address the potential unlawfulness of storing the data
- Deleting data to comply with GDPR requirements because searching, identifying, and classifying personal information and then applying for reconsent is a labor- and resource-intensive task

A top 5 global hospitality firm revealed to IDC at a workshop that it had deleted 20 years of loyalty data because of concerns over legality of data processing under GDPR. Those that have the resources and an appetite for consent-based processing are finding the reconsenting process lengthy and depletory to their business data. For instance, a top European financial services organization said to IDC said that it sought to obtain consent from its customers and obtained only a 60% success rate.

In IDC's opinion, these strategies of data deletion or <100% consent results in wasted opportunities. IDC believes organizations need to balance their risks and opportunities and adopt GDPR-compliant pseudonymisation technologies.

GDPR-friendly pseudonymisation for data processing is applicable in the following scenarios:

- Consent is not practical, or may undermine the business
- Statistical analysis to identify broad trends or general conclusions
- Retention of personal data under strict policies for compliance with industry-specific regulations such as healthcare or banking data regulations

Data enablement/security start-up Anonos collaborated with storage vendor Hitachi Vantara to launch a solution called SaveYourData to create a legal and technical foundation for legitimate interest processing using privacy-compliant pseudonymisation. The solution offers a deep-freeze state for existing personal data repositories without violating GDPR principles. This helps organizations avoid data deletion, blanket encryption, or reconsent exercises.

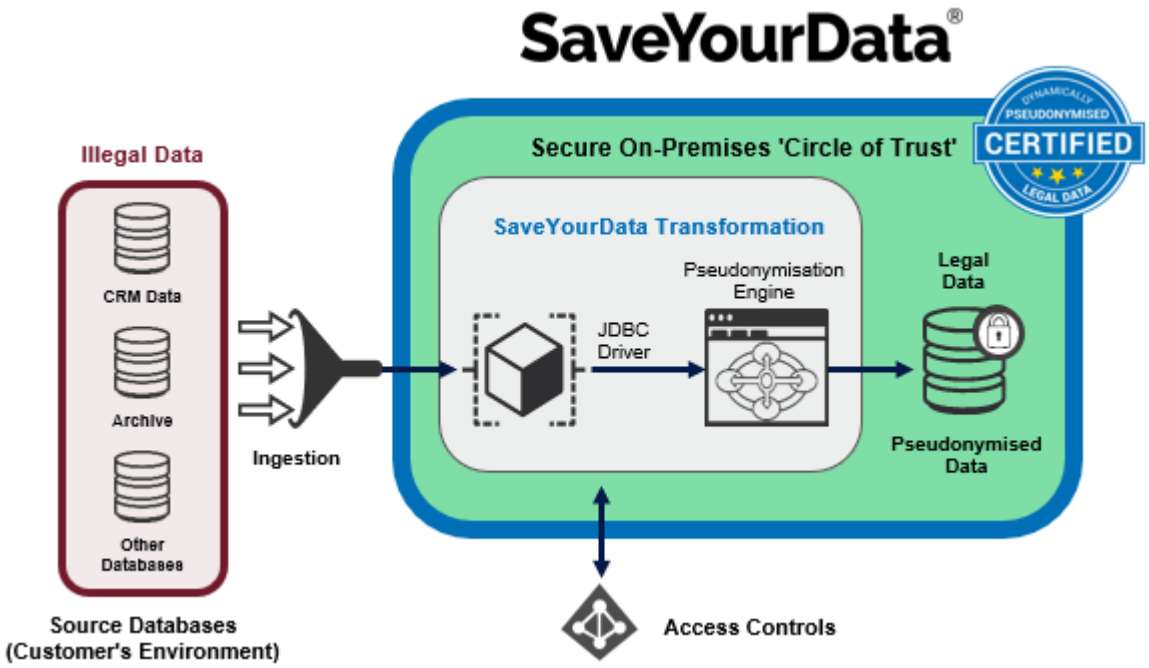
SaveYourData provides a means to safely and legally save existing personally identifiable data — putting it in deep-freeze — while enterprises implement solutions to address analytics processing issues to comply with GDPR.

The SaveYourData software is installed on-premises and pseudonymises database files containing personal data. Companies process a source file (database) containing personal data to convert it into a pseudonymised file where personal data values are replaced by randomized data. Alongside the pseudonymised file, it also creates an index file that contains the information necessary to map the randomized data with the original data to execute data analytics tasks

The technology pseudonymises datasets or databases into a deep-freeze state to give organizations the opportunity to implement technical and operational measures to use data lawfully because even pseudonymised data is still within the scope of GDPR, and data breaches may still be subject to fines.

FIGURE 3

SaveYourData Data Flow Architecture



Source: Anonos, 2018

SaveYourData software relies on patented dynamic pseudonymisation technology that:

- **Assigns dynamic pseudonyms to replace sensitive data**. Dynamic pseudonymisation is a technology that assigns a new or different token to the same personal data value used in another instance. This helps prevent the identification of data subjects, unlike static tokenization where a person can still be re-identified by combining pieces of information in varied data sets because of the same static token – this is called identification through "mosaic effect."
- **Works with the OS kernel to convert the sensitive data values into pseudonymised data**. By pseudonymising data with SaveYourData, organizations can break from the initial deadlock of possessing data collected using pre-GDPR general consent (which is now illegal under the GDPR and with respect to which no "grandfather" or "savings" clause allows ongoing legal possession or use) to legitimate interest (non-consent) legal basis, so that deletion is not required. It is a first step to safeguarding existing data in a compliant manner and gives businesses time to evaluate their data processing projects.

Anonos' SaveYourData, combined with Anonos BigPrivacy dynamic pseudonymisation, leverages the company's patented dynamic pseudonyms algorithm that prevents the recombination and correlation of static pseudonyms that can lead to unauthorized re-identification of data subjects, directly or indirectly, breaching regulatory requirements.

The role of Anonos BigPrivacy dynamic pseudonymisation in the enterprise journey to becoming GDPR-compliant while executing analytics projects is explained in IDC's *Anonos: Helping Businesses Become Data-Driven Without Compromising GDPR Compliance Obligations* (IDC #EMEA43641318, March 2018) and *Europe's Political Leaders Put Ethics at the Heart of AI Strategy* (IDC #EMEA43324118, April 2018).

ADVICE FOR ANONOS

What differentiates Anonos' SaveYourData is its patented dynamic pseudonymisation technique. Another highlight is that it is certified as a GDPR-compliant pseudonymisation technology. The solution (version V 1.0.1.) has been assessed according to the EuroPrivacy certification scheme as complying with European GDPR requirements.

The certification by EuroPrivacy using the "Privacy Flag" certification scheme, developed under a research project co-funded by the European Commission and Switzerland, highlights that the software meets the GDPR requirement of data protection against scenarios required for compliant Pseudonymisation to prevent a situation where "...even if the key is not revealed, a malicious actor may be able to identify individuals by combining indirect identifiers in the pseudonymous database with other available information", as is the case with static tokenization.

The certification body concluded that non-conformities identified in the initial assessment of the software have been addressed. But to ensure optimal performance under GDPR, EuroPrivacy recommends that the software be enriched to enable selective pseudonymisation (enabling users to select what files should be pseudonymised) because currently, SaveYourData enables randomization of all the data contained in a source file. Although this helps prevent the risks of direct or indirect identification in another adjacent file, such indiscriminate randomization stops enterprises from using non-personal data contained in the adjacent file, which is usually possible through pseudonymisation mechanisms. EuroPrivacy states that the availability of the separate Anonos BigPrivacy platform enables differentiated access and data processing of pseudonymised data using complementary tools to increase the value of the software.

Based on the demos IDC has seen, SaveYourData software is simple to use and helps organizations buy time to implement their GDPR compliance plans and data analytics road maps for the long term. Priced per terabyte, the software leverages existing access controls set by the data controller and is database-agnostic.

Launched in September 2018, SaveYourData comes at a time when organizations have moved beyond the initial minimum viable compliance to operationalizing compliance and assessing opportunities around analytics and optimizing investments in data lakes. The software launch also highlights how Anonos has maintained the momentum of innovation and continuously engaged with large enterprises in heavily regulated sectors to understand challenges around broader digital transformation and data-driven initiatives. The software is a valuable addition to Anonos' flagship BigPrivacy platform that takes source data and converts it into a pseudonymised format that de-identifies data by replacing personal data with dynamically tokenized variants. IDC commends Anonos' strategy to develop SaveYourData as an independent software that does not require enterprises to invest in Anonos' separate BigPrivacy platform. This strategy is likely to open more doors for Anonos to then upsell its flagship technologies.

As part of the go-to-market strategy, Anonos has teamed up with a storage and analytics vendor Hitachi, which recently rebranded itself to offer data-driven solutions across cloud, datacenters, and edge IT. IDC believes Hitachi, with its large installed base especially in the retail and financial services sectors, is likely to provide start-up Anonos with the exposure and entry to large enterprises necessary to fully capitalize on the six years of R&D underlying SaveYourData and BigPrivacy technology.

Moving forward, Anonos needs to forge alliances with multiple storage, security, software-as-a-service (SaaS), and data management vendors to develop a rich ecosystem. It also needs to continue raising awareness of the benefits and challenges around different pseudonymisation techniques and how it directly aligns with the requirements of data protection by design and default (DPbDD) in GDPR. IDC data shows that investments in GDPR-related technologies for traditional security and storage will peak between 2017 to 2019, and Anonos is well-positioned to make the most of this window of opportunity, as well as the expanding role for privacy-compliant technology, if it can raise awareness levels.

Given the time-sensitive nature of the one-off right to transform data that is otherwise illegal under GDPR into a new legal format under the Consent Requirements, organizations should evaluate their options and take appropriate action.

LEARN MORE

Related Research

- *Actifio Strengthens Its Data-as-a-Service Vision with App-Centric Approach in the Multicloud and DevOps Era* (IDC #EMEA44399818, October 2018)
- *ETSI Releases ABE Standards and Claims GDPR Compliant Data Transfer Protection* (IDC #IcEMEA44232718, August 2018)
- *Large Italian Bank Adopts Cohesity to Make Secondary Storage Strategic for Analytics, Test&dev, and Time to Market, While Also Modernizing and Automating Backup* (IDC #EMEA44153618, July 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)

Synopsis

This IDC Market Perspective analyzes Anonos' SaveYourData, a EuroPrivacy-certified technology to meet the legal and technical requirements for pseudonymisation under GDPR. The software comes at a time when enterprises are evaluating whether to delete, encrypt, or reconsent data for analytics purposes.

"Successfully implementing Big Data analytics, ML, and AI inevitably bring significant business, technical, and ethical challenges. Not least among these are the absolute requirements that personal data are always appropriately protected and that privacy and security by design is adopted from the outset," said Martin Whitworth, research director, European Privacy and Data Security at IDC."Pseudonymisation, in particular, is a much-misunderstood technique that can be successfully employed to help to manage data subjects' risk while enabling data utility, and to provide support for legitimate interest processing. Choosing the appropriate protection methods for individual data sets and use cases is a non-trivial exercise and requires that we also have a means of safeguarding our current accumulated data stores, in a privacy-compliant manner, prior to implementation."

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