# Structured Data Archiving for Application Retirement

A Guide to Structured Data Archiving and Decommissioning your Legacy Applications

# Structured Data Archiving for Application Retirement

## Introduction

Organizations routinely look to retire little-used, obsolete, or outdated applications. According to a recent survey conducted by Gartner Group[1], managing the technical debt associated with legacy applications is one of the top three priorities for IT leaders worldwide. The key drivers for this are:

- Reduce legal, regulatory and operational risks
- Reduce on-premises application infrastructure costs
- Reduce IT operational costs
- Cloud adoption
- Mergers and acquisitions

The issue most organizations face when contemplating application retirement is what to do with the structured data within these applications: delete it along with the retiring application or archive it for business continuity, regulatory compliance, and eDiscovery requirements. Deleting the data without using it to drive analytical insights and AI models, knowing if it's subject to regulatory requirements or potentially involved in litigation (eDiscovery/spoliation) is a considerable risk. The challenge with archiving structured data is ensuring that potentially sensitive data is securely managed, and that content can be accurately searched and presented in a viewable and usable format.

Most legacy application retirement candidates are on-premises, third party, or in-house developed back-office applications (such as CRM, ERP, financial and HR applications) that generate structured data sets. In most cases, the data is stored in a non-standard file format in a custom storage repository. Data search and retrieval is dependent on the application. Reliance on the individual application makes it much more difficult to simply copy the structured data to a common/centralized data repository for later search and reference.

## Structured Data Archives

Organizations that decide to move data to an archive need to consider how the product controls the migration to ensure it is optimized to handle structured data archiving (SDA). Specifically, the SDA solution must maintain data relationships and access controls while enabling users to search, view, and report on the data. SDA solutions include active database archives and enterprise archives (Figure 1).

### Active Database Archives

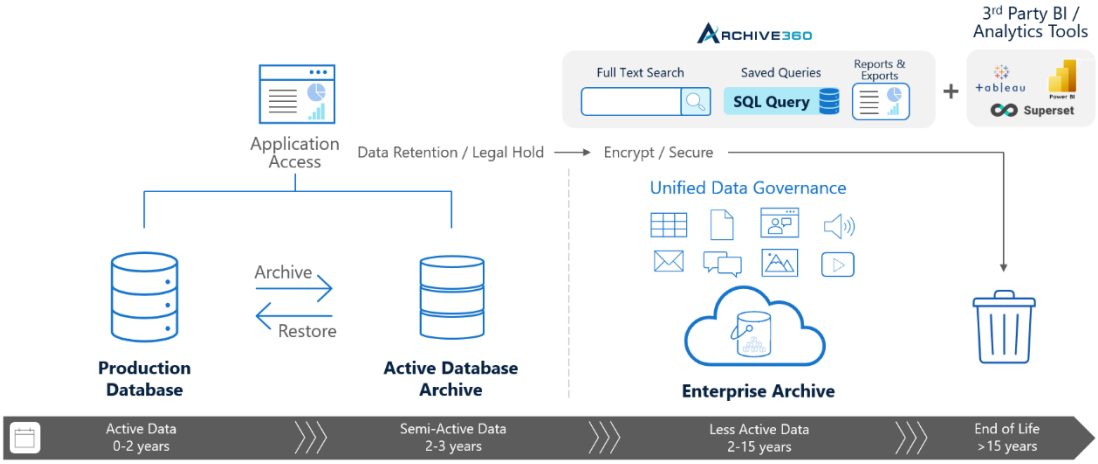
Real-time analytics changed what we considered to be active data. In the past, the most common archive decision point was “last modification date”. Data that is X years old can be archived, everything else must stay in production. Real-time analytics changed that decision point. Instead of last modification date, the key archive decision point is “what data will need to be quickly analyzed”. Active database archives help organizations improve the performance of their enterprise applications, reduce maintenance and overhead costs, and meet their SLAs by removing semi-active data from the production system but making that data readily available to users of the production system should it ever be needed. Active database archives replicate the production system’s data model to be able to quickly restore the data to the production system. The stricter performance and infrastructure requirements usually make active database archives more costly than enterprise archives.

### Enterprise Archives

An enterprise archive, on the other hand, is used to preserve inactive and compliant data, make the data available through the archive if needed, and manage its lifecycle. This type of structured data archive is best suited for application decommissioning. Structured data, such as transactions, and associated

unstructured data, such as contracts, are managed as business objects with metadata or information about the data, including its table, relationships, and access controls, so users can either search or run SQL queries against the data and analyze it like they did in the production system.

### Structured Data Archiving (Figure 1)




## Why archive structured data?

Enterprise Archiving solutions provide a way to cost-effectively store and manage structured data sets for later reference, regulatory compliance, litigation response, and data consolidation because of a merger or acquisition.

**Regulatory Compliance:** Many organizations have structured data applications that contain data subject to regulatory retention requirements for extended periods. For example, SEC Rule 17a-4 requires broker-dealers to preserve their books and records for no less than 3years. The Sarbanes-Oxley Act requires publicly traded companies to retain financial records and reports for seven years. Nearly all industries have various retention requirements – Financial Services, Pharmaceuticals, Energy, Healthcare, Public Sector, etc. Regulatory retention requirements extend beyond defining a period. They often include being able to:

- find and furnish the data promptly when requested.
- establish the data's integrity or authenticity.
- protect the data from loss.
- hold organizations accountable for how their data is managed (internally & by third parties).
- provide an audit trail of all activities associated with the data (e.g.- access, modification, deletion)

Organizations unable to comply with legal discovery requests or their regulatory obligations risk fines, such as the privacy law fines listed in Table 1, and/or sanctions including:

- dismissal of all claims.
- taking measures to remedy the resulting prejudice.
- payment of expenses, including attorney's fees, incurred by other parties.
- monetary awards to the opposing party.

In an age of artificial intelligence, machine learning, cybersecurity threats and data breaches, data has become more valuable and data ownership and privacy has become increasingly important. Data privacy

regulations have redefined how organizations must manage their data. Data is no longer simply controlled by the people or organizations who collect it. According to GDPR and other privacy regulations, individuals have ownership rights over their personal data. They have the right to determine who can possess their personal data, the right to access that data, the right to dictate how that data is used, the right to dispose of that data, and the right to get value from that personal data. As a result, organizations are increasingly obligated to understand what data they have, not just to meet regulatory requirements but to meet new customer expectations regarding the handling of their personal data and to protect their organization's reputation. Another data ownership factor to consider is the rise of data sovereignty laws. Countries have always controlled National Security related information, but with an increasing focus on data privacy, they are now looking at data types that impact their national interests, including data associated with their citizens, infrastructure, and financial markets. Having data subject to data sovereignty regulations trapped in a legacy system that doesn't comply with the law is one more reason to retire the system.

Organizations can no longer afford to let data age-out in legacy systems. Particularly when systems are no longer supported, haven't received security updates, or don't provide the controls necessary for effective governance to meet data privacy, security, or sovereignty requirements, and are at risk of a data breach. Structured Data Archiving enables organizations to cost-effectively retire their legacy applications and meet legal discovery and regulatory requirements.

## Consumer Data Privacy Law Comparison (Table 1)

| Data protection rights | EU (GDPR) | California (CCPA & CPRA) | Virginia (VDPA) | Connecticut (CDPA) | Canada (CPPA) |
|---|---|---|---|---|---|
| Effective Date | May 2018 | Jul. 2023 | Jul. 2023 | Jul. 2023 | Jun. 2022[1] |
| Right to be informed | ☑ | ☑ | ☑ | ☑ | ☑ |
| Right to access | ☑ | ☑ | ☑ | ☑ | ☑ |
| Right to correct | ☑ | ☑ | ☑ | ☑ | ☑ |
| Right to delete | ☑ | ☑ | ☑ | ☑ | ☑ |
| Right to data portability | ☑ | ☑ | ☑ | ☑ | ☑ |
| Right not to be profiled[2] | ☑ | ☑ | ☑ | ☑ | ☑ |
| Right to data minimization | ☑ | ☑ | ☑ | ☑ | ☑ |
| Deadline for response | **1 month** (+2 month extensions) | 10 business days to confirm receipt<br>**45 days** to respond<br>(+45 day extension) | **45 days** (+45 day extension) | **45 days** (+45 day extension) | **30 days** |

| | | | | | |
|---|---|---|---|---|---|
| Fines | **4% global revenue** or **€20 million**, whichever is higher | **$2,500 per violation** or **$7,500 per violation** for intentional violations & violations involving children's data | **$7,500 per violation** | **$5,000 per violation** | **3% of global revenue**, or **$10,000,000 CAD** whichever is higher |

[1] Re-introduced in Bill C-27. Not enacted yet.
[2] Otherwise known as the right to opt-out of automated decision making.

**Operational Readiness:** Regulatory retention, privacy, and security requirements are important reasons for organizations to keep legacy application data, but that data can also continue to have operational value. If the legacy application contains data related to existing long-term customers, partners, and employees, keeping that data available could be important to managing the relationship effectively. Also consider the opportunity costs associated with not understanding (or leveraging) the data in legacy systems that could be useful or harmful. The data may be important for identifying business events, patterns, and trends, creating insights, and improving business decisions. For example, an organization's HR department recently switched to a cloud-based human resource management system (HRMS) and decommissioned their legacy HRMS. They noticed employee absenteeism was on the rise. Since they archived employee data from their legacy HR application prior to its retirement, they could use diagnostic analytics to find out why employees are missing work more often and develop strategies to reverse the trend. AI could also be developed leveraging the legacy employee data to conduct predictive analysis, identify turnover risks and prescribe steps to mitigate those risks.

**Litigation Preparedness/eDiscovery:** Another important reason to archive data from retired applications is for eDiscovery purposes. Relevant data need only be kept and secured with a litigation hold if the company anticipates future litigation or is involved in a lawsuit that could include data from the retiring application. Not properly managing structured data poses the same risks inherent in not managing emails or documents. The hefty costs associated with producing electronically stored information during litigation is no different for structured data as it is for emails or any other form of communication. Many General Counsels recognize that this legacy data could also bolster a future case, so keeping it available is a good strategy.

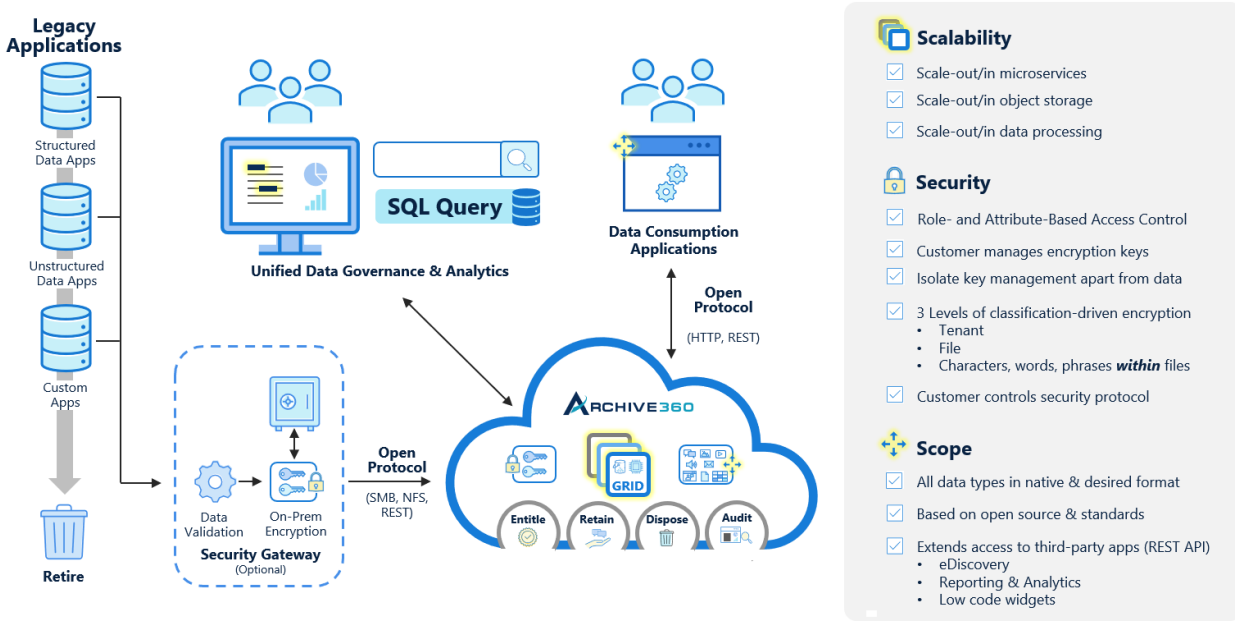
**Mergers and acquisitions:** As organizations merge, data consolidation from competing/duplicate-use applications is often a corporate strategy to reduce costs and increase productivity. By consolidating this data to a centralized archive (with both structured and unstructured data), the company can eliminate duplicate infrastructure and improve the efficiency/productivity of business user access and ongoing retention policy management. Additionally, for inactive structured data still required for regulatory compliance or eDiscovery, archiving the data is the most efficient and cost-effective solution.

## Archive360 and Structured Data Archiving

Archive360's Unified Data Governance platform is a native Azure dedicated SaaS solution that provides organizations the ability to archive legacy system data as well as continually onboard inactive/compliant data from active systems, classify, store (in native format), secure, manage, and find both structured ***and*** unstructured data in their own Azure tenant. The Archive360 platform utilizes the Azure dedicated cloud model to provide users more direct control and customization over their retired data while applying retention/disposition policies, full legal case management, and the highest levels of data security in the

industry. Additionally, Archive360's Unified Data Governance platform allows customers to closely control costs by changing storage tiers and adjusting their application performance based on changing needs.

## Archive360 Legacy Application Retirement




When considering Structured Data Archiving (SDA) alternatives, the solution you select should:

- **Accommodate all types of data from all types of systems to facilitate data governance, discovery, and usage.**
  Instead of relying on administrators and application access controls to manage data and access, implement a common framework and governance platform. The framework enables you to provide consistent data naming, structures, and standards. The data governance platform directly applies the framework to data from systems across your organization and enables your legal and compliance teams to automatically manage data risks while providing secure access to business teams and facilitating data discovery and usage. With the Archive360 Unified Data Governance platform, its class-based approach provides tremendous data management flexibility. Rather than using a single, overarching construct, your Archive360 framework can grow with the addition of each new data source. This allows organizations to more granularly manage the storage, privacy, security, retention, and disposition requirements of all your data. Data can be collected from all types of systems including communication, collaboration, file system, content management, and structured business applications. The platform centralizes control of archived data storage, entitlement, retention, disposition, legal holds, reporting, and eDiscovery. In addition, Archive360 enables analytics, making data available for AI/ML and business intelligence applications.
- **Scale to process an ever-increasing volume of data and accelerate data migration, discovery, management, and analysis.**
  Multi-tenant SaaS archives need to balance data processing resources across all their tenants. They can limit the flow of your data into and out of their platform.as well as limit search

performance, particularly when large data sets are involved. Unlike multi-tenant SaaS archives, Archive360 gives you complete control over how quickly your data is processed. You can scale data ingestion, search, and export performance to meet your schedule and cost requirements.

- **Enable you to easily meet your data privacy, security, residency, legal hold, retention, and disposition requirements**.
  Automatically analyze, migrate, tier, and replicate data from operational systems to our governance platform, in accordance with your policies. Data can be collected from all types of systems without disrupting users, stored in its native format, and secured while ensuring accessibility based on user entitlements and business need (e.g., eDiscovery, analytics). Our unique, dedicated SaaS platform gives you greater control of your data's storage and processing costs, storage location, ingestion and search performance, security protocols, encryption keys, and user entitlements than any other solution on the market. Policy-driven data governance ensures your data is managed consistently, efficiently, and cost-effectively to meet your organization's legal, regulatory, and business requirements.

- **Give you all the tools necessary to follow the principle of least privilege and balance data access and data protection**.
  Deployed in your dedicated Azure cloud tenant, you retain ultimate control over your tenant security. You can deploy your preferred security applications for monitoring, vulnerability scanning, and logging, integrating directly with your SIEM for control not found in other cloud archiving solutions. Our platform provides role-based and attribute-based data access control (such as case status, project schedule, IP address, etc.) and SHA-256 level encryption at multiple levels (tenant, file, and specific content – field, word, or phrase) to ensure the right people or applications have the right access to your data, under the right conditions.

- **Demonstrate data integrity and defensible data management.**
  Data integrity refers to the trustworthiness and accuracy of data throughout its lifecycle. Data integrity is important when evaluating the reliability of information in investigations, litigation, and business analysis. Archive360 enables customers to demonstrate data integrity and compliance in several ways. During data migration, Archive360's platform can securely migrate data 20 times faster than traditional migration tools with full fidelity and a complete chain of custody. Once in the platform, all data can be securely stored in tamperproof WORM storage and consistently controlled according to retention, access, and management policies compliant with applicable laws and regulations. Once the retention period is complete, data disposition is reviewed and approved, producing a certificate of disposition. All activity associated with archived data is recorded in a complete audit trail.

## About Archive360

Archive360 is the unified data governance company transforming how organizations identify, collect, manage, and act on their data. Businesses and government agencies worldwide rely on the security, scalability, and scope of our cloud-native platform to address their increasing data governance obligations across growing volumes of disparate data. With Archive360, our customers are eliminating data silos, securing data access, increasing data insights, while reducing cost and risk. Archive360 is a global organization that delivers its solutions both directly and through a worldwide network of partners.

Archive360 is a Microsoft Cloud Solution Provider, and the Archive2Azure™ solution is Microsoft Azure Certified. To learn more, please visit https://www.archive360.com.



[1] Technical Debt Reduction Roadmap for Large Enterprises, July 2, 2020, Gartner, Inc.

Microsoft Partner
Microsoft
Gold Application Integration
Gold DevOps
Gold Application Development
Gold Cloud Platform
Gold Security

A360WP0011123