

Deloitte.

 databricks

TRANSFORMING PUBLIC SECTOR EFFICIENCY:

*Leveraging modern data management
with Deloitte and Databricks*



Public sector organizations face increasing pressure to deliver more with fewer resources, making cost optimization through effective data management crucial. Deloitte and Databricks have collaborated to offer robust solutions that harness big data, advanced analytics, and machine learning to streamline operations, enhance decision-making, and uncover cost-saving opportunities. This collaboration leverages Deloitte's deep industry experience and Databricks' unified data analytics platform to deliver transformative results.

Data serves as a critical asset for public sector organizations, empowering informed decision-making and enhancing service delivery. However, efficiently managing and processing this data through an enhanced extract, transform, load (ETL) framework is essential to unlocking its full potential.

The role of ETL in today's data management landscape

ETL processes are necessary for integrating, cleansing, and preparing data for analysis. Without efficient ETL processes, organizations cannot fully leverage their data for decision-making and operational improvements. ETL processes transform raw data from various sources into a structured format ready for analysis, essential for identifying inefficiencies, predicting future trends, and allocating resources effectively.

Efficient ETL processes can enhance data management, creating an AI-ready data state that supports advanced analytics and AI applications. Modern data platforms integrate disparate data sources, providing a comprehensive view of organizational data. This enables public sector entities to identify inefficiencies, predict future trends, and allocate resources more effectively. Conversely, outdated ETL tools often lead to data duplication and increased costs.

Challenges of traditional ETL

Traditional ETL processes can be costly for several reasons:

- 1 Complexity and volume of data:** Public sector organizations often handle vast amounts of data from various sources, such as legacy systems, databases, and external data feeds. Integrating and transforming this data into an interoperable data format requires significant investment in computational resources, data management experience, and sophisticated tool procurement.
- 2 Scalability issues:** Traditional ETL tools may struggle to scale efficiently as data volumes grow, leading to performance bottlenecks and necessitating additional hardware and software investments.
- 3 Manual intervention:** Traditional ETL processes often require manual coding and intervention for data transformations, error handling, and data quality checks, increasing labor costs and the potential for human error.
- 4 Maintenance and upgrades:** Maintaining and upgrading traditional ETL systems can be resource-intensive, requiring ongoing investment in skilled personnel and technology as data sources and business requirements evolve.
- 5 Licensing costs:** Many traditional ETL tools come with high licensing fees, adding to the overall cost of data management, especially for public sector organizations operating under tight budget constraints.

Deloitte's AI Ready Data Foundation (AIRDF) services

Deloitte's AI Ready Data Foundation (AIRDF) provides a comprehensive framework to prepare, integrate, govern, and secure data for advanced analytics and AI applications—enabling organizations to enhance data quality and compliance.

By leveraging AIRDF, public sector organizations can streamline their ETL processes, reduce the time and resources required to manage data, and create a robust environment for data-driven decision-making.

Leveraging Databricks for enhanced data analytics

The Databricks Data Intelligence Platform, integrated with AIRDF, delivers advanced tools and scalable infrastructure for modern data analytics and machine learning. Leveraging Apache Spark™, Databricks combines data engineering, science, and analytics within a single platform. Its lakehouse architecture merges the strengths of data lakes and warehouses, enhancing efficiency, reducing costs, and supporting AI-ready data.

Driving efficiency with Databricks

The Databricks Data Intelligence Platform also offers several key features that drive efficiency and value realization:



Cluster management: Auto-scaling and auto-termination features help ensure resources are used only when needed.



Data caching: Reduce repeated data reads to save on compute costs.



Spot instances: Utilize spot instances for non-critical workloads to reduce costs.



Monitoring and alerts: Track resource usage and costs for proactive resource management.



Job scheduling: Run workloads during off-peak hours to save costs.



Photon engine: Accelerate query performance and reduce costs by optimizing SQL workloads.



Storage optimization: Use Delta Lake for efficient data storage and management.

Accelerating data conversion with ConverTech

A component of AiRDF, ConverTech by Deloitte addresses common data conversion challenges such as high costs, data inconsistency, and project delays. Powered by Databricks, ConverTech reduces procurement, setup, integration, and maintenance expenses, while its built-in data quality features enable early issue profiling and seamless integration from various legacy systems. Its ability to automatically scale infrastructure supports consistent performance with large data volumes, making it a cost-effective and reliable choice for data conversion.

*By combining Deloitte's experience in data engineering, cloud migration, and analytics with ConverTech's advanced features and integration with the **Databricks Data Intelligence Platform, organizations gain a comprehensive solution that improves data quality and can reduce total cost of ownership (TCO) by up to 40%.***

Real-world impact: Public sector use cases

The following use cases highlight how public sector organizations have achieved significant efficiency gains, cost reductions, and operational improvements with Deloitte and Databricks:



FEDERAL HEALTH AGENCY

CHALLENGE: A federal health agency struggled to process pharmaceutical application data due to the variety of documents and datasets involved. The labor-intensive review process led to inefficiencies and delays. As data volumes grew, the existing system became inadequate. Integrating historical data with new submissions added further complexity, highlighting the need for a smarter, more scalable solution.

DELOITTE SOLUTION: Databricks was used to centralize the agency's data by configuring over 14 tenant workspaces, developing a Universal Parser for data extraction, and setting up automated data ingestion frameworks. Data governance processes and data catalog models were applied to maintain data quality. CI/CD pipelines with Gitlab and Jenkins were established for automated deployments, and AWS Data Migration Services facilitated the migration of data to AWS RDS Oracle instances.

OUTCOME: The agency saw significantly improved data management and analytics, processing nearly 300 million records daily and ingesting approximately 1,000 data pipelines, feeding over 1,000 reports and dashboards. The optimized data ingestion process reduced design errors by 60% and handled a 98% increase in nightly data volume. The open data lake architecture processed 10 times more data during peak periods. Agency resources spent less time on data management and more time generating insights to enhance the efficiency of pharmaceutical evaluation and research activities.



FEDERAL SECURITY AGENCY:

CHALLENGE: A federal security agency was facing disparate and continuously growing data sources with a need for federation of data to enable cross-agency data sharing and coordination. Lack of visibility across key client entities and reliance on manual processes, delayed data reflected in systems.

DELOITTE SOLUTION: Deloitte utilized Databricks for large-scale, near-real-time replication and ingestion of data, built data transformation pipelines using Apache Spark and developed a three-layer data architecture. Implemented automated data pipelines, granted data access to users with enforced data governance standards.

OUTCOME: Improved efficiency in identifying and matching entities across agencies, enhanced ability to perform historical analysis, enabled senior leadership to make key operational decisions, improved data sharing and service application response times, and provided flexibility to shift resources to address evolving mission priorities.





LARGE STATE HEALTH EXCHANGE PROGRAM: DELOITTE'S DATABRICKS TRANSFORMATION

CHALLENGE: A state health exchange was experiencing significant performance and stability issues with its data warehouse, leading to restricted access and delayed decision-making. Persistent problems included unstable data load jobs, inconsistent data quality, and incomplete data coverage.

DELOITTE SOLUTION: Databricks was utilized to stabilize and optimize data pipelines, improve data quality, and integrate missing data sources. The solution involved uncovering and fixing major architectural flaws and design issues in the data load jobs, making them more efficient. The ETL processes were redesigned and rewritten on Databricks to enhance performance and stability.

RESULTS DELIVERED BY DELOITTE:

- **Faster Insights:** Deloitte's Databricks-enabled solution reduced daily data load times from over 11 hours to approximately 6.5 hours, ensuring timely delivery of critical analytical reports.
- **Dramatic Stability Gains:** Within six months, data warehouse support ticket volumes dropped by 90% as a direct result of Deloitte's enhancements.
- **Scalability and Agility:** The new Databricks-enabled solution can now accommodate 5–10 times greater data volume, with system changes delivered up to three times faster (monthly releases instead of quarterly).
- **Business Value:** Senior leadership now relies on Deloitte's optimized solution to make operational decisions based on timely, high-quality data—unlocking agility and confidence across the organization.

Embracing data-driven efficiency

The Deloitte and Databricks alliance offers advanced data management and analytics capabilities tailored to public sector needs. By leveraging these solutions, public sector leaders can transform data into actionable insights, drive efficiencies, and achieve significant cost savings.

Considerations to start your journey:

To drive effective data transformation, begin by thoroughly assessing your current data management practices to identify inefficiencies and improvement opportunities. Engage with Deloitte and Databricks specialists to customize solutions that address your organization's unique needs. Established frameworks such as Deloitte's AIRDF and ConverTech can reduce costs and improve data quality for your data foundation as well as help you to scale advanced analytics and AI for your organization.

By pursuing these goals, public sector leaders can harness the power of modern data management to drive efficiency, reduce costs, and deliver impactful results for their communities.

READY TO GET STARTED?

Dave Thomas

Principal
GPS AI & Data
davidthomas@deloitte.com

Harsha Goonewardana

Solution Architect
GPS AI & Data
hgoonewardana@deloitte.com

Harsha Musunuru

Managing Director
GPS AI & Data
hmusunuru@deloitte.com

Vijaya Krishna Peddinti

Specialist Leader
GPS AI & Data
vipeddinti@deloitte.com

Emily Cole

Alliance Manager
GPS Databricks
emcole@deloitte.com

This publication contains general information only and Deloitte is not, by means of this publication, rendering accounting, business, financial, investment, legal, tax, or other professional advice or services. This publication is not a substitute for such professional advice or services, nor should it be used as a basis for any decision or action that may affect your business. Before making any decision or taking any action that may affect your business, you should consult a qualified professional advisor. Deloitte shall not be responsible for any loss sustained by any person who relies on this publication.

As used in this document, "Deloitte" means Deloitte Consulting LLP, a subsidiary of Deloitte LLP. Please see www.deloitte.com/us/about for a detailed description of our legal structure. Certain services may not be available to attest clients under the rules and regulations of public accounting.