Maintaining Your Data Privacy Lifecycle with IBM Data Privacy Passports

AN ENTERPRISE MANAGEMENT ASSOCIATES® (EMA™) WHITE PAPER
PREPARED FOR IBM
BY DAVID MONAHAN
SEPTEMBER 2019

EMA™
IT AND DATA MANAGEMENT
RESEARCH | INDUSTRY ANALYSIS | CONSULTING

Table of Contents

Executive Summary 3

Why the Current State of Data Protection is Lacking 4

Data Protection vs. Data Privacy 5

Challenges of Maintaining Data Protection and Privacy 5

Maintain Data Control Everywhere with Data Privacy Passports 6

- Overview of Information Rights Management Using Data Privacy Passports 6
- How it Works: Overview 6

Business Use Cases 7

- Protecting Structured Data On-Premises, in the Cloud, or Other Shared Environments ... 7
- Data Segmentation and Brokering 7
- Data Access Control for Compliance 7
- Embedded Data Retention Policies 7
- Data Deduplication 7

EMA Perspective 8

About IBM 8

EXECUTIVE SUMMARY

Data protection has been a best practice since the formation of secrets...data privacy, not so much. Privacy-related data is seen as an increasingly valuable resource for marketers driving amplified collection and sale.[1] Those collecting privacy-related data were free to use it not only for their declared business purpose, but also as a revenue source, with many organizations selling their customer data to the highest bidder with little concern for how it was being used.[2] Securing the stored data was not the highest priority in many organizations, leading to mass thefts.[3] These practices, along with inadequate data security, led to massive data thefts, supporting increasing identity theft rates over the last 15 years.[4]

The good news for consumers is that the tide has shifted. Privacy legislation is being enacted globally, with far-reaching effects on businesses collecting personal data. Though good news for consumers, this can be a tough pill to swallow for organizations collecting personal information and not prepared to properly protect it.

With the wave of expanding privacy legislation like the European Union's GDPR, California's CCPA, and Japan's APPI, to name a few, data protection is a significantly broader term, holding more weight than it has in the past and requiring immediate attention.

This paper will discuss the current inadequacies of maintaining data privacy with traditional encryption and how the introduction of IBM's Data Privacy Passports technology can meet the ever-burgeoning data privacy legislation demands. Data Privacy Passports can even exceed the current data protection best practices in most organizations.

1 25 Mind-Blowing Statistics on the State of Data-Driven Marketing 2018
2 How Businesses are Collecting Data (and What They're Doing With it)
3 Stopping the Data Breach Epidemic
4 Identity Theft: Facts & Stats 2018-2019

WHY THE CURRENT STATE OF DATA PROTECTION IS LACKING

Today, data protection is a serious business. Millions of dollars are spent yearly to maintain the confidentiality of intellectual property and sensitive employee and customer data. Organizations deploy access management solutions, mandate strong passwords, develop role-based identity management, and leverage multifactor authentication to control access to systems, applications, and both structured and unstructured data. Beyond that, they use encryption to further protect the data at rest, in transport and in processing, and technologies like data loss prevention (DLP) and cloud access security brokers (CASB) to watch for data movement and data requests that are against policy. Yet, data is still being compromised.

Research has identified that over 90% of successful attacks stem from phishing attacks,[5] which exploit the identity of the victim and/or vulnerabilities in the target application or system at the point of attack. Due diligence is often missing in areas of identity and access control as it relates to data. Adherence to best practices is lacking in many industries. While the financial industry and certain governmental departments mandated stringent data controls and encryption beyond the minimum requirements, it gets less thorough outside that, with many companies and industries having far looser data access governance and only encrypting the most sensitive internal and customer data as they are required to by law.

Conversely, only 4% of the records breached since 2013 have been encrypted,[6] but the opportunity for theft of data at rest is huge compared to the opportunities for theft of data in transit. In review, the mega-breaches have all been from data at rest. The setup required is significantly less complicated, attacks are easily replicated for scale, and the attack vector does not need to be focused on a single person or web property to be successful.

The difficulties in protecting data at rest come from cost, effort, data compatibility, and data control. Users need not do anything to gain encrypted communications and organizations need only buy a digital certificate and install it on the application to be protected once every few years, with that interval totally at their discretion based on the purchase. In contrast, the process for encrypting and maintaining data at rest is seen as far more costly for the solutions and has significantly higher difficulty with key distribution and management. The latter is something of a misconception, with modern encryption solutions like IBM's pervasive encryption, but it is still a widespread perception that inhibits adoption. Some applications do not react well to having their data encrypted. If they cannot recognize or validate the data type, they stop operating or refuse to accept the data. Data compatibility affects only a proportionately small number of legacy applications, but because it is sometimes difficult to detect ahead of time or solve after the fact, it is a common roadblock. This is also more an issue of failing to ensure proper requirements validation and testing prior to purchasing a solution.

The third part of the equation is data control. This is a growing concern because of a fundamental shift in business operations requiring a much higher instance of data sharing. EMA 2018 research identified that 90% of organizations have teams that share regulated or sensitive data externally with vendors, partners, or customers for work on a monthly or more frequent basis, and 44% share that data one or more times daily.[7] Fifty-seven percent of organizations are highly concerned with the risk this brings to their organization,[8] and 46% believe those teams experience unauthorized data leakage often to very often due to a lack of controls on the data once shared.[9]

The last issue needs an enhanced technology approach because it cannot be addressed entirely by encryption. While encryption can absolutely protect against unauthorized access, it lacks the ability to fully protect shared data. First, if someone makes a mistake in allowing access to the information (e.g., through an autocomplete or other typographical error) they are reliant on that person to destroy the information once notified. By that time, the damage may be done. Once shared legitimately and the trusted person decrypts it, data protection then falls to the new data custodian, whether the original sharer wants it to or not. That person is then free to share at their own discretion, out of sight and control of the sharer.

Data protection needs another step and that seems to be recognized by those sharing data. Forty-four percent of research respondents said they had funded and were already searching for a means to put better controls on shared data. An additional 35% had a project identified to start in the next year.

5 SANS Institute
6 https://www.breachlevelindex.com
7 EMA 2019 Security Megatrends
8 EMA 2019 Security Megatrends
9 EMA 2019 Security Megatrends

DATA PROTECTION VS. DATA PRIVACY

Today, data protection and data privacy are frequently used synonymously. This is truer with people who work outside of IT and security, but is still used erroneously across roles. The confusion is understandable because the difference can be subtle unless dutifully explained. Once people are shown the difference it usually sticks with them, so it is important to address the difference in the context of this paper.

Data protection is concerned with the triad of confidentiality, integrity, and availability to secure data against unauthorized access for any type of sensitive or confidential data. Data privacy is specifically focused on the subset of personally identifiable information (PII). It is directed at authorizing access and who has the authority to grant and revoke that access. Data protection focuses on technologies to secure access, whereas data privacy focuses on the legalities, policies, and procedures that govern the use of PII when it is shared.

Organizations drive data protection to maintain business advantage and differentiation. Consumers drive data privacy through representative legislators to protect personal information they have shared with the organizations in the course of obtaining products and services, or behavioral information that is being collected by service providers (like Internet access and search engines).

Until recently, the approaches to managing both protection and privacy were focused on identity and access management and encryption. However, they have different objectives, so the technology for protection has been insufficient to maintain strong and persistent privacy. Also, existing constructs do not enforce end-to-end data protection and governance. They rely on trusted parties to appropriately maintain data and provide no construct or information that can verify whether trusted parties remain faithful to agreements on data handling and exposure.

Prior to stringent privacy legislation, both datasets were a revenue source and/or differentiator for businesses. Personal data may be maintained internally, but in many cases, subsets of the data were sold and traded with no regard for or influence by the data subject. Intellectual property (IP) was also maintained or sold at the sole discretion of the business data owners. Since the expansion of privacy legislation, it is becoming more difficult to collect and sell privacy-related information without the data subject's consent. Fines for noncompliance are becoming more significant, making data collection and holding riskier, thus driving greater scrutiny on balancing risk and reward on the part of the collectors. Both are targets of theft, but the reduced revenue opportunities for privacy-related data make maintaining that data a greater liability with reduced financial gain. That, in turn, incentivizes companies to reduce the amount of privacy data they collect and maintain. Organizations are still highly incentivized to maintain all the intellectual property they can create or acquire.

CHALLENGES OF MAINTAINING DATA PROTECTION AND PRIVACY

All types of sensitive data and information are targets of theft. Where data is located only minimally changes the protection requirements for normal business operations. Exceptions to this include sensitive data placed in hostile environments, like battlefields or traveling executives whose computing devices have resident IP when in foreign countries.

When using a single public or private cloud provider or on-premises applications, data should generally remain within the perimeter or move out of it only via predetermined paths. Access management is used as the primary control with privacy, finance/payment, and IP being encrypted. If the data attempts to leave the cloud, a CASB or DLP solution can often aid in identifying and/or stopping the infraction.

Hybrid cloud and hybrid multi-cloud environments use similar tools with the addition of established data communication conduits between the applications and clouds.

Community cloud can leverage the same controls as the other clouds. The change here is the data sharing. All data is not necessarily shared openly amongst all members within the community. It varies by relationship, role and responsibilities, contractual agreements, and as new data types/classifications are introduced into the community.

The key factor for reliably protecting data is controlling who should have access and how the data is supposed to move and be distributed or shared within the environment. If the data stays resident within the company-owned or leased environment, the risk to the business exists but is significantly less than if it is transitory in and out of the environment. If the monitoring and defensive controls are in place to watch for it moving through the boundaries of authorized locations, applications, or user accounts, then the tools and processes for protection are similar. However, if the data is shared, the tools and processes for maintaining protection as bound by the privacy policies are significantly different, driving the need for data protection tools that are persistent with the data. Only this way can the original data owner/custodian continue to control access even after the data leaves the environment. Normal encryption schemes do not provide this level of protection.

MAINTAIN DATA CONTROL EVERYWHERE WITH DATA PRIVACY PASSPORTS

IBM Data Privacy Passports (DPP) is designed to extend data access controls beyond the controlled environment and maintain it throughout the data lifecycle. With DPP, the data policy owner not only controls access, but also controls the lifespan of the data. IBM Z's inherent security, scalability, and resilience provide the optimum delivery platform for Data Privacy Passports, enforcing policy controls on any data regardless of its origination or destination. The following section discusses the features and components that accomplish this.

Overview of Information Rights Management Using Data Privacy Passports

Modern encryption has been used successfully to protect data for many years. So long as the encryption keys are properly protected and all parties remain trustworthy, encryption protects the data without fail. IBM introduced pervasive encryption to extend the protections from encryption across the system, so if the systems or applications leveraging it are compromised, the encryption would still be effective in protecting the data. However, pervasive encryption suffers from the same limitation that other encryption schemes have. Neither traditional encryption nor pervasive encryption are able to protect data once management of the encryption policy leaves the data owner/custodian's control.

How it Works: Overview

DPP is currently only supported for deployment in the IBM Z environment—specifically the brain of the system, called the Data Privacy Passports Controller. While it is only deployable on Z, it operates to protect any structured data located on-premises or in the cloud throughout its desired lifecycle. DPP uses world-standard AES256 encryption, extending data protection and data access management beyond normal encryption use cases. It is like other data protection tools in that it applies the data security policy to the data at the source. Unlike other common-use tools, it enforces the policy at the receiving end with the data so the data owner maintains persistent and continuous control regardless of where the data resides.

The IBM Z Data Privacy Passports Controller acts as a proxy to intercept the JDBC call for Db2, IBM DVM for sequential data. The controller is called through a standard Apache Hive driver and uses ordinary protocols to hook into the authentication tool of choice, such as LDAP or AD, so no additional identity management database must be maintained.

The Passports Controller maintains the policy as constructed by the data owner/custodian. It also maintains the encryption/decryption keys and the logic control engine for managing access and requests. To start enforcing a new policy or allow access to changed source data, the Passports Controller must be able to communicate with the originating database and the endpoint receiving the copied data. If either is not true, no transaction is permitted. To allow actions on shared data protected by existing policies, the Passports Controller must be able to communicate with the endpoint hosting the shared data. If a data copy is accessed locally on the endpoint where it resides, no communication with the Passports Controller is necessary and the policy remains in force.

While policies can be created at the table level or the attribute/column level to facilitate custom data views, enforcement is performed at the field level. Policies can be updated dynamically. Updated enforcement is applied to the Passport Controller the next time it processes a Trusted Data Object (TDO). The TDO is a wrapper that holds the shared data and maintains some policy information as metadata. The policy stays with the data copy throughout its lifespan. Data lifespan can be expressed in the policy destroying keys at the specified date, rendering the data inert. The data lifetime can be extended at any time prior to expiration.

Since the policy is maintained with the data, it is always in full force. Should the data owner wish to change the policy in any way, including revoking access totally, the Passports Controller does not attempt to delete data remotely. It merely destroys the encryption keys required for access to the trusted data object, rendering the data inaccessible from that point forward. A new dataset with the updated policy can be downloaded by those who still have access. Their own encryption key for access can represent each user or group of users. This representation is based on the way the policy is defined. If the access policy is defined by an LDAP or AD group, that group is issued a key. If it is defined using individuals, each individual is issued a separate key. If any group is modified the access authentication changes, with new members being granted access to the key(s) and retired members being denied.

Different keys can protect different tables. Should a member authorized to access one table attempt to join data with a table that he/she is not allowed to access, the join is denied. Should multiple members authorized to access different tables attempt to collude to join tables, the join is denied.

BUSINESS USE CASES

When DPP is being used to control data shared with third parties, it must be accessible for communications from the third party when accessing protected data in a TDO. As with most security devices, it is recommended that it not be generally accessible from the Internet, so third-party verification should be done via a VPN or other policy-controlled gating mechanism.

Protecting Structured Data On-Premises, in the Cloud, or Other Shared Environments

If sensitive structured data is shared with anyone, internal or external to the business, DPP delivers granular, dynamic, and revocable data control to the data owner/custodian regardless of where the data starts or ends. So long as the Passports Controller is reachable, the data can be accessed per policy.

Data Segmentation and Brokering

For circumstances in which a data owner is engaged with one or more third parties, the granular policy control offers the ability to control which parties can use and combine which data. This can be highly useful in controlling date flows in blind and double-blind testing formats and in maintaining privacy during largescale statistical analysis. No one party has all of the data or has access to enough personal data to break privacy. Data expiry is also useful for any time-bound shared data, as the policy can render the data inert at the time of expiration of lease or contractual agreements.

Data Access Control for Compliance

Compliance allows the data to be used for any authorized purpose without violating privacy. The policies are documented and logging can be used to validate enforcement. Change control logging can be used to validate when the policies were enabled and enforced.

Embedded Data Retention Policies

Each industry has its own regulations and best practices for data retention. Policies can be constructed to destroy the use keys for the data, rendering it inaccessible at the time of expiry.

Data Deduplication

If DPP is applied to all data at the source when it is created, the policies not only control who can make copies, but all undesired copies can be destroyed. Another interesting benefit of the TDO implementation is that only one copy of the TDO is required while providing diverse access if the policy specifies different data views for alternate roles, since the policy is resident with the data.

EMA PERSPECTIVE

Sixty-six percent of organizations EMA surveyed said they are looking to make a technology investment to reduce data exposures.[10] Most organizations are driven to data at rest encryption in the traditional form because of a lack of options, and they would rather move forward incrementally than not at all to show due diligence. However, information rights management via DPP is by far the better choice for any area in which data is to be shared, either internally or externally.

Using Data Privacy Passports, data control remains in the hands of the data owner/custodian who is most knowledgeable about who should have what types of access. It is no longer thrust on the security team, which can become a layer of abstraction for constructing data policies and thus an unwitting accomplice to data risk. As a best practice, a defined data lifespan is recommended at the start of sharing. The lifespan can be as long or as short as the data's profile permits. If it exists at the beginning, then if the recipient goes offline for an extended period, the data will become inaccessible by default when the time expires. It will remain offline until the endpoint reconnects to the Internet and gets the policy and TDO updates, granting renewed access.

DPP is the first information rights management tool to be applied to structured data outside the primary data environment. IBM has developed a data protection ecosystem based on the Z platform that will aid organizations in meeting the ever-increasing data protection and data privacy requirements. With pervasive encryption, IBM can protect any data at rest on the Z platform in storage or in processing. With their newly released Data Privacy Passports, IBM Z extends the protection to structured data located on-premises, hosted, or in any cloud configuration. With Data Privacy Passports, the structured data need not reside or originate on the Z platform. Just as importantly, the data can be protected through its entire lifecycle, both in its original location and wherever it is shared, with the original data owner/custodian retaining complete control of who has access to all shared copies—including the data lifespan.

ABOUT IBM

For more information about IBM, visit https://www.ibm.com.

10 EMA 2019 Security Megatrends

About Enterprise Management Associates, Inc.

Founded in 1996, Enterprise Management Associates (EMA) is a leading industry analyst firm that provides deep insight across the full spectrum of IT and data management technologies. EMA analysts leverage a unique combination of practical experience, insight into industry best practices, and in-depth knowledge of current and planned vendor solutions to help EMA's clients achieve their goals. Learn more about EMA research, analysis, and consulting services for enterprise line of business users, IT professionals, and IT vendors at www.enterprisemanagement.com or blog.enterprisemanagement.com. You can also follow EMA on Twitter, Facebook, or LinkedIn.

Corporate Headquarters:
1995 North 57th Court, Suite 120
Boulder, CO 80301
Phone: +1 303.543.9500
Fax: +1 303.543.7687
www.enterprisemanagement.com

3878.082019