R&D Outsourcing

25 2023

Data management

The relevance of the topic of data management (Data Governance) is growing every year. Indeed, the need to organize processes aimed at improving the efficiency of collecting, processing, storing and using data as a valuable asset is already obvious to almost all companies. Much has been said about the benefits that companies can get from having well-designed data management processes, and many organizations have already begun implementing this initiative. At the same time, organizations often make similar mistakes that negatively affect the pace of implementation and the effectiveness of the created data management processes. Svetlana Bova, Chief Data Officer of VTB Bank, tells about what these mistakes are, how to avoid them and what questions the organization should find answers in the process of implementing Data Governance in the material prepared for TAdviser.

Data Quality Management
Main article: Data quality management

The definition of data quality is formulated as a generalized concept of data usefulness, formalized in a certain set of criteria. For corporate data of management information systems, it is customary to distinguish the following six criteria: relevance, accuracy, consistency, timeliness, accessibility and interpretability. For each criterion, a set of key performance indicators (KPIs) is determined and practices are developed to improve them (more).

Data visualization
Main article: Data visualization

Data leaks
Main article: Data leaks

Data protection
Main article: Data protection

Improving Data Management Models
Main article: Improving data management models

Disclosure, use and sale of data
2022: How to benefit from simple and secure data exchange
Thanks to advances in data sharing technologies, in the first half of 2022, there is an opportunity to buy and sell potentially valuable information in highly efficient cloud marketplaces. By combining this data with a new array of privacy technologies, such as Fully Homomorphic Encryption (FHE) and Differential Privacy, encrypted data can be shared and computed without the need for initial decryption. This provides the emergence of new opportunities: the exchange of data while maintaining security and confidentiality. All this contributed to the emergence of new promising trends. Sensitive data stores, servers around the world, due to privacy issues or regulatory requirements, are starting to generate value for enterprises in the form of new business models and opportunities. In 2022, a growing number of organizations are expected to explore the possibility of seamless and secure data sharing, and opportunities that will help them monetize their own

information assets and fulfill business goals using other people's data.

This data sharing trend is gaining momentum. A Forrester Research survey released in March 2021 [1] found that more than 70% of data and analytics decision makers are expanding their ability to use external data, with another 17% planning to implement in 2022.

Moreover, the global FHE market alone is growing at an annual rate of 7.5% and is expected to reach $437 million by 2028. In 2022, healthcare and finance are the sectors that lead the most FHE research.

What explains this growth? Simply put, data acquires value when it is shared. Gartner predicts that by 2023, organizations that promote data sharing will outperform their competitors on most business metrics [2].

The following scenarios illustrate data sharing in action:

Using aggregated data to securely achieve common goals. Even competing organizations will be able to work together towards common goals, such as deepening customer insights or detecting fraud patterns across the sector.
Increase efficiency and reduce costs. Data providers no longer need to provide hardware to enterprises, maintain databases, and create Application Programming Interfaces (APIs). The client only needs to press a button to access anonymous, curated data feeds. Inside the enterprise, encrypted data makes working with artificial intelligence and machine learning safer, and compliance auditing easier.
Expansion of cooperation in the field of research. Sharing fundamental or early results can accelerate critical research initiatives without compromising hard-won competitive advantage.
Protection of intellectual property. Ultra-sensitive data, such as AI training data, can be stored in open clouds but still be more secure.
Encryption of data in motion. In the fields of high-frequency trading, robotic surgery, and smart manufacturing, sensitive data must

s be transmitted quickly. FHE allows users to access critical data quickly without encryption keys.
Opportunities like this to monetize data through sharing and aggregation can offer many competitive advantages to first-timers. Then, competitors, seeing that the “first movers” are effectively using technology, will also want to redesign their business with an organizational structure based on the use of data and artificial intelligence.
But, again, unlocking this potential requires a different kind of data management—this time adding innovative technologies and methods that free information assets from traditional constraints to privacy and security.

The data trend in 2022 has three main dimensions: how to take advantage of opportunities, ease of use, and privacy.

New business models and opportunities
Shared data can create shared opportunities and new business models. As the data sharing trend continues, Deloitte expects more organizations to engage in “data collaboration” to solve common problems and exploit mutually beneficial benefits, operational and research opportunities.

In addition, the ability to securely exchange data with external data management service providers can help organizations streamline data management processes and reduce associated costs.

The following possibilities are being considered that can be opened up by data exchange:

"Industry vertical marketplaces". Even the fiercest competitors often face common challenges that are best solved through collaboration. For example, food industry suppliers: if everyone could anonymize sensitive sales and supply data and aggregate it for analysis, perhaps these suppliers could unlock the "mystery" of supply and demand. Or banks in developing regions could combine anonymous credit data to create an interbank credit risk assessment system. Or one of the biggest opportunities: Can pharmaceutical researchers and physicians working within a secure ecosystem combine data to understand how to bring life-saving innovations to market faster?
"Partners in the value chain". Many manufacturers and retailers buy consumer data from third-party data brokers, but as is often the case, quality data is not enough to make the right decision. What if partner systems in the value chain—from suppliers to manufacturers to marketers—combined data about their customers to create a more granular picture of demand?
"Let someone else train the AI model." AI models are often considered highly sensitive forms of intellectual property. Because they usually fit on a flash drive, they also pose a high security risk, which is why many organizations have traditionally done their own simulations in-house. Thanks to encryption technologies, this may soon change. By protecting simulation data, CIOs can securely outsource simulation and AI training to third parties.
Data providers optimize delivery. On data exchange platforms, it will be possible to buy access to real-time market or logistics data as simply as pressing a button. Data providers will no longer need to provide an API or submit files.

Link to article:: Data management