

DATA BLACK HOLES

ARE DATA SILOS UNDERMINING DIGITAL TRANSFORMATION?

MANAGEMENT SUMMARY



Authors

Jacqueline Bloemen
Timm Grosser

This study was prepared by BARC, an independent market analyst firm. It is available free of charge thanks to the generosity of Ab Initio, Collibra, Dataiku and Denodo.



INTRODUCTION

The existence of data silos is nothing new. Data-producing applications were once isolated systems. They were built to at least partially automate a specific subtask of a business process. The transactional data was stored in isolated data sets and initially served only one purpose, namely, to document the transaction that had taken place.

Over time, enterprises realized that data is worth more. Utilization of operational data for enterprise management helps to gain insights into the current state of the business and supports fact-based decision-making. This has been important for decades. However, the operational data stored in data silos was not suitable for this task. Many companies therefore built a data warehouse to consolidate their operational data silos.

In the age of digitalization, more extensive data and analytics requirements have emerged for which the data warehouse was not sufficiently designed. Data-based insights are being used to automate decisions. The goal is to make business processes faster, more efficient and less vulnerable to risk. Analytics-driven insights are also expected to drive business innovation. Thus, alternative data architecture concepts have emerged, such as the data lake and the data lakehouse. Which data architecture is right for the data-driven enterprise remains a subject of ongoing debate.

This study evaluates current thinking around data silos and addresses several questions:

- What are the implications of data silos for the data-driven enterprise?
- What are the main challenges companies are facing due to data silos?
- Which approaches are being adopted to break down the barriers of these data and knowledge silos?

Participants around the world were polled, with most responses coming from Europe. We examined the current approaches of companies of different sizes from various industries. For deeper insight, we also analyzed the answers according to data management maturity. Participants were asked to rate the skills and competencies in the handling of data in their company compared to their main competitors. This allowed us to gain a better understanding of what “best-in-class” companies are doing to overcome their data silos in comparison to “laggards” (see demographics).

We hope this survey contributes to your company's efforts to overcome your existing data black holes. Please do not hesitate to contact us if you have any questions

Jacqueline Bloemen and Timm Grosser, June 2021

MANAGEMENT SUMMARY



RECOMMENDATION

Decentralized data preparation is both an opportunity and a risk. Avoid an uncontrolled proliferation of data silos. Create transparency about decentralized data preparation processes and where they negatively impact business efficiency and effectiveness. Identify how you can optimize central data provisioning processes to increase the usability of the data provided.



RECOMMENDATION

Establish business responsibility for data and its strategic development. Make sharing knowledge about data and its use in the company mandatory and create corresponding processes. In the digital enterprise, distributed data sets are not only a matter of course, but a necessity. Therefore, logical, cross-functional accountability for data is indispensable.

01

DATA BLACK HOLES: THE HIGH COST OF SUPPOSED FLEXIBILITY

Data warehouse, data lake, data lakehouse: the main paradigm for provisioning data for analytics is a centralized approach. However, business users often resort to individual data preparation because centralized data sets do not meet their needs, they trust their own data sets more, or it enables them to get results faster. Consequently, despite all attempts at centralized data provisioning, most companies are unable to reduce the number of data silos they have. On the contrary, their proliferation continues to increase. As the number of silos continuously grows, finding and understanding the right data and preparing it for a specific purpose becomes even more complex and time-consuming.

02

DATA SILOS PREVENT DIGITAL TRANSFORMATION

Having key personas tied up with elementary data problems instead of working on the digital future of the company is the biggest challenge caused by the state of current data landscapes. This is a vicious circle. Their capacity is tied up in maintaining existing data problems instead of solving them. In turn, they do not have the time to work on data-driven evolutions, both for themselves and to support others. This limits business agility and velocity as well as the innovation of the data-driven enterprise.

MANAGEMENT SUMMARY



03 ARCHITECTURE AND TECHNOLOGY HELP BALANCE CENTRALIZED AND DECENTRALIZED DATA REQUIREMENTS

Predominantly centralized data architecture is not the silver bullet of the digital future. Physical data silos will neither be avoided nor completely dissolved. Therefore, an exclusively centralized approach to data integration is doomed to fail. Instead, distributed data must be linkable as needed. Architecture and technology play a significant role here. They help overcome the limitations of physically distributed data sets more efficiently by supporting intuitive and on-demand exploration, preparation and (advanced) analysis of data.

04 ORGANIZATIONAL SILOS WEIGH HEAVIER THAN DATA SILOS - OVERCOMING THEM IS A CULTURAL JOURNEY

Unified business terminology, not centralized data management, is the main success factor for the digital enterprise. Sharing data, analyses and data-based insights across business units only delivers value if they are understandable to others. However, this requires a uniform (data) language. But because business units have been managed as silos for many years, these language barriers are primarily organizational. Data silos are often just a consequence of this. Accordingly, a lack of communication is the most encountered cultural challenge when it comes to improving the data landscape.



RECOMMENDATION

Plan a logical architecture that ensures combinability of physically distributed data. Populate this architecture with smart technologies such as data cataloging, data virtualization and AI-augmented data preparation that help overcome the boundaries of physical data sets.



RECOMMENDATION

When it comes to ending the data silo dilemma, change requires bold actions. Shaping data strategy and goals and transforming the behavior of people in the company will not succeed without executive vision and leadership. Secure the support of your company management for your data strategy and convince them that they must actively lead the digital transformation. Employees need guidance based on clear strategic goals and what becoming a data-driven business means in practice.

BARC – MAKING DIGITAL LEADERS

BARC – BUSINESS APPLICATION RESEARCH CENTER

BARC (Business Application Research Center) is one of Europe's leading analyst firms for business software, focusing on the areas of data, business intelligence (BI) and analytics, enterprise content management (ECM), customer relationship management (CRM) and enterprise resource planning (ERP).

Our passion is to help organizations become digital companies of tomorrow. We do this by using technology to rethink the world, trusting data-based decisions and optimizing and digitalizing processes. It's about finding the right tools and using them in a way that gives your company the best possible advantage.

This unique blend of knowledge, exchange of information and independence distinguishes our services in the areas of research, events and consulting.

Research

Our BARC studies are based on internal market research, software tests and analyst comments, giving you the security to make the right decisions.

Our independent research brings market developments into clear focus, puts software and vendors through their paces and gives users a place to express their opinions.

Events

Decision-makers and IT industry leaders come together at BARC events. BARC seminars in small groups, online webinars and conferences with more than 1,000 participants annually all offer inspiration and interactivity. Through exchange with peers and an overview of current trends and market developments, you will receive new impetus to drive your business forward.

Consulting

In confidential expert workshops, coaching and in-house consultations, we transform the needs of your company into future-proof decisions. We provide you with successful, holistic concepts that enable you to use the right information correctly. Our project support covers all stages of the successful use of software.



Germany
BARC GmbH
Berliner Platz 7
D-97080 Würzburg
+49 931 880 6510
www.barc.de

Austria
BARC GmbH
Hirschstettner Straße 19 / 1 / IS314
A-1200 Wien
+43 660 6366870

Switzerland
BARC Schweiz GmbH
Täferstraße 22a
CH-5405 Baden-Dättwil
+41 56 470 94 34

Rest of the World
+44 1536 772 451
www.barc-research.com

SPONSOR PROFILE: AB INITIO

ABOUT AB INITIO

For over 25 years, Ab Initio has focused on processing and managing corporate data. Many of the largest companies in the world work with Ab Initio to solve the challenges of turning big data into meaningful data. Ab Initio offers a complete solution for reading, writing, processing, and querying data no matter whether it is located in the latest cloud database or in legacy mainframe data stores or anywhere else in the corporate ecosystem.

You can run your operations on the cloud, on premises, or in any combination. Whether you want to run across mainframes, Unix/Linux boxes, or containers with Kubernetes, Ab Initio does it all. You develop your code once and deploy it where you need it.

Ab Initio's innovative visual development paradigm and fully integrated product suite eliminate the most pernicious obstacles to building high-quality software. Ab Initio's visual approach makes it easy to see the flow of logic and data through the system. We eliminate the process of trans-

lating logic diagrams into programs; the diagrams are the programs. Ab Initio's powerful semantic discovery capabilities accelerate data enrichment and automate adding business meaning to physical data sets. Ab Initio data catalog capabilities enable data virtualization, so you can easily find the data you need when you need it.

Ab Initio software plays well with other software packages, including statistical analysis and machine learning modules. Ab Initio's extensive debugging and automated testing capabilities, combined with Ab Initio's unparalleled ability to rapidly make major changes to the software, helps users get systems into production quickly and enables continuous integration and continuous deployment pipelines. It is no surprise that working with Ab Initio reduces software development time by as much as 75% compared with traditional software development approaches.



Ab Initio Software Germany GmbH
Landsberger Straße 302
80687 München

Tel: +49 89 90405-800
Mail: Solutions@abinitio.com

SPONSOR PROFILE: COLLIBRA

ABOUT COLLIBRA

Collibra is the Data Intelligence company. We accelerate trusted business outcomes by connecting the right data, insights and algorithms to all data citizens. Our cloud-based platform connects IT and the business to help companies trust their data, democratize access and empower the innovation to transform their business. Global organizations choose Collibra to unlock the value of their data and turn it into a strategic, competitive asset.

Collibra acts as the system of record for data, delivering end-to-end Data Intelligence to accelerate digital business transformation and helping companies to:

- Discover, understand and access trusted data when they need it, so they can generate impactful insights that drive business outcomes
- Establish a shared business language for data assets and maintain it over time so that teams can trust data and drive digital use their data to improve their business transformation
- Map relationships between data to show how

it moves from system to system and how data sets are built, sourced and used, providing complete, end-to-end lineage visualization

- Operationalize and manage data privacy policies across the privacy lifecycle, and scale compliance for multiple regulations from a single system

Collibra has a diverse global footprint, with offices in the U.S., Belgium, Australia, Czech Republic, France, Poland and the U.K. For more information, visit collibra.com.



Collibra

Collibra

1 Broadway, 31st Floor
New York, NY 10006 – USA
info@collibra.com
www.collibra.com/BARC

Mr. Paul Dietrich / Area Vice President, DACH;
paul.dietrich@collibra.com

SPONSOR PROFILE: DATAIKU

ABOUT DATAIKU

Dataiku is the world's leading AI and machine learning platform, supporting agility in organizations' data efforts via collaborative, elastic, and responsible AI, all at enterprise scale. At its core, Dataiku believes that in order to stay relevant in today's changing world, companies need to harness Enterprise AI as a widespread organizational asset instead of siloing it into a specific team or role.

To make this vision of Enterprise AI a reality, Dataiku is the only platform on the market that provides one simple UI for the entire data pipeline, from data preparation and exploration to machine learning model building, deployment, and monitoring, and everything in between.

Dataiku was built from the ground up to support usability in every step of the data pipeline and across all profiles, from data scientist to cloud architect to analyst. Point-and-click features allow those on the business side and other non-coders to explore data and apply AutoML in a visual interface. At the same time, robust coding features (including interactive Python, R, and SQL note-

books, the ability to create reusable components and environments, and much more) make data scientists and other coders first-class citizens as well.

The commitment to openness and flexibility in Dataiku doesn't stop there. Because each company's path to Enterprise AI looks different, Dataiku supports the creation of a spectrum of applications, whether that means building out a self-serve analytics platform or fully operationalized AI integrated with business processes.

Dataiku was founded in 2013 and has grown exponentially since. In December 2019, Dataiku announced that CapitalG — the late-stage growth venture capital fund financed by Alphabet Inc. — joined Dataiku as an investor and that it had achieved unicorn status, valued at \$1.4 billion. In August 2020, the company announced a \$100 million Series D round led by Stripes with Tiger Global Management joining the round as investors. Dataiku currently employs more than 600 people worldwide between offices in New York, Paris, London, Munich, Sydney, and Singapore.



Dataiku
We Work
Neue Rothofstraße 13-19
60313 Frankfurt am Main
Germany

<https://www.dataiku.com/home/contact-us/>

SPONSOR PROFILE: DENODO

ABOUT DENODO

Denodo is the leader in data virtualization providing agile, high performance data integration, data abstraction, and real-time data services across the broadest range of enterprise, cloud, big data and unstructured data sources at half the cost of traditional approaches. Denodo's customers across every major industry have gained significant business agility and ROI by enabling faster and easier access to unified business information for agile BI, big data analytics, Web and cloud integration, single-view applications, and enterprise data services.

The Denodo Platform offers the broadest access to structured and unstructured data residing in enterprise, big data, and cloud sources, in both batch and real-time, exceeding the performance needs of data-intensive organizations for both analytical and operational use cases, delivered in a much shorter time frame than traditional data integration tools.

The Denodo Platform drives agility, faster time to market, and increased customer engagement by delivering a single view of the customer and operational efficiency from real-time business intelligence and self-serviceability.



Denodo
Karlstraße 10
Munich, 80333
Germany

E-Mail: info.emea@denodo.com
Website: www.denodo.com

Phone:
Germany +49 89 599 904 50
UK +44 20 3196 4710
Spain +34 912 77 58 55
France +33 1 42 68 51 27
Italy +39 2 72546395

AUTHORS



JACQUELINE BLOEMEN

SENIOR ANALYST

Jacqueline Bloemen is a senior analyst with a major focus on data & analytics strategy and culture, architecture and technology, governance and organization. She is an author and speaker and has been advising companies of various sizes and industries for over 35 years. Currently, her research and consulting activities focus on the transformation to becoming a data-driven company.



TIMM GROSSER

SENIOR ANALYST

As a senior analyst, Timm Grosser has been advising domestic and international companies of various sizes and industries in the areas of BI, data management and analytics for more than 10 years. During his time as a consultant, he has designed numerous solutions in BI/big data strategy, organization, architecture and tool selection with customers or in the BARC test lab. He is a frequent speaker at conferences and seminars as well as the author of numerous industry articles and market studies.

Germany

BARC GmbH
Berliner Platz 7

D-97080 Würzburg
+49 931 880 6510

Austria

BARC GmbH
Hirschstettner Straße 19
/ I / IS314

A-1220 Wien
+43 660 6366870

Switzerland

BARC Schweiz GmbH
Täferstraße 22a

CH-5405 Baden-Dättwil
+41 56 470 94 34

Rest of the World

www.barc-research.com
+44 1536 772 451

