REPORT REPRINT

Architectural data control: turning privacy requirements into a blessing, not a curse

JANUARY 22 2019

By Paige Bartley

Most organizations view data privacy and data protection regulation as a burden. But the core data management capability required for compliance – granular data control – is also necessary for proactive leveraging of data. Data control and architectural optimization is a strategic opportunity.

Consumer awareness of privacy, and corresponding outcries over privacy infringement, are having tangible effects on businesses. Data breaches, unauthorized third-party sharing of data, and the lack of data portability have all become pain points for today's digital consumer, eroding trust in brands.

In the EU, the General Data Protection Regulation (GDPR) aspired to set a 'gold standard' for the management and processing of personal data, and many other jurisdictions soon forged mirroring policy. With global regulatory requirements proliferating, many businesses are having trouble keeping up. The challenge is not one of simply increasing security or meeting individual technical requirements, but rather one of improving and maintaining core data-control capabilities. For many organizations, GDPR was a wakeup call, highlighting the painful truth that their data management architecture was siloed and ineffective. Data privacy, and the ability to achieve it, requires foundational data control capabilities that have downstream benefits to other data-driven initiatives.

451 TAKE

Data privacy, and the regulatory obligations associated with it, are frequently perceived by the enterprise as a net burden in both cost and time. However, there lies immense business opportunity in the requirements to control personal data. Any data-driven regulation, such as GDPR, is a mandate to gain better control of informational assets. Better control of data, ultimately, has downstream benefits for the entire enterprise IT ecosystem and those that depend on it – driving increased data quality and providing more relevant inputs to applications such as self-service analytics tools. Both regulatory compliance and effective leverage of data share the common requirement of granular data control, which needs to be addressed at the architectural level.

Regulatory requirements are often approached by the enterprise as a 'checkbox' list of technical requirements that must be fulfilled to obtain compliance for a particular mandate. This approach is neither scalable nor sustainable in a world where privacy and data protection regulation is proliferating, and regional rules are defined by their idiosyncratic nuances.

Data privacy as a concept is largely agreed upon, but the specific rules for ensuring its achievement can vary widely based on jurisdiction. For the enterprise, the implementation of a stand-alone compliance tool for each new regulation that evolves is simply not economically feasible. As a methodology, it is also prone to creating data silos and challenges with data integration. What organizations must increasingly do, as regulations become more numerous and complex, is focus on key data protection and data privacy principles, which are shared across regulatory frameworks.

The common denominator of all data-driven regulation is the requirement for complete, granular control of data within the enterprise IT ecosystem. Organizations cannot protect or provide privacy controls for data if they cannot quickly and consistently locate data, identify and resolve duplicates, accurately associate personal information with identities, and enforce policies. Both structured and unstructured data must be controlled with the same rigor, as today's regulatory definitions of personal data increasingly encompass sources of information that are textual and social in nature, rather than just traditional structured identifiers such as credit card numbers, ID numbers and phone numbers.

Siloed architecture has long been a barrier to this goal of unified control for data. Different applications and repositories each have different capabilities for search and data policy management. Lack of integration prevents the enterprise from obtaining a unified view and access of informational assets. Not only does this create major challenges for meeting regulatory requirements, it also severely limits an organization's ability to leverage data for insight. In 451 Research's Voice of the Enterprise: Data and Analytics survey in late 2018, 'accessing and preparing data' was the most commonly reported barrier to using data platforms and analytics, with 19% of respondents reporting it as the most significant barrier. In this sense, both reactive compliance capabilities and proactive use of data are two sides of the same data control coin.

Privacy requirements as an opportunity, rather than a burden

Today's data privacy and data protection requirements, then, should be viewed by the enterprise as an opportunity to optimize data management architecture from the ground up. For many organizations, particularly those that were in traditionally unregulated industries, the attempt to comply with GDPR revealed the true lack of data control that was endemic to IT ecosystems.

That doesn't mean that organizations hadn't felt the pain of insufficient data management capabilities before. Poor data quality, difficulty in gaining a single view of customers, and unnecessary duplicative knowledge worker effort all have been symptoms of this underlying data control disorder. Regulatory mandates simply served as the external impetus for many businesses to seriously reconsider and reassess their data management practices. As data increasingly becomes the enterprise's most valuable asset, it also becomes its biggest risk factor. Complete control of data, both structured and unstructured, is the foundational requirement for the enterprise to defensibly derive value from information.

A shift from data quantity to data quality

The enterprise, in the big-data era, has a persistent phobia of any externally imposed restriction that is perceived to reduce the access to, or the collection of, data. If big data is good, all data must be better – at least the reasoning goes. Key evolutionary trends in computing, such as separation of storage and compute, have made it feasible to collect and store a dizzying array of data sources indefinitely. Most organizations today have a data 'hoarding' philosophy because the economics of modern storage allow for it, and the perception that the data might eventually become useful at some point is pervasive.

Unfortunately for the enterprise, this methodology is directly contradictory to the data minimization principles that are widely shared across many of today's data protection and data privacy regulations. Thus, modern compliance requirements are typically perceived as a major point of friction to the business's overarching strategic objectives, which are to essentially collect and analyze as much data as technologically feasible. It is assumed that the consumer, given more autonomy and control over their data, will share less information. Data minimization principles are assumed to reduce the volume of data available for analysis. Less data, it is assumed, is always a strategic disadvantage.

While this perception is pervasive and is based on a kernel of truth – consumer controls for privacy do, objectively, reduce the volume of data available for unrestricted analysis – it overlooks several inherent benefits. The data control mechanisms that are a core requirement for compliance have broad downstream benefits for data-driven initiatives such as self-service analytics, because they facilitate the administration of granular data access permissions and allow the enterprise to better understand which informational resources are most relevant and representative of the pressing business questions that need to be answered. With strong data control capabilities, the effects of silos are also minimized, resulting in the ability to aggregate and analyze diverse data sources in a more contextual way.

Data privacy and data protection mandates, effectively, shift the balance of power back from data quantity to data quality. Inherently, the strong data control required for compliance fortifies data quality initiatives, making it easier to identify and resolve duplicate and near-duplicate data. The ability to accurately associate diverse data sources with individual identities, critical for fulfilling data subject rights, is essentially the same construct as customer 360° initiatives. While the data available for unrestricted analysis and processing may be lower in theory, the enterprise will likely find that the pressure to resolve issues with silos will ultimately make more quality data available for analysis.

Trust as a driver of profitability

When consumers or data subjects are given more choices over the use of their data, they may indeed choose to restrict the volume of information that they provide. However, when given these choices and autonomy, trust is fostered. When a trusting relationship is built, more accurate information is volunteered over time. Less obfuscation behavior, such as providing junk email addresses, occurs. Consumer trust, in turn, is correlated with more profitable lifetime relationships, lower churn, and more positive word-of-mouth presence in the market.

As consumers become more aware of the value of their data, and are increasingly given regulatory rights to take control of it, trust will become a competitive differentiator for organizations. Strong data management practices, and the corresponding ability to swiftly fulfil consumer requests for data control and access, are the bedrock of this relationship. Data management, too, is inherently tied to data security capabilities; consumers are acutely wary of headline data breaches. Once a trusted relationship is established, consumers are more willing to selectively and voluntarily share accurate personal data in exchange for perceived valuable benefits, such as special offers and highly personalized recommendations. The crux here is that consumers, to be motivated to share their accurate personal data in the privacy regulation era, must believe they are getting something of equal value in return – the relationship is transactional.

Current enterprise data analysis trends and the intensifying regulatory landscape bring the issue of consumer trust to the forefront. 451 Research's Voice of the Enterprise: Data and Analytics survey in late 2018 found that consumer behavior data is still the most popular data source for analysis, with 58% of respondents reporting that their organization analyzes it. These organizations are at inherently higher risk of running afoul of data privacy and protection regulations' restrictions on the processing of personal data, and are additionally at risk of raising the ire of increasingly privacy-aware consumers.

Achieving data control, from the ground up

Data control, in summary, is the common requirement for both reactive compliance and proactive data leverage capabilities. It is also, ultimately, essential to building the trust with consumers that drives long-term profitability. If the enterprise is to strategically fulfil compliance requirements while maintaining the ability to competitively maximize the insight it derives from data, to the extent allowed by regulations, it must optimize its data management architecture and strive toward a unified view of data.

This runs contrary to the procurement methodology that many organizations have taken when faced with new regulatory requirements. It is tempting to break down regulations into a list of their respective technical requirements, and purchase a specialty tool (or set of tools) capable of helping achieve them. This ultimately is a temporary approach that doesn't address the much more difficult and enterprise-wide challenge of underlying architectural optimization. Worse, a narrow 'solution' approach to compliance can exacerbate existing architectural challenges, spawning additional silos and making data integration more difficult.

For long-term competitive viability in the data protection and privacy era, organizations today must begin striving toward data control that is achieved from the bottom up via architectural optimization, rather than top down with individual tools. This isn't just a technology problem; it is a people and process challenge as well. The control of data, today, has many more diverse stakeholders than the past, when IT was primarily responsible for architectural and platform decisions. Communication is key, and alignment of objectives at the highest level (with accompanying executive sponsorship) is critical. While beyond the scope of this report, this procedural challenge is something we will continue to address in our research throughout the year.