
White Paper

Data Sourcing: The Overlooked Data Capability

By Robert Grant Beauchamp

Contents

Introduction	2
History of Sourcing Data	4
Data Sourcing as a Corporate Capability	8
The Three Capability Tiers of Data Sourcing	9
The Time for Data Sourcing has Arrived	10
Making the Case for Data Sourcing	12
Summary	14
About the Author	15

White Paper

Data Sourcing: The Overlooked Data Capability

By Robert Grant Beauchamp

Introduction

The clarion call is clear. Data is the key to the future. Everything from IOT to machine learning to health management not only depends on data but also generates seemingly limitless amounts of data, which in turn can be leveraged for new purposes.

But if you listen closely, one also hears, sometimes in hushed tones, sometimes with a cynical joke or snide comment, and sometimes in the very silences of what remained unspoken, another theme.

That theme is that the data is seemingly not that good or valuable, or more specifically it is very difficult to work with, if available at all. Valuable, actionable data is hard to find and identify, it is time consuming to get into usable shape for analytics, it is not uniformly defined or formatted making true interoperability still a dream. And as for universal access and information sharing across different entities . . .

Evan as we are in a new age of data collection, distribution and usage organizations are still grappling with their own internal use of their own data for their own purposes.

Every system upgrade, new web portal, new EDI trading partner, new report request, and new security audit exposes this problem. It is often only overcome with exhausting facilitated sessions, long nights and damaged personal relationships to get the data into the state it needs to be in, to be, at a minimum, usable. And often this work is lost or forgotten until a new need arises and the painful process is started all over again.

Compounding the difficulty of formatting data into usable forms is a new breed of extremely smart BI analysts and data scientists who are continually dismayed at how much work they must do to the data before then can even begin the analysis of that data.

Skills, knowledge and tools are racing into the future, but their potential is being blunted by the overall state of the data in many corporate computing environments.

Purpose

This clearly indicates that there is a large hole in the ability of these organizations to deliver on their future.

There is an important comprehensive business and IT *capability* that organizations must identify, understand, implement and master.

This capability is called **Data Sourcing**. It involves building an organization's capability to find, identify, understand, classify, map and share its data. It also clearly demonstrates how that data flows through, and is transformed by, the enterprise computing environment.

The purpose of this document is to provide:

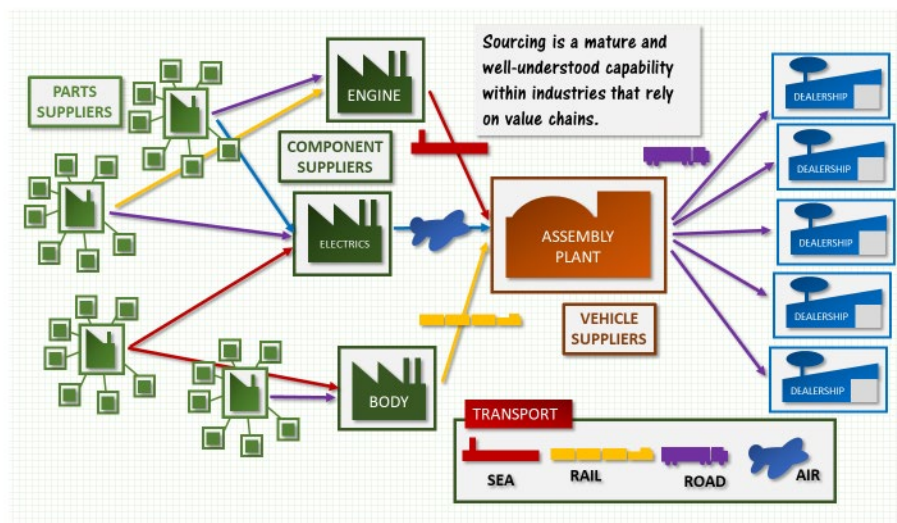
- A high-level overview of the history of sourcing data in an enterprise environment.
- An introduction to Data Sourcing as a corporate capability.
- A rationale for making Data Sourcing a corporate priority.

History of Sourcing Data

What is Sourcing

Most of us are familiar with the concept of a supply chain and how it works in the manufacturing of automobiles. A core capability of each channel partner is the ability to source and procure what that partner needs to add value to the chain.

- When sourcing, buyers evaluate parts on quality, availability and cost; choosing the best supplier to meet the need.
- Successful sourcing requires intimate knowledge of suppliers, specifications, schedule and how parts are to be used.
- Complexity is no excuse for inefficiency. The more complex sourcing requirements become, the better the organization needs to be at it.

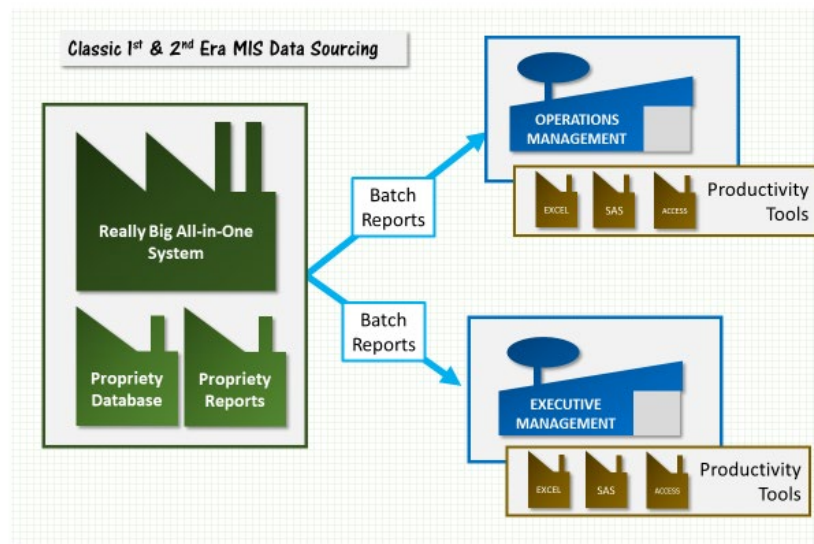


The complexity of a supply chain is no excuse for not getting it done right.

Classic 1st and 2nd Era MIS Data Sourcing

To understand data sourcing, it helps to go back in time. In the first and second eras of information management there was usually only one major source of data, and that was the primary business system of the organization. This business system usually had a proprietary database and a proprietary report generator.

Data analysis consisted primarily of reviewing batch or 'green bar' reports so named for the type of paper they were printed on. Later, electronic versions of these reports became available so that staff could load the data into personal productivity tools such as excel and access and early data analysis tools such as SAS. Despite only one primary source of data, even then, database maintenance, along with generating accurate timely reports and deleting obsolete ones, was often difficult for some organizations.



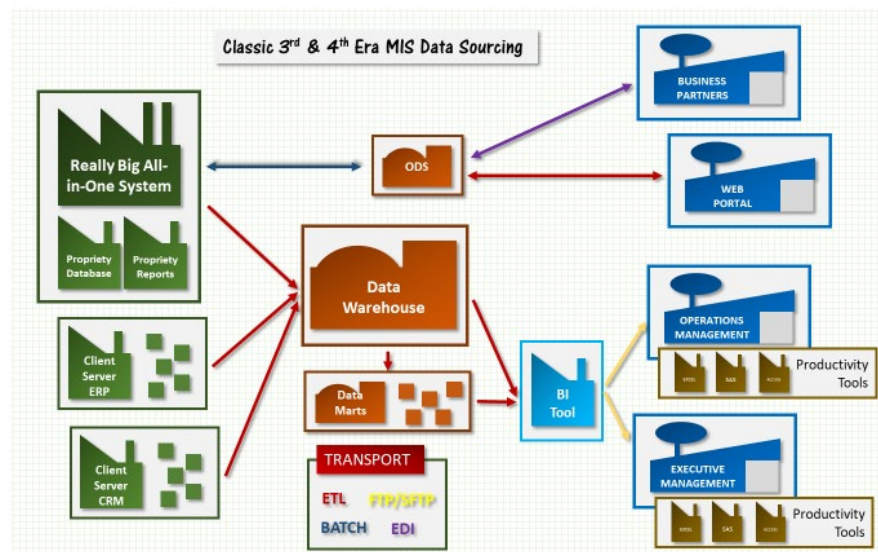
Even in simpler times, getting information to where it was needed was hard.

3rd and 4th Era MIS Data Sourcing

The third and fourth eras saw the introduction of client server systems and more specialized systems such as CRM and ERP, significantly increasing the volume of data produced by an organization. The desire to utilize data from multiple sources and offer it up in a reliable, consistent manner led to the development of data warehouses and data marts.

Dedicated Business Intelligence tools were leveraged to take advantage of this newly available data. These eras also saw the introduction of new methods of moving and transforming data such as EDI and ETL.

Overall, the dynamic changed from *pushing* data to stakeholders to that of stakeholders *pulling* data to meet their operational and analytical needs.



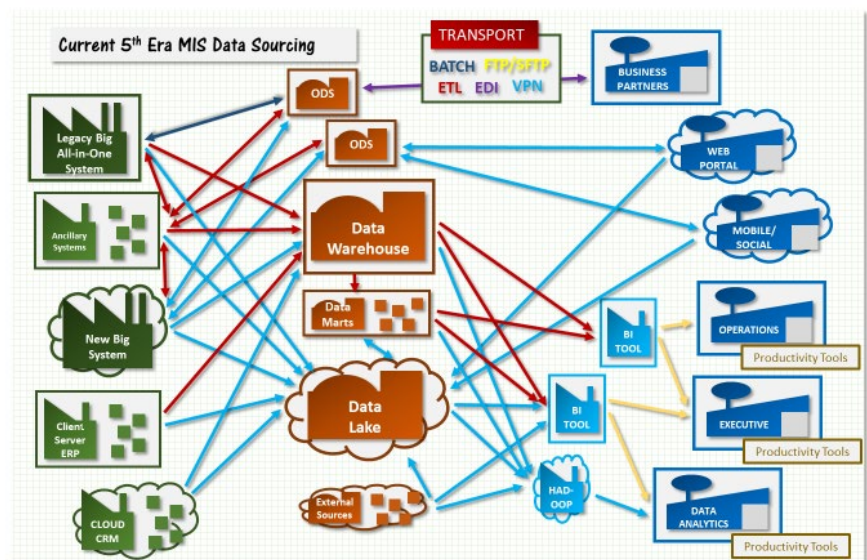
Tensions were often high between the suppliers of data and the consumers of data.

Current Era Data Sourcing

In the current era data sources, transport methods, stakeholders and sophisticated data analysis tools have proliferated. Information systems management has achieved the complexity of the most complex industry supply chains, but it has not, unfortunately, achieved their efficiency and effectiveness.

Locating, preparing and moving data, along with managing the stress on available computing resources, for many organizations, is still a haphazard and often painful process, often leading to frustrating bottlenecks, inconsistent data quality, and unreliable reporting.

The need for efficient and effective data-related capabilities is greater than ever.



Information value chains have become as complex as any found in industry.

Data Sourcing as a Corporate Capability

What do we mean when we talk about a capability

Sourcing data is an *activity*, organizations do every time they populate a web portal, run a report from a data warehouse or analyze unstructured data from a data lake.

But to be considered *capable* indicates a level of competence, performance and consistency. If words like haphazard, inconsistent, unreliable, difficult, frustrating are being used to describe your data and your data environment, it is time to

Sourcing data is an activity organizations do every time they populate a web portal, run a report from a data warehouse or analyze unstructured data from a data lake.

ACTIVITY

Haphazard Processes
High Stress
Unpleasant Surprises
Extensive Data Prep
Inconsistent Results
Ongoing Maintenance Issues
Lost Opportunities
Back-End Quality Control
Uneven Management

CAPABILITY

Repeatable Processes
Routine
Clear Expectations
Reduced Data Prep
Reliable Results
Reduced Maintenance Issues
New Discoveries
End-to-End Quality
Comprehensive Management

To upgrade an ad-hoc activity to a routine capability requires putting the talent, skills, resources, tools, techniques and processes in place to enable it.

© 2018-2019 ROBERT BEAUCHAMP CONSULTING

Frustration on the part of data consumers is often directly related to the capableness of an IT function.

The Three Capability Tiers of Data Sourcing

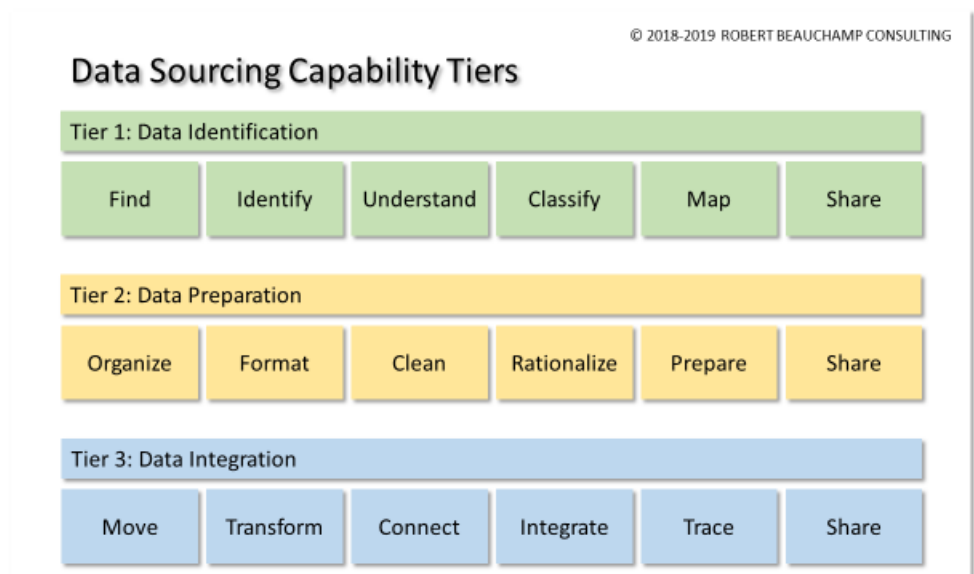
There are three tiers to data sourcing, each tier building on the one prior. They are data identification, data preparation and data integration.

Each of these tiers, in turn, have capabilities that support them. The organization must be able to find data, identify it, understand it, classify it, map it and, most importantly, share this information across the organization so it can be used by anyone in the organization.

To make the data usable the organization must be able to organize it, format it, clean it, rationalize it prepare it and share this prepared data for use by the entire organization.

In order to leverage the data, the organization must be able to move it, transform it, connect it, integrate it, and trace its path so that it can be integrated into the organization's systems and tools.

Data Sourcing also goes far beyond the data in the organization's own data environment. An organization must have the ability to source data from multiple external data providers *as effectively as it sources its own data*.



You can rate your organization on each of these capabilities. How well does your organization perform them and how many problems are caused by not performing them well?

The Time for Data Sourcing as a Capability has Arrived

<p>The problem</p>	<p>Companies have been dealing with inadequate data for a long time, this is nothing new. Organizations have been trying to get their data houses in order for decades.</p> <p>So why haven't they? It certainly wasn't because the problem was hidden. On the contrary it was painfully obvious almost every day. It certainly wasn't due to the lack of thought leadership. Data Management, Master Data Management, Data Quality, Data Stewardship and similar important concepts and disciplines have been part of the common data lexicon and toolkit for years.</p> <p>To put it bluntly; it just wasn't important enough. Business operations, customer relations, and even financial reporting were negatively impacted, yet many organizations chose to muddle through. Smoothing things over, fixing things on the back end.</p> <p>Not to say nothing wasn't attempted. Countless initiatives have been launched by committed data warriors, enlightened management and talented consultants. Many to good effect, but many others simply fading into insignificance due to a lack of attention and resources, overwhelmed by the everyday priorities of running a business.</p> <p>The pain never reached the level that made it worth the time, money, effort and amount of enterprise-level coordination and cooperation to do what needed to be done.</p> <p>It simply wasn't a strategic or existential problem. There was no compelling need. Things stayed much the same and life went on.</p>
<p>So, what's different now?</p>	<p>Why should organizations focus time, energy and resources on raising data sourcing to the level of a corporate capability?</p> <p>Because now it has become aa strategic and even existential problem.</p> <p>Several significant forces are finally tipping the balance, causing corporate leadership to understand that their data and how they manage it is strategic. <i>It is necessary for their success if not for their very existence as a business.</i></p>

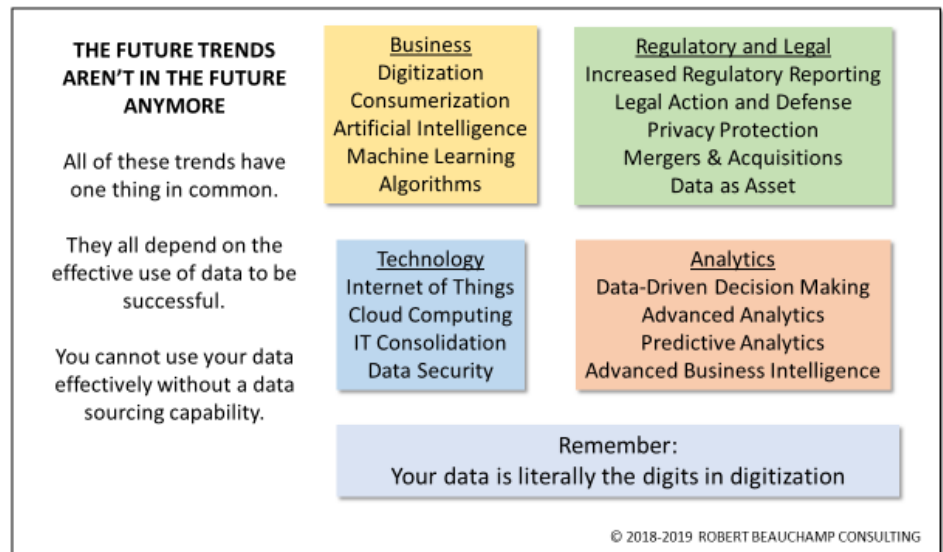
Many organizations are beginning to grapple with these forces and some are already neck deep in them. But what is significant is that these forces are complex, they often overlap, and they are all happening *now*.

These forces include:

- Digitization and Digital Transformation
- Data as Asset
- Data-Driven Decision Making
- Consumerization
- Internet of Things
- Cloud Computing
- Advanced Analytics and Business Intelligence
- Artificial Intelligence, Algorithms and Machine Learning
- Robotics
- Increased Regulatory Scrutiny and Reporting Requirements
- Data Security and Privacy Protection
- Preparation for, and response to, legal actions related to data
- IT Consolidation resulting from Mergers and Acquisitions

For those organizations just entering the fray, they are quickly coming to the realization that ***if you can't use your data properly you simply can't do these things.***

And it can be asserted that you can't use your properly if you lack a Data Sourcing capability.



An organization may not get to decide when the optional becomes conditional.

Making the Case for Data Sourcing

Key Messages

It might be hard to believe, but there was a time, and not so long ago, that there was no formal project management, quality assurance, portfolio management or agile development. These disciplines came on line as the need for their capabilities became painfully obvious.

And it also may be hard to believe each was met with significant resistance in many organizations. The need for Data Sourcing as a capability will also be, if it is not already, painfully obvious. And if organizations are true to form there will probably be some resistance.

It is important to remember and remind our colleagues of the value of Data Sourcing as a corporate capability.

Gets your data where it is needed, in the condition it is needed, as fast as possible.

Today it is all about speed, speed, speed. There is no other single capability you can add to your organization that lets you leverage your data faster and cheaper than data sourcing.

Bring your data and your data environment out from the shadows.

How can you use what you don't know you have? How can you turn your data into a corporate asset usable by everyone, if it is hidden behind layers of technology, jargon and organizational turf wars? Know what you own.

Answer the tough questions

When it comes to audits, data breaches and litigation, demonstrating a mastery of your data and data environment goes a long way to protecting your reputation when you are managing a crisis. When the going gets tough, the tough have actionable data.

Resilience

All the leading experts say we are experiencing profound technological change at levels never seen before. It's been called the fourth industrial revolution or the second machine age. Nonetheless, no matter how the technology changes, how indeed the world changes, it will require you to be able to leverage your data more than ever before.

Data Driven Decision Making

While it would seem to be obvious, and unfortunately for many organizations it is not, the value of your data-driven decisions is directly related to the value of your data.

You're already doing it, so why not do it well

Unlike many corporate enterprise initiatives that start from the top down with a blank slate, your organization has many of the capabilities necessary for data sourcing already in place. You just need to do them better and in a more intentional manner.

Don't stumble on your way to the future

Fortune favors the bold if your data is valuable. Whatever forward-thinking strategic initiatives you will engage in, whether digitization, advanced analytics or artificial intelligence, you don't want to be held back because you don't know your data or your data environment. Get ahead of the game and reap the benefits of your strategic initiatives sooner.

Data Sourcing leads to knowledge and mastery over, what in the future, will probably be the single most important asset your company owns.

Summary

It should be no surprise to anyone that we are on the cusp, and probably beyond the cusp, of a foundational change in culture, society and business.

Driving that change are technologies such as Artificial Intelligence, the Internet of Things, and Big Data analytics. The impact of this change on organizations can be seen in trends such as Digital Transformation, Increased Regulatory Requirements and Data as an Asset. No matter the technology or trend, however, they all have one thing in common. They utilize data.

Therefore, the ability to make the best use of data is key to success in this new world. An organization will not be able to utilize its data in the best possible way without a data sourcing capability, any more than a manufacturer can deliver the affordable high-quality automobiles the market demands today without mastering its sourcing capability.

It simply can't be stated often enough, if your organization wants to be good at any of these forward-looking technologies and trends it must be good at data sourcing.

Remember, your data is literally the digits in digitization. Adoption of a comprehensive enterprise-level Data Sourcing capability is a prerequisite for any plans for an existence in the business world of the future.

About the Author

Robert Grant Beauchamp is a consultant, architect, and former CIO with a proven record of helping organizations understand and improve their data quality, data environments and data-based capabilities. As a systems integrator, Robert has successfully introduced and implemented data-related technologies such as BI, EDI, ETL, data warehousing and three-tier architectures.

If you would like to learn more about data quality or would like help improving data quality within your organization, connect with him on LinkedIn

Robert has filled the roles of computer journalist, tech writer, business analyst, marketing communications manager, business architect, project manager, program integrator, program manager, account manager, data security consultant, solutions architect and trusted advisor, including:

- Thirty years of business and information technology experience including over five years as Chief Information Officer of a rapidly growing health plan.
- Proven track record of successfully strategizing, developing and implementing enterprise-level business and technology initiatives in the health care and financial services industries.
- An experienced and practiced consultant with the ability to work with C-level executives to develop strategy, assess capabilities, manage risk, and offer solutions that can be successfully implemented in an organization's unique environment.
- A proven communicator well versed in public speaking, meeting facilitation, webinars, journalism and video.
- A proven history of developing and implementing successful service offerings for a major IT consulting firms including Y2K, HIPAA Security and Privacy and HIPAA Electronic Transactions.

Currently Mr. Beauchamp is championing a capability-based approach to data quality. He is a leader in the adoption, education, and implementation of Data Sourcing as a corporate capability. He is in the process of writing a book on Data Sourcing.