International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Mainframe Modernization as a Catalyst for Democratized Data and Participatory Analytics

Ashish Prakash Khandelwal

Abstract—Mainframe systems have proved to be highly stable and are being used by organizations in the banking industry, insurance industry, government, and other large-scale manufacturing systems as destination repositories, data processing systems, and mission-critical datasets. The main benefits of the mainframe use are predictability, safety, and performance, but the old architecture has limited information availability and scale of analytical ability, which can be viewed as a breadth of access to a few technical specialists. The growing requirement to use data to make decisions at all levels of the organization has resulted in the strategic concern of procuring legacy data in an accommodating manner. This article explains the capacity in which the existing mainframe migration (hybrid cloud structures, data virtualization, change data capture, and domain-oriented data ownership models) can create democratized access to data and participative analytics. Such modernization will enable the broader adoption of analytics, less dependence on qualified expertise on the mainframe, and accelerate innovation. The paper also addresses the impediments, governance, and strategic considerations that companies should consider in case they desire to apply the concept of inclusive and scalable analytics structures and simultaneously keep the integrity of data, its compliance, and the maintenance of business operations.

Keywords—*data democratization, mainframe migration, hybrid cloud architecture, analytics accessibility, legacy data modernisation, inclusive analytics*

I. INTRODUCTION

Mainframe systems have formed the basis of large computers around the world and have, over the decades, served as robust transaction processing units, archiving and storage units, as well as a stable data management unit. These systems remain the repositories of large amounts of institutional data accumulated through decades of service, which are crucial to the history of organizations, compliance, and organizational stability. Nonetheless, the design of conventional mainframes, which are optimized to operate under batch processing with centralized control and proprietary interfaces, has led to the existence of tightly coupled data silos that are ill-suited to form part of current analytics networks. Consequently, analytical functions tend to be isolated to a few experts in an organization, and business users, domain analysts, or operational managers are naturally left out of direct data access. This informational inequality in information retrieval takes away agility and democratization of knowledge-based decision-making.

The emergence of cloud computing, distributed data models, and scalable analytics models has transformed the technology landscape in a drastic way. A substantial number of organizations are currently undertaking the process of mainframe modernization in order to close this disconnect: without affecting the mission-critical transactional character of mainframes, migrating data and workloads to hybrid platforms or cloud solutions. By encouraging this form of migration, the enterprises have a goal to release a legacy data value, blend historical information with the latest flows of data, and expose more analytics tools and insights. By so doing, they provide the baseline to democratize data, namely the ability to have self-service analytics, real-time reporting, and participation across organizational functions.

The article explores the role of the current mainframe migration in opening up democratized data access and inclusive analytics. It explores both technical foundations, including data virtualization, change data capture, a hybrid cloud store, and domain-oriented ownership of the data, and implications of organizations and governance. The objective of the article is to offer a thorough insight into data transformation by industry reports and academic work to elucidate how migration can transform data into an inclusive resource.

Amravati University, India
ashish.p.khandelwal@gmail.com

II. Background: The Legacy Mainframe Challenge.

Traditional mainframes have been designed to be as reliable, transactional, and data-integrity conscious as possible, and their flexibility and access are frequently sacrificed. In mainframes, data is usually stored in legacy relational systems or hierarchical/network databases or proprietary file systems. Pathways of access are strictly regulated, and an extraction into analytics needs expertise and might need manual ETL (Extract/Load) operations or batch exports. These operations are cumbersome, sluggish, and prone to error. Consequently, access control to mainframe data has traditionally been restricted to an infrequently updated group of professional operators or IT experts, which is, in actuality, entrenching a pyramid structure of information access.

This structure presents a variety of challenges:

- Data invisibility: Business units with minimal technical knowledge have no visibility of legacy data, nor can they query it directly; they need intermediaries who pull out and convert data on demand, contributing to latency and bottlenecks.
- Analytical backlog: Data teams work under a load of ad-hoc requests for data, which decreases the possible capacity to do strategic analysis or innovation.
- Little use of historical data: Years of past data are not available to current analytics, machine learning, or predictive modeling and are basically squandering historical information.
- Uneven access: Analytical power is controlled by a small technical elite, making it unattainable to have more extensive organizational membership and the creation of insight.

Therefore, mainframes offer reliability and security, yet their architecture prohibits organizational agility, data democratization, and inclusive analytics capability.

III. Technical Innovations to facilitate Democratized Analytics.

The contemporary migration approaches tend to address the deficiencies of the legacy by proposing architectural solutions that can separate the data accessibility and mainframe constraints to reduce the disconnection between the legacy storage and the advanced analytics. Key innovations include:

3.1 Hybrid Cloud Architectures and Distributed Data Platforms.

The hybrid cloud system will allow the organization to have the transactional workloads stored in the mainframes and replicate or migrate the data to the cloud-based version, which is more efficient in analytics [1]. The combined model enables elastic storage and compute, which are bound to scalable engines of analytics, as opposed to creating some unreasonable overpressure on the mainframe resources without breaking operational continuity [2]. Cloud-based distributed data platforms enable new storage (data warehouses, data lakes) and compute (support large-scale analytics, reporting, machine learning, and self-service business intelligence applications) (that is, can be used by non-technical users).

3.2 Virtualization of Data and Logical Data Abstractions.

Virtualizing data provides a layer of abstraction between the old data set and exposes it to normal interfaces (e.g., SQL, APIs) without necessarily having to copy the whole data set or to physically migrate it (e.g., into the new system) [5]. Virtualization gives the user the capability to query data that has mainframe origins as though it had been ordered in an existing relational database or cloud-based database. This also reduces heavyweight ETL processes and enhances controlled access as well as governance control. Virtualization can also be applied to a hybrid architecture where real-time streaming and past archives are combined, whereby one can query transparently across different sources of data.

3.3 Real-Time Streaming Pipelines and Change Data Capture (CDC).

The updates on mainframe databases in real time, or close to real time, of transactional updates on mainframe databases are streamed to cloud-based data lakes or analytics systems using CDC systems. CDC does not use periodic batch exports but simply describes incremental changes and continuously pushes them downstream. With the help of streaming systems, this approach may be implemented to make sure that analytics environments are kept up-to-date with low latency [5]. This means that business users and analysts are able to use new data, and this leads to the ability to have real-time dashboards, event-driven analytics, and responsive decision support. which are core components of democratized, inclusive analytics.

3.4 Domain-Shared Data (Data Mesh)

The data mesh is a new form of organization that intends to decentralize owner teams and formation and treat data as a product of domain teams and not a cumulative central database [4]. The creation of a data mesh architecture during mainframe migration would allocate the load evenly within organizational regions (finance, operations, customer service, compliance, etc.) such that all regions would be able to maintain, expose, and control data sets as necessary. Combined with the phenomena of virtualization and hybrid constructions, data mesh allows access to self-service, domain ownership, and scale-controlled data, allowing numerous consumers to transact with data without sacrificing control and compliance.

.

IV. Migration to Inclusion: Democratization and Impact on Organizations.

Migration of the mainframe is not just an exercise of a technical nature, but it embodies the restructuring of access to information, regulation, and organizational culture. In the implementation of the described innovations, migration opens up inclusive analytics with far-reaching consequences:

4.1 Dismantling Access Gatekeepers.

Having the data abstracted and disclosed through modern interfaces (APIs, SQL, BI tools), the migration will cease to depend on the services of a few experts. Real-time access to datasets, dashboards, and analytical apps can provide business individuals, analysts, and domain experts with an organization-wide way of creating insight by direct access to data. This reduces the bottlenecks and ensures decisions are made faster and the distribution of the analysis potential is more balanced [6].

4.2. Powering self-service analytics.

Hybrid cloud infrastructures, together with virtualization, facilitate self-service BI and machine-learning pipelines, which can be employed by non-technical stakeholders without having an in-depth understanding of the infrastructure. Such flexibility brings analytics out of a centralized, IT-driven activity and makes it more of a collaborative, cross-functional process and democratizes not only access but also authorship of data-driven insights [3].

4.3 Data Value Unlocking Historical Data.

Most organizations have decades of archived information stored on their mainframes, transactional records, audit trails, historical operations records, and so forth that have never been accessed for years due to access and format limitations. Migration opens this historical data, allowing longitudinal analyses, trend modeling, predictive maintenance, compliance audits, and machine learning to operate more effectively on rich data (historical data) otherwise resting in a dormant state. This ensures the institutional memory and conversion of legacy archives to assets of strategy [10].

4.4 Nurturing Innovation and Agility in the Organization.

With data available, tools of analysis implemented, and teams of domain experts enabled, there is more responsiveness and agility across organizations. They will be able to implement new analytics-based services, experiment with machine learning, react fast to market dynamics, and create a data-driven culture. Migration, therefore, becomes a strategic facilitator of innovation, but not a technical upgrade.

V. Implementation Strategies: Pathways to Inclusive Analytics.

Enterprises might adopt various approaches to migration depending on the maturity and volume of data and the risk tolerance. The most common ones are re-platforming, refactoring, and virtualization; they are usually combined to be adjusted and provide resistance to risks.

Migration Method	Description	Inclusivity impact
Replat forming	Migrate data and schema to cloud-native storage with a minimal code adjustment.	Faster access to old data, moderate analytics promptness.
Refactoring + Re-engineering	Rewrite or refactor distributed system applications into microservices and a modular design.	Complete cloud integration, high analytics, and data agility
Data Virtualization	Present mainframe data over data virtualization layers/API gateways without duplicating the data	Disruption is minimized; applicable to analytics and reporting.

Table 1—Typical Migration Techniques and Their Inclusive Effects.

Migration strategy: It is a suggestion that a consistent migration strategy proceed in stages, starting with read-only virtualization / migrating archival data and then verifying access control and governance processes and migrating or refactoring active transactional data. This gradual method minimizes threat and allows the data to be democratized early enough.

A conceptual hybrid architecture of inclusive analytics made possible by migration is presented in Figure 1 (below):

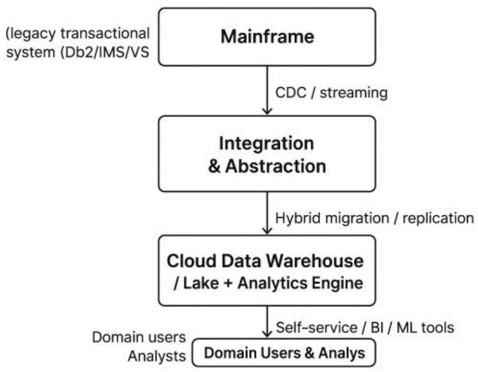

Figure 1 - Conceptual Hybrid Data Architecture Post-Migration.

This architecture separates transactional computing and analytics, provides real-time data mining, and offers several entry points to users in the enterprise, representative of the ethos of democratized data access.

VI. RISKS, GOVERNANCE AND CHALLENGES.

Although a promising idea, mainframe migration and democratized analytics come with a row of issues and risks, which organizations need to handle carefully:

6.1 Data Migration and Data Integrity Risk.

Legacy systems can be characterized by a complex data model, interdependent modules, and hidden business logic. Refactoring or re-platforming can either create accidental dependencies, data mappings, or orphans. Even the virtualization strategies cannot be overlooked, as data semantics must be preserved, and legacy constraints are to be considered [9]. Unsuccessful migration could lead to the disruption of business continuity.

6.2 Security, Compliance and Access Control.

An increase in data access leads to the possibility of unauthorized exposure, usage, or data loss, particularly when working with sensitive personal information, financial documents, or regulated information. To maintain a high level of governance, organizations should adopt strong measures of web security, role-based access control, encryption, audit, data classification, and compliance standards so that the democratization efforts do not jeopardize the secrecy or regulatory standards [8], [7].

6.3 Change Culture and Organization.

To shift the model of a gatekeeper to a self-service data culture, organizational buy-in as well as training and change management is needed. The decentralization might not be welcomed by legacy IT teams; the data literacy may be questioned by domain users; governance might be viewed as a process with bureaucracy. Migration cannot succeed in providing culture-related democratization benefits without adequate stakeholder engagement, training, and cultural change programs [6].

6.4 cost of Infrastructure and Operational Overhead.

Cloud platforms are scalable, but they may also be associated with additional costs in the form of storage, data egress, and compute costs, maintenance of streaming pipelines, and all these are to be considered with great care. Moving huge volumes of old (data) and executing pipelines on a routine (scheduled) basis, along with self-service analytics, may strain the budget when not thought through carefully [4], [10].

VII. STRATEGIC RECOMMENDATIONS FOR EFFECTIVE DEMOCRATIZATION

The following strategic guidelines must be taken into account by the organizations that are going to democratize the idea of data by moving it to the mainframe depending on the technical opportunities and challenges:

- Move at the scheduled rate: Move the first virtual archive or read-only data and a substantial and consistent flow of active data; do not rush migration and do not use advanced changes, but instead push it on through tactical steps of constraint validation to constrained active data systems.
  - Adoption of a hybrid cloud architecture: Retain the mission-critical workloads within a mainframe and recreate the data within cloud platforms to apply in analytics; retain stability and accessibility at the same time [2].
  - Structured, managed, strong data engineering format: role-based access, encryption, audit logs, metadata category, and data authorities should be constrained and unlocked; data access should precede the opening up of the data access to the masses [8], [7].
  - Promote organizational data literacy: Educate domain users and analysts, foster collaboration among workers across domains, and foster a culture of data as being a shared resource and not a treasure trove [6].
  - Measure performance and costs: Measure storage and compute usage, the latency in pipelines, query loads, and compliance costs; periodically modify the strategy to render the strategy viable [4].
  - Adopt domain-based data ownership (data mesh): Domain teams must be able to run and self-publish their data; this brings about responsibility and accountability and governance decentralization [5].
  -

VIII. GENERAL IMPLICATIONS: EQUITY, INCLUSION, AND INNOVATION.

The migration of mainframes to democratize access to data does not only transform the enterprise analytics but also powers relations within the organization, the feeling of inclusion, and transforms innovation:

- Operational Equity: The access by non-technical workers as well as domain professionals to information assists organizations in distributing the capacity of analysis more evenly and reducing the necessity to rely on expert conveyors [6].

- Releasing Historical Value: Decades of archived information make it possible to do more historical analysis, longitudinal research, and policy-making that are data-driven—the memory of the organizational institution and releasing latent value [10].
- Quick Innovation: Aided by the data that is available and scalable analytics infrastructure, organizations are able to use machine learning, predictive analytics, and real-time dashboards, which leads to faster decision-making, new services, and competitive advantage [3], [2].
- Cultural Transformation: Data is turned into a mutual company asset, which evolves openness, collaboration, and life-long learning. It is beneficial because this cultural shift assists one in being agile, flexible, and adaptable to the rapidly evolving digital context.

IX. CONCLUSION

Enterprise data storage and transaction processing have relied on the mainframe systems long enough. Nevertheless, their structural inflexibility has frequently limited data access and data analytics inclusivity. Hybrid cloud migration and data virtualization, change data capture, and domain-oriented data ownership offer an excellent model of modernization of legacy systems and a promising pathway to democratized analytics. Through these advancements, organizations will be able to convert these old data silos into appreciative, usable, and utilizable data ecosystems. The shift allows a wider range of people to engage in analytics, increases the speed of innovation, has a history of institutional data, and has an organizational culture based on common insight as opposed to centralized gatekeeping. However, the migration and democratization processes need proper planning, good governance, cost management, and change management in cultures. Those organizations that make such commitments will enjoy the advantage of more equitable data access, stronger agility, and long-term digital growth. Mainframe migration thus must be considered not only as a technical enhancement but also as an enabler of analytics inclusiveness, the change of the organization, and long-term innovation.

REFERENCES

[1] IBM, What is Data Modernization?, IBM White Paper, 2025.

[2] Microsoft, Modernize Mainframe Data to Azure — Reference Architecture, Microsoft Azure Architecture Center, 2024.

[3] T. Bukhari, M. Ghani, and H. Mir, “Cloud-Native Business Intelligence Transformation,” Int. J. Sci. Res. Humanit. Soc. Sci., 2025.

[4] NASSCOM Community, Mainframe Modernization Leveraging the Power of Hadoop Framework, 2025.

[5] K. Waehner, Mainframe Integration with Data Streaming: Technical Report, 2025.

[6] TDWI, How Enterprises Can Democratize Mainframe Data, TDWI Report, 2022.

[7] ModLogix, Mainframe Data Modernization: No Data Left Behind, White Paper, 2024.

[8] BMC Software, Why Migrate Mainframe Data to Hybrid Cloud—and Why Now?, 2025.

[9] M. Fahmideh et al., “Challenges in Migrating Legacy Software Systems to the Cloud,” 2020.

[10] M. D. Assunção et al., “Big Data Computing and Clouds,” J. Parallel Distrib. Comput., 2015.