

STRATEGIC PERSPECTIVES SERIES

# The Business Value of a Data Catalog

A Boomi Perspective

By Dave Wells

May 2020

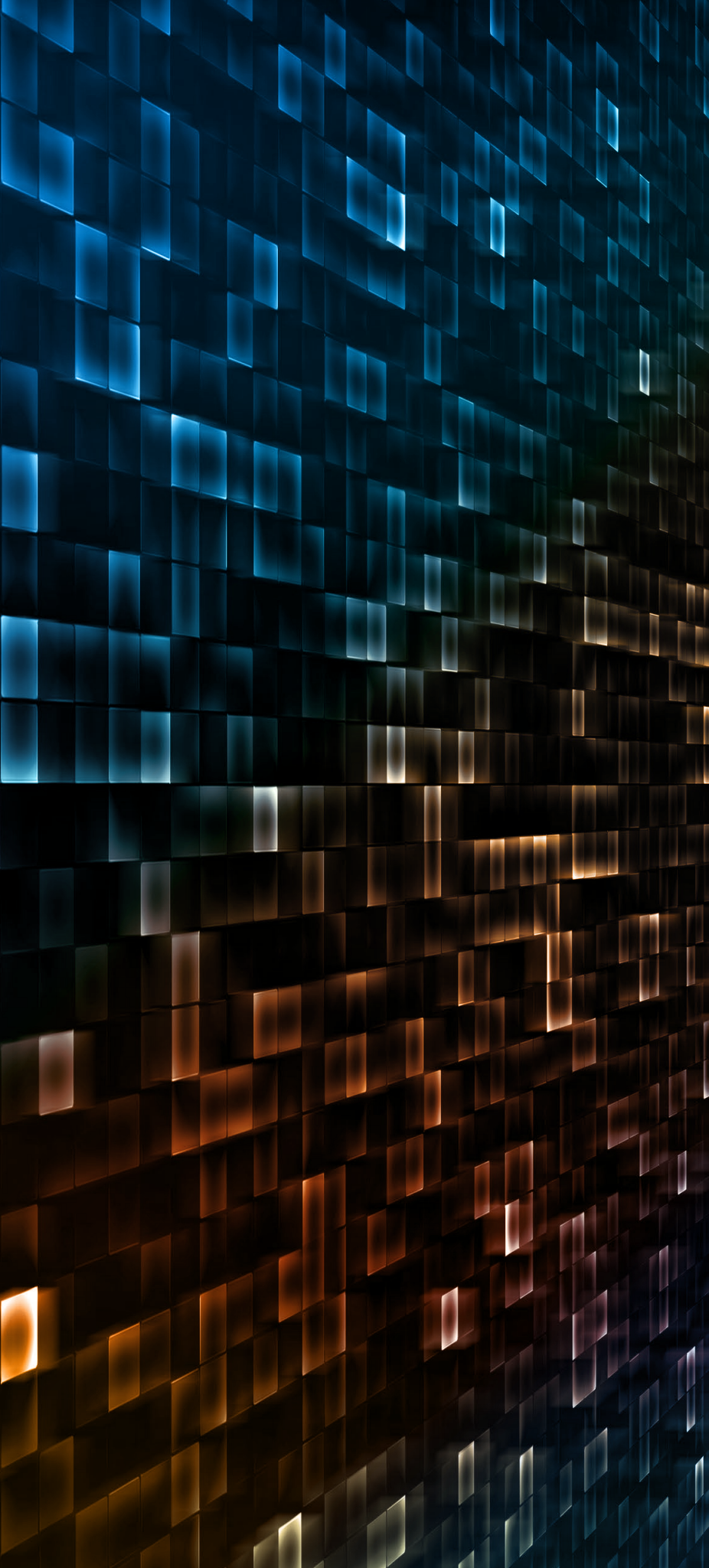


Research Sponsored by

**boomi**



This publication may not be reproduced or distributed  
without prior permission from Eckerson Group.



# Table of Contents

<b>Introduction</b> .....	3	<b>Boosting the Value of Business Data</b> ....	11
<b>The Value of a Data Catalog</b> .....	5	<b>About the Author</b> .....	12
<b>A Data Analysis Example</b> .....	5	<b>About Eckerson Group</b> .....	13
The Scenario .....	6	<b>About Boomi</b> .....	14
Comparing Time and Cost of Analysis ....	9		
The Impact on Business Outcomes.....	9		
<b>Data Catalog Return on Investment</b> ....	10		
Direct Financial Impact.....	10		
Additional Savings .....	10		
Value from Governance.....	10		
Workplace and Cultural Value .....	11		





# Introduction

**Over the past several years, the difficulties of data management have steadily intensified.** That's been due in large part to the complexities of big data, cloud hosting, self-service analytics, and tightening regulations. As a result, effective data management has become a top priority for most organizations. But getting there can be a challenge. One essential tool for overcoming these challenges is a data catalog.

Data catalogs were originally introduced to help data analysts find and understand data. Before then, most data analysts worked blind — that is, without visibility into data-set contents, or their quality and usefulness. That meant analysts had to spend much of their time finding data, understanding data, and re-creating data sets that already existed. With the advent of data catalogs, these issues were resolved, as data catalogs offered analysts a new way to manage their data inventories and expose their data sets.

Since then, data catalogs have added new functions, becoming more popular and

increasingly important. Modern data catalogs continue to meet the needs of data analysts, but their reach has expanded. Today they're also vital components of data stewardship, data curation, and data governance.

By touching nearly everyone who works with data, data catalogs create value in new ways, including:

- Faster analytics
- Easier collaboration and reuse
- Increased trust in data
- Direct answers to business questions with support for natural language queries

This eBook is the first in a series exploring the data catalog's business case. It offers a data-analysis example, both with and without the use of a data catalog. This eBook also demonstrates how a data catalog can earn its keep by showing a strong return on investment (ROI).

*Data cataloging is an investment that goes beyond software licensing and includes the cost of people, time, and projects.*

This eBook is complemented by the two other titles in the series: “The Real Cost of Reinventing the BI Wheel” and “The Value of Trusted Data and the High Cost of Uncertainty.” The complementary titles dig deeper into the ways a data catalog can add value for organizations building their data analytics ecosystems.

For those who are just getting started with data cataloging, consider the value opportunities described in this eBook as fundamental concepts when making the business case for a data catalog. It’s an investment that goes beyond software licensing and includes the cost of people, time, and projects. Those who are already underway with data cataloging can employ these opportunities to maximize the reach and impact of the data catalog. Those

working to sustain and grow data cataloging can measure data catalog success with metrics such as:

- Data analysis time
- The ratio of time spent in data preparation versus analysis
- Data catalog return on investment
- Frequency of shared data and shared analysis
- Levels of trust in data and analysis

These metrics can even be applied to measure and manage the adoption of data cataloging based on a maturity model with dimensions of agility, quality, speed, collaboration, and trust.



## The Value of a Data Catalog

A data catalog is an investment with real costs in money and time. Beyond buying the software, an organization that builds a data catalog must also hire and train people and populate the catalog with metadata. Companies building a data catalog often struggle to describe its business case. They also have a hard time answering a common yet important question: What will be the data catalog's ROI? Some organizations will use anecdotal cost justifications. These focus on a frequently described situation: Analysts spend 80 percent of their time finding and preparing data, but only 20 percent of their

time analyzing it. Another benefit is a higher quality of analysis. It's listed when analysts can find and use the best data sets, instead of just using those that are easy to find. Yet another advantage is centralizing and sharing "tribal" knowledge; this prevents a loss of knowledge while also increasing collaboration, communication, and sharing.

While these are good examples of how data cataloging can deliver value, they're soft, anecdotal, and non-quantitative. They provide no guidance when the goal is to estimate or calculate a data catalog's ROI.

*A difficult but important question: How is ROI for a data catalog determined?*

## A Data Analysis Example

We'll begin with a data-analysis scenario that illustrates the process of working with and without a data catalog. There are three options for completing an analysis project. Option 1: analysis through an IT project. Option 2: self-service analysis without a data catalog. Option

3: self-service analysis with a data catalog. As we explore each option, we'll also note the time and cost of analysis — quantitative variables that provide the foundation for calculating ROI.



# The Scenario

Meet Kate, the manager of the online sales line-of-business for a midsize retailer with annual e-tail revenue of \$216 million. Kate is evaluating product-bundling as a way to maximize the value of each customer sale, and she needs to make decisions about product bundling and pricing. Kate knows that bundled items sell at a lower per-unit price than do the same items when sold separately. But she has questions about the optimum number of items in a bundle, the right mix of items, and the cost savings for bundled purchases. Kate seeks the sweet spot where price and sales volume combine to produce the highest revenue.

## OPTION 1 – AN IT PROJECT

Kate’s first option is to request that her company’s IT group build a bundled-sales simulator. Then she can use this simulator to explore the sales volume and revenue impacts of different bundling options. This approach will have the following timeline:

WEEK 1	Develop and submit project request
WEEKS 2–3	Wait in project queue
WEEK 4	Requirements gathering and definition
WEEKS 5–9	Design and development
WEEK 10	Testing and debugging
WEEK 11	Bundle scenario simulations
WEEK 12	Bundle and pricing decisions made with moderate uncertainty

In other words, Kate will have to wait three months from when the analysis need is first identified before she can finally implement bundling. Then, after operating bundled sales for a few months, Kate will want to evaluate their performance. Continuing to work through

IT, she will request a report of bundled sales performance showing trends in sales volume and revenue generation. Then she’ll experience:

- Some delay while waiting in the project queues
- A few days of requirements definition
- Two to three weeks of design and development
- A few days of testing

Finally, Kate will be able to view and evaluate the performance of bundled sales, comparing the actual and simulated results. She may then choose to revise and refine bundling and pricing, or she may request modifications to the bundle simulator.

## OPTION 2 – SELF-SERVICE WITHOUT A DATA CATALOG

Kate's second option is to perform the analysis using a self-service data visualization and analysis tool, but without a data catalog. Although Kate is comfortable with Excel, she recognizes that the analysis she needs is too complex for that software. So Kate decides to step up to Qlik, which already has some presence in her company. Here's the timeline for this approach:

DAY 1	Acquire Qlik and learn the basics
DAY 2	Frame the problem and determine data needs
DAYS 3–5	<div>Search for data<ul style="list-style-type: none"><li>• Ask IT if they can provide any of the data needed</li><li>• Ask colleagues and work the tribal knowledge network</li><li>• Find some “close enough” data through tribal knowledge</li><li>• Use some spreadsheet data from a past project (although not quite the right fit and somewhat dated, it is easy to acquire)</li><li>• Wait for IT to deliver a data extract</li></ul></div>
DAYS 6–9	<div>Data preparation<ul style="list-style-type: none"><li>• Load the data into Excel</li><li>• Manually blend the data from different sources in a single workbook</li><li>• Add formulas to calculate new variables for analysis</li><li>• Determine needs for additional data</li></ul></div>
DAYS 10–11	Find and acquire additional data
DAYS 12–13	<div>Continue data preparation<ul style="list-style-type: none"><li>• Manually blend in newly acquired data</li><li>• Format for Qlik compatibility</li></ul></div>
DAY 14–15	Data visualization and analysis
DAY 15	Bundle and pricing decisions made with some uncertainty

With self-service, Kate can make decisions and begin to implement bundling in just three weeks instead of 12. Once the software is implemented, Kate can monitor her bundled sales performance. Then, to continue working with Qlik, she'll need to:

- Find and acquire data for sales-performance monitoring
- Build the data-preparation workflow for her sales-performance dashboard
- Build the sales-performance dashboard
- Make the sales-performance dashboard operational

Finally, Kate can continuously monitor the performance of bundled sales, comparing the results to her goals. If there's a gap, she may then refine her bundling and pricing strategy. And as new questions arise, Kate may also modify the dashboard.

### OPTION 3 – SELF-SERVICE WITH A DATA CATALOG

Kate's third and preferred option is performing the analysis using a self-service data visualization and analysis tool, this time with a data catalog. But this option is not entirely under her control; it's available to Kate only if her company has already acquired and implemented a data catalog. Assuming that has happened, Kate can use a combination of Boomi Data Catalog and Preparation along with Qlik to quickly find and prepare data, perform the needed analysis, and get on with her decision-making and bundling implementation. For this approach, here's Kate's timeline:

DAY 1	<ul style="list-style-type: none"><li>• Acquire Qlik and learn the basics</li><li>• Frame the problem and speculate about data needs</li></ul>
DAY 2	<ul style="list-style-type: none"><li>• An hour or so to find the data<ul style="list-style-type: none"><li>– Search using the Boomi Data Catalog</li><li>– Find similar analyses conducted by others</li><li>– Find trusted data</li><li>– Build project specific data preparation workflows</li></ul></li><li>• Visualize and analyze<ul style="list-style-type: none"><li>– Deliver prepared data in the right format for Qlik analysis</li><li>– Visualize and analyze iteratively</li></ul></li></ul>
DAY 3	Bundle and pricing decisions are made with confidence in the quality of data and analysis

With this approach, a mere three days after Kate has identified the need for analysis, she's able to make decisions and start to implement bundling. As the implementation proceeds, she'll monitor the bundling initiative daily. To do this she can:

- Revisit Boomi to find and prepare data for monitoring
- Use Qlik to build her monitoring dashboard

Kate can also actively monitor and manage bundles and pricing with confidence. She can use data cataloging, data preparation, and self-service analysis tools to quickly find answers as new questions arise and the bundling program evolves.



# Comparing Time and Cost of Analysis

Reviewing these three options, it's clear that Kate gets what she needs much faster with self-service analysis than with an IT project. And she saves even more time by using self-service analysis in conjunction with the Boomi platform for data catalog, data discovery, and data preparation.

The cost savings are equally important. Kate's annual salary plus overhead is \$120,000. The average cost of an IT analyst in her company is \$100,000. Working with these numbers, we can calculate the direct cost of analysis for each option:

*Analysis cost using Boomi for data cataloging and data preparation is only 4 percent of IT project cost!*

	OPTION 1	OPTION 2	OPTION 3
Time to analyze	3 months	3 weeks	A few days
Cost to analyze	\$17,300 <sup>1</sup>	\$3,460 <sup>2</sup>	\$690 <sup>3</sup>

<sup>1</sup> 9 weeks of IT analyst @ \$100,000 annual salary

<sup>2</sup> 50% of Kate's time for 3 weeks @ \$120,000 annual salary

<sup>3</sup> 50% of Kate's time for 3 days @ \$120,000 annual salary

The difference? The cost of analysis with Boomi is only about 20 percent of what it would be using self-service analysis without it. And it's only about 4 percent of what the project would cost with IT. The time saved is similarly impressive. Working with Boomi takes only about 25 percent of the time of analysis without it, and less than 4 percent of the time it would take with IT.

## The Impact on Business Outcomes

But the effect on the analysis process is only the beginning of the story. A data catalog's greatest business impact often appears in revenue generation. Let's continue with Kate. Assume that her product bundles generate a modest 0.5% increase in revenue; also recall that her line of business, prior to the bundling, had generated \$216 million a year. At this rate, Kate's bundles will grow revenue by \$1.08 million (\$216 million x 0.5%) a year.

Using these figures, let's look at each option's lost revenue opportunity. Let's start with Option 1, which had a three-month delay to implement bundling. Using Kate's figures, that represents a lost revenue opportunity of \$270,000. Similarly, Option 2 had a three-week delay. That represents a lost revenue opportunity of \$124,600. Option 3, because it's so rapid, has essentially no lost revenue, allowing Kate's company to realize the maximum revenue opportunity.

Comparing Option 3 (data analysis with a data catalog) to Option 1 (working with IT), the combined cost savings and revenue realization total \$286,610 (\$270,000 + \$17,300 - \$690). The combined cost savings and revenue realization for Option 3 vs. Option 2 (data analysis without a data catalog) total a considerably lower \$127,370 (\$124,600 + \$3,460 - \$690).

# Data Catalog Return on Investment

## Direct Financial Impact

At this point, we've seen an example of the direct financial impacts of a data catalog for one analysis use case — a single analysis need in a single line of business. Assuming this example is a representative and typical use case, we can project annual return on a data catalog investment based on the anticipated number of use cases per year across the entire enterprise.

Your numbers, of course, will likely vary, depending on the size of your enterprise, its annual revenue, and the frequency of analytics use cases. Still, by using the pattern and calculations illustrated here, you can plug in your own data and assumptions, and then calculate a reasonable estimate of direct data catalog value for your organization. Your ROI calculations should also consider the costs of catalog implementation, operation, and administration — including software licensing, staffing, training, and data curation.

NUMBER OF USE CASES PER YEAR	SHIFT FROM IT ANALYSIS PROJECTS TO SELF-SERVICE WITH BOOMI	SHIFT FROM SELF-SERVICE WITHOUT A DATA CATALOG TO SELF-SERVICE WITH BOOMI
100	\$28,661,000	\$12,737,000
200	\$57,322,000	\$25,474,000
500	\$143,335,000	\$63,685,000
1000	\$286,610,000	\$127,370,000

## Additional Savings

The waste and inefficiency from redundant and overlapping analysis in a large corporation can be difficult to quantify, but they're clearly substantial. Business and data analysts frequently reinvent analysis previously performed by others simply because they're unaware that those analyses exist. To remedy this situation, the Boomi platform offers discovery capabilities that bring visibility to existing analysis and dashboards. When analysts can see the analysis work already done by their colleagues, their redundant efforts shrink, saving time and cost. The organization also benefits from the shared knowledge of business dynamics and insights.

## Value from Governance

The Boomi Data Catalog's integrated data governance and security features add value in two important areas: quality improvement and risk avoidance. Cataloging key metrics ensures that they're calculated consistently, improving the quality of analysis and engendering trust in the data and analytics. Similarly, protecting privacy and securing sensitive data helps organizations comply with data-protection regulations. Similarly, active security measures help to reduce the risk of data corruption and loss. While it's difficult to quantify the value of these intangibles—increased trust and confidence, compliance violations avoided, and data breaches prevented—the benefits are undeniable.



## Workplace and Cultural Value

A data catalog can deliver human, cultural, and financial value. That includes reducing workforce friction. In many organizations, the interactions described in Options 1 and 2 can cause friction, as delays in business initiatives can postpone desired business results. In these cases, business urgency can collide with IT overload, raising tensions and damaging working relationships. The problem can be further compounded by time-sensitive information and analysis needs. For example, imagine that Kate's objective is a product-bundle promotion that targets Mother's Day shoppers. Now imagine that she experiences a

delay of weeks or even months. Any benefits of her timed promotion could easily be lost.

Similarly, marketers commonly complain about social-media sentiment that may have positively or negatively driven sales, especially when it can take weeks or even months to find the anomalies. That's far too late to address the issue or respond to the sentiment.

In contrast, analysis that's timely can help strengthen working relationships, build a collaborative culture, and deliver faster, more profitable business outcomes.

## Boosting the Value of Business Data

The business value of Boomi Data Catalog and Preparation is real. Quantifiable benefits such as revenue realization and cost savings help to calculate and deliver an ROI. But that's only part of the story. Less readily quantified benefits of trust, confidence, and avoided risks are just as important. Whether your current analytics processes rely on IT projects, self-service analysis without

cataloging, or some combination of the two, you're sure to enjoy tangible benefits by adopting the Boomi platform. To learn more about the transformative power of Boomi's data cataloging and preparation technology, download our companion eBooks: "The Real Cost of Reinventing the BI Wheel" and "A Compelling Case for Data Cataloging." ■





## About the Author

**Dave Wells** is an advisory consultant, educator, and research analyst dedicated to building meaningful connections throughout the path from data to business impact. As an educator he has written dozens of courses and taught hundreds of classes about data warehousing, data modeling, data architecture, and business intelligence for professional organizations such as Dataversity, eLearningCurve, and TDWI.

Dave works at the intersection of information and business, driving value through analytics, business intelligence, and innovation. More than 40 years of information management experience combined with over 10 years of business management create a unique perspective about the connections among business, information, data, and technology. Knowledge sharing and skills building are Dave's passions, carried out through consulting, speaking, teaching, and writing.



## About Eckerson Group

Our passion is increasing the value our clients receive from managing their data. We demonstrate this passion in many ways—by working directly with our consulting clients to solve their current problems, by sharing our best practices with our education and training clients, and by looking into the future at the upcoming best practices with the support of our research clients.

Eckerson Group is a research and consulting firm focused on serving the needs of business intelligence (BI) and analytic leaders in Fortune 2000 organizations worldwide. Our seasoned team of consultants, educators, and researchers each have more than 20 years of experience in the field. They are uniquely qualified to help business and technical leaders use data and technology to drive better insights and actions.

Eckerson Group's three lines of business work together synergistically to provide complete support for BI and analytic leaders who want to succeed in with business intelligence, data governance, advanced analytics, data warehousing, performance management, and big data.

Get More Value From  
**Your Data**



Our **consulting** team helps BI and analytic leaders assess their current capabilities, develop a BI/analytics strategy, create a roadmap to deliver data-driven applications, and design data architectures, data models, dashboards, governance programs and Centers of Excellence to execute their strategy.

Our **research** provides objective, unbiased investigation of important data analytics topics, evaluation of vendor products in key market segments, and market trends in the adoption of new technologies and approaches by BI and analytic leaders.

Our **education and training** provides continuous learning for enterprises that are data-dependent and data-driven. As the complexities of data and analytics multiply, education provides the skills and competencies to keep your team out ahead.

As a complimentary service, Eckerson Group has pioneered a unique assessment tool to help BI and analytic leaders gauge the capabilities of their team to deliver business value. It assesses their readiness to apply specific technologies or methodologies. Assessments are delivered both online and in person.

Eckerson Group was founded by Wayne Eckerson, a long-time thought leader in the BI and analytics field who has a passion for helping business and technical teams drive positive change in their organizations by strengthening their leadership skills and data management effectiveness.



## About Boomi

Boomi, a Dell Technologies business, quickly and easily unites everything in your digital ecosystem so you can achieve better business outcomes, faster. Boomi's intelligent, flexible, scalable platform accelerates your business results by linking your data, systems, applications, processes, and people.

Harnessing the power of the cloud to unify everything inside and outside of a business, Boomi gives more than 10,000 organizations around the world and across industries the ability to future-proof their data management strategies. For more information, visit [www.boomi.com](http://www.boomi.com).