# CAPCO

THE CAPCO INSTITUTE JOURNAL OF FINANCIAL TRANSFORMATION

DATA INTELLIGENCE

Data entropy and the role of large program implementations in addressing data disorder

SANDEEP VISHNU | AMEYA DEOLALKAR
GEORGE SIMOTAS

# DATA ANALYTICS

50TH EDITION | NOVEMBER 2019

# THE CAPCO INSTITUTE

## JOURNAL OF FINANCIAL TRANSFORMATION

RECIPIENT OF THE APEX AWARD FOR PUBLICATION EXCELLENCE

# CONTENTS

## DATA MANAGEMENT

## DATA ANALYTICS



## DATA INTELLIGENCE

# DEAR READER,

Welcome to the milestone 50th edition of the Capco Institute Journal of Financial Transformation.

Launched in 2001, the Journal has covered topics which have charted the evolution of the financial services sector and recorded the fundamental transformation of the industry. Its pages have been filled with invaluable insights covering everything from risk, wealth, and pricing, to digitization, design thinking, automation, and much more.

The Journal has also been privileged to include contributions from some of the world's foremost thinkers from academia and the industry, including 20 Nobel Laureates, and over 200 senior financial executives and regulators, and has been co-published with some of the most prestigious business schools from around the world.

I am proud to celebrate reaching 50 editions of the Journal, and today, the underlying principle of the Journal remains unchanged: to deliver thinking to advance the field of applied finance, looking forward to how we can meet the important challenges of the future.

Data is playing a crucial role in informing decision-making to drive financial institutions forward, and organizations are unlocking hidden value through harvesting, analyzing and managing their data. The papers in this edition demonstrate a growing emphasis on this field, examining such topics as machine learning and AI, regulatory compliance, program implementation, and strategy.

As ever, you can expect the highest caliber of research and practical guidance from our distinguished contributors, and I trust that this will prove useful to your own thinking and decision making. I look forward to sharing future editions of the Journal with you.

Lance Levy, **Capco CEO**

# FOREWORD

Since the launch of the Journal of Financial Transformation nearly 20 years ago, we have witnessed a global financial crisis, the re-emergence of regulation as a dominant engine of change, a monumental increase in computer processing power, the emergence of the cloud and other disruptive technologies, and a significant shift in consumer habits and expectations.

Throughout, there has been one constant: the immense volume of data that financial services institutions accumulate through their interactions with their clients and risk management activities. Today, the scale, processing power and opportunities to gather, analyze and deploy that data has grown beyond all recognition.

That is why we are dedicating the 50th issue of the Journal of Financial Transformation to the topic of data, which has the power to change the financial industry just as profoundly over the coming 20 years and 50 issues. The articles gathered in this issue cover a broad spectrum of data-related topics, ranging from the opportunities presented by data analytics to enhance business performance to the challenges inherent in wrestling with legacy information architectures. In many cases, achieving the former is held back by shortcomings around the quality of, and access to, data arising from the latter.

It is these twin pillars of opportunity and challenge that inform the current inflection point at which the financial industry now stands. Whilst there is opportunity to improve user experiences through better customer segmentation or artificial intelligence, for example, there are also fundamental challenges around how organizations achieve this – and if they can, whether they should.

The expanding field of data ethics will consume a great deal of senior executive time as organizations find their feet as they slowly progress forward into this new territory. In my view, it is critical that organizations use this time wisely, and do not just focus on short-term opportunities but rather ground themselves in the practical challenges they face. Financial institutions must invest in the core building blocks of data architecture and management, so that as they innovate, they are not held back, but set up for long-term success.

I hope that you enjoy reading this edition of the Journal and that it helps you in your endeavours to tackle the challenges of today's data environment.

Guest Editor
Chris Probert, **Partner, Capco**

# DATA ENTROPY AND THE ROLE OF LARGE PROGRAM IMPLEMENTATIONS IN ADDRESSING DATA DISORDER

**SANDEEP VISHNU** | Partner, Capco
**AMEYA DEOLALKAR** | Senior Consultant, Capco
**GEORGE SIMOTAS** | Managing Principal, Capco

## ABSTRACT

Clutter is a highly pervasive phenomenon. Homeowners are very familiar with this occurrence as their acquisitions grow to fill available space. Closets, garages, basements, and many areas not in obvious sight become dumping grounds for things that do not have immediate utility or a logical place in the house. Now think of a scenario where the volume, velocity, and variety of goods entering the house goes up by several orders of magnitude in a very short period of time. The house will simply start to overflow with articles strewn wherever they can fit, with little thought given to order, use, and structure. Enterprises face a similar situation with data as volumes have grown dramatically over the last two to three years. Organizational reluctance to retire or purge data creates overflowing repositories, dark corners, and storage spaces full of outdated, unseen, and difficult to access information – i.e., data clutter. Temporary fixes only add layers to the problem, creating additional waste, maintenance challenges, damage, inefficiency, and improvement impediments. All these factors drive **data entropy**, which for purposes of this paper is defined as the tendency for data in an enterprise to become increasingly disorderly. Large programs are often data centric and surface data clutter issues. This paper explores the concept of data entropy in today's world of rapidly expanding data types and volumes entering an organization at exponentially higher speeds, and how large program implementations can be used as catalysts to address data clutter and modernize the data supply chain to streamline data management.

## 1. INTRODUCTION

Over the last few years, data generation has risen exponentially causing organizations to immensely accelerate their ability to store, process, and use data. Every visionary company in the world is working towards leveraging data to differentiate themselves, provide better customer experience, and fuel growth. In large financial institutions, these strategies often clash with legacy systems and architectures, which can accommodate incremental increases in utilization but are inadequate when faced with exponential growth.

With competition from fintech startups and consumer expectations on the rise, financial institutions are leveraging data from an increasing number of data sources to feed an expanding set of applications that provide insight and value to customer and stakeholder segments. In addition, banks face changing regulations, new privacy laws, and a growing need to integrate with third-parties, all of which place additional demands on their data infrastructure.

As organizations launch initiatives that significantly alter business and IT operations, they increasingly face complications and risks driven by data complexity, which in turn surfaces challenges commonly faced by organizations tackling strategic change, including:

- Derived data elements created for specific solutions that need to be continuously maintained over time as other data structures evolve

- Use of temporary data structures and workarounds that become permanent components of the technology ecosystem
- Extra project work required to develop data for specific applications, which further adds to overall data complexity.

Simply put:

**Data complexity = *f*(derived data elements, number of systems, number of independent data stores, data uses)**

For instance, a large program – like CCAR (comprehensive capital analysis and review) or FDIC370 – will have data elements required by the business shared across multiple systems, databases, and uses as shown in Figure 1. The same data element lives in multiple databases with different names and may be transformed by each user as needed.

Adding new or changing existing lines of business impacts applications, databases, and users. Business, systems, and data have a tightly interconnected relationship – for example, business process enhancements trigger changes in the existing IT infrastructure. Every business element required in the process resides in systems, databases, and consumer use-cases as one or many data elements. Every new application, data store, or use-case increases overall complexity. Figure 2 provides an example of key drivers needed to meet FDIC370 requirements and how they impact data elements, systems, databases and users.

**Figure 1:** Requirements drive data complexity

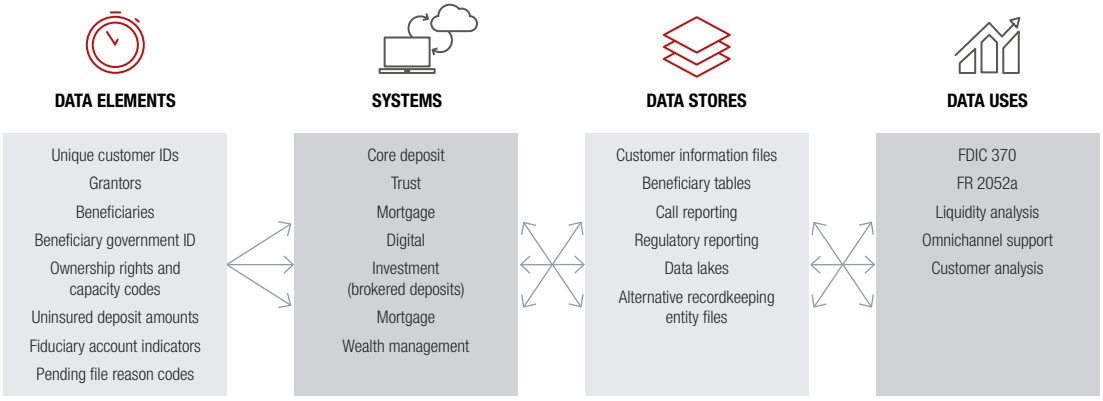



**Figure 2:** Business and process elements drive data complexity and FDIC 370 costs for specific covered institutions

| | DATA ELEMENTS | SYSTEMS | DATA STORES | DATA USES |
|---|---|---|---|---|
| NUMBER OF DEPOSIT ACCOUNTS | ✓ | ✓ | ✓ | ✓ |
| DISTINCT CORE SYSTEMS | ✓ | ✓ | ✓ | ✓ |
| NUMBER OF LEGAL ENTITIES AND GEOGRAPHIES | | ✓ | ✓ | ✓ |
| SWEEP ACCOUNTS | ✓ | | | ✓ |
| BUSINESS, PRODUCT, AND ACCOUNT FEATURES | ✓ | | ✓ | ✓ |

Source: FDIC 370 Recordkeeping for Timely Deposit Insurance Original Rule Making Commentary

Unless proactively managed, each increase in complexity drives up program cost, compliance risk, and inefficiencies. Much like that homeowner, "boxes" of independent data pile up over time and start to clutter the overall environment.

The good news is that these challenges can be mitigated by using a variety of emerging practices. These can be as simple as formulating a common framework for data assessment and lineage to very significant architectural changes designed to allow individual applications to plug into centralized services and data sharing pipelines. The key to effective data governance is proactive and constant vigilance against the build-up of data complexity. Like that nagging feeling when you open a certain closet in your house, the organization needs to have a sense of when things are starting to get out of hand.

Often a large program like CCAR or FDIC370, can trigger a good old-fashioned spring cleaning and prompt the organization to implement leading data governance and management practices.

## 2. DATA GOVERNANCE AND MANAGEMENT CHALLENGES

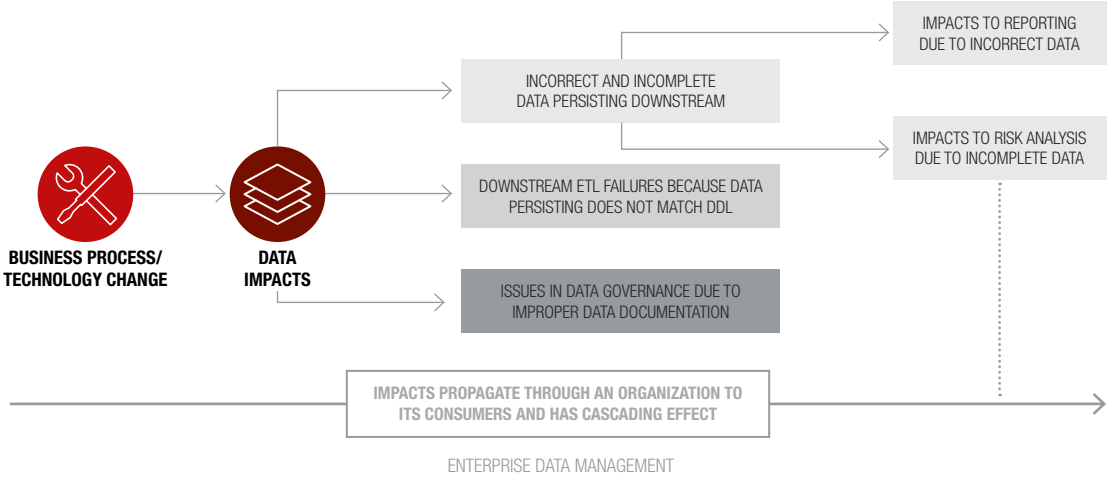
Large financial services organizations (FSOs) face data challenges largely as a result of constantly changing business objectives, regulations, and evolution in technology. As data moves through systems and processes in an organization, complexity often grows. Data discrepancies in a single node of a network could have a cascading effect throughout an organization, as illustrated in Figure 3.

These situations occur naturally at different points in the data pipeline. Data management is complex and a change to even a single data point can create cascading challenges across the organization and data stores. When this occurs frequently, multiple changes pile up and interact – like the growth of items over time in a basement or garage.

Efficient data governance and management with well documented metadata, continuously maintained data dictionaries, data access policies, and data retention policies can help avoid these issues. However, organizations keep growing with large business programs that have timelines and goals. When timing considerations become paramount, data governance takes a back seat, which adds to data clutter instead of addressing it. Just like a spring-cleaning project, organizations need to establish thresholds and monitoring that can signal when data complexity has reached a level that requires action.

We have identified three of the biggest drivers of change in an organization that make maintaining good data governance difficult: (1) incomplete M&A integration, (2) business process changes, and (3) technology inconsistencies (e.g., old data structures from outdated systems and applications, mixed coding schemas). Large programs that implement the above-mentioned organizational changes often work around data governance practices, thereby creating data inefficiencies and increasing data complexities, when they could be used as catalysts to solve data disorder. Inefficient data governance in turn adds data complexity, which makes data governance and management challenging, thereby creating a vicious cycle. Data complexities associated with each driver or change are discussed in greater detail below.

**Figure 3:** Cascading impacts of data



## 2.1 Incomplete M&A integration

Mergers and acquisitions are common in the financial services industry. These are increasing with large banks acquiring fintech startups and data providers for a competitive edge. Such a transaction not only integrates businesses but also IT infrastructures and data. This includes integration of data governance policies, data management principles, retention policies, and metadata management, to name a few. In general, M&As increase complexity of data and by extension data entropy.

M&As would bring 'n' new systems and the underlying data into an organization. During M&As, data integration is sometimes short-changed, as aggressive deadlines and resource shortages increase the pressure on business and IT personnel. Frequently, data conversion and integration becomes an afterthought and does not get handled effectively – the business believes that IT is on point and IT believes that it is a business activity. This lack of clarity on activities and accountabilities can have significant consequences for a business. Incomplete integration of data leaves M&As with a big liability, which could impact every strategic initiative for the new combined entity.

Incomplete integration of data prevents comprehensive analysis of data from both organizations and the establishment of a common governance framework. Lack of insight on available data becomes an issue when combined data from both organizations is required. Temporary fixes will often be made to meet such requirements. While not optimal, this approach may sometimes be required to meet deadlines. Organizations make major investments that can often not wait for all data issues to be resolved before realizing gains. One fix could be to create a temporary data store with data load processes that extract data from different sources, transform data as required, and load it into the store. Incomplete analysis of data stores in both organizations would create new, redundant, or unnecessary datasets, which would further increase complexity. For example, when two banks merge, temporary regulatory reporting data sets need to be established for FDIC 370 purposes along with several related compliance requirements, such as call reporting, CCAR, etc. As a result, 'm' independent data stores are added to the overall complexity.

Organizations going through a M&A process are expected to have common data elements with information on their customers and business. Both organizations have a considerable number of derived data elements for their business processes. Performing detailed analysis of data from both organizations can help identify common data elements and reduce the number of derived data elements in the new combined organization. Reducing the number of derived elements 'd' would help control the data complexity which can now be expressed as:

**Data complexity = *f*(d + derived data elements, n + number of systems, m + number of independent data stores, data uses)**

These issues create risk of delay in the successful implementation of the venture, risk of having bad or incomplete data, and will take a toll on the cost and number of resources required to execute the venture. It is important to bear in mind that the ability to meet the objectives of a strategic acquisition or merger will depend greatly on the combined data from both the organizations. Data can become a huge asset, offer significant insight, and serve as a source of competitive advantage. However, this value can only be realized if the organization succeeds in efficiently integrating and managing the data.

## 2.2 Business process changes

The financial services industry has undergone significant disruption over the last decade or so, with increased regulation, continuous technology innovation, and changing customer preferences. To adapt, financial institutions have evolved business processes to not only efficiently manage the existing portfolio of products and services, but to also incorporate new products, new consumers, and new business rules. However, these changes have major impacts on data governance and management.

Data governance often gets short shrift when executing changes that are vital for the overall success of the organization. A financial services company with several business units and products like credit cards, housing finance, personal banking, and wealth and asset management will likely have frequent changes in business processes in each of these units. However, not all of them will follow all relevant data governance and data management procedures. For example, FDIC 370 may require source system changes and modifications to data structures within individual applications to achieve compliance, and an ongoing process for handling change to maintain compliance. Business process changes have the potential to add new data in the organization, which means an addition of new data uses 'u' and derived elements 'd'. Addition of new use-cases and derived elements adds to the existing problem of incomplete traceability from data sources to data consumers. As mismanaged data from each

business transformation program keeps accumulating, data duplication and overall disorder starts to creep up – back to data entropy.

To meet strict delivery timelines with resource constraints, business transformation programs will often cut corners. The organizational strategy of governing data assets to efficiently provision data for downstream consumers may be ignored and a solution for provisioning data on the fastest and cheapest route can get implemented. This is common in organizations where project sponsors, stakeholders, data stewards, and consumers are not aligned on the enterprise data strategy or enforcement is poor. This misalignment will result in data solutions being implemented in silos without leveraging enterprise architects and data architects, thereby creating inefficient workarounds. Siloed implementations that do not involve data governance teams that are responsible for managing metadata and maintaining the data glossary result in incomplete documentation. Siloed solutions can also create unnecessary data transformations implemented as a workaround to provide required data quicker. This will add an additional complexity of 'e' derived data elements. A business transformation process will now transform the complexity to:

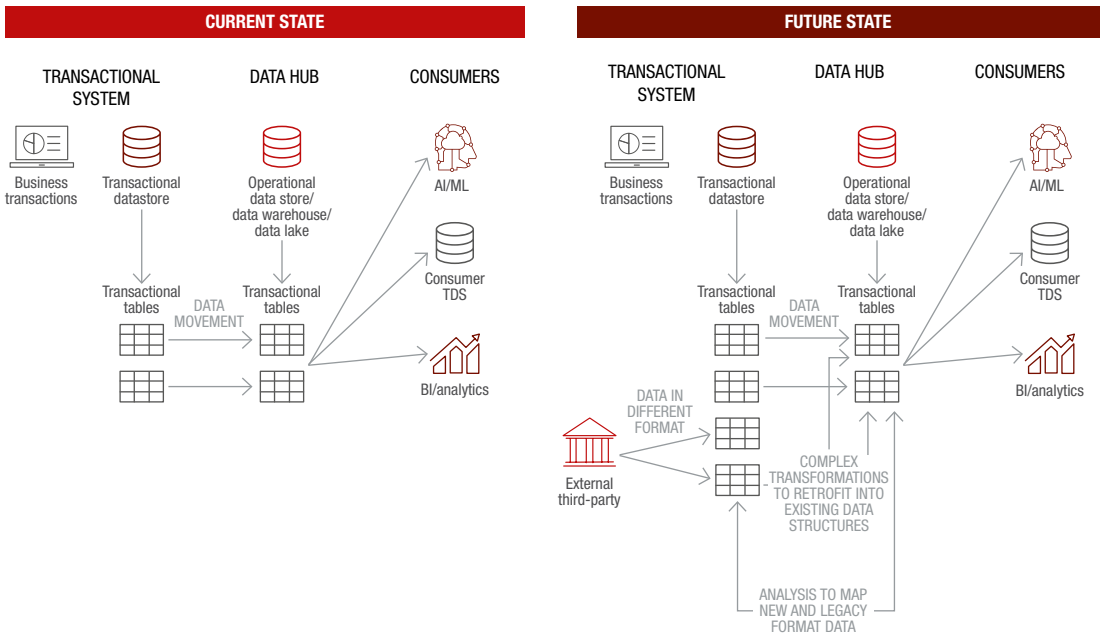
**Data complexity = *f*(d + e) derived data elements, number of systems, number of independent data stores, u + data uses)**

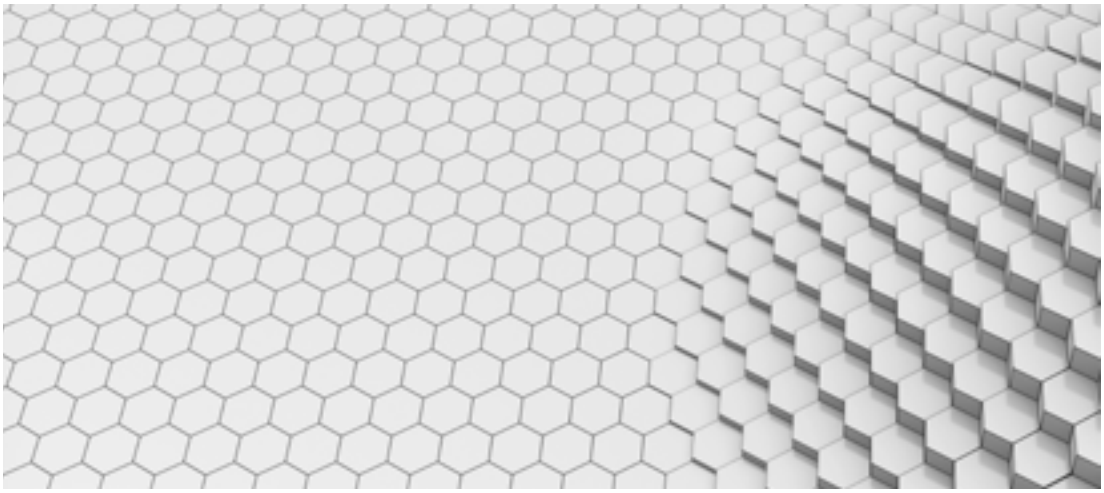
Business process changes implemented without an efficient solution for managing data change are a constant threat to an organization's data governance framework. The main challenge for the governance and management team will be to maintain a delicate balance – on one hand, enforcement of policies will delay the execution of the change and on the other, not implementing policies will create more problems in the future.

## 2.3 Technology inconsistencies

Structured data, unstructured data, big data, machine learning (ML), blockchains, and all the other emerging technologies are now buzzwords in every organization. Rise in social media and IoT (Internet-of-things) generated data has exponentially increased the ability to improve intelligence on customer preferences. This data is ingested and persists in many different formats, requiring a variety of technology solutions to process, organize, analyze, and present. What adds to the chaos is that advanced technologies required to ingest and process this data have to integrate with legacy architectures and code bases. Having volumes of data and advanced technologies like big data, ML, and data lakes is of no use if existing applications in an organization cannot consume this data. A large organization typically has hundreds of applications and it is unrealistic to expect that they will immediately switch to new data formats and subscribe to advanced data provisioning technologies.

**Figure 4:** Retrofitting new data into existing data structures





As the systems increase by 'm', they process more data in different formats, which increases demand for data resulting in provisioning of more data stores 'n', in turn increasing the number of derived elements 'd', thereby increasing the number of consumers 'u'. This brings the data complexity to:

**Data complexity = *f*(d + derived data elements, m + number of systems, n + number of independent data stores, u + data uses)**

Data ingested in different formats will now have to be transformed to fit legacy data structures. This is generally a substantial and resource intensive mapping exercise, which is complex to begin with and further compounded by incomplete data dictionaries and loosely modeled databases. Complicating matters further is the compatibility of technology solutions that enable data movement. Large organizations have several legacy IT components that may not work well with newer technology solutions. Substantial re-coding and re-architecting may be required to make things work seamlessly. These technology inconsistencies are a challenge to the data governance structure in an organization. Since there is no unambiguous solution, a lot of harmful workarounds can proliferate across the enterprise. Figure 4 illustrates a business process change where external data is added for better business decisions. When consumer adoption is not possible within the timeline and budget, a workaround is executed for data to be retrofitted to existing data structures. While this may superficially be quicker, it will likely turn out to be more expensive in the long run.

Retrofitting adds several transformation rules to derive existing data elements from new data sources. This adds 'e' elements to the overall complexity:

**Data complexity = *f*(d + e) + Derived data elements, m + number of systems, n + number of independent data stores, u + data uses)**

A complex data ecosystem relies on good metadata documentation and data dictionaries, which clearly define data elements and how they relate to each other. Data governance suffers when metadata and data dictionaries are not managed and documented, which is more likely to occur as complexity increases.

## 3. IMPACTS TO DELIVERY OF LARGE PROGRAMS

The importance of establishing and maintaining good data governance should have become apparent by now.

We have seen how data can turn into a major liability if it is not well managed. Threats to governance are plenty and not easily avoided. An organization needs to treat data governance as an ongoing activity, which gets stronger with every business initiative, merger, or technology implementation. Enterprise operations drive data entropy and large program implementations offer the potential to move towards better data order.

### 3.1 Quality issues and costs for on-going data quality teams

Inefficient data governance and management increases data quality issues, which surface while consuming and analyzing data across use-cases. For example, institutions may be looking to utilize a larger pool of customer data for up-sell, cross-sell, retention, and win-back purposes. Identifying and creating a unique ID for each individual account holder,

beneficiary, and beneficial funds owner is required under FDIC 370, forcing covered institutions to address this issue. Most institutions still do not have a holistic single view of the customer. However, a large implementation like FDIC 370 can make it a core requirement and develop the capability for enterprise-wide use (Figure 5).

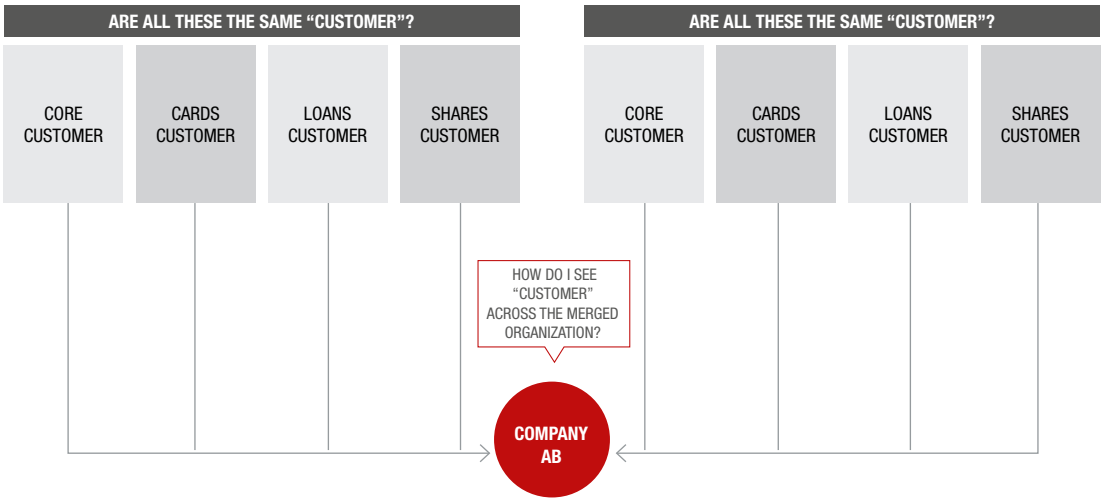
It should be apparent how bad data governance can hamper a venture that depends on good quality customer data. If good governance and management structure is not established between the two merged firms, this task will run into issues like:

- Different physical names for the same data element
- No data dictionary to identify the correct data elements from the new organization
- Erroneous results due to the use of incorrect data elements
- Incompatible metadata resulting in duplication of effort and resources lost in matching metadata between the two organizations
- Inconsistent data standards between the two organizations requiring resources to develop data transformation services
- Resources spending time in identifying the correct data to use instead of deriving insights.

As seen in the example above, identifying good quality data is of utmost importance for successful execution of a business venture. Organizations have established data quality teams that have infrastructures in place to ensure completeness and accuracy of data. When regular data flow is disrupted because of a business transformation (as detailed in Section 2.2 above) or integration with new technologies (as mentioned in Section 2.3 above), then the data quality infrastructure needs to adapt or change. It is sound data governance practice to do this, and most organizations do a good job of implementing data quality checks on major data stores and systems of record. However, it is the temporary data objects that can cause data quality nightmares. These temporary data objects create data duplication, which confuses the data consumer and may result in the use of incorrect data.

Data disorder also happens when new technologies and data standards are integrated. New data elements from a well-structured data model are mapped to legacy databases, which is done purely based on business definitions that may or may not be well documented. Often, organizations try to forcefully retrofit in order to serve data consumers that may not be receptive to change. Such instances exacerbate data quality issues arising from incorrect mapping, change in data batch job frequencies, change in valid values, and rounding versus truncating, to name a few. This causes data quality teams to spend extra time in analysis by navigating complex data mappings and system changes to identify the source, which requires additional resources and increases cost.

**Figure 5:** Incomplete M&A integration




Source: Joss (2016)[1]

[1] Joss, A., 2016, "The role of data in mergers and acquisitions," Informatica, December 16, https://infa.media/2oYDpdz

### 3.2 Project inefficiencies

A project is ideally based on a business vision or regulatory mandate. The ability of a vision to be executed depends largely on the impacts to current data, which further impact processes and consumers. The roadblocks and costs associated with mitigating these impacts have a huge influence on the scope of the project. Such initiatives often face large complications and risk, mostly in relation to data.

As mentioned in section 2.2 above, the relationship between business, systems, and data is highly connected and linked. Projects often do not prioritize data as much as the other two. Data typically comes into picture when the project is progressing with full steam, but then hits a bump caused by data quality issues and data governance workarounds. Project teams have to then rework some of their timelines, deliverables, and objectives, thereby creating inconsistencies and inefficiencies. The other approach that teams often take is to build workarounds – and we are back to data entropy. Figure 6 provides an example (from FDIC 370) of how legacy data complexities increase overall program implementation costs.

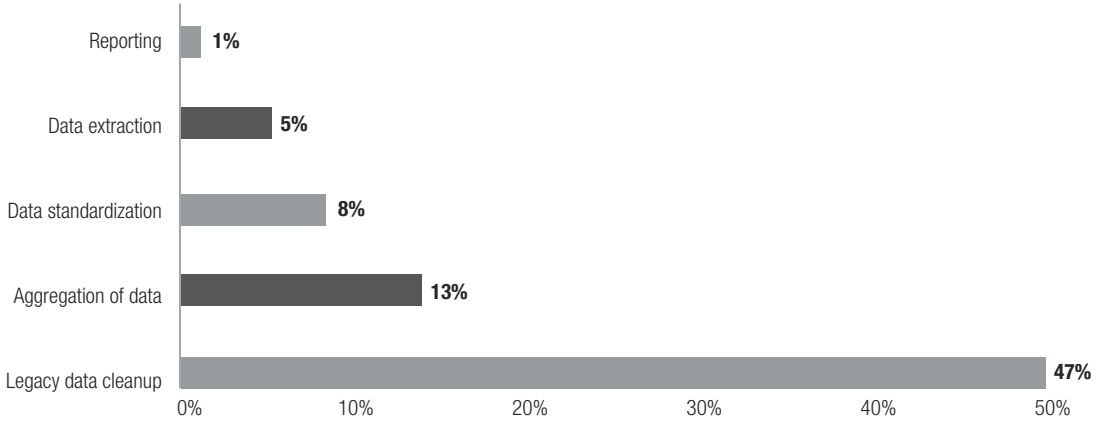
In addition to cost, program inefficiencies also come in the form of slower time to market. Inefficient data governance increases complexity and these complexities create roadblocks in implementation of large programs. Increased data complexity requires greater analysis and more time spent on data lineage, including sources and transformations. IT implementation teams face challenges with complex transformation logic and increased cost of data changes in multiple data stores that may have similar data.

## 4. LARGE PROGRAMS – A CATALYST FOR CHANGE

Clutter of our personal belongings is generated when we acquire new possessions and decide we need our old belongings as well. Surprisingly, it is these acquisitions that often trigger a need to clean-up our old belongings. Similarly, in the data world, a large program disrupts data governance and creates clutter, but it also presents an opportunity to reduce data entropy and drive data agility.

We have seen how large programs increase complexity and disrupt data governance, which in turn increases the net complexity that impacts delivery of large programs. Complexity is not a binary phenomenon, instead it operates on a continuum. Some level of data complexity is necessary to support business objectives and enable operational agility. An organization needs to establish guidelines or guard rails to indicate when data complexity is becoming worrisome and needs to be addressed. Establishing risk, cost, and benefit thresholds will help an organization determine when initiatives are material enough to warrant broader data management considerations beyond the needs of the specific program or project.

**Figure 6:** Data complexity drives average FDIC 370 implementation costs




FDIC 370 costs ~$1 dollar per deposit account to implement

75 cents out of each dollar spent on FDIC 370 implementation costs are driven by data complexity

50 cents out of each dollar spent on FDIC 370 implementation costs are driven by legacy data issues

Source: FDIC 370 Recordkeeping for Timely Deposit Insurance Original Rule Making Commentary

The solution then becomes straightforward. If complexity is driven by the number of systems, data stores, derived data elements, and information uses then an organization should leverage individual initiatives that reduce the impact of those variables. This can be achieved through:

- **Eliminating outdated or partially used applications and migrating to a single, common platform:** it may seem counterintuitive that the savings from keeping smaller legacy systems out of scope would be overwhelmed by future, hidden data management costs.
- **Re-architecting systems around a common data backbone:** to enable individual applications to leverage centralized services and isolate older systems and data structures. This is akin to many cloud implementations where application datasets leverage common utilities such as customer masters in an on-demand fashion to maintain data consistency while enabling flexibility.
- **Conducting periodic data cleaning:** at times, projects and M&A integrations are not fully completed, as resources become scarce and enterprise focus shifts elsewhere. Just as those junk drawers also need to be periodically cleaned, organizations should complete projects and remove temporary workarounds or fixes and eliminate those data "loose ends" that end up permanently in the back of the closet.
- **Establishing and utilizing common data definitions:** this becomes critical when managing related regulatory reporting regimes such as FDIC 370, CCAR, call reporting, and 2052a. Similarly, common datasets across business units and lines of defense are often overlooked and can be streamlined during specific initiatives to the long-term benefit of the organization.
- **Capturing data in source systems using common data definitions:** this reduces the amount of data derivation required, enables faster system migrations as data anomalies are limited, and supports easier maintenance of centralized data pipelines.
- **Creating simple frameworks for data assessment and lineage:** will strengthen overall data management.

## 5. CONCLUSION

Large programs present an opportunity to implement leading practices in data management. An organization that instills a culture where data is seen as an enterprise asset will be successful in ensuring that every large program contributes to the enhancement of the organization's data ecosystem. Large programs come armed with budget, resources, executive support, and a mandate for change. Enforcement of data governance and upholding standards can go a long way in managing complexity in large programs.

### Case study: Resolving enterprise data clutter through FDIC 370 implementation

The FDIC began work on a new rule for Recordkeeping for Timely Deposit Insurance Determination shortly after the financial crisis. After resolving IndyMac and facilitating the sale of Wachovia to Wells Fargo in 2009, the FDIC recognized that the largest banks had too complex of a technology and data environment to enable efficient bank takeover in the event of a failure. As an answer, the FDIC shifted the burden for maintaining information and developing an application to all large financial institutions to calculate deposit insurance, report on beneficial ownership, and quickly make funds available to depositors.

To comply with FDIC 370, Covered Institutions (CIs) are required to create a unique identifier for each customer, assign the appropriate FDIC ownership code to each account, confirm that supporting documentation exists to support these classifications, and run the deposit insurance calculation by aggregating ownership across these categories. In addition, FDIC 370 banks need to be able to quickly ingest information from third-parties to complete insurance calculations within a short time after failure. Finally, banks need to produce reporting that supports annual certification of IT capabilities by the CEO/COO. This data driven compliance effort has exposed covered institutions to many legacy data challenges.

Large banks have complex data environments. CIs have had to integrate data from a variety of source systems, establish unique customer IDs across platforms, map data into centralized data stores, and create new data outputs derived through the assignment of FDIC ownership codes and deposit insurance calculation. This rule has driven institutions

to create new data stores, complete master customer record initiatives, and remediate legacy data from prior mergers and across the enterprise. However, CIs have also used the requirements of FDIC 370 to address data complexity and position the organization for future opportunities.

Some institutions, in preparing for FDIC 370 have linked its requirements to core deposit transformation initiatives. FDIC 370 requires banks to collect new and updated information at the time of customer onboarding, account opening, and maintenance for both. Linking these compliance requirements to changes in customer or deposit operations enables the bank to achieve business enhancement while meeting the compliance requirements for improved data. In addition, streamlining data capture processes reduces data variability enabling stronger analysis of customer activity. This enables cleaner data to feed product level analytic processes and customization of client specific offers.

FDIC 370 requires banks to be able to ingest data from third-parties on individual account holders and beneficiaries. This has prompted CIs to leverage standard data structures for bringing information into the deposit insurance calculation processes. Standardization of data inputs into the centralized calculation engine also enables these banks to connect with other internal systems in a streamlined manner.

In many cases, FDIC 370 banks have grown through acquisition. As a result, the CI has inconsistent data structures as legacy account and customer setups were often not integrated into common data structures. Supporting documentation, such as signature cards, were never scanned into imaging systems. Data analysis has highlighted the existence of these anomalies leading to systematic or manual data remediation and customer outreach. In some cases, AI use-cases have been identified to accelerate the clean-up processes. Addressing FDIC 370 with an AI toolset has enabled the institution to experiment with emerging technologies and use them to address specific business and compliance needs.

CIs have established common data definitions for both operational and other regulatory reporting requirements. The annual summary reporting and certification requirements align to CCAR, Call Reporting, and 2052a. Progressive institutions have directly linked FDIC 370 to these other efforts and aligned data elements and specific aggregations of information to enhance all related programs.

One question FDIC 370 banks have addressed is the degree to which data should be developed and assigned at a source system level or be derived later. A key element in the data stream is the mapping of accounts to FDIC Ownership Rights and Capacity codes. These codes are FDIC based ownership categories that banks have not previously needed to maintain. In most cases, these codes are derived based on combinations of account tiles, relationships, customer types, and product indicators. In some cases, banks need to derive data elements to do this mapping, such as an indicator if the trust is revocable or irrevocable. Some banks have chosen to completely derive this data. However, other institutions are pushing these business rules into core deposit systems and will display the ORC assignment as the account level. This enables front line staff to aid customers in understanding insurance coverage and managing account types to maximize this benefit.

Most FDIC 370 CIs have completed some form of customer alignment. This ranges from data quality initiatives to reduce the number of duplicate customer records to complete redevelopment of customer master files. FDIC 370 requires banks to be able to uniquely identify each customer and tie all accounts to each individual. This has included linking non-core systems such as wealth or trust into these efforts. Completing customer related master data management opens the door to future omnichannel services and enables the institution to gain a better view of customer activity to link to future growth opportunities.

In general, FDIC 370 has served as a catalyst for reducing data clutter and improving data management. Institutions that had more proactively managed their data environment have had an easier time implementing FDIC 370 requirements. Just like the homeowner who periodically cleans out and organizes the basement or garage, organizations that have embraced leading practices in data management have found themselves better prepared for significant projects and are able to manage data clutter and data entropy.

## 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.

## WORLDWIDE OFFICES

**APAC**
Bangalore
Bangkok
Hong Kong
Kuala Lumpur
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**

CAPCO