

CAPCO

THE RISE OF DATA COLLABORATION



INTRODUCTION

The age of COVID-19 has reinforced the importance of making data-driven decisions for financial institutions. As the use of tools for data management continues to increase, organizations are seeking ways to consolidate and streamline access. Nevertheless, organizational blockers such as silos and integration of complex technology layers still exist, which makes it harder for users to obtain the data they need in a timely and efficient manner. Data fabric solutions have emerged that, “accelerate insights by automating ingestion, curation, discovery, preparation, and integration from data silos,”¹ to try to solve for this. Despite the advances made by introducing data fabric technology, several pain points still linger. These include the following:

CHALLENGE	WHY?
Challenges with third-party data integration	<ul style="list-style-type: none">• Absence of secure integration capabilities• Governance controls and security concerns• Lack of partners and business sponsors
Lack of a central platform to access multiple data sources	<ul style="list-style-type: none">• The proliferation of end-user technologies due to frustration with excessive centralization of capabilities• Vendor-focused solutions that require vertical integration
Disjointed governance controls and implementation	<ul style="list-style-type: none">• Evolving standards and regulations• Continued changes in data consumption behavior
Inability to holistically understand the data ecosystem and landscape	<ul style="list-style-type: none">• Disparate data sources and outdated data governance policies around acquisition• Inclusion of non-traditional sources such as unstructured data

Recent developments in technology have made significant strides in addressing these challenges. In 2020, the expectation is that the following trends will help alleviate some of the stress from these pain points:

- Data exchanges
- Unified data environment
- Augmented data management
- Knowledge graphs

These elements effectively complement the advances made through the implementation of Data Fabrics. One can define the combined approach of these solutions as ‘Data Collaboration.’ This approach to data management focuses on empowering the end-user while providing automated solutions for integration and accessibility. As highlighted in the below diagrams, traditional data management solution architecture revolves around ETL (extract, transform and load) from data sources into a data warehouse or a data lake. Data collaboration leverages data fabric technology to eliminate these technical nuances, allowing users to access data in a much quicker and more efficient manner. The circumstances surrounding COVID-19 have pushed this need further, accelerating this defining trend for advancements in data management in 2020.

1. Yuhanna, Noel. The Forrester Wave™: Big Data Fabric, Q2 2018. Forrester, 12 June 2018

COMPARISON OF TRADITIONAL DATA MANAGEMENT FRAMEWORK VS. DATA COLLABORATION

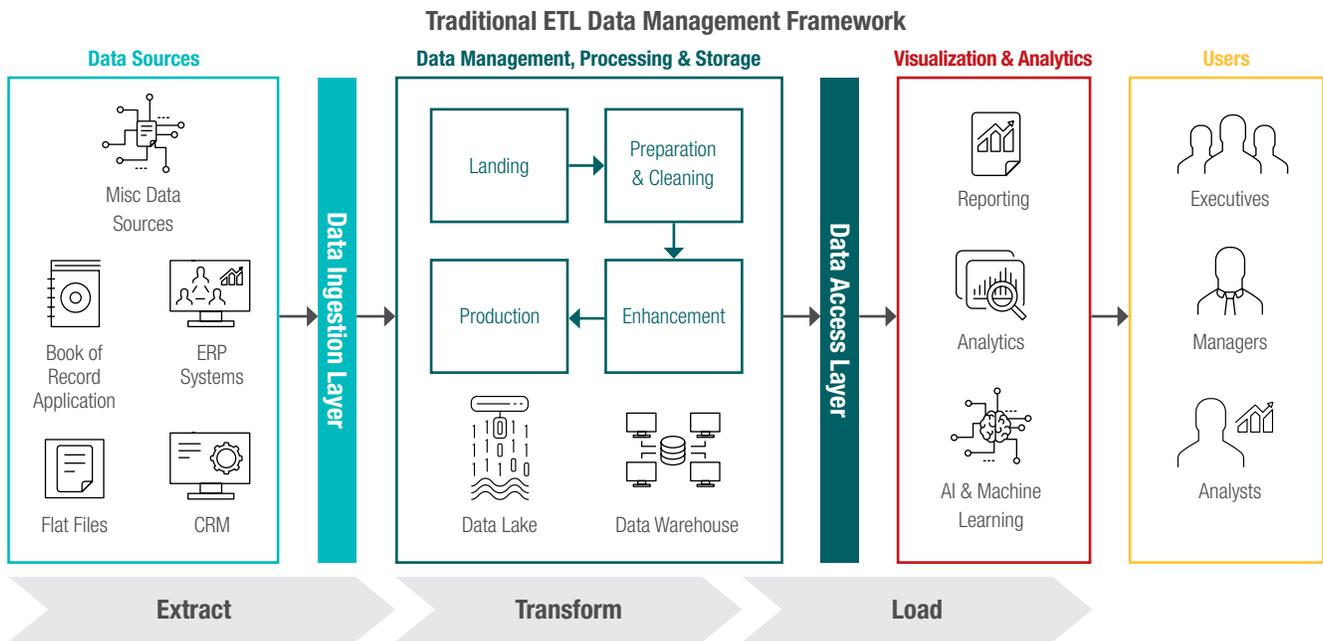


Exhibit A - Traditional ETL approach for Data Management

Data Fabric and Data Collaboration

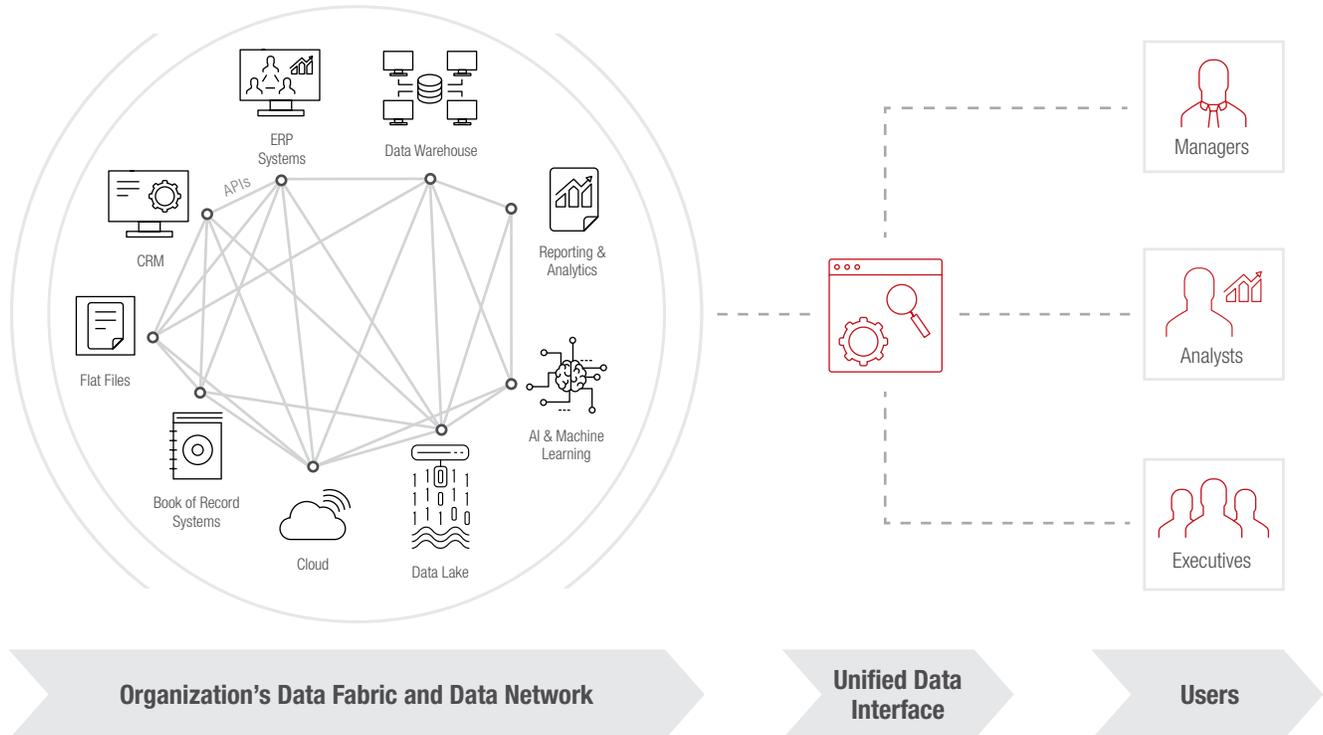


Exhibit B – New approach to Data Management using Data Fabric and Data Collaboration

2020 DATA MANAGEMENT TRENDS

Recent trends in data management reflect the need for more data collaboration solutions. Companies are now expanding the scope of their strategy to be even more holistic, enabling data analysis across all areas of business and extending to external sources. For example, many firms are leveraging sources from Johns Hopkins University and various government bodies for COVID-19 related data. Subsequently, this has led to the federated, hybrid model of centralized governance to become a preferred operational model. This model tends to be more empowering for business groups who are driving solutions independently within an enterprise setting. As data fabric technology becomes more adopted across the architecture, a 'network effect' is created that continuously streamlines integration and enables operational gains and efficiencies. Capco has identified the five trends highlighted below as the most prominent topics for Data Management at financial services organizations for 2020.

DATA EXCHANGES

Aggregators of third-party data are not new to the financial services community. However, the impact of COVID-19 has ushered in a new era of awareness around the value of data and analytics. Cloud providers such as Amazon are now offering full access to related datasets to support research and development through integrated Data Exchange platforms and SDKs.² Companies like Apple³ and Google⁴ are also offering free access to insights and data. Financial services organizations can now leverage this information to calibrate risk models better and prepare responses. Although the open data community has been growing in strength for many years, the COVID-19 pandemic accelerated the role of data exchanges into the spotlight and revealed their importance.

The primary challenges in this area involve integration and security. Companies need to find ways to ensure customer data is protected and that regulatory concerns can be addressed. Additional work is required to establish data exchange providers as operational partners.

UNIFIED DATA ENVIRONMENT

As data sources proliferate both internally and externally, the challenge of integration is now more evident than ever before. Nevertheless, data fabric solution providers are delivering meaningful solutions. Vendors are now accelerating the holistic view of data management by offering data lineage tools and API/ODBC connectors to connect sources from across the entire data ecosystem. This resolves the longstanding challenge of establishing a unified environment to run queries and access data. By leveraging data fabrics, companies can eliminate the need to consolidate sources on isolated platforms. Pipelines can become flexible and span across multiple cloud instances or databases. In turn, users can access data elements true to their lineage and pull directly from their sources. The result is an environment that consolidates all data storage platforms, simplifies operations, and removes the overhead complexity of integration.

AUGMENTED DATA MANAGEMENT

One of the key benefits of leveraging data fabric technology is the ability to automate processes that were previously complex and time-consuming. Vendors are now introducing solutions that can identify related fields and intelligently connect data elements. Activities like change management reduce in complexity from an

-
2. "The Data Exchange: Moving Data between 1-2-3, DBase, and Other Popular Programs." AWS, Amazon, 1990, aws.amazon.com/data-exchange/covid-19/?cards.sort-by=item.additionalFields.order&cards.sort-order=asc.
 3. "COVID-19 - Mobility Trends Reports." Apple, www.apple.com/covid19/mobility
 4. "COVID-19 Community Mobility Report." Google, www.google.com/covid19/mobility

ETL perspective. More importantly, solutions are now capable of automatically tagging metadata elements and intuitively tracking end-to-end data lineage – offering tremendous value from a governance perspective, since you can now streamline these manual tasks. Moreover, you can design data quality rules using machine learning models that you train to identify incorrect patterns in data elements. All these features are important for reducing operating costs and empowering business users to deliver data-driven solutions.

KNOWLEDGE GRAPHS

Another trend emerging in 2020 is the advent of knowledge graphs to visualize data ecosystems, which are a meaningful extension of data fabric technology in the sense that users can now see the benefits of integration. They also enable the usage of graph databases more cohesively. Data elements can be connected via semantic lookups or through defined relationships. Network diagrams capable of delivering a holistic view of all data sources can be constructed and leveraged for additional insights and analyses. Vendors such as Stardog and AnzoGraph are leveraging semantic methods to associate linked entities and visualizing them in a way that is consumable for end-users. Knowledge graphs are a useful mechanism to formalize a more comprehensive structure to ecosystems by leveraging ontological methods. This development will surely have implications down the road on everything from data models to governance.

DATA COLLABORATION

Data collaboration builds from data fabric, one of the key trends in data management from 2019⁵. Data fabric technology creates and integrates a data network of an organization's internal data assets and empowers the sharing of the data in a simplified and easy-to-use environment. Data collaboration takes this a step further by having a single interface where any user, from analysts to executives, will be able to access, create, maintain, and consume the data using the same platform. The single interface eliminates the complexity of integrating data from multiple sources through the traditional ETL method, which is very labor-intensive. This simplification of the data ecosystem enables collaboration so that data can be used for its intended purpose.

Just like the world and financial services industry, the data space is constantly evolving. The data collaboration model combines the current emerging trends and ultimately helps organizations streamline the traditional model of moving and copying data (potentially creating data silos) to allow users to have a centralized platform to manage and use the data for its intended purpose. Empowered organizations will be able to leverage this concept to react faster and stand up solutions in a more nimble and agile way.

5. Yuhanna, Noel. The Forrester Wave™: Big Data Fabric, Q2 2018. Forrester, 12 June 2018

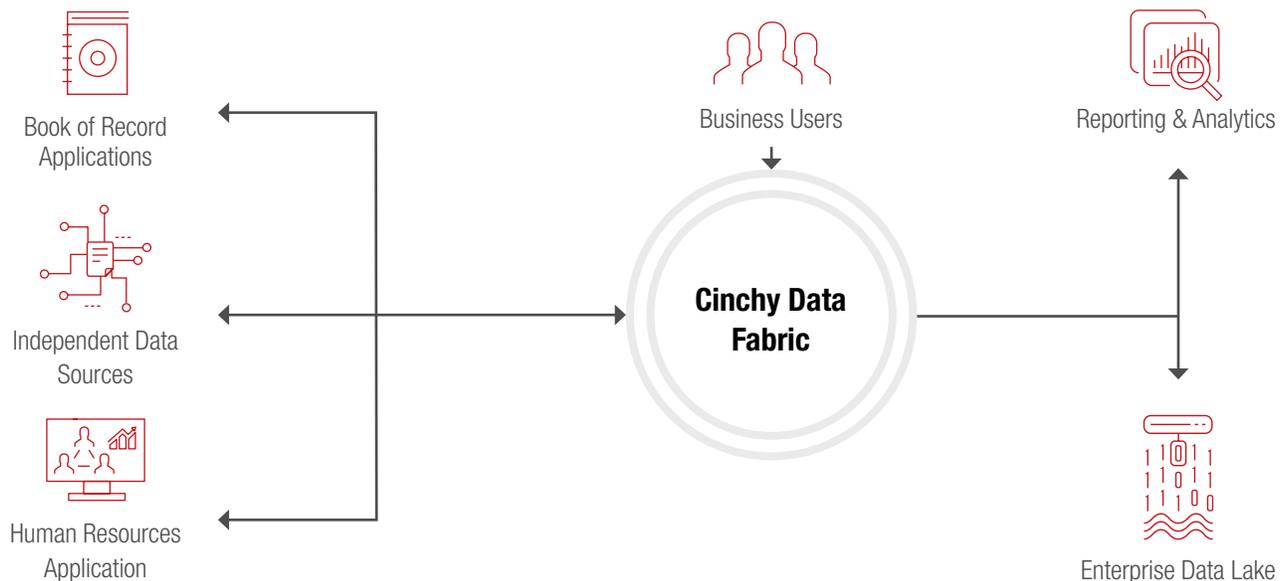
USE CASE FOR DATA COLLABORATION

Cinchy's data collaboration platform, recently listed in Gartner's Cool Vendors in Data Management report⁶, is based on the vision of eliminating data silos and empowering users to take control and collaborate on their data. The tool allows users to accelerate and streamline ETL processes by easily connecting to different data sources and eliminate time spent moving and copying data between applications through an interconnected architecture. This helps establish a platform to deliver a unified data environment. Users can govern and control their data using a single source of truth, without worrying about the duplication and replication of this data, replicating the augmented data management capabilities. Cinchy's platform leverages data fabric technology and combines this with rich user experience, data protection features, and a solution building environment to provide a more robust platform for the enterprise.

DATA COLLABORATION IN WEALTH MANAGEMENT

The following use case demonstrates how a data collaboration tool like Cinchy, can help a Tier 1 Canadian bank speed up its solution delivery within its wealth organization master data solution.

To begin, the bank's wealth division had issues with visibility into their internal organization. The bank's data about their front-office advisors, reporting structure, finances, and on-boarding processes were in fragments across the bank, without a source of truth. This resulted in issues with integrating data from multiple legacy systems, as well as operational challenges with its data. Cinchy was brought in as a data collaboration platform to address the underlying data issues and deliver multiple new solutions at speed and scale to solve the organization's master data challenges.



6. "Cool Vendors in Data Management", Gartner, <https://www.gartner.com/document/3984805>

Cinchy was used to connect to and manage multiple data sources within the bank. The portal acted as a unified data environment that became a one-stop-shop of where the organization data was being held and stored. Though there were multiple sources for this data, Cinchy was used to connect to the data easily. Business users were able to leverage the augmented data management capabilities through the Cinchy platform by defining the business rules used to master the data. Lastly, through Cinchy's capabilities to export data through its API's, the mastered organizational data was used as input to other processes such as reporting and analytics, and clean data was exported into the bank's enterprise data lake. Cinchy has a built-in visualization piece to showcase its data network through knowledge graph capabilities, which was used to integrate with other datasets used within the Cinchy environment.

Ultimately, deploying data collaboration technology through Cinchy enabled the bank to deliver the solution 50 percent faster than initially promised using a more traditional method; this represented the time and cost savings, which impacted the bank's bottom line while also meeting business objectives. Dan DeMers, CEO and founder of Cinchy, states, "Do more with less is no longer just an expression; it's an existential requirement! As a result, financial institutions everywhere are accelerating the adoption of data fabric technology. As the most advanced data fabric, Cinchy is already helping some of the world's biggest banks move toward a zero-integration environment where data is seamlessly connected, automatically protected, and available for real-time data collaboration. Cinchy's Data fabric has allowed customers to meet these new demands of their organization by delivering projects in half the time and half the cost."

CONCLUSION

The year 2020 will see the acceleration of this combination of data fabrics and data collaboration technologies. Many organizations are still struggling to realize the value of their big data stacks and platforms, which they've invested heavily in. The traditional ETL and ELT models of ingesting and integrating data have become a tedious process, with confusion around data lineage emerging due to the sheer volume of data sources in the modern age. Building multiple instances of this data and perpetuating replication creates additional data silos, which repeats the problems of multiple sources of truth from the past. The industry's shift to focusing on data eliminates many of these pain points for large organizations.

The emerging category of data collaboration on top of the data fabrics empowers organizations and their users to leverage their data without the burden of redesigning and overspending on data management solutions. Many enterprise-level vendors are now leveraging data fabrics as a key part of their offerings. Market conditions in 2020 are leading the industry towards simplifying process and delivery through data collaboration. Companies are refocusing on solutions that allow them to reduce technical complexity and react quickly.

Capco marries its domain focus in financial services with our deep expertise in the data space, to be a trusted partner for our clients to navigate emerging technologies. Reach out to Capco for further information about data fabrics and data collaboration and see how we can help.

AUTHORS

James Musgrave, Senior Consultant

Aaron Choi, Senior Consultant

ABOUT CAPCO

Capco is a global technology and management consultancy dedicated to the financial services industry. Our professionals combine innovative thinking with unrivalled industry knowledge to offer our clients consulting expertise, complex technology and package integration, transformation delivery, and managed services, to move their organizations forward.

Through our collaborative and efficient approach, we help our clients successfully innovate, increase revenue, manage risk and regulatory change, reduce costs, and enhance controls. We specialize primarily in banking, capital markets, wealth and asset management and insurance. We also have an energy consulting practice in the US. We serve our clients from offices in leading financial centers across the Americas, Europe, and Asia Pacific.

To learn more, visit our web site at www.capco.com, or follow us on Twitter, Facebook, YouTube, LinkedIn and Instagram.

WORLDWIDE OFFICES

APAC

Bangalore
Bangkok
Gurgaon
Hong Kong
Kuala Lumpur
Mumbai
Pune
Singapore

EUROPE

Bratislava
Brussels
Dusseldorf
Edinburgh
Frankfurt
Geneva
London
Paris
Vienna
Warsaw
Zurich

NORTH AMERICA

Charlotte
Chicago
Dallas
Houston
New York
Orlando
Toronto
Tysons Corner
Washington, DC

SOUTH AMERICA

São Paulo

WWW.CAPCO.COM



© 2020 The Capital Markets Company. All rights reserved.

CAPCO