

# Trends in Data Architecture:

## A DATAVERSITY® 2017 Report

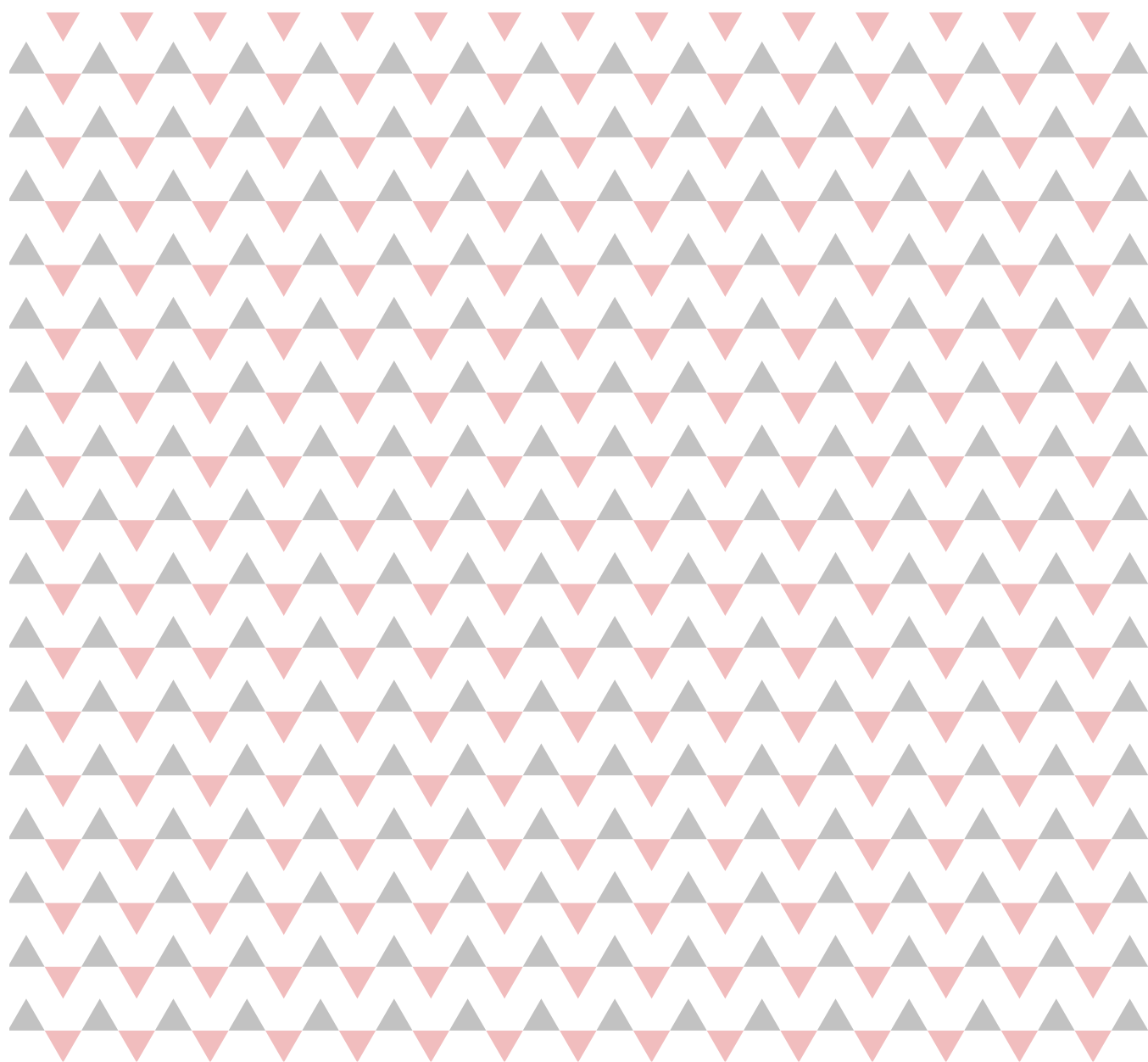
---

by Donna Burbank and Charles Roe



I D E R A

 DATAVERSITY®



# Contents

Executive Summary

Research and Demographics

What is Data Architecture?

Recent Changes in Data Architecture

Emerging Trends in Data Architecture

Business Drivers & Organizational Roles

- a. “What are your main goals & drivers for implementing a Data Architecture? [Select all that apply]
- b. “What role(s) are typically responsible for creating a Data Architecture? [Select all that apply]”
- b. “How are Data Architecture decisions made in your organization?”
- d. “How do you get trained in Data Architecture? [Select all that apply]”

Platforms & Environments

- a. “Which of the following data sources or platforms are you currently using? [Select all that apply]
- b. “Which of the following do you plan to use in the future that you are not using currently? [Select all that Apply]”

Cloud Computing

- a. “Are you currently leveraging the Cloud in your Data Architecture?”
- b. “If you answered Yes to Question 14, what were your reasons for moving to the Cloud? [Select all that apply]”
- c. “Whether you answered Yes or No to Question 14, what are your concerns regarding moving data to the Cloud? [Select all that apply]”

Big Data Ecosystems

- a. “Are you currently using a Big Data platform?”
- b. “For what use cases are you using Big Data? [Select all that apply]”
- c. What are your concerns, if any, about using Big Data? [Select all that apply]”

## **Open Source & Open Data**

- a. “Are you currently using any Open Source technologies in your Data Architecture?”
- b. “Are you currently using Open Data Sets?”

## **Blockchain Technologies**

- a. “Are you currently using or considering Blockchain technology? Which of the following best represents your organization’s view of Blockchain?”
- b. “For which use cases are you using or considering use of Blockchain technology? [Select all that apply]”

## **Master Data Management (MDM)**

## **Data Virtualization**

## **Data Design & Modeling**

- a. “Which methods of Data Modeling do you use in your organization? [Select all that apply]”
- b. “What types of models and diagrams do you use in your Enterprise Architecture? [Select all that apply]”c. “Do you currently have a Metadata Management effort in place?”
- d. “What are your current main use cases for Metadata? [Select all that apply]”

## **Conclusion**



# Executive Summary

In today's data-driven economy, the definition and scope of Data Architecture is changing and evolving at a rapid pace. Data Architecture is as much a business decision as it is a technical one, as new business models and entirely new ways of working are driven by data and information. Blockchain, for example, has the opportunity to revolutionize the way the industry works with financial transactions. Big Data platforms now allow organizations to process and leverage information on a scale not possible with earlier technologies, allowing for new discoveries and the ability to process information beyond the scale of typical relational systems. Internet of Things (IoT) technologies provide the opportunity to harness information that was once considered simply “noise” or “exhaust,” such as sensor data, log files, and more.

With such precipitous change and the vast number of technologies available, it can be a daunting task for the average organization to build a comprehensive Data Architecture. This report looks to demystify a number of the current trends, provide practical insights into today's modern Data Architectures, and provide insights into what might be the best possible choices for your organization.

This paper is an analysis of a DATAVERSITY® 2017 Survey on the latest trends in Data Architecture. Some of the primary findings of the survey include:

- Data is increasingly seen as a business asset, as well as a technological one.
- More business stakeholders are involved with data, driving the need for self-service options, as well as collaboration and governance across roles.
- Many foundational Data Architecture components still remain popular:
  - » Over 95% of respondents are actively using Data Modeling.
  - » Over 78% of respondents are either already involved in Metadata Management or plan to be in the future.
  - » Data Governance is a key driver for many data initiatives (approximately 60%).
  - » Reporting & Business Intelligence is a key business goal for organizations (approximately 70%).
  - » Relational Databases are by far the most common platform in use (over 85%).
- At the same time, new technologies and approaches are taking hold:
  - » Cloud migration is a key driver with over 75% of respondents currently

implementing a Cloud strategy or planning to in the future.

- » Over 73% of respondents are currently using Open Source technologies, or are planning to in the future.
- » Over 70% of organizations are currently implementing Big Data technologies or planning to in the future, with Data Science & Discovery and Reporting & Analytics listed as the main use cases.
- Many of the technologies that have been hot discussion topics in the industry have yet to take a strong foothold in actual implementation:
  - » Blockchain has a current adoption rate of approximately 8%.
  - » Columnar NoSQL databases have a current adoption rate of only 12%.
  - » Only 12.5% of respondents are currently using Data Virtualization techniques.
  - » Internet of Things (IoT) data is in use by approximately 17% of organizations.
- We will analyze these results in the remainder of the paper and investigate other key findings that affect today's data-driven organization.



# Research and Demographics

**Research Scope:** The survey asked a number of questions about the many different architectures that enterprises are using or plan on using in the future. It began with three open-ended questions concerning the respondents' definition of Data Architecture, changes over the next three to five years, and what they saw as the primary emerging trends within the industry. It then broke the scope of Data Architecture down into multiple specific topic areas in order to gain a more well-defined and comprehensive view of the contemporary Data Architecture landscape.

The survey included 30 questions:

- General Demographics (four questions)
- Data Architecture Overview (five questions, three open-ended)
- Goals & Business Drivers (two questions)
- Platforms & Environments (two questions)
- Cloud Computing (three questions)
- Big Data (four questions)
- Open Source & Open Data (two questions)
- Blockchain (two questions)
- Master Data Management (one question)
- Data Virtualization (one question)
- Data Design & Modeling (four questions)

Most of the questions also contained an extra area for written comments, and those comments will be discussed when relevant throughout this paper.

**Principal Demographics:** The three main demographics questions (outside of contact information) asked about the respondents' job function, industry, and company size.

The largest percentage in terms of job function [Figure 1] identified themselves as working with Data and/or Information Architecture at almost 44%. A broader list of roles included:

- Data and/or Information Architecture: 43.9%
- Information/Data Governance: 18.9%
- Business Intelligence and Analytics: 9.8%

- IT Management: 6.1%
- Data Science/Data Scientist: 4.5%
- Database Administration: 4.2%

Some other noteworthy answers within the open-ended comments included Process Engineering, Scrum Master, Taxonomist, and Cognitive Insurance Leadership.

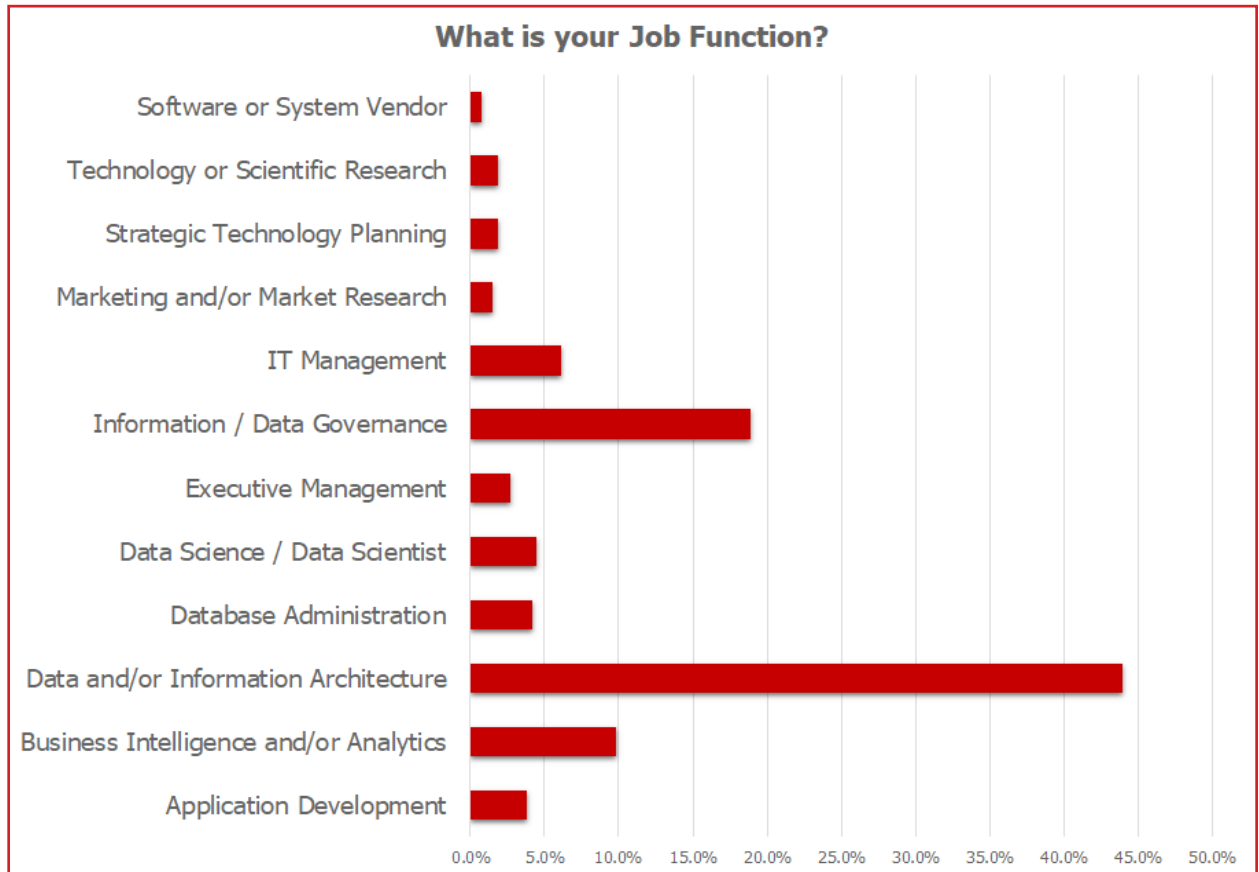


Figure 1

The survey respondents represented over forty different industries, with the largest percentages including [Figure 2]:

- Technology: 12.2%
- Finance: 10%
- Consulting: 9.3%
- Government: 7.5%
- Insurance: 7.5%

- Healthcare: 6.8%

Other notable industries included Banking, Education, Manufacturing, Communications, along with write-in industries such as Oil & Gas, Life Sciences, Legal, Human Resources, Travel, Pharmaceutical, and Lottery.

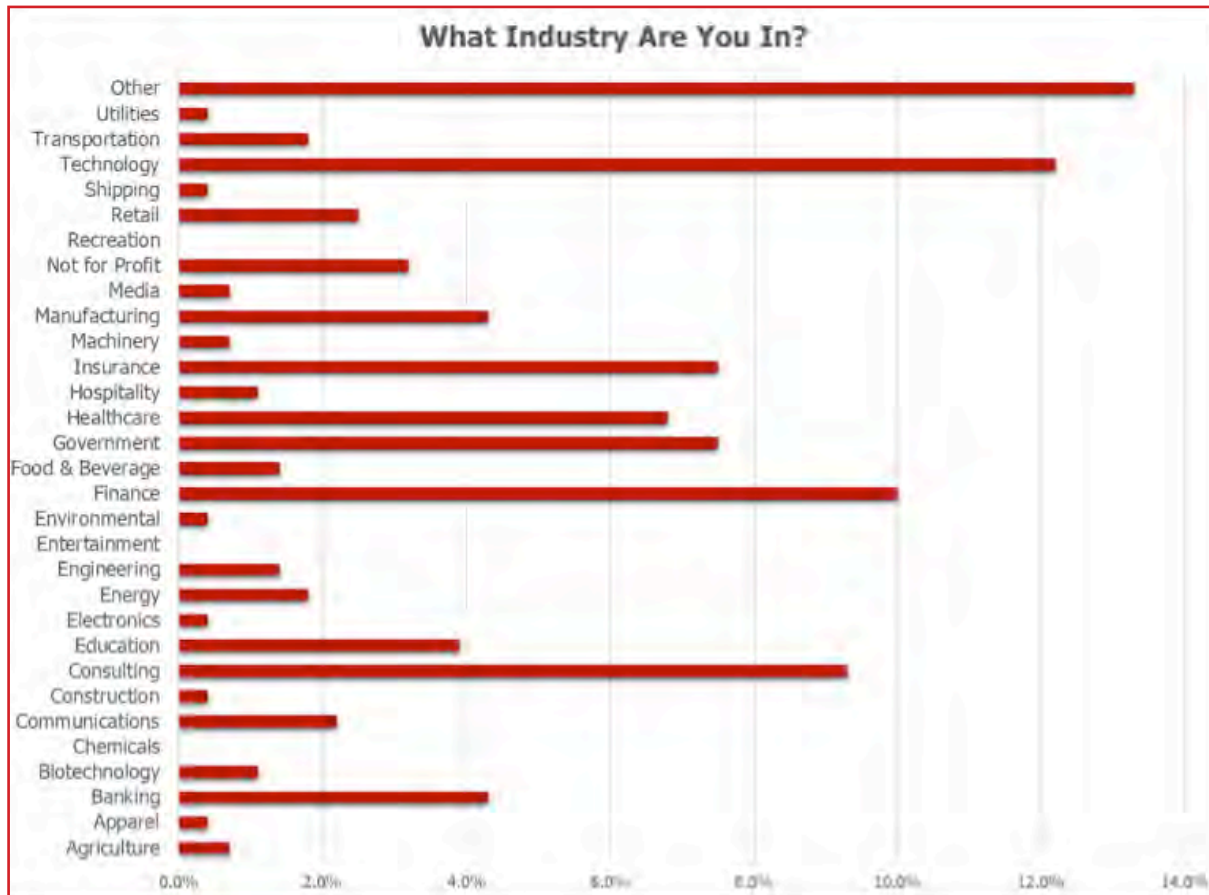


Figure 2

The final demographics question asked about company size [Figure 3], and the results were well-balanced between the various ranges:

- 10,000 – 50,000: 21.5%
- 101 – 1,001: 16.8%
- 1,001 - 5,000: 16.8%
- Less than 10: 11.1%

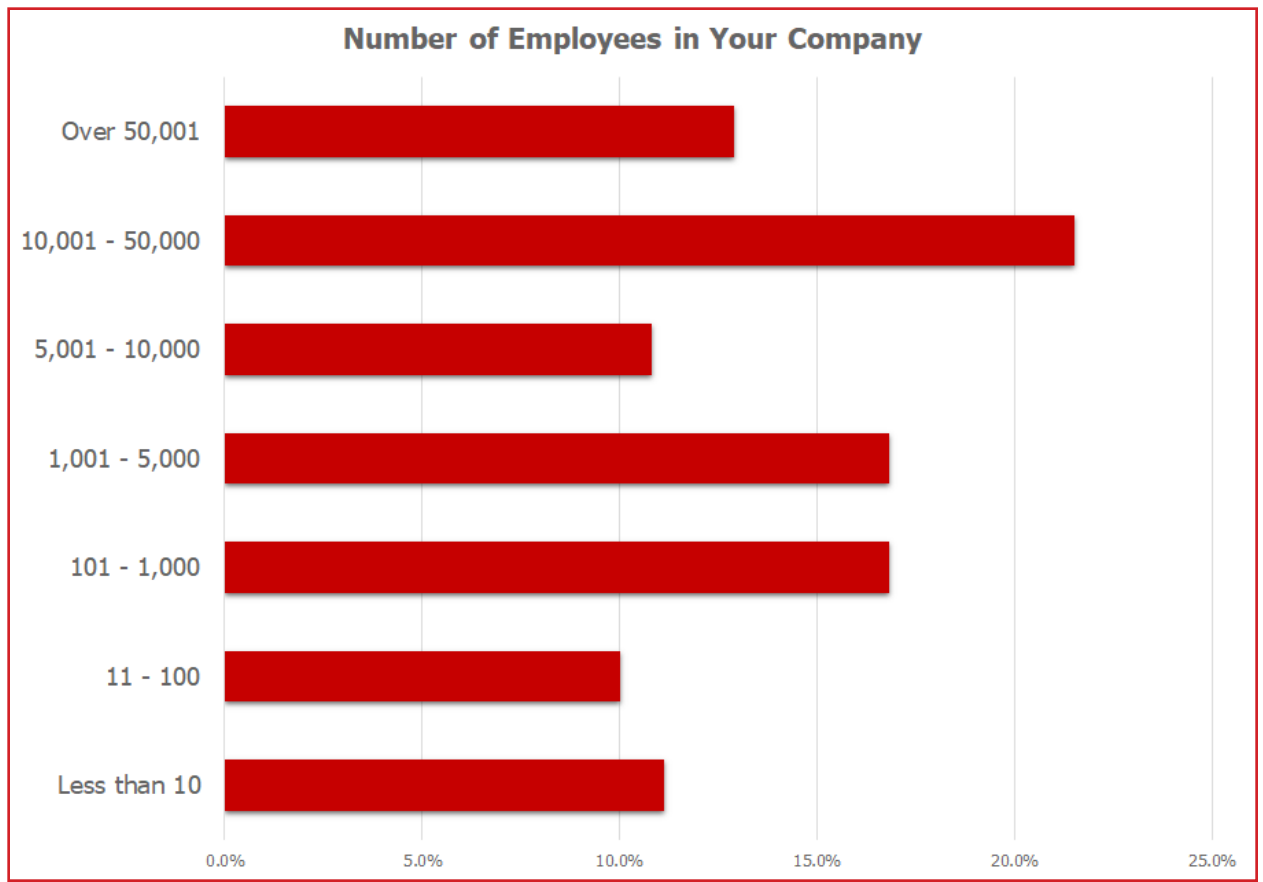


Figure 3





# What is Data Architecture?

To begin a deeper analysis, it is first necessary to come to a consensus regarding what Data Architecture is and what it entails. As a working definition, we are using the DAMA International definition from the Data Management Book of Knowledge:

“Data Architecture is fundamental to data management. Because most organizations have more data than individual people can comprehend, it is necessary to represent organizational data at different levels of abstraction so that it can be understood and management can make decisions about it.”

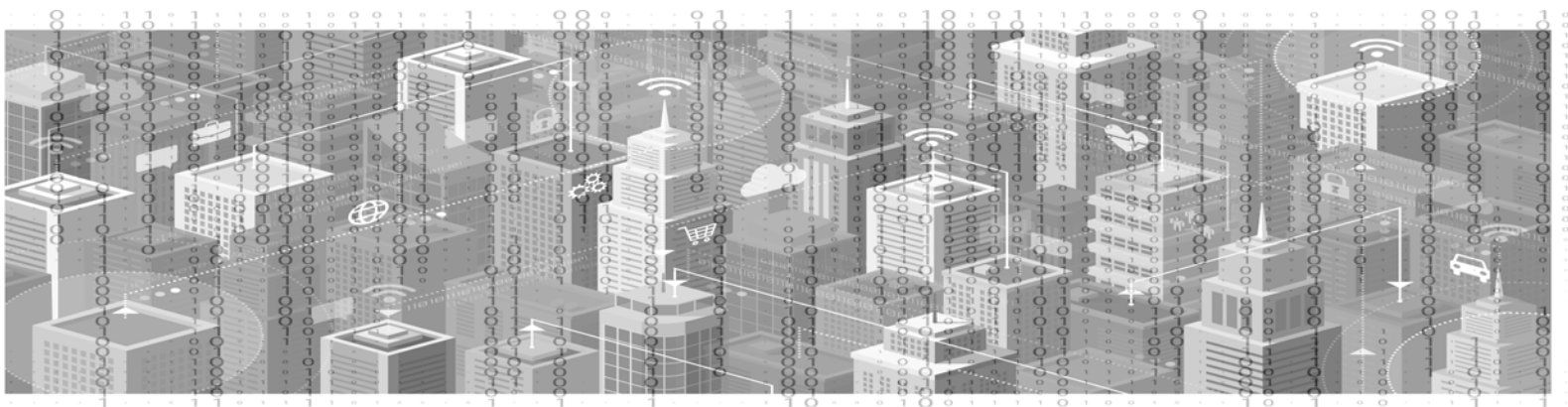
[And] “Data Architecture artifacts includes specifications used to describe existing state, define data requirements, guide data integration, and control data assets as put forth in data strategy. An organization’s Data Architecture is described by an integrated collection of master design documents at different levels of abstraction, including standards that govern how data is collected, stored, arranged, used, and removed. It is also classified by descriptions of all the containers and paths that data takes through an organization’s systems.”

The survey began with an open-ended question to our respondents:

- “What is your definition of Data Architecture?”

We started with this to gain an essential high-level view of the topic from multiple perspectives and to see any commonalties and possible discrepancies. The primary focus of most definitions was in-line with the DAMA definition above, including:

- alignment of IT and business systems.
- data models and associated structures.
- components that comprise such functions as security, capture, storage, integration, arrangement, use and overall management of an enterprise’s data assets.



One of the key elements throughout most of the definitions, though, is the necessity for Data Architecture to support the business, as business needs are the principal focus from which Data Architecture provides its value. A few of the definitions given were:

Data Architecture is...

***...the practice of examining the enterprise strategy, and identifying the key Data Integration points that need to be enabled to execute that strategy, and laying out a roadmap for creating the core capabilities to deliver those integrations, so that companies can leverage data as a strategic asset.***

- “...the practice of examining the enterprise strategy, and identifying the key Data Integration points that need to be enabled to execute that strategy, and laying out a roadmap for creating the core capabilities to deliver those integrations, so that companies can leverage data as a strategic asset.”
- “...the science of assessing where your company data is, and the art of designing the best future way of storing data across the enterprise.”
- “...the discipline to help businesses manage their data in the most effective, secure, compliant, and profitable way.”
- “...everything a business does to ensure information is accurate and available to facilitate valid business purposes, as defined by regulatory/legal requirements, stakeholder value propositions, and customers (internal/external).”
- “...an integrated set of specification artifacts that define strategic data requirements, guide integration of data assets, and align data investments with business strategy. Enterprise Data Architecture is the subset of Enterprise Architecture, including those artifacts describing data and information.”
- “...the blueprint that looks at the entire data landscape and governs the data lifecycle. It will describe the language to be used by stakeholders. This includes the required people, processes, tools, technologies, and artifacts.”
- “...rules, policies, & standards governing (in no particular order): 1. Data requirements documentation for application systems; 2. Database design & modeling; 3. Metadata documentation, management, & dissemination/access; 4. Data collection, use, & integration across the enterprise; 5. Data Quality & Governance definition & enforcement; 6. Data retention compliance & archival; 7. Master Data Management.”

Those definitions, while varying in scope and detail, provide a clear snapshot of the importance of Data Architecture for business success. A well-designed architecture is of prime importance to make sure that data remains trustworthy and manageable in the new world of digital and business transformation.



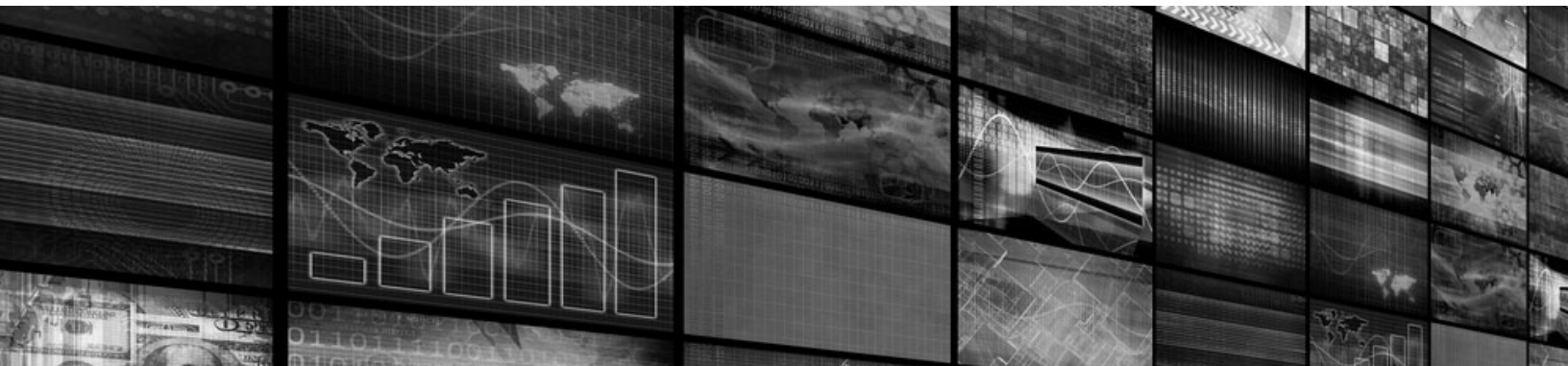
# Recent Changes in Data Architecture

We asked our respondents to discuss what they felt were the most recent developments changing the landscape of Data Architecture in an open-ended question:

- “What has changed most in Data Architecture in the past 3-5 years?”

While the answers varied considerably in focus, common themes included:

- Data seen as a strategic asset.
- Understanding the importance of data and focusing on business drivers around data.
- More attention and involvement from business users.
- Support of the fundamentals that drive better Data Quality, Data Governance, Metadata Management, Master Data, and other more traditional practices.
- A global perspective and how data needs are different across regions. There were a number of comments concerning the upcoming GDPR requirements.
- Volumes of data increasing and the associated technologies to support those volumes such as Big Data systems, Cloud scalability models, Microservices, and Advanced Analytics.
- New types of architectures and data platforms to support a variety of use cases (beyond relational) like ever-expanding NoSQL systems, IoT, Big Data, and Data Lakes.
- The growth of Agile practices to aid in the increasing speed of modern business.
- The volatility and disruption of so many longstanding practices due to the changes currently occurring.



# Emerging Trends in Data Architecture

With the rapid changes occurring in the industry, we wanted to get a sense of what survey respondents felt would be the next “big thing” with the question:

- “What do you see as the next emerging trend/s in Data Architecture?”

The answers were quite balanced between focusing on some of the newest technologies and expected growth areas such as Artificial Intelligence, Machine Learning, Data Virtualization, and Cloud Computing. In addition to these new technologies, respondents also expected a continued emphasis to be placed on core technologies that help companies make sense of and manage their data, such as Metadata, Models, Glossaries, and Data Dictionaries.

Many mentioned Data Governance as a key driver moving forward, as regulations and compliance requirements continue to play ever bigger roles. The most comprehensive agreement centered on the need to better align with business drivers, either formally through business models, ontologies, and others, or more informally by working with stakeholders to center all of an organization’s architectures on the organization’s goals. Such alignment also included continued growth and development of self-service tools to aid business users in gaining greater access to data. Several themes emerged across the collective responses.



## Data as a Strategic Business Asset

- “I believe there will be a continued growth in Information Architecture as businesses determine exactly how they turn information into business knowledge.”
- “More business ownership and involvement.”

## Metadata Management

- “Tracking Metadata and lineage across heterogeneous architectures.”
- “Disciplines of Data Modeling and true Metadata Management. These activities must be automated as much as possible but should also be closely monitored -- the bottom line is that Data Quality is #1. Big Data, Data Science, Analytics, Machine Learning etc. are all useless when the quality of the data is poor.”
- “More automation of Metadata Management.”

## Self-Service

- “The volume of data the business is requiring will require more tools to enable them to do more Self-service Analytics.”
- “Less IT control and more business self-service.”
- “Self-service data procurement.”

## Artificial Intelligence

- “The increasing speed with which AI can be used not only to analyze data predicatively, but actually automatically determine how it should be defined, organized, and managed to support the types of analysis desired.”
- “Artificial Intelligence will begin to play out, creating the need for smart technologies.”
- “More automated and intelligent analysis across this growing volume of data, to continue to improve business outcomes -- denser and greater storage capacities, faster data transfer rates, AI-driven Machine Learning and cognitive algorithms that can even replace the need for Data Scientists by more quickly analyzing multiple parameters to identify potential correlations and the parameters that drive them.”

“

*...there will be a continued growth in Information Architecture as businesses determine exactly how they turn information into business knowledge.*

”

## Data Governance

- “I see a big need for the implementation of Data Governance principles to support an organization’s ability to access, catalog, and potentially transform information into business knowledge.”
- “The understanding that data is an asset to the organization and as much emphasis in quality, consistency, accuracy, and usability that goes into the core products of the company need to go into the trust of the data that is used to support it.”

## Data Modeling

- “A refocus on the Enterprise Data Model and unlocking its true potential.”
- “Data models that support data visualization. Data Architects with the ability to explain the ‘story’ the data tells.”

## Agility

- “More and more agility, with periodic updates and publishing of artifacts on an as-needed basis, or at least every x number of Sprints. More diagnostic modeling, reverse engineering what is there after a series of sprints to document the ‘as is’ and to find areas for improvements.”

## Convergence of Technologies & Roles

- “More of a holistic perspective including data movement, Cloud provider, and social sites that contain data about the organization.”
- “Converged databases, universal memory, and Blockchain.”
- “Connecting Data Governance with Master Data Management. Building a suite of software to cover all Data Management capabilities, or connecting the best of breed Data Management solutions together into one solution for the customer.”
- “Model curators for not just data artifacts, but all architecture artifacts. Less focus on being a ‘Data Architect’ and being more generalized as a system engineer that architects the data, but other systems as well.”
- “RDBMS products taking on more and more NoSQL features, potentially making it easier to get the best of SQL and NoSQL data stores .”
- “Merging of RDBMS, DW, Graph, Document, XML, JSON into a relation-based structure, framework, architecture.”

## Cloud

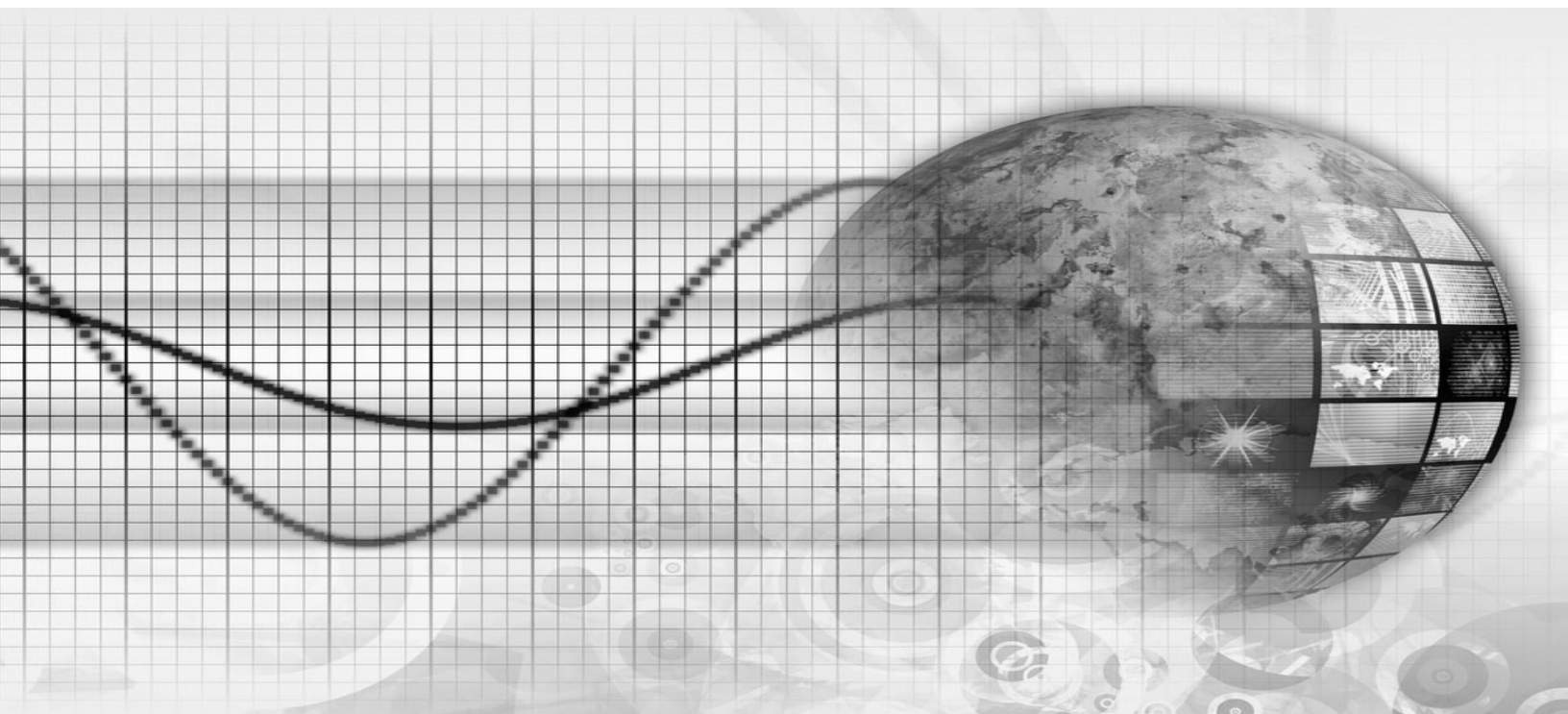
- “Cloud and the Internet of things colliding for the good of business, consumers, etc.”
- “Moving more data/databases to the Cloud.”

## Increasing Volume & Velocity of Data (Big Data)

- “Big Data Architecture taking a much stronger role (and attention) than it is now... if that’s possible!”
- “Variety in availability of data – Data Lakes, Virtual databases, connected data (company to company, government to government, etc.).”
- “With the 4 Vs of data, methods to quickly identify the critical data and methods to sustain.”
- “Big Data, The Internet of Things, Data Lakes, NoSQL.”

## Graph Databases

- “Monetization of data (in legal) Graph technology.”
- “Ascendancy of the graph!”





# Business Drivers & Organizational Roles

To gain a better understanding of where Data Architecture exists today in terms of organizational roles and responsibilities, decision making processes, and goals and business drivers, the survey asked four questions. Each is discussed and analyzed below.

“

*While innovations in technologies and architectural options continue to grow, the core need remains rather simple and direct: to better understand an organization through its data.*

”

## A. “What are your main goals & drivers for implementing a Data Architecture? [Select all that apply]”

As more organizations see data as a strategic asset, and with the drive towards Digital Business Transformation on the rise (38% of respondents), the need to analyze and understand core data assets continue to be a key goal. While innovations in technologies and architectural options continue to grow, the core need remains rather simple and direct: to better understand an organization through its data.

Top responses notably centered around reporting and discovery of information as business users look to increase control of their data assets, with Reporting & BI at 68%, followed by Data Science & Discovery at 52.7%, and the specific need for Self-Service BI at 49.1%.

With the need for reporting comes the related need for Data Governance & Regulatory Compliance, as organizations understand that the data being reported upon must be valid, accurate, and accountable. It is significant that the second most popular response after Reporting & Business Intelligence was its companion activity, Data Governance & Regulatory Compliance. As data is seen as more of a strategic asset, there is an increasing need to manage it accordingly.

While organizations continue to focus on the “tried and true” needs of reporting and related accountability for data, newer technological approaches are also beginning to take hold, including Big Data, Cloud, and AI.

In response to the question: “What are your main goals & drivers for implementing a Data Architecture?”, notable answers included [Figure 4]:

- Reporting & Business Intelligence: 68.0%
- Regulatory Compliance & Data Governance: 59.2%
- Data Science & Discovery: 52.7%
- Self-Service BI: 49.1%
- Digital Business Transformation: 38.5%

Emerging technologies, while certainly not as ingrained as the items above, show signs of adoption:

- Big Data Use Cases: 33.1%
- Cloud Migration: 21.9%
- Artificial Intelligence and/or Machine Learning: 18.9%

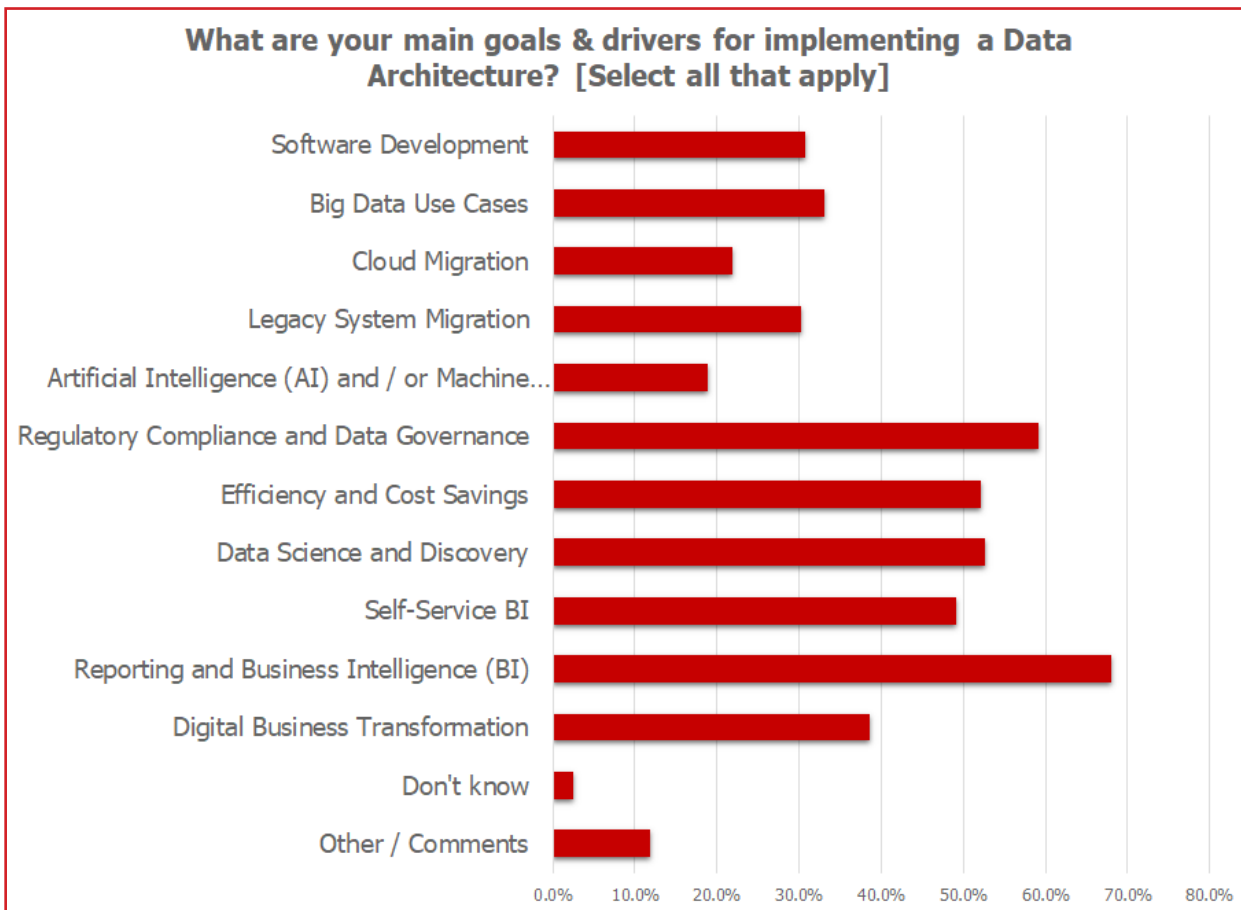


Figure 4



**B. “What role(s) are typically responsible for creating a Data Architecture?  
[Select all that apply]”**

With a greater business focus on data and a wider range of technologies associated with Data Management, it is not surprising that there is a concomitant rise in the diversity of roles responsible for developing a Data Architecture. The role of Data Architect leads the responses (at 90% of total respondents), but the diversity of other selections indicates that these roles are assisting in the creation of a Data Architecture alongside the Data Architect or, in some cases, instead of the Data Architect. Enterprise Architect was the second most popular response at 65.3%, which may indicate the wider context in which data is placed in today’s organization, including process, business capability, and organizational considerations.

The involvement of business-centric roles shows the increasing importance of data to business users, with Business Architect (51.2%), Data Governance Officer (50.0%), and Business Stakeholder (32.9%) showing significant responses.

Collaboration across roles is clearly critical in building a Data Architecture, and one respondent commented that: “Most important are the architects, governance, and DBAs, but Data Quality is also important. The primary ingredients are C-level support and collaboration.”

The push to become data-driven organizations requires collaboration at all levels, and Data Architecture is no different. But a Data Architect should be leading such an effort, because there are core Data Architecture principles that must be understood by a skilled practitioner with proper training in data fundamentals.

This is an encouraging development. In DIY homebuilding, many homeowners can do remodeling and repair projects, but typically an architect or skilled contractor lays the foundation to ensure stability. It’s a similar concept with data. The Data Architect is needed for the foundation, but other roles can and should take an increasingly active role in building out the architecture.

Some of the core results of the question (respondents could select more than one choice) were [Figure 5]:

- Data Architect: 90%
- Enterprise Architect: 65.3%
- Business Architect: 51.2%
- Data Governance Officer: 50.0%
- Database Administrator: 37.6%



- Business Stakeholder: 32.9%
- Data Scientist: 27.6%

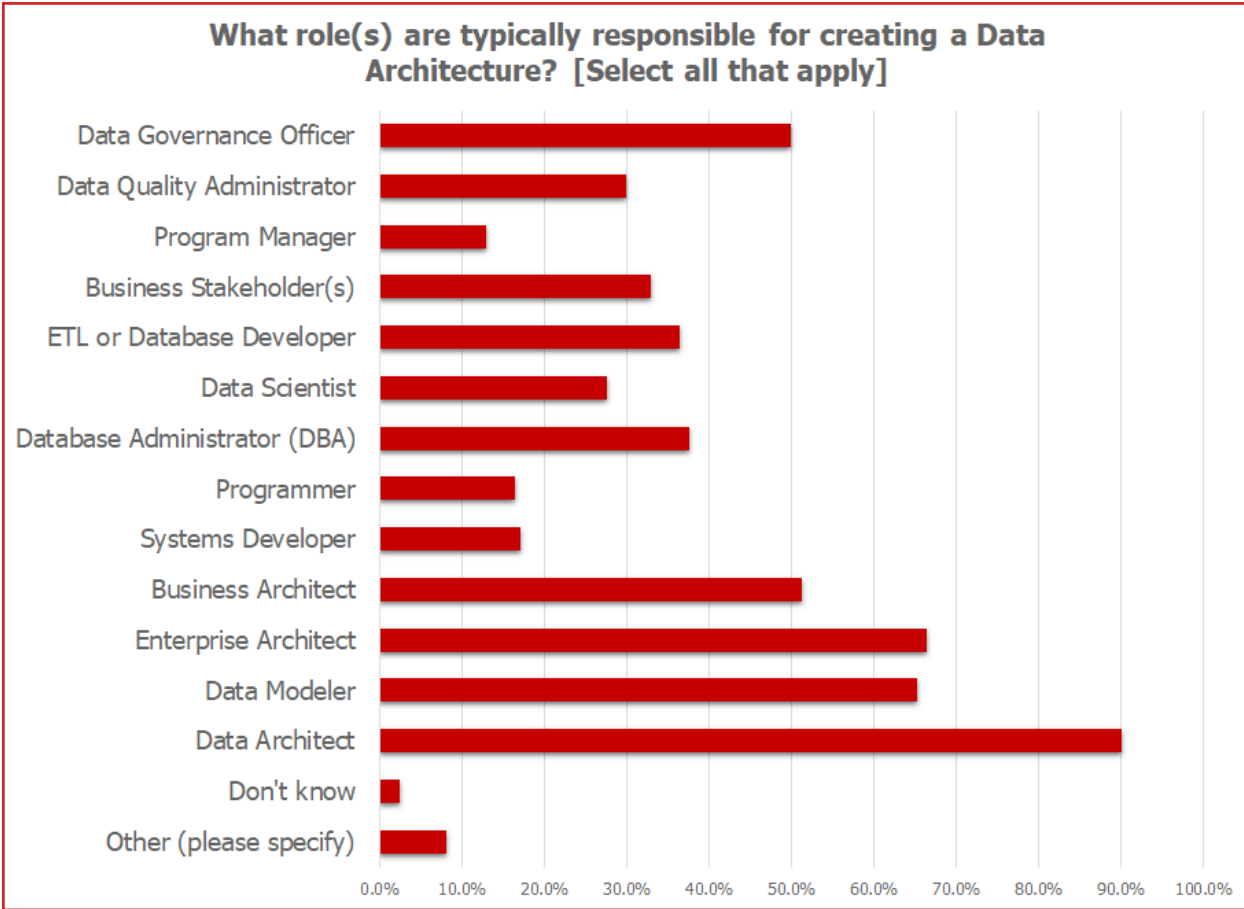


Figure 5



### C. “How are Data Architecture decisions made in your organization?”

As the previous questions indicate, there is a changing dynamic in terms of which roles are involved in driving a Data Architecture strategy. Not surprisingly, then, the results of this question demonstrate the wide range of Data Architecture approaches and use cases occurring within different organizations. There is actually a fair amount of balance, with “project-level design decisions” being the main answer at 34.1%.

Many of the individual comments indicated that organizations are moving towards a centralized approach, but the process is still on-going. There were also a number of others who indicated their organization is moving in the opposite direction.

Several consultants who replied said that such processes – in terms of determining what sort of model makes the most sense for their given enterprise and organizational structure, use cases, corporate culture, available technologies, and C-level priorities – were a significant part of their effort in assisting clients. Truly, Data Architecture implementation and design is not a “one size fits all” process and must be tailored to each individual enterprise and its drivers, needs, and available resources.



The main results of the question were [Figure 6]:

- Project-level design: 34.1%
- Centralized decision-making team: 25.9%
- Ad hoc, on a case-by-case basis: 21.2%
- Other: 12.4%

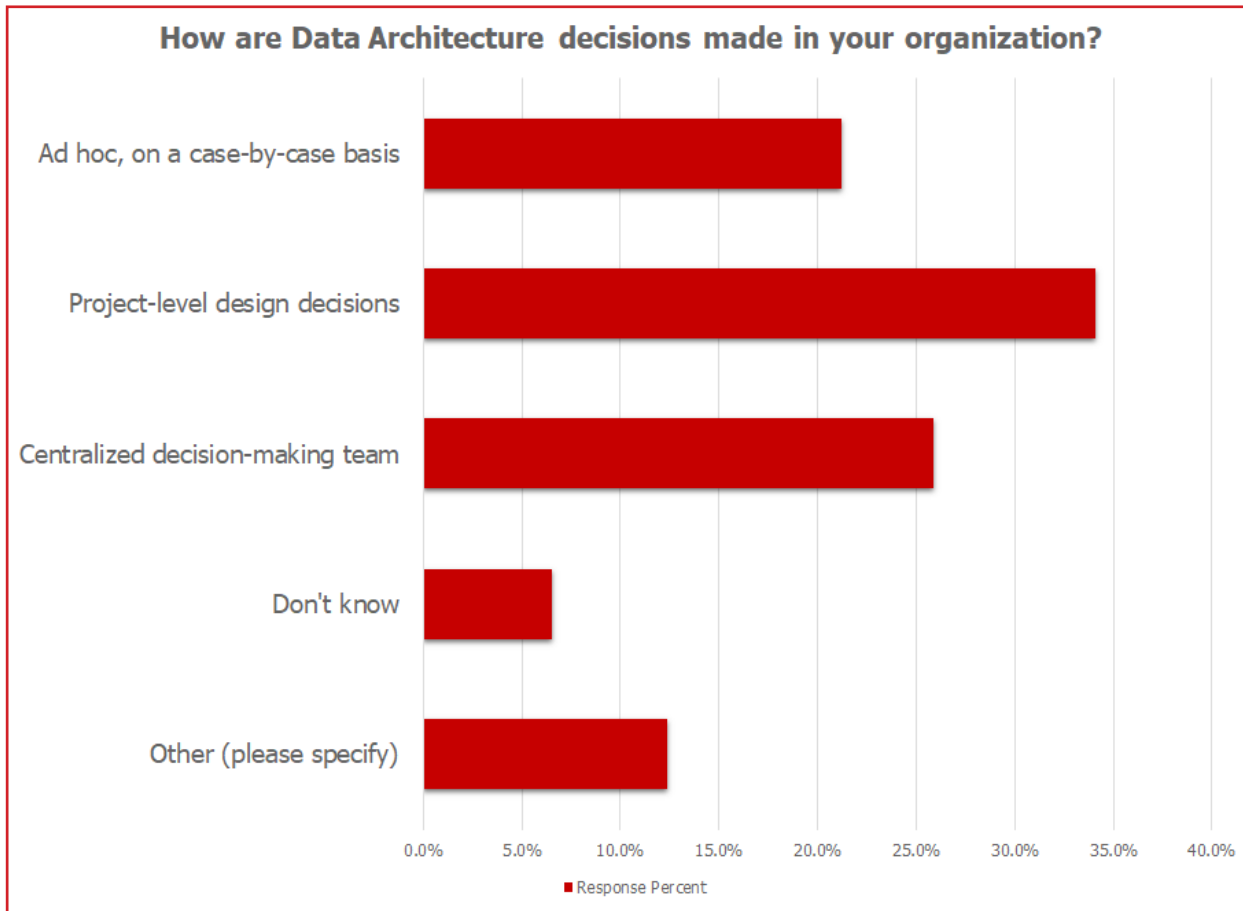
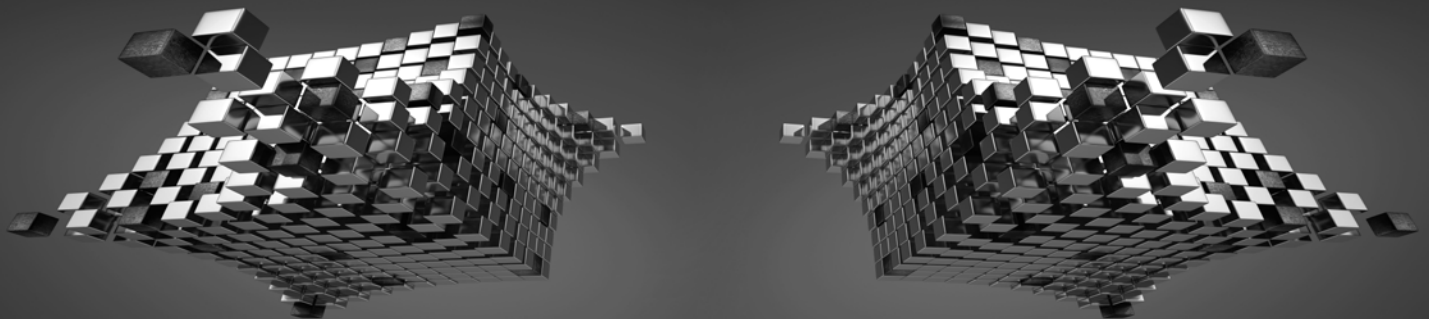


Figure 6

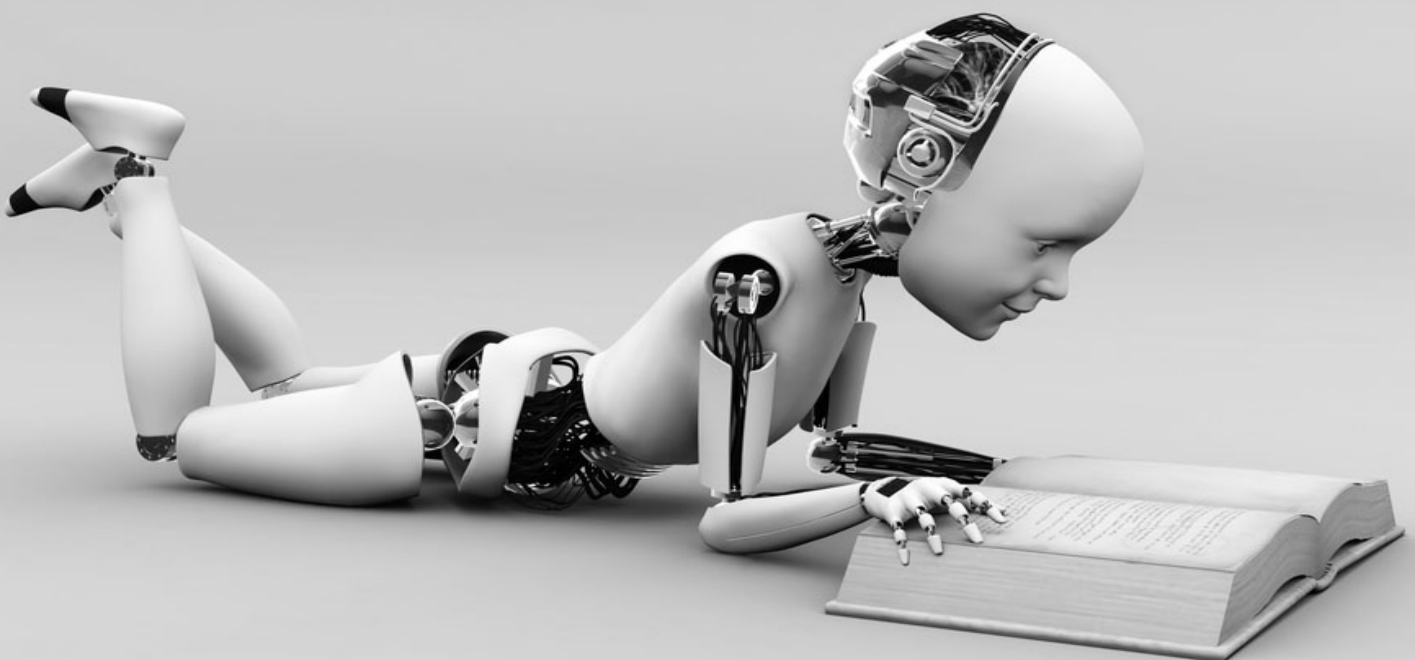


#### D. “How do you get trained in Data Architecture? [Select all that apply]”

With the rapid rise in demand for data-centric skills and the variety of roles involved in Data Management, it is notable that only 14% of respondents indicated that they have formal university training in the field. Most respondents receive information from blogs and Web resources (69.2%), as well as books on Data Architecture (64.5%). Other popular responses were tutorials from vendors (42.6%), training courses outside of university work (41.4%), and mentoring with an experienced professional (40.2%). With the rapidly changing technological advances, this makes a lot of sense. Web and other resources are necessary to keep track of the ever-changing industry.

A significant percentage of respondents (46.7%) claimed that they received no training, and just “started doing.” With the matrixed approach to building a Data Architecture that was described in an earlier section, this is not surprising, as business and other stakeholders rely on experienced Data Architects for the more technical and foundational aspects. But this figure is concerning if these respondents are the key developers of a Data Architecture, as foundational information science principles should be at the core of any Data Architecture.

Another notable response was the high percentage of training from Data Architecture vendors (42.6%). While it can be seen as a positive trend that vendors are offering such training, particularly regarding their unique technical solutions, receiving training solely from a vendor perspective can run the risk of tying one’s architecture too closely to a vendor’s messaging and sales strategy. A balanced mix of education should be the goal—to understand both core architectural principles as well as the unique features and functions of individual tools.



This question allowed respondents to select more than one answer, since it is unlikely practitioners are learning from only one source. The most significant results were [Figure 7]:

- Blogs, screencasts, various sources from the Web: 69.2%
- Books on Data Architecture: 64.5%
- No training, just started doing: 46.7%
- Tutorials and other materials from Data Architecture vendor: 42.6%
- Data Architecture courses outside of university work: 41.4%

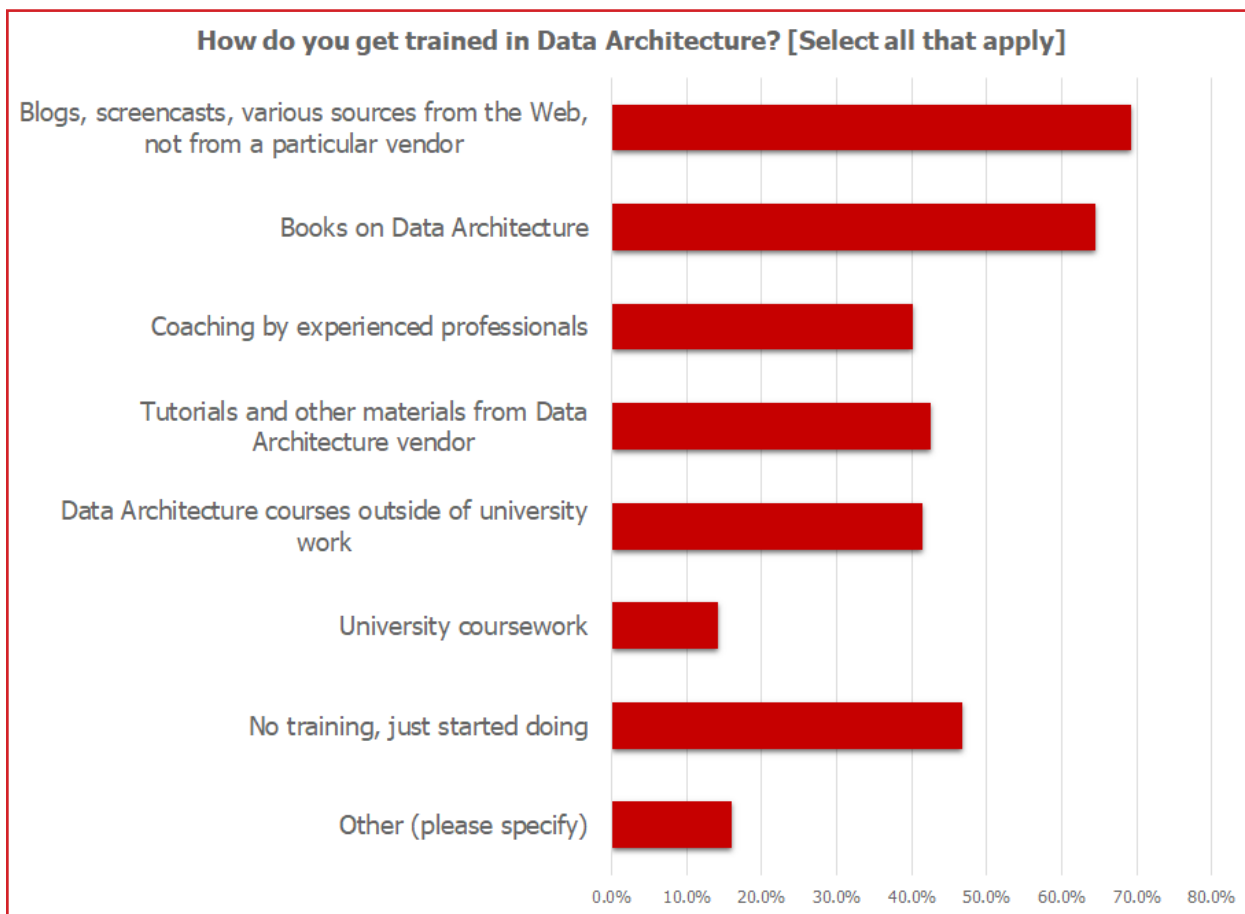


Figure 7



## Platforms & Environments

Organizations of all sizes, especially those that have been around for years, use a multitude of diverse architectural designs to suit their varying needs. It is not uncommon to have a large organization with hundreds of different applications, platforms, legacy systems, new technologies in development and/or production, and a host of systems that are either only nominally used, forgotten, or in the interim space between archiving and retirement.

*Even with the emergence of new technologies, the relational database continues to be the tried and true workhorse of the organization.*

When looking into the range of systems that enterprises are using for their Data Architectures, it was important to include the most possible choices – spreadsheets and legacy systems such as COBOL, Cloud platforms and Big Data ecosystems, relational and non-relational (NoSQL) databases, and many others. The following responses indicate the wide range of technologies in use today.

**A. “Which of the following data sources or platforms are you currently using?  
[Select all that apply]**

The results below clearly demonstrate that relational technologies are by far the largest percentage in use – either on-premises (85.5%) or in the Cloud (51.2%). Even with the emergence of new technologies, the relational database continues to be the tried and true workhorse of the organization.



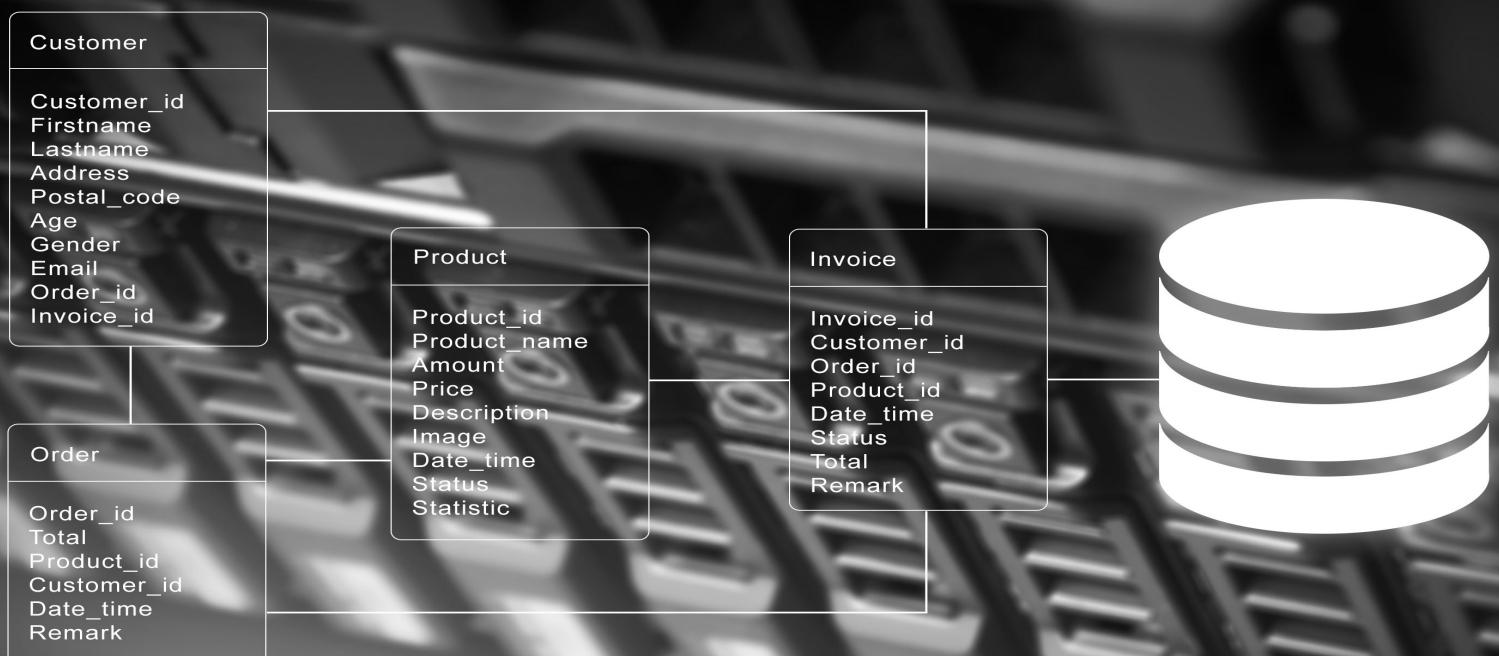
It is also clear that organizations are looking to other platforms to augment relational databases, notably Big Data (e.g. Hadoop), NoSQL, and more emergent technologies like AI, Machine Learning, and IoT. The survey split the many NoSQL data stores/models into separate categories, with Key-Value as the top choice. When looked at as a group, the use of Key-Value, Document, Columnar, and Graph data stores is quite a significant number at 70.5%.

Not surprisingly, spreadsheets are still a ubiquitous data source in the organization (66.3%), due to their accessibility and ease of use, particularly for the growing number of business users interested in data and data analysis.

Packaged applications such as ERP and CRM also play a key role and are a particular challenge. While they often store the most critical data – for example, around customers and accounts – they are typically very complex under the hood, and their architectures very “black box,” difficult to understand, and difficult to integrate into a larger Enterprise Architecture. There are tools and models that can help, but it is still a challenge for many.

The high percentages for JSON and XML show the need for data exchange between users and platforms in today’s heterogeneous data ecosystem, and the particular need for Web-based exchange.

Legacy systems such as mainframe and COBOL show no signs of going away, as organizations need to understand the structures in order to migrate to new platforms. In many cases those systems are still running the organizations in a stable, successful manner. The old adage of “if it ain’t broke, don’t fix it” seems apt when discussing legacy systems – they still do the job they are supposed to, and do it well enough to keep around.



Some of the results of this question were [Figure 8]:

- Relational On-Premises Database: 85.5%
- Spreadsheets: 66.3%
- Packaged Applications (e.g. ERP, CRM. Etc.): 56%
- Relational Cloud-based Database: 51.2%
- XML and JSON: 52.4% and 48.2%
- Legacy Systems (e.g. Mainframe, COBOL, etc.): 44.6%
- Big Data Platforms (e.g. Hadoop Ecosystem): 42.2%
- NoSQL Database (Key-Value; Document; Columnar; Graph): 25.3%; 20.5%; 12%; 12.7%
- Internet of Things (IoT) Data: 16.9%

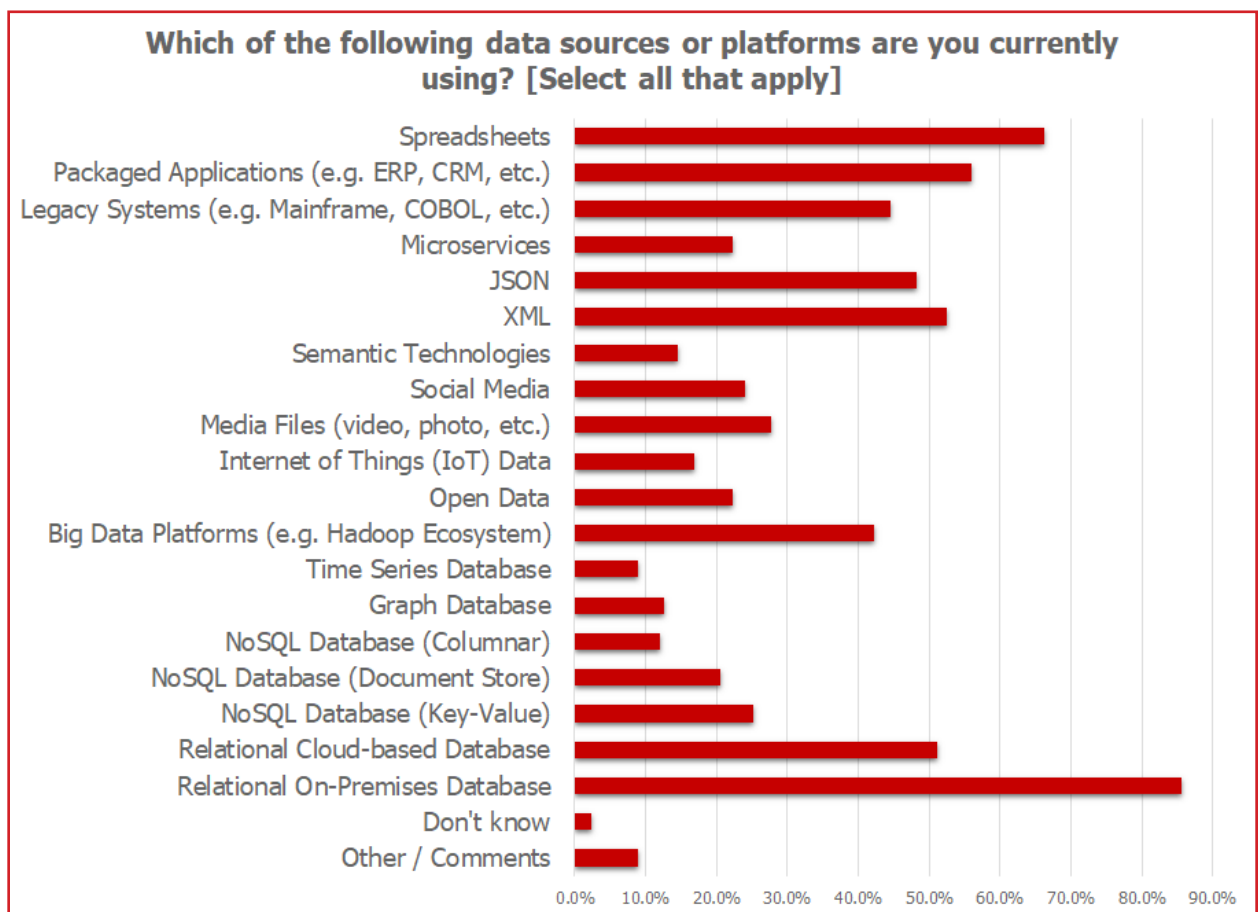


Figure 8



**B. “Which of the following do you plan to use in the future that you are not using currently? [Select all that Apply]”**

Future plans for data-centric technologies give insight into the growing trends in the industry. Not surprisingly, Big Data Platforms were the largest response, as organizations look to integrate higher volume and more diverse data sources into their analytic environments.

As noted in the previous section, relational databases continue to be popular among most organizations, although the trends are clearly shifting from on-premises to Cloud as organizations look to take advantage of the scalability and cost-effectiveness offered by these solutions.



The other leaders in new technologies include Graph Databases, which are seeing demand from organizations looking to create enterprise knowledge graphs with the flexibility of the graph model. Other key drivers include social media connections and fraud detection. Real-Time /Streaming Databases are also seeing a growth in interest. As more organizations look to become data-driven, the need for real-time data is increasing in order to provide data at the speed of business. Some of the specific results from this question were [Figure 9]:

- Big Data Platforms (e.g. Hadoop Ecosystem): 32.1%
- Relational Cloud-based Database: 29.6%
- Graph Database: 22.6%
- Real-time/Streaming Database: 22%
- Internet of Things (IoT) Data: 18.2%

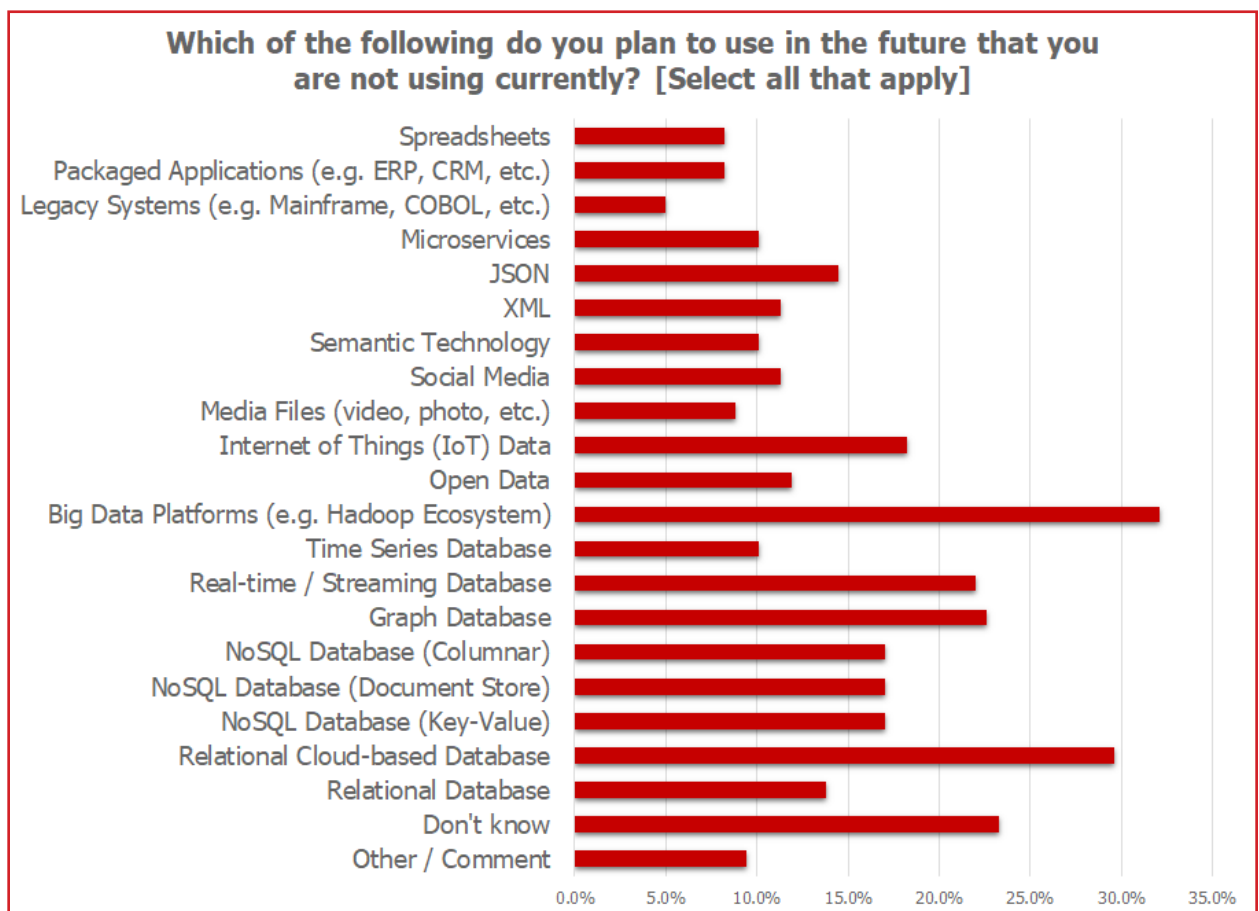


Figure 9

Several respondents included comments within this section. One comment in particular highlights the close relationship between business drivers and data-centric technologies that was highlighted earlier in the paper: “It can be whatever realizes my business information needs from a business operational, management, or strategic perspective.”



# Cloud Computing

The growth of Cloud Architectures has been expanding for many years now. There are many different models such as hybrid, private, and various configurations of public Clouds. Such models are dependent upon the needs of a particular organization. Big Data, IoT, and the need for rapid scalability and low-cost storage are key drivers for the growth of the Cloud space.

“

*Cloud adoption is clearly on the rise with over 75% of respondents currently implementing a Cloud strategy or planning to in the future.*

”

The question of Cloud Computing was covered in three survey questions.

## **A. “Are you currently leveraging the Cloud in your Data Architecture?”**

Cloud adoption is clearly on the rise with over 75% of respondents currently implementing a Cloud strategy or planning to in the future. The primary trend at this time is focused on a hybrid Cloud model with 47.7% of organizations using such a model. The hybrid model utilizes the many benefits of Cloud Computing, including scalability, on-demand control, and flexibility, while keeping some of the more sensitive data on-premises. As organizations pick the applicable use cases for Cloud vs. on-premises and leverage their relative strengths, such models will likely continue to expand.





Such implementations are not based on “all or nothing,” but rather “best of both worlds.” To save on costs for the storage of “cooler” data for backup and archiving, Cloud Architectures are certainly clear leaders. Only 4.3% of respondents said all their data was in the Cloud, while 17.7% said they are not using the Cloud now and have no plans to in the future [Figure 10].

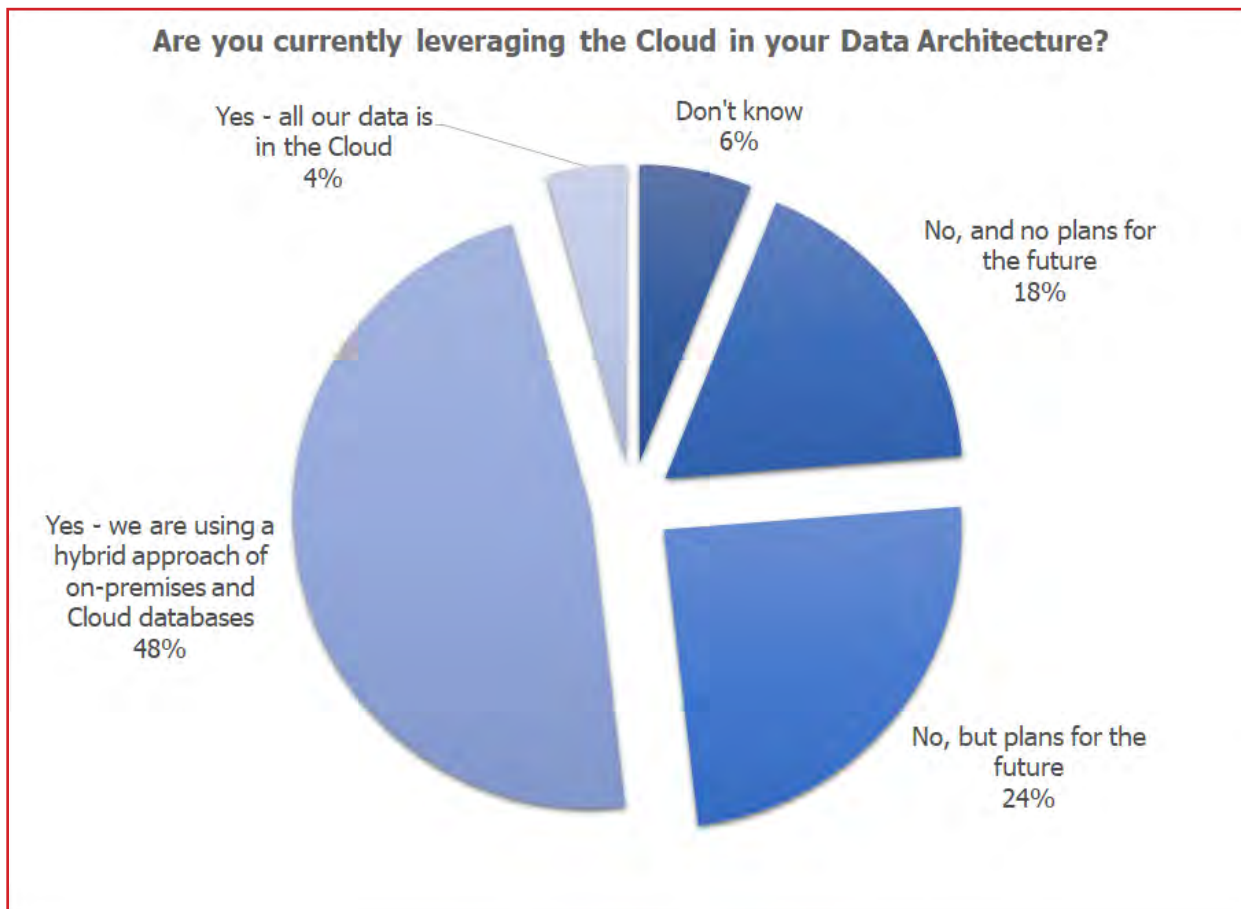


Figure 10

**B. “If you answered Yes to Question 14, what were your reasons for moving to the Cloud? [Select all that apply]”**

Scalability is the most prominent driver for Cloud adoption, which aligns with respondents’ earlier statements that growth in data volumes is one of the biggest trends in recent years. As data volumes continue to expand, storage costs also grow. Therefore, organizations are putting more thought into the issue of retention, storage, and archiving strategies, and the Cloud offers a suitable model for high-volume or infrequently accessed data.

In terms of expenditures, Cloud costs are often considered Opex (operating expenditures) instead of Capex (capital expenditures), which is appealing for many organizations trying to gain executive support for moving to such a model.

The top three percentages for this question were [Figure 11]:

- Scalability: 74.3%
- Cost Savings: 60%
- Variability in data usage patterns throughout the year: 18.1%

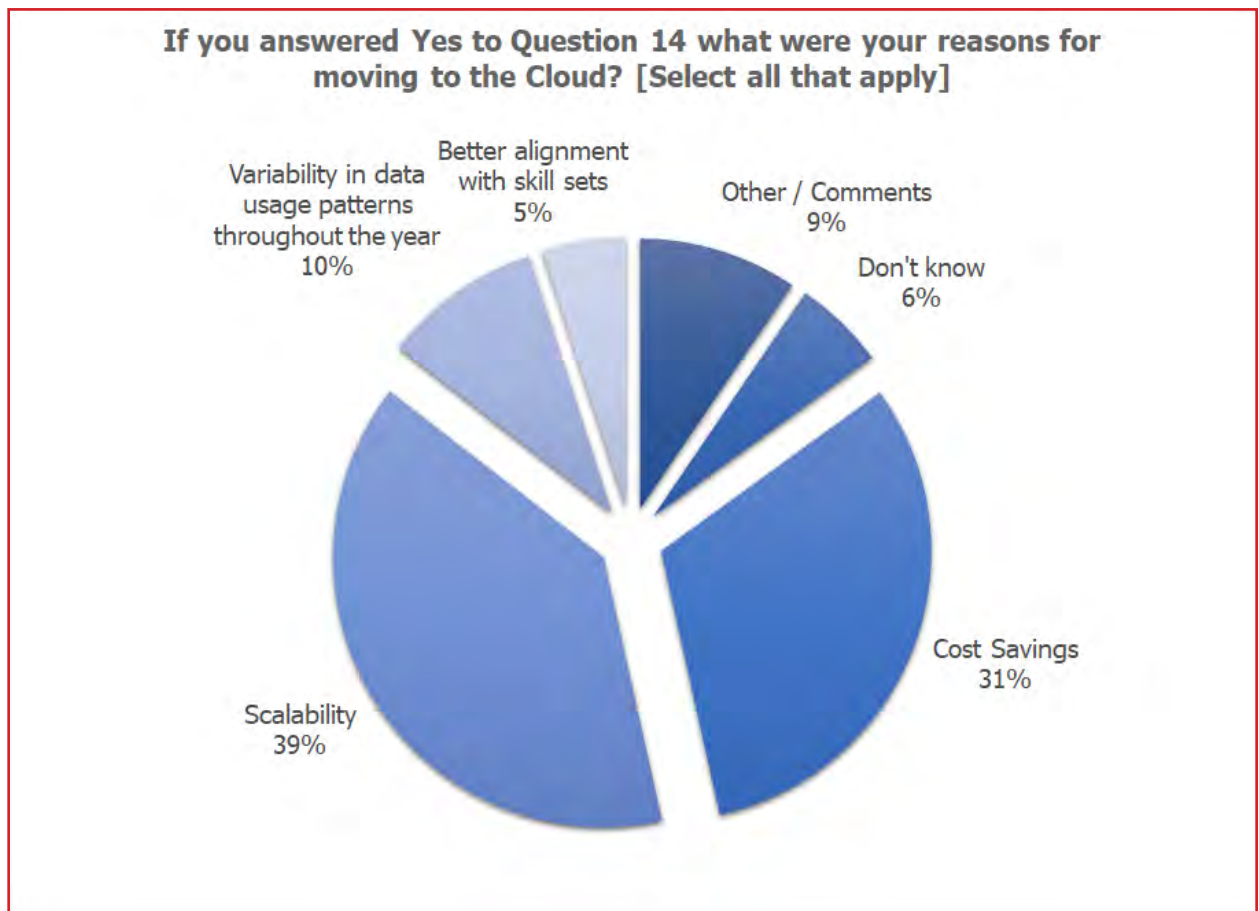


Figure 11

Several respondents commented on this question, and some of the more revealing remarks included:

- “Feel safer on an expert provided Cloud that can manage emerging security risks. That said, fully understand that security risks continue to be very real wherever our data is.”
- “Information integration and interoperability with other applications.”
- “Keeping up to date with technology upgrades: in some cases, there is not so much choice about when to upgrade.”
- “Access to data during natural disaster or other emergency situation.”
- “Data localization, Data privacy, ability to move to a different Cloud provider later.”

**C. “Whether you answered Yes or No to Question 14, what are your concerns regarding moving data to the Cloud? [Select all that apply]”**

We followed this question up with a deeper look into the concerns of people moving to the Cloud. The top three responses were [Figure 12]:

- Security and Privacy concerns: 77.2%
- Skills Required: 32.7%
- Features not available in Cloud vs. On-Premise: 29%



Security remains a key concern for many considering a move to the Cloud, although some respondents to the previous question indicated that they actually felt safer in the Cloud. Similarly, while many felt that lack of skills hindered their move to the Cloud, many others felt that the Cloud helped relieve the burden of in-house skills. Clearly, there is a dependency on specific in-house skills, and there is no “one size fits all” approach.

Another key concern is the latency of vendors to provide feature parity in their current Cloud-based offerings, which is likely to mitigate over time as vendors continue to enhance their Cloud solutions.

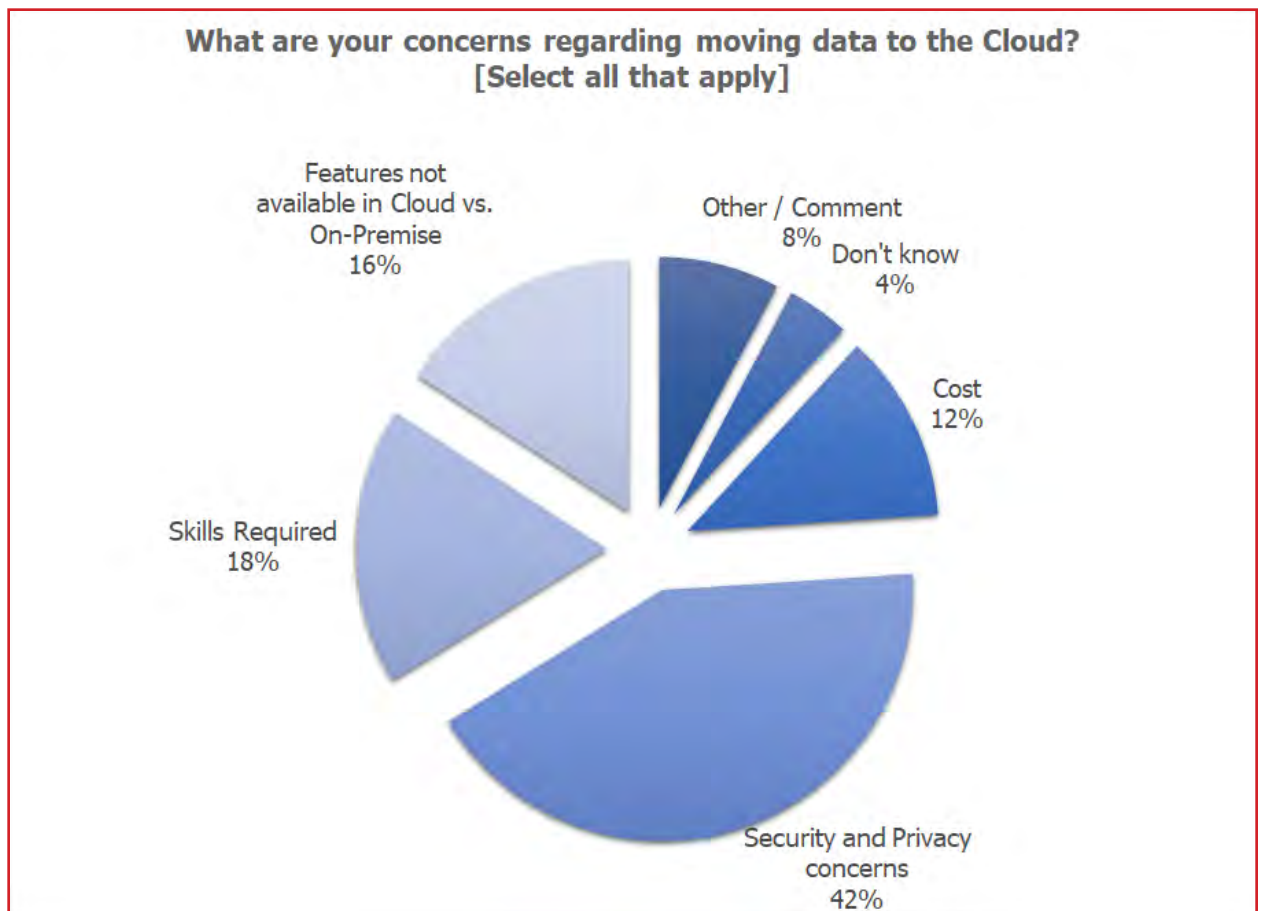


Figure 12



# Big Data Ecosystems

As the volume and variety of information continues to grow, many organizations are looking to Big Data solutions to provide a flexible, cost-effective way to manage these disparate sources for new business insights. A company can store literally exabytes of data relatively inexpensively compared to past costs. It is not always a simple task, however, to interact with that data and perform effective data discovery through Data Science and Machine Learning techniques while combining stored data with legacy and current relational-transactional data and other assets in Data Warehouses at the speed that the modern marketplace demands. This is where the comments of previous sections concerning the continued focus on Data Governance, Data Quality, Metadata Management, and Master Data Management all come into play: all the Big Data in the world is useless if it can't be reliably governed and leveraged.

This section of the survey asked four questions in relation to Big Data and Big Data technologies.



### A. “Are you currently using a Big Data platform?”

Not surprisingly, the “Yes” answer garnered the largest percentage by far at 42.9%. Many organizations not currently using Big Data solutions are considering them in the future at 27.6%, leading to a total of 70.5% of organizations who are either currently using or planning to use these solutions. At the same time, however, a large percentage of organizations are still not using Big Data solutions, with “No, and no plans for the future” receiving 21.5%, indicating that this technology is not a fit for every organization and use case [Figure 13]:

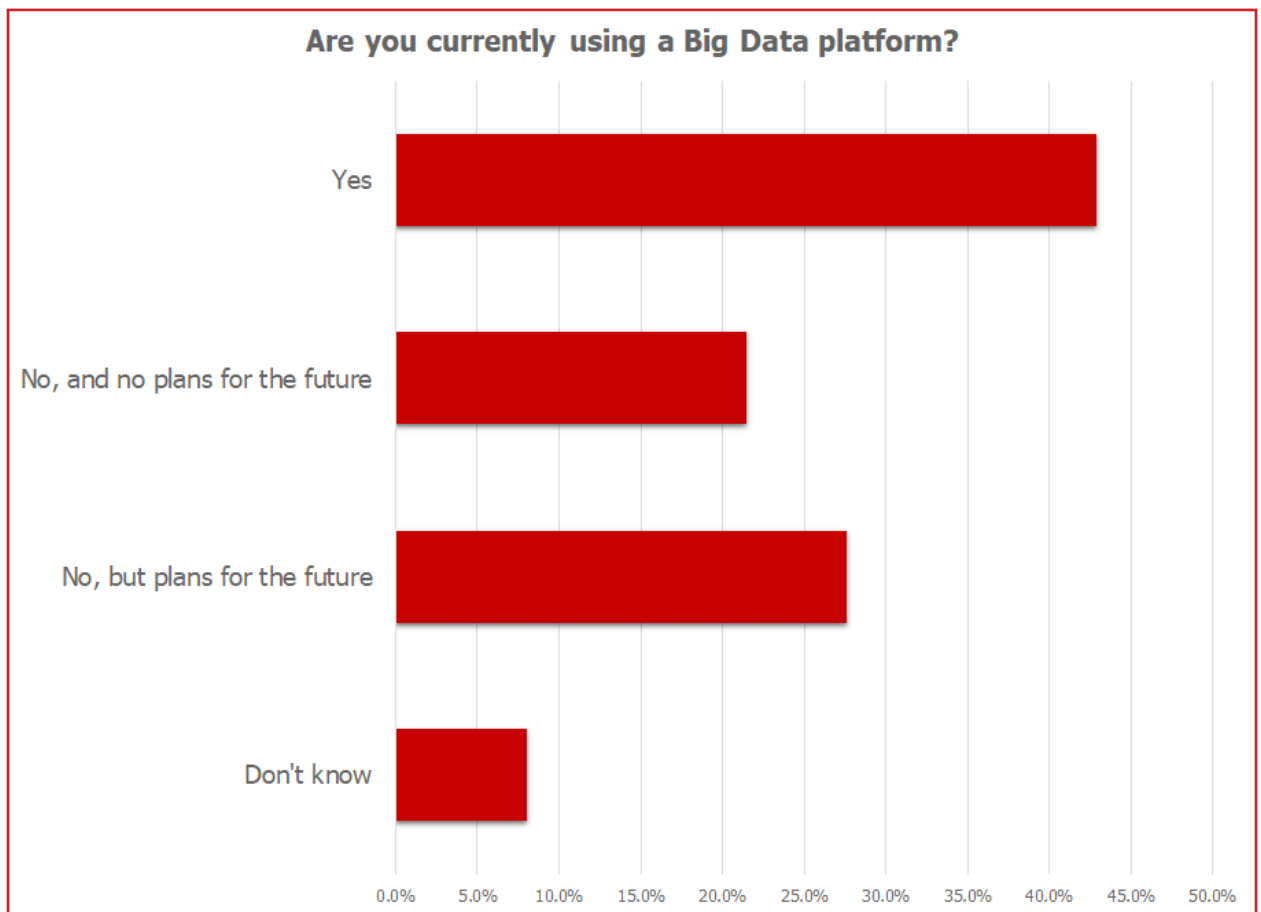


Figure 13

### B. “For what use cases are you using Big Data? [Select all that apply]”

Data Science, Reporting, Analytics, and Exploration lead the list of use cases, as organizations look to discover new business insights from their disparate data sources. More organizations are seeking to better understand their diverse range of

data assets across a wide variety of sources, from relational database assets to videos, social media, IoT data, spreadsheets, emails, and many more [Figure 14]:

- Data Science and Discovery: 58.6%
- Reporting and Analytics: 55.5%
- “Sandbox” data exploration or testing: 43%

If we look back to the results of Figure 4 on “main goals and drivers for implementing Data Architecture,” these answers are in direct concordance. Reporting/Business Intelligence (68%) and Data Science/Discovery (52.7%) were two of the top choices mentioned. Along with those top choices, emergent trends like Machine Learning/AI garnered a fairly high percentage at 28.1%, as did Storage/Cost Savings at 26.6%.

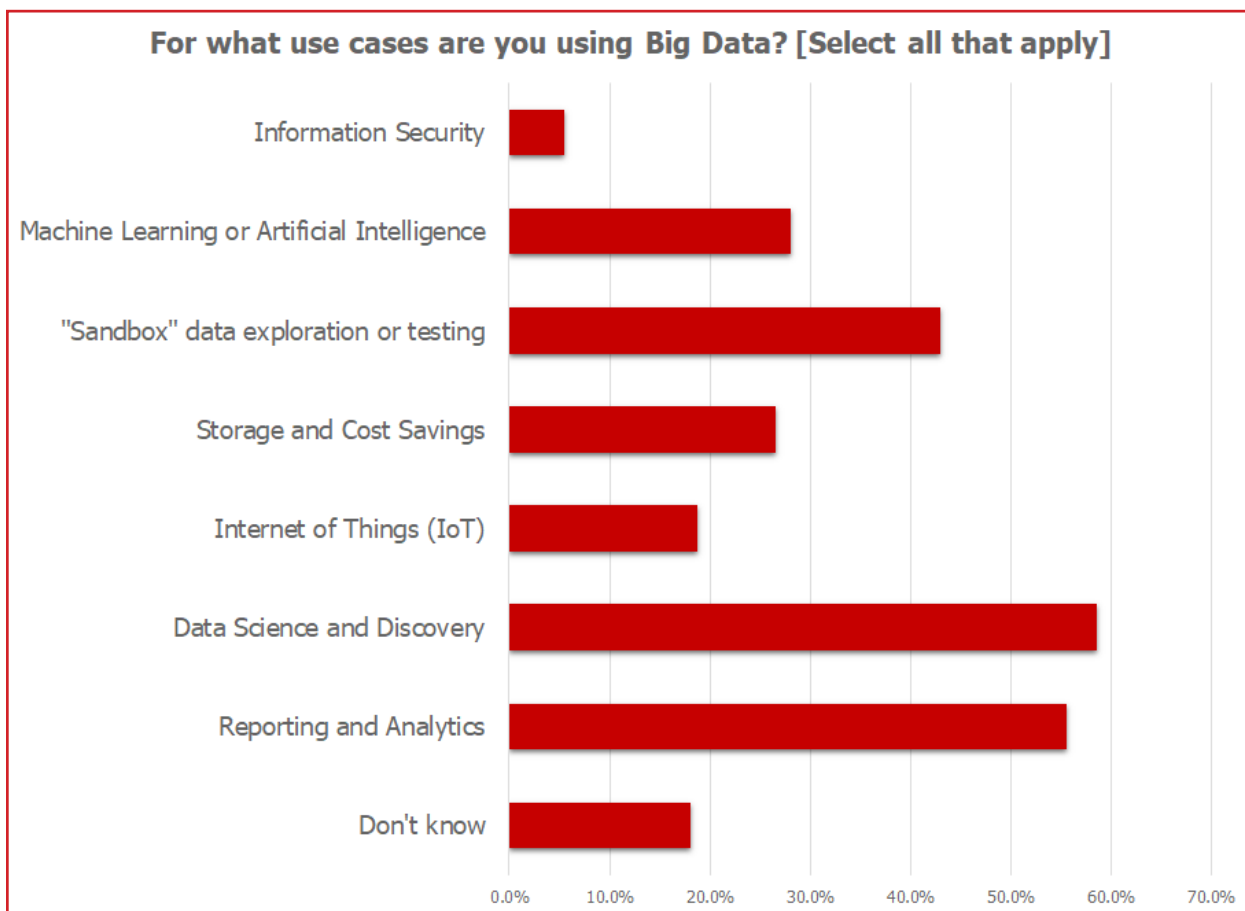


Figure 14

### C. What are your concerns, if any, about using Big Data? [Select all that apply]"

The movement towards Big Data solutions is evident. If an organization is not already experimenting with such technologies or doesn't already have one or more in production, most are considering such options. However, many concerns still exist about Big Data technologies that need to be addressed by Open Source developers and commercial vendors to better aid in the growth of the market, including [Figure 15]:

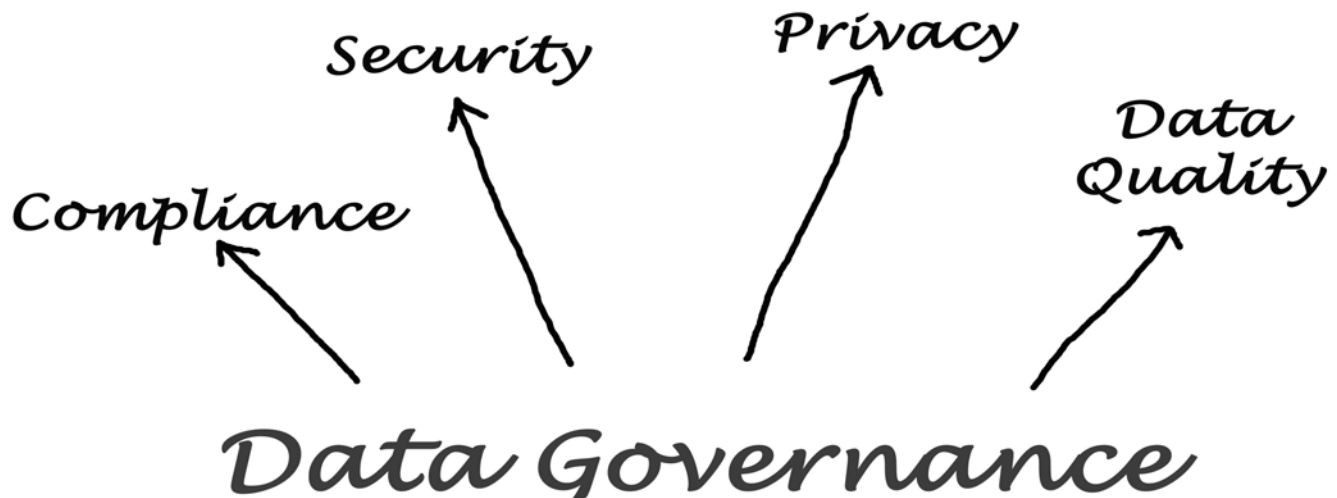
**Data Governance and Data Security are key issues for both Big Data and Cloud implementations. This a key challenge of modern Data Architecture—managing high volumes of data from disparate sources in a governed manner.**

- Complexity of solution: 39.9%
- No skills in-house: 36.5%
- Security: 35.1%
- No use case: 26.4%
- Cost: 20.3%

A few of the open-ended comments included such issues as:

- Other database platforms can support very large data sets, so the amount of data has to be tremendous in order to justify Big Data.
- ROI difficult to justify.
- Concerns of quality, governance/provenance, etc.
- Difficulty in integrating with existing systems.
- Seen as hype.

Data Governance and Data Security are key issues for both Big Data and Cloud implementations. This a key challenge of modern Data Architecture—managing high volumes of data from disparate sources in a governed manner.



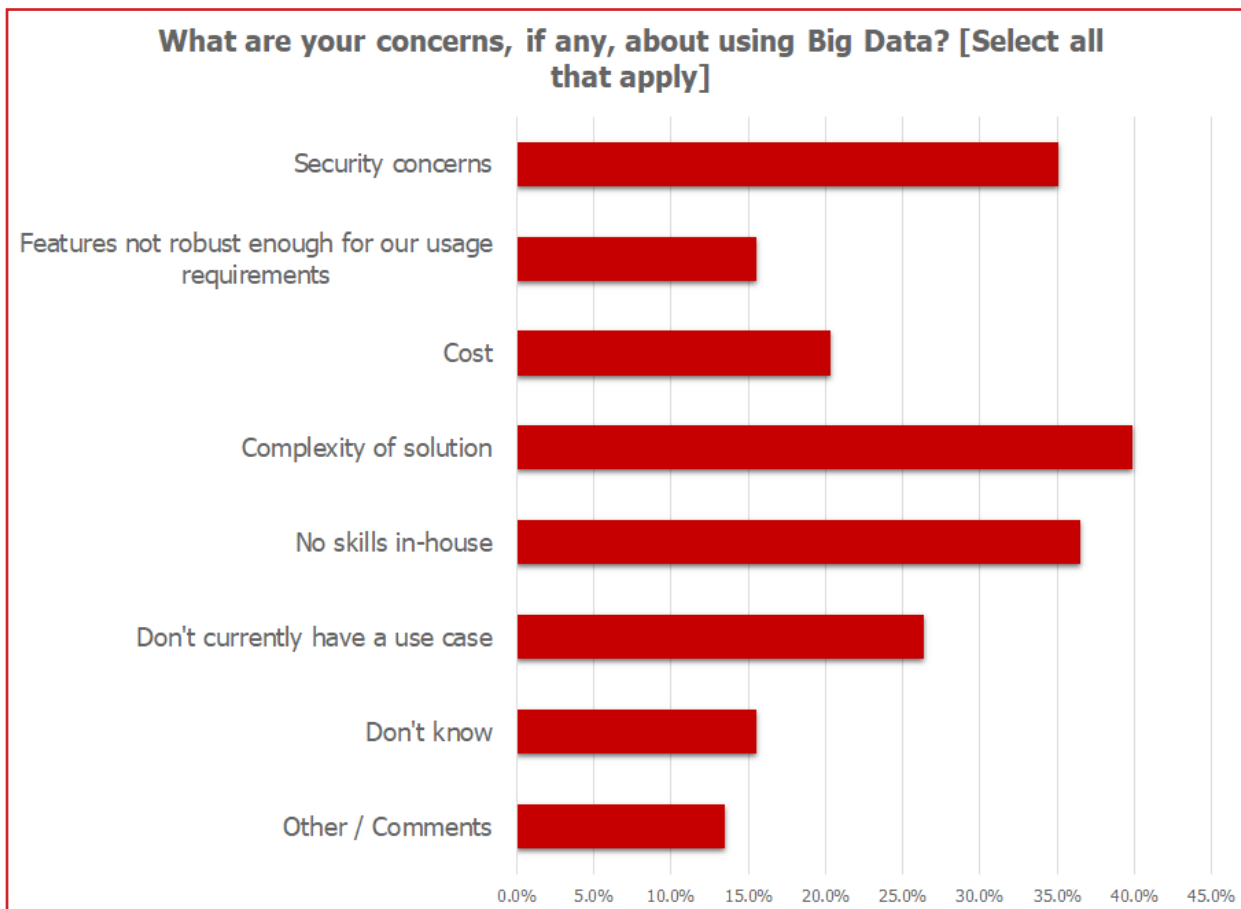


Figure 15

#### D. “Are you currently implementing a Data Lake?”

The final question of this section discussed the issue of Data Lakes, an emerging technology similar (at least conceptually) to the Data Warehouse, but built with Big Data in mind. Where Data Warehouses are a long-standing, well-known, and relatively stable technology built primarily for relational data structures, Data Lakes are relatively new in the data space.

Data Lakes are typically focused more on unstructured and/or raw data, though they can be used to store all forms of data, while Data Warehouses tend to work with structured or processed data. Data Lakes typically use a schema-on-read versus a schema-on-write model, are designed for the low-cost storage needs of Big Data assets, and strive to be highly agile with the ability to configure them on the fly, versus the fixed configuration common to Data Warehouses. Both technologies should be built around a required business need and desired benefit, and both need proper Data Governance policies and procedures to be effective.

Data Lakes endeavor to bring more flexibility, scalability, ease of organization, and the

ability to look across diverse and high-volume data sets to answer business questions through the work of Data Scientists, Predictive/Prescriptive Analytics, Machine Learning/AI models, etc., but as has often been the case with implementations done without proper due diligence, Data Lakes can quickly become Data Swamps.

The top three answers from respondents concerning Data Lakes were [Figure 16]:

- No, we do not have a use case for a Data Lake: 32.5%
- Yes, in conjunction with a Data Warehouse (DW): 22.9%
- Yes, as its own solution (apart from a DW): 14.6%

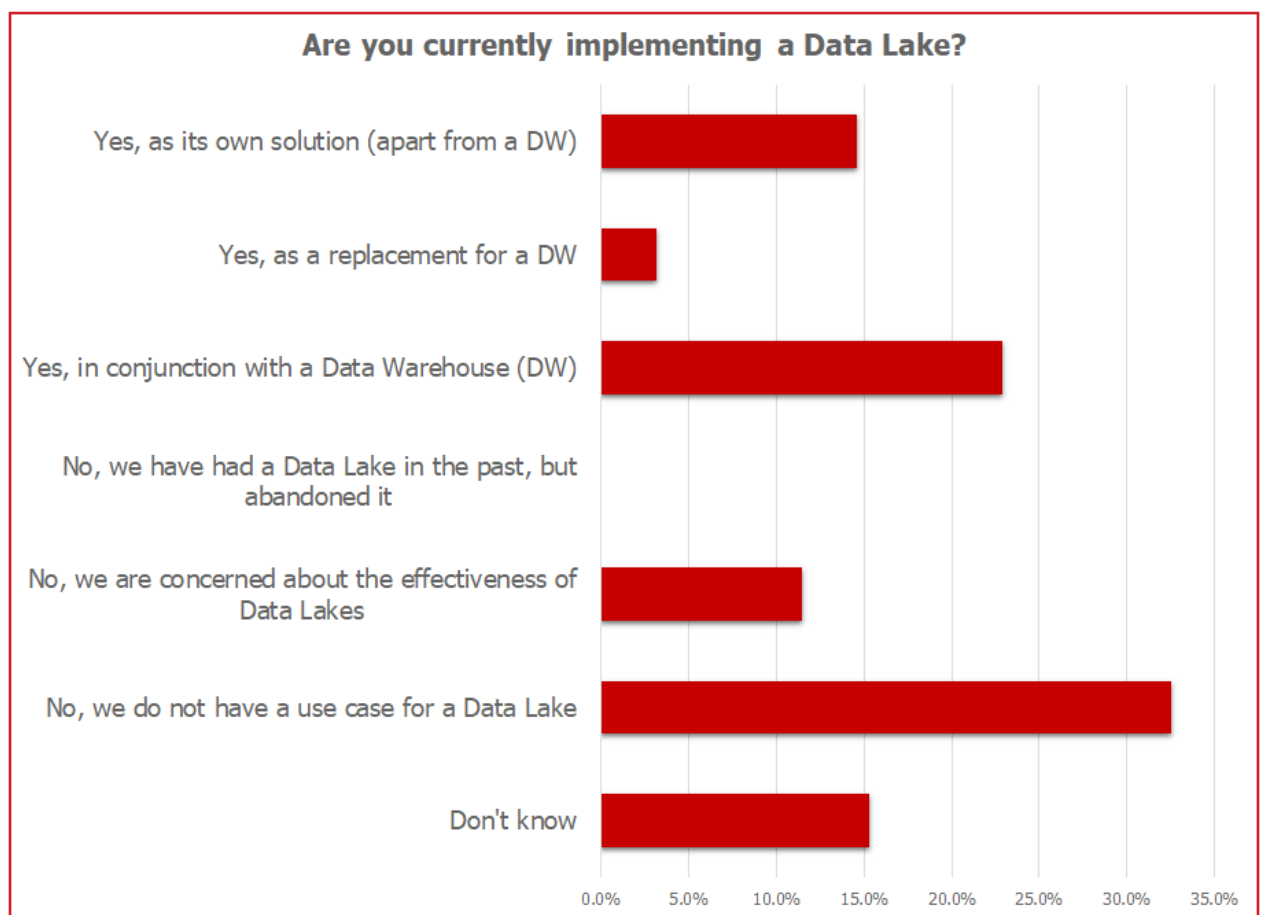


Figure 16



A few of the more notable comments from respondents about this question were:

- “Really an extension of traditional DW. Issue is whether a Data Lake can really cover all the types and sources of data that an enterprise would want to cover through an integrated Data Architecture strategy.”
- “We see more ‘Data Swamps’ than Data Lakes. This is due to the following major issues: 1. Lack of understanding of why a Data Lake is necessary (what is the business problem that it solves and how much of a cost savings does solving the business problem in this way provide?). 2. Poor Data Quality of data source(s). Data Warehouses usually have specific ETL routines to ensure that the source data is clean or prepared to the Data Warehouse model specifications. Big Data is just a lot of data that is usually poor quality. 3. Not enough knowledge or not proper organization infrastructure to support a Data Lake (or even a Data Warehouse). 4. Belief in magic solutions (we can build a Data Lake in two weeks and get a bonus).”
- “Potential replacement to the DW. We are currently running through use cases now.”





## Open Source & Open Data

The use of Open Source software has been a growing trend in the Data Management industry, as organizations look to cost savings and flexibility with the Open Source approach. Projects supported by the Apache Foundation such as Cassandra, Hadoop, Hive, Spark, Lucene, Pig, Storm, and a long list of others have been developed and are in production at enterprises worldwide. Many are now packaged together with other applications and sold as commercial platforms with varying uses, from Big Data processing to Advanced Analytics to data storage to search & query. Open Source platforms allow organizations to test and develop new tools without the overhead costs of buying off-the-shelf products, but also often require more in-house developers to implement and lack the support of commercial products.

*The majority of respondents (52.1%) said they were using Open Source technologies, while a significant percentage (21.5%) indicated they had plans to in the future.*

### A. “Are you currently using any Open Source technologies in your Data Architecture?”

The majority of respondents (52.1%) said they were using Open Source technologies, while a significant percentage (21.5%) indicated they had plans to in the future. Only 10.4% said they had no plans to use Open Source solutions in the future [Figure 17]:

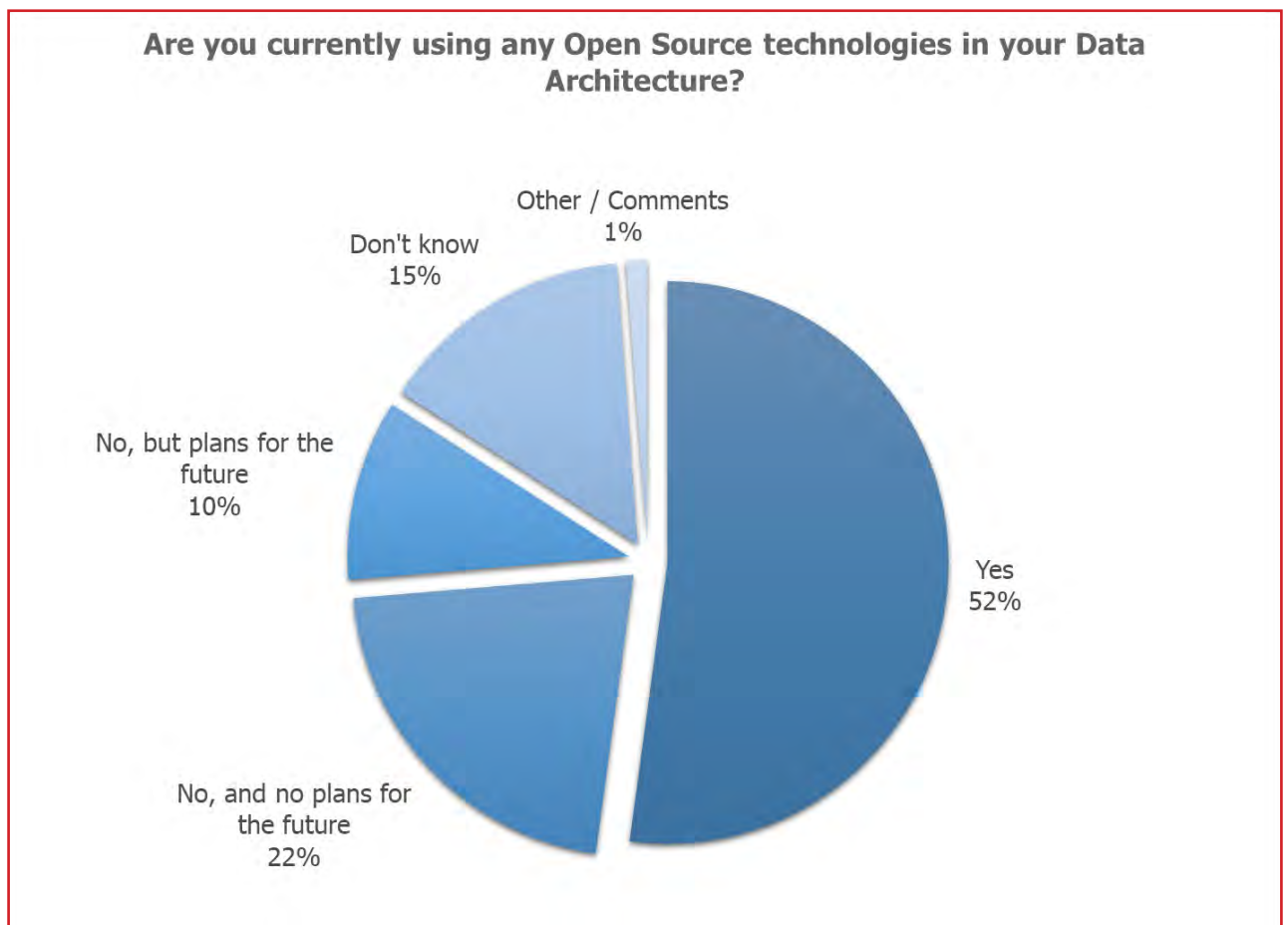


Figure 17

## B. “Are you currently using Open Data Sets?”

Open Data are data sets provided by organizations, governments, and even individuals for free public use. The Open Data movement allies closely with the Open Source movement, as well as Open Government, Open Hardware, Open Content, and other movements that support transparency and free use of information. Some of the biggest Open Data providers are various government agencies, along with healthcare organizations, scientific communities, and educational institutions. A few well-known Open Data sets include:

- Landsat satellite imagery
- 1000 Genomes Project
- Common Crawl webpage index
- Multimedia Commons
- IRS 990 filings on AWS
- USA Spending on US Government expenditures



Other topics in the Open Data movement include climate, agriculture, consumer-related topics, energy, finance, manufacturing, and public safety. As data becomes an ever more crucial aspect of modern marketplace success, we wanted to know if organizations are employing these free data sets themselves. The results show that while Open Data has not yet taken a widespread foothold, nearly 46% (in total) of respondents are currently using or plan to use Open Data as either a publisher or consumer in the future [Figure 18]:

- No, not currently using: 36%
- Yes, we are consuming Open Data sets: 18%
- Yes, both consuming and publishing Open Data sets: 8.1%
- Yes, we are publishing Open Data sets: 3.1%
- No, but planning to in the future: 9.9%

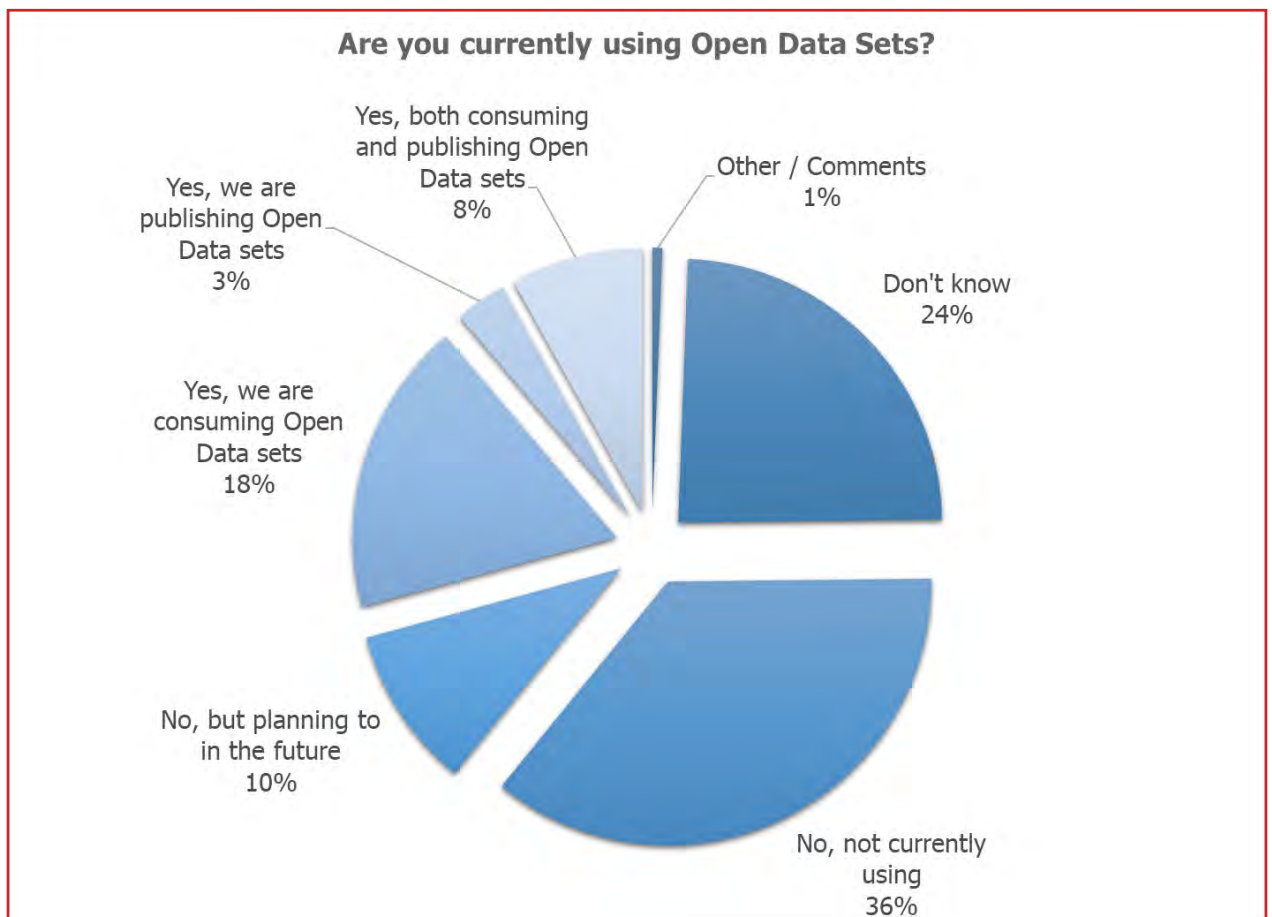


Figure 18

# Blockchain Technologies

Blockchain is most known for its association with Bitcoin and other allied cryptocurrencies, although Blockchain technology is considerably more than that. Blockchain is at its core a distributed database technology based on “blocks” that are reputed to be more secure than other database types. According to [Blockgeeks](#):

“Information held on a blockchain exists as a shared — and continually reconciled — database. This is a way of using the network that has obvious benefits. The blockchain database isn’t stored in any single location, meaning the records it keeps are truly public and easily verifiable. No centralized version of this information exists for a hacker to corrupt. Hosted by millions of computers simultaneously, its data is accessible to anyone on the internet.”

Such a shared and reconcilable system provides robustness to the database and stops problems such as single points of failure and control of the data by singular individuals/bodies.

The benefits of Blockchain are therefore many, especially for legal and financial institutions. According to the [Harvard Business Review](#):

“With blockchain, we can imagine a world in which contracts are embedded in digital code and stored in transparent, shared databases, where they are protected from deletion, tampering, and revision. In this world, every agreement, every process, every task, and every payment would have a digital record and signature that could be identified, validated, stored, and shared. Intermediaries like lawyers, brokers, and bankers might no longer be necessary. Individuals, organizations, machines, and algorithms would freely transact and interact with one another with little friction. This is the immense potential of blockchain.”



The survey asked two questions concerning Blockchain technologies to help us gain more clarity about where organizations are in terms of their understanding of this emergent and clearly important trend.

**A. “Are you currently using or considering Blockchain technology? Which of the following best represents your organization’s view of Blockchain?”**

The results here are not surprising due to the recent focus on Blockchain as a technology for possible general use. Clearly, many organizations are aware of Blockchain and are starting to learn more, but very few are actually employing such technologies in production [Figure 19]:

- We know what Blockchain is, but are not considering it seriously at this time: 26.5%
- We know what Blockchain is, and are interested in learning more about its benefits and uses: 17.4%
- We are using or actively pursuing adoption of Blockchain technologies: 7.7%

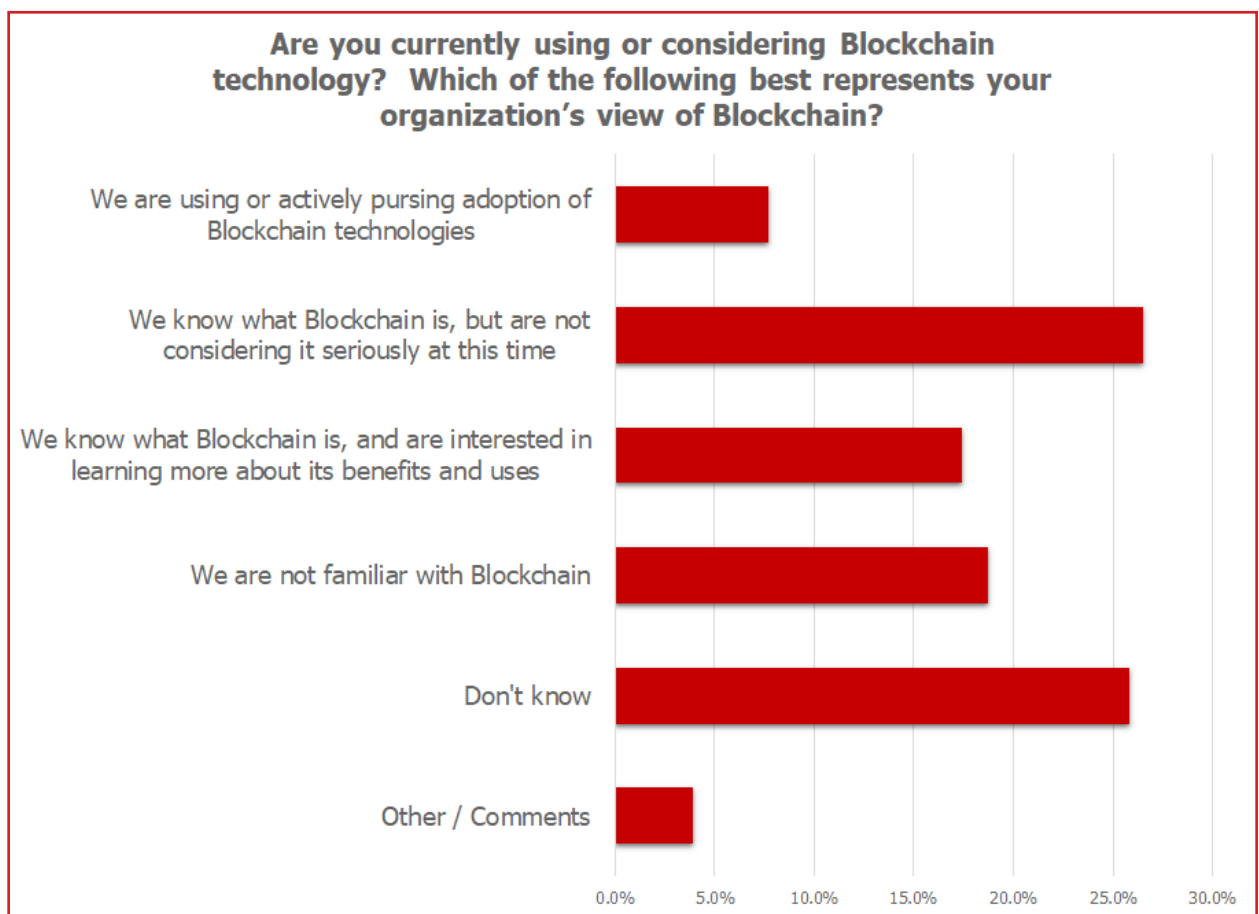


Figure 19



**B. “For which use cases are you using or considering use of Blockchain technology? [Select all that apply]”**

The current primary use cases for Blockchain include Record Management (14.6%), Peer-to-peer transactions (14.6%), Regulatory Compliance & Audit (13.8%), and Identity Management (13.1%) [Figure 20]:

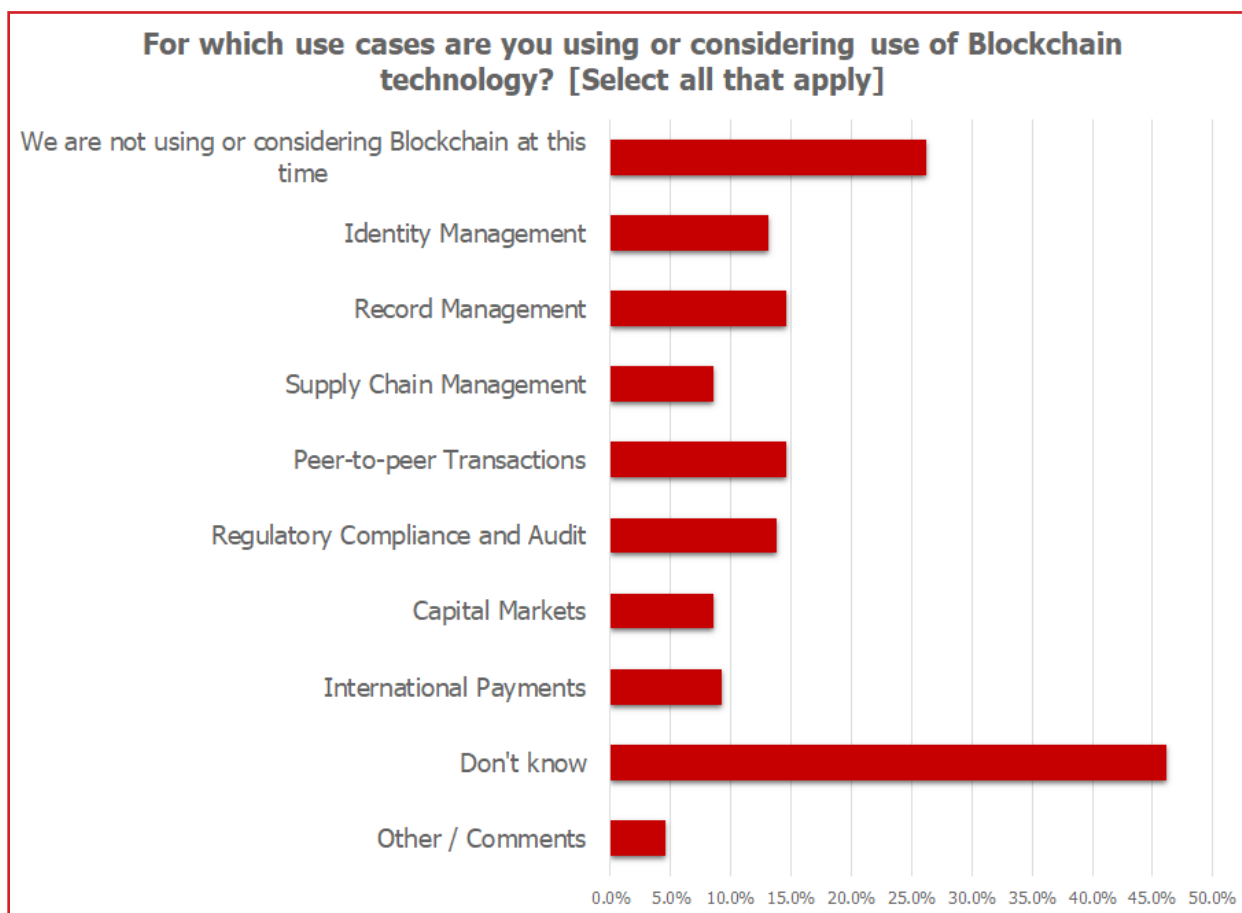


Figure 20

The open-ended comment section included answers such as:

- Property Conveyancing, Mortgage Settlements, and Settlement Payments.
- Managing and rewarding development of intellectual property.

# Master Data Management (MDM)

As more and more organizations struggle with obtaining a single, consistent view of core data such as Product, Customer, Vendor, and Location data, Master Data Management (MDM) is seeing a rise in implementation. Master Data Management is closely linked to other key initiatives discussed in this paper, such as Data Governance and Data Quality, but is a practice in and of itself made up of diverse activities such as matching rules, golden record survivorship rules, and data migration and/or virtualization.

**Notably, over 65% of respondents are actively pursuing an MDM strategy in various stages of maturity.**

Notably, over 65% of respondents are actively pursuing an MDM strategy in various stages of maturity [Figure 21]:

- We are in the process of beginning MDM, but have not yet fully implemented: 33.5%
- We are actively using MDM, but are in our initial stages (lower level of maturity): 23.2%
- We are actively using MDM technology and are at a high level of maturity: 9.7%

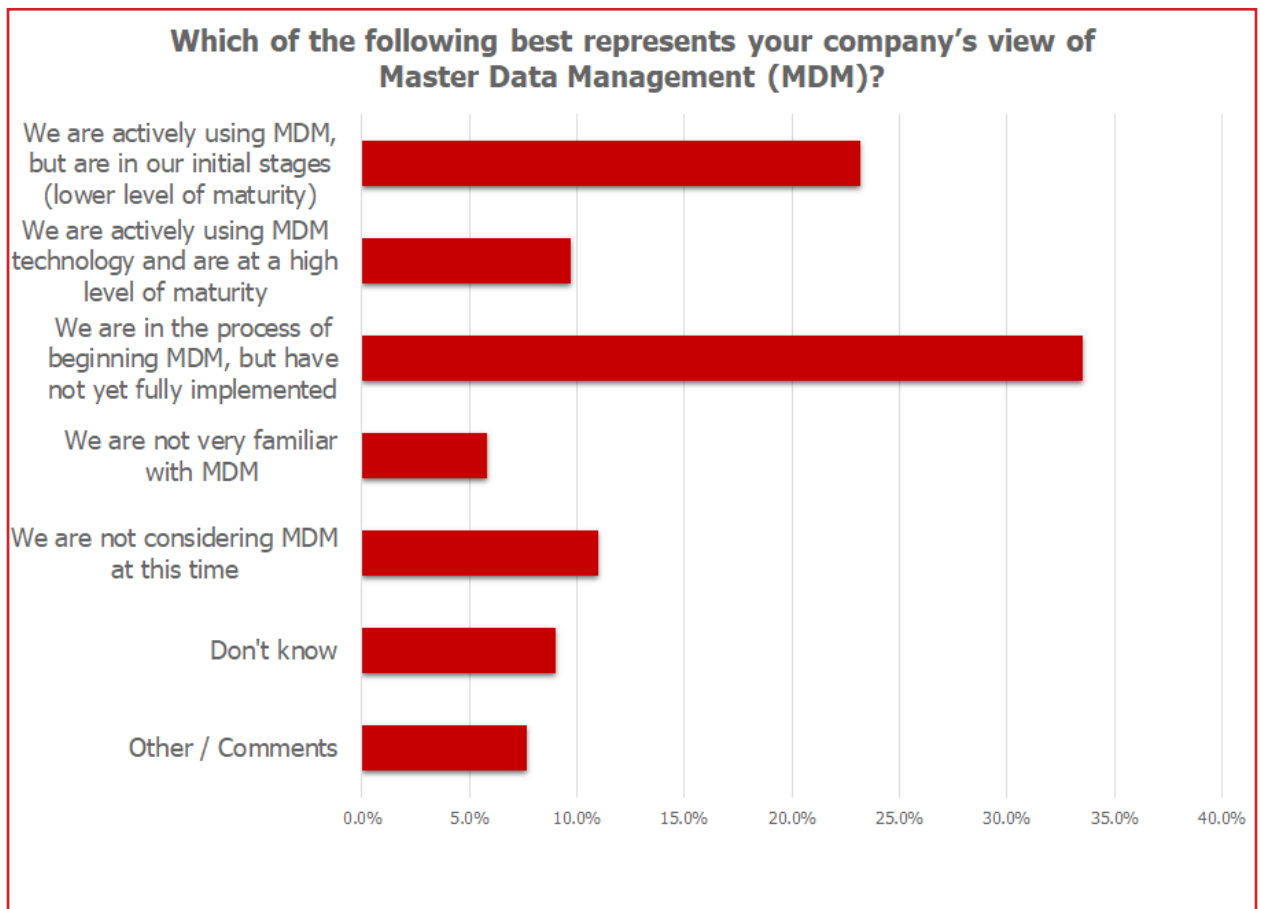


Figure 21



# Data Virtualization

Data Virtualization (DV) allows for the management, delivery, and use of an organization's data assets through the “virtual” presentation of those assets. Data Virtualization combines data from various sources, structures, and possible locations into a single abstracted data layer or “virtual” space without requiring data movement or ETL into a common data store.

The question of Data Virtualization was also discussed in a DATAVERSITY® 2013 report titled [\*The Utilization of Information Architecture at the Enterprise Level\*](#). In that paper, we asked respondents, “Which of the following best represents your company's view of Data Virtualization?” The responses included not planning on implementing at that time (28.6%), using or actively pursuing DV at that time (18.8%), and lacking familiarity with DV (32.3%) [Figure 22]:

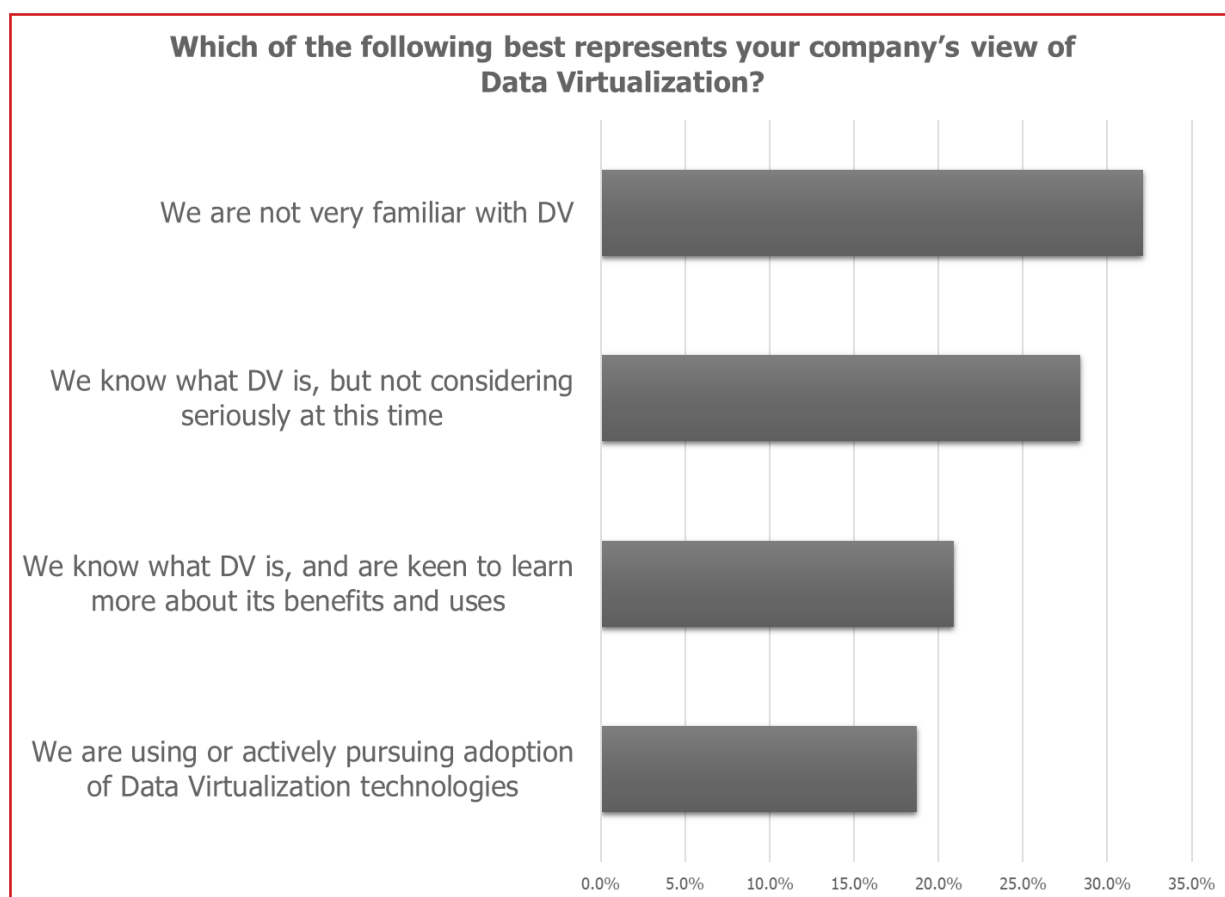


Figure 22

We therefore decided to ask a similar question again this time around, since Data Virtualization is part of the wider Data Architecture world. The responses show that while there is now more DV adoption and general understanding, it remains a difficult and complex topic that is not well understood at a wider level [Figure 23]:

- We know what DV is, but are not considering seriously at this time: 22.4% (down from 28.6% in 2013)
- We are actively pursuing adoption of DV technologies, but are in the early stages: 17.8% (down from 18.8% in 2013)
- We are not very familiar with DV: 15.1% (down from 32.3% in 2013)
- We are using DV technologies: 12.5% (not directly asked in 2013)
- We know what DV is, and are interested in learning more about its benefits and uses: 11.2% (down from 20.3% in 2013)



There were several open-ended comments that discussed the fact that, for the successful adoption of DV technologies, it is first necessary to have a strong Data Architecture in place and a well-governed system that promotes collaboration between all the enterprise stakeholders who will be using such a system.

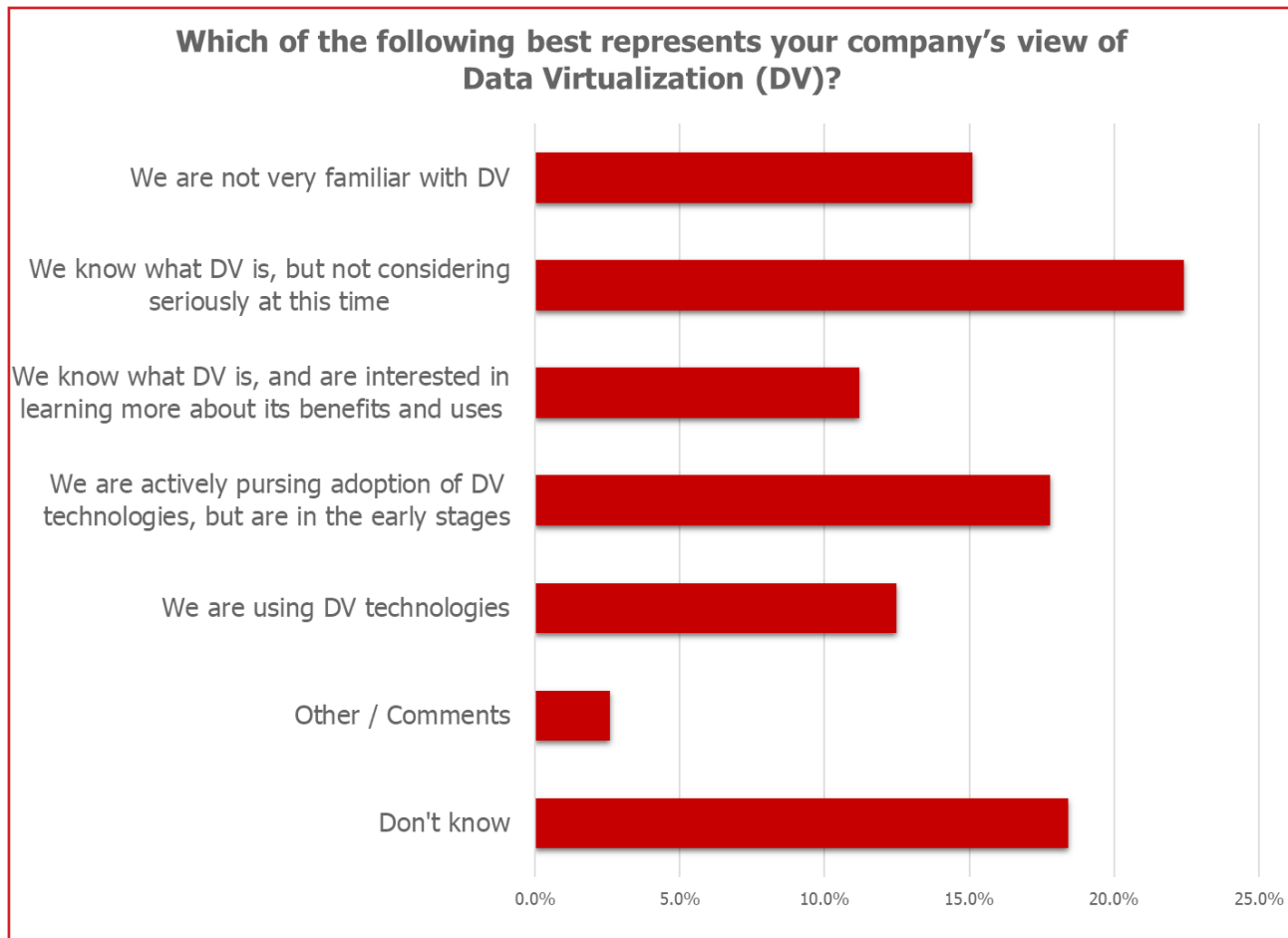


Figure 23



## Data Design & Modeling

“

*Notably, over 96% of respondents indicated that they were engaging in Data Modeling activities, which is a significant, positive response.*

”

The final piece of this Data Architecture trends report covers an important set of associated topics: Data Modeling, Metadata, and Data Design. These are all core practices and concepts to building an effective and stable Data Architecture. Even in the new world of hybrid relational and non-relational designs, the ability to understand the way that data and information flow through an organization is paramount to successful business practices and processes. Notably, over 96% of respondents indicated that they were engaging in Data Modeling activities, which is a significant, positive response.

### A. “Which methods of Data Modeling do you use in your organization? [Select all that apply]”

The clear majority of organizations are using ER modeling, which aligns with the high percentage of organizations using relational databases (85.5%) as discussed in Figure 8.

There are also large numbers employing Data Warehouse modeling techniques with either a Kimball- (53.9%) or Inmon-style (27%) schema. This supports the information discussed in previous questions that demonstrated a high demand for reporting and analytics to better understand and gain insights from an organization's data.

The popularity of Data Flow Diagrams (DFD) at 50% aligns with the existence of many disparate data sources and the need for integration. DFDs are critical to understanding how data flows across systems. This ties into the confirmed importance of Data Governance – which covers where and how the data is used, and how it flows – as well as Data Integration efforts such as Data Warehousing and MDM.





Other notable percentages include Hierarchical (e.g. XML) styles (34.9%), Object-relational designs (25%), Key-Value Models (21.1%), and UML (26.3%). Only 3.3% of respondents said that they are not doing any Data Modeling at all [Figure 24]:

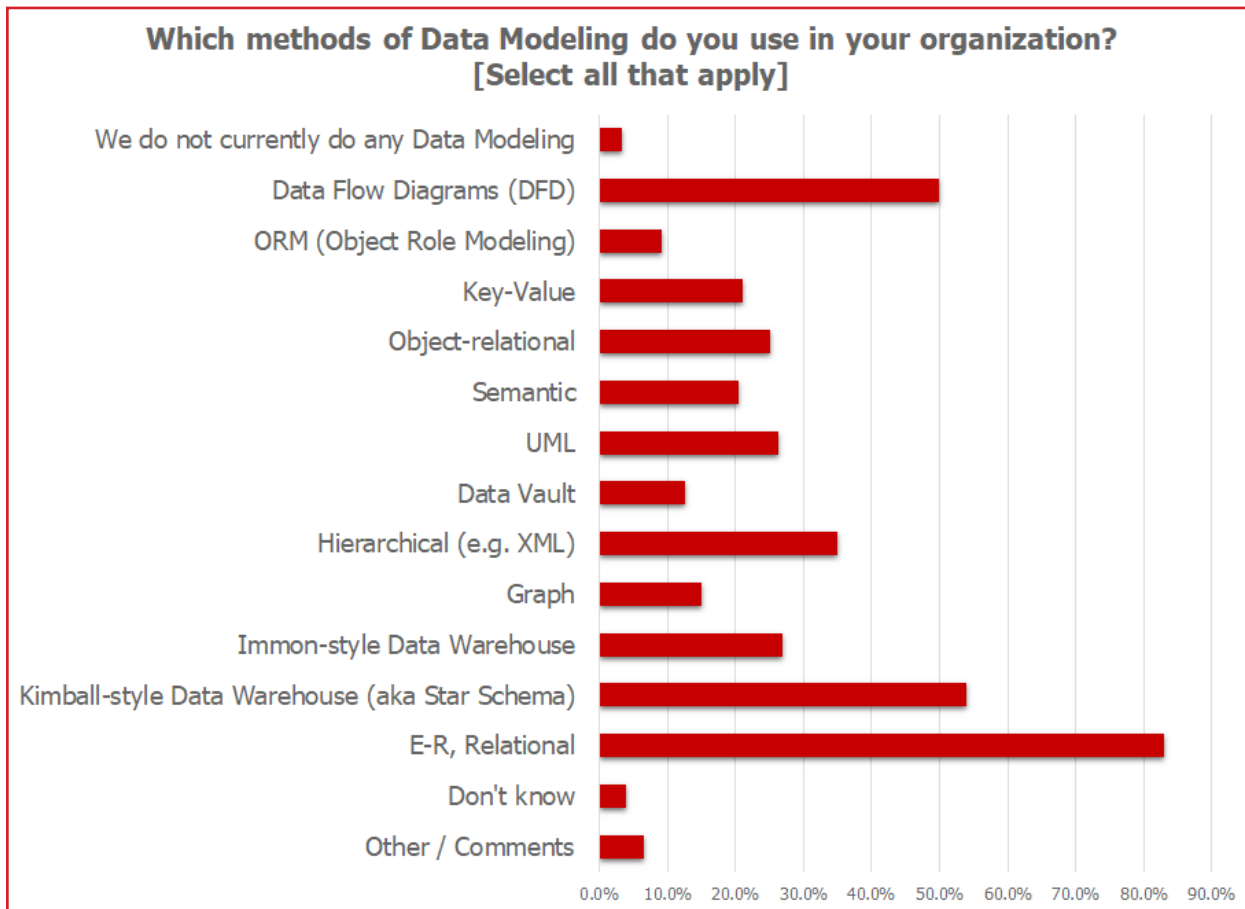
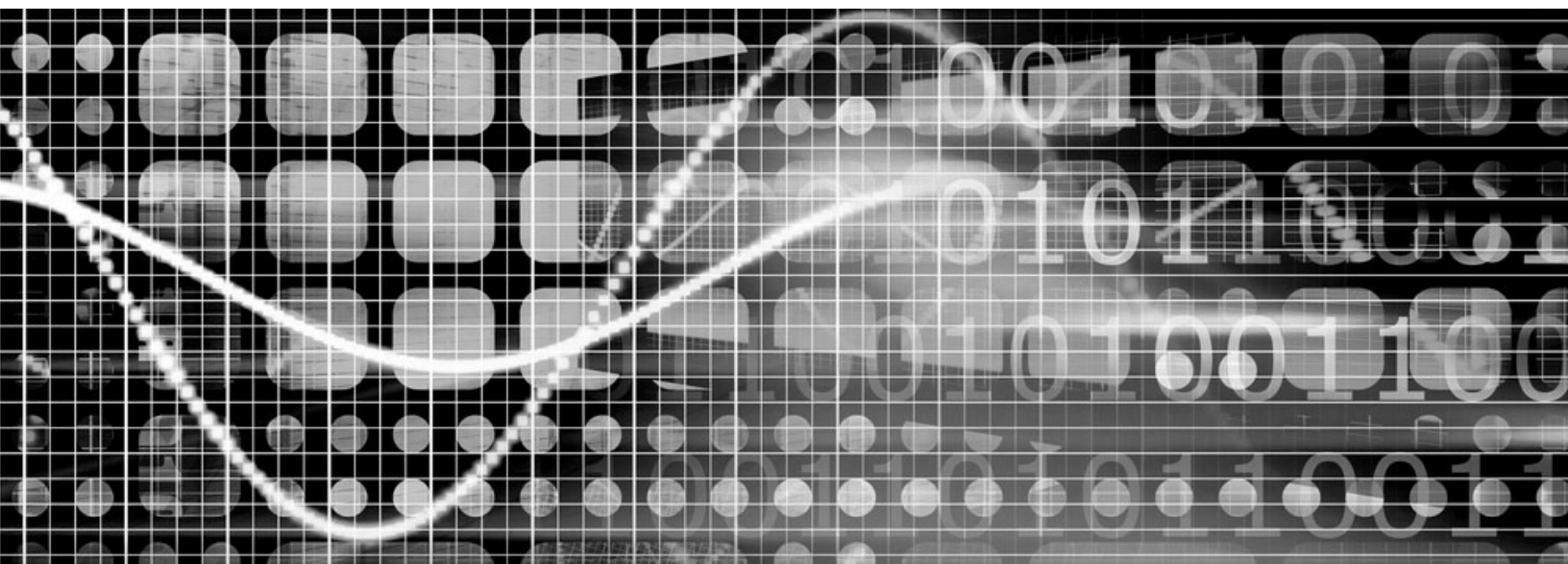


Figure 24



**B. “What types of models and diagrams do you use in your Enterprise Architecture? [Select all that apply]”**

As organizations increasingly link data to their business-centric initiatives, integrating data with a wider Enterprise Architecture is more and more common. Many of the most commonly used models are business-centric models that help with understanding diverse data sets and the interactions between them. Top responses included:

- Logical Data Models: 72.5%
- Conceptual Data Models: 67.8%
- Physical Data Models: 67.1%
- Business Process Models: 63.8%
- Data Flow Diagrams: 63.1%

The high popularity of Logical and Conceptual Data Models is not surprising, given the increasing number of business users looking to understand and leverage enterprise data. Physical data models are of course popular as well, as they are used to create and maintain the physical data asset inventory of relational and other storage mechanisms.

As business initiatives drive Data Management initiatives, Business Process Models are used, often along with CRUD matrices, to understand how data is used by critical business processes (Created, Read, Updated, or Deleted). Data Flow Diagrams are similarly used to understand the movement and interaction of data-centric systems across the enterprise.

“  
The high popularity  
of Logical and  
Conceptual Data  
Models is not  
surprising, given the  
increasing number  
of business users  
looking to understand  
and leverage  
enterprise data.”





These business-centric models are generally used in conjunction with each other to provide a comprehensive view of the enterprise data asset for a variety of use cases and stakeholder types [Figure 25]:

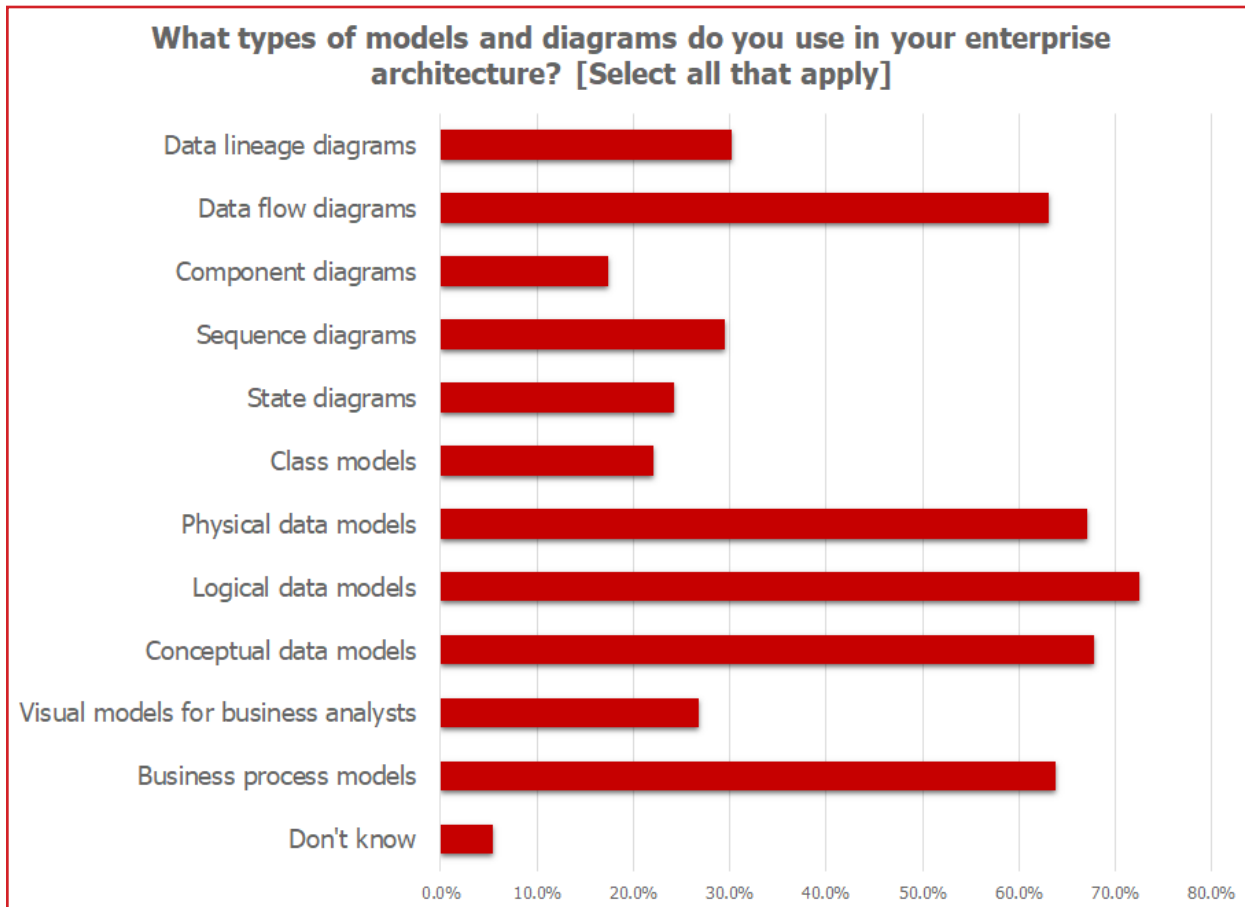
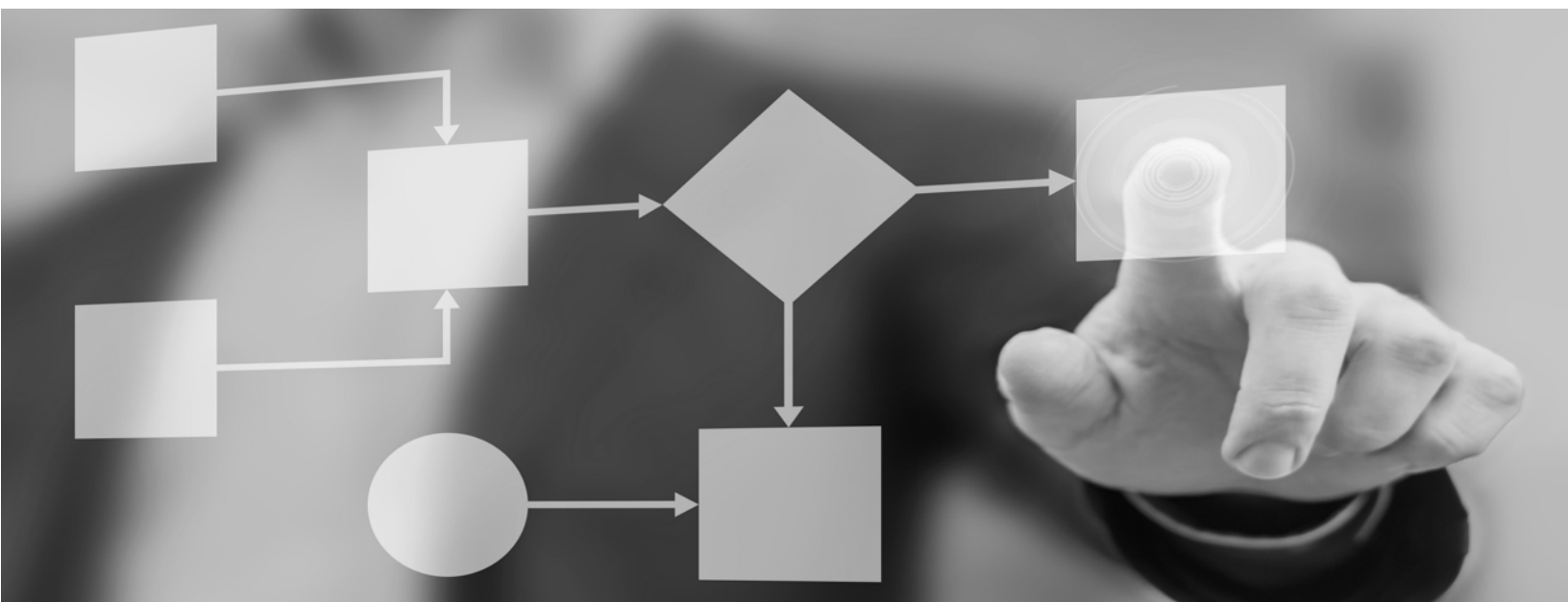


Figure 25



### C. “Do you currently have a Metadata Management effort in place?”

Metadata can be defined as data in context, or the “who, what, where, why, when, and how” of data.

That specific definition was also used in the [\*Emerging Trends in Metadata Management: A DATAVERSITY® 2016 Report on the Top Business & Technical Drivers for Metadata\*](#) and is still applicable today. In that report we discussed numerous different trends occurring within organizations concerning Metadata Management. A number of those previous findings align well with the information presented in this paper, including:

**“Metadata Management continues to be an important element in the management of data and the creation of a stable and trustworthy Data Architecture for both business and technical stakeholders.”**

- Two-thirds of respondents said that Metadata is more important now than it was ten years ago.
- Business users were the largest group of Metadata consumers at almost 80%.
- Big Data Platforms ranked as the most wanted future Metadata asset for organizations.
- Data Governance was the most prevalent current Metadata use case, along with Data Quality Improvement, Data Warehousing, and BI Reporting.
- Master Data Management, Big Data Analytics, Data Science, Semantic Technologies, Regulation and Audit, IoT Management, and Efficiency and Agility all rank as significant future planned Metadata use cases.

Clearly, based on the numbers of the earlier report analyzed in terms of the information presented here, Metadata Management continues to be an important element in the management of data and the creation of a stable and trustworthy Data Architecture for both business and technical stakeholders.

The vast majority—over 78%—of respondents for this report said that they are either already using Metadata Management or plan to in the future [Figure 26]:

- Yes, we have an enterprise-wide effort: 30.2%
- Yes, for individual projects only: 22.8%
- No, but we have plans for the future: 25.5%
- No, and we don't have any future plans at this time: 11.4%

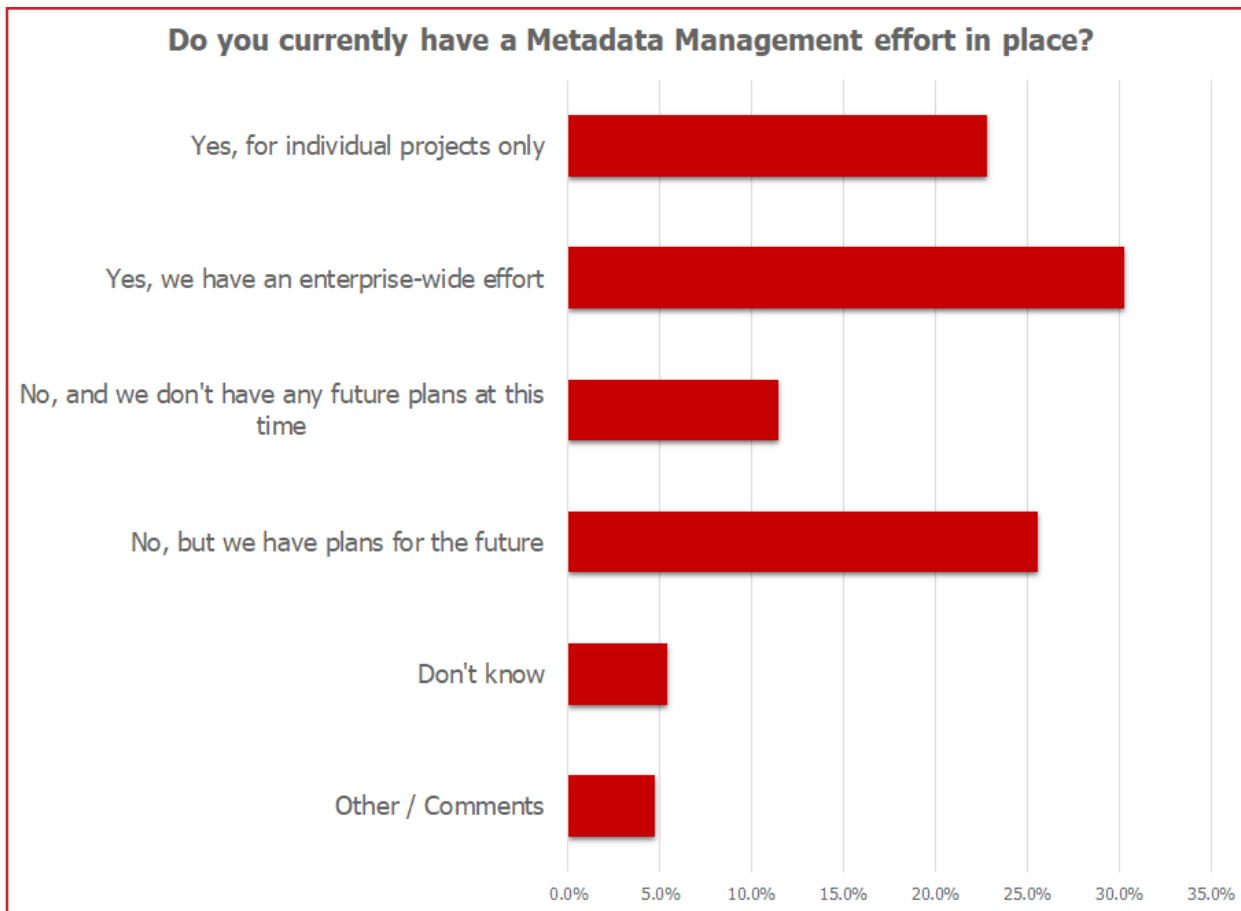


Figure 26

**D. “What are your current main use cases for Metadata? [Select all that apply]”**

This question was also asked in the 2016 Metadata Management Trends Report, and the percentages are telling. Many of the options have shown a considerable increase in importance. The themes already discussed throughout this paper are shown to be true for Metadata Management as well, such as the need to better govern data as an asset, better Data Quality, better Data Governance, and the ability to report and gain clear insights from data. Governance is not only determined by business drivers, but also by the need for regulation and audit controls. Also, Metadata is seen as an important driver of efficiency and agility.

Some notable responses for this question are listed below. The 2016 report and this 2017 report is Figure 27.

- Data Governance: 78.2% (in 2016, 64%)
- Data Quality Improvement: 62.7% (in 2016, 52.2%)

- Data Warehousing and Business Intelligence Reporting: 62% (in 2016, 53.2%)
- Regulation and Audit: 31% (in 2016, 22%)
- Efficiency and Agility: 26.1% (in 2016, 26.3%)
- Big Data Analytics: 23.9% (in 2016, 23.7%)

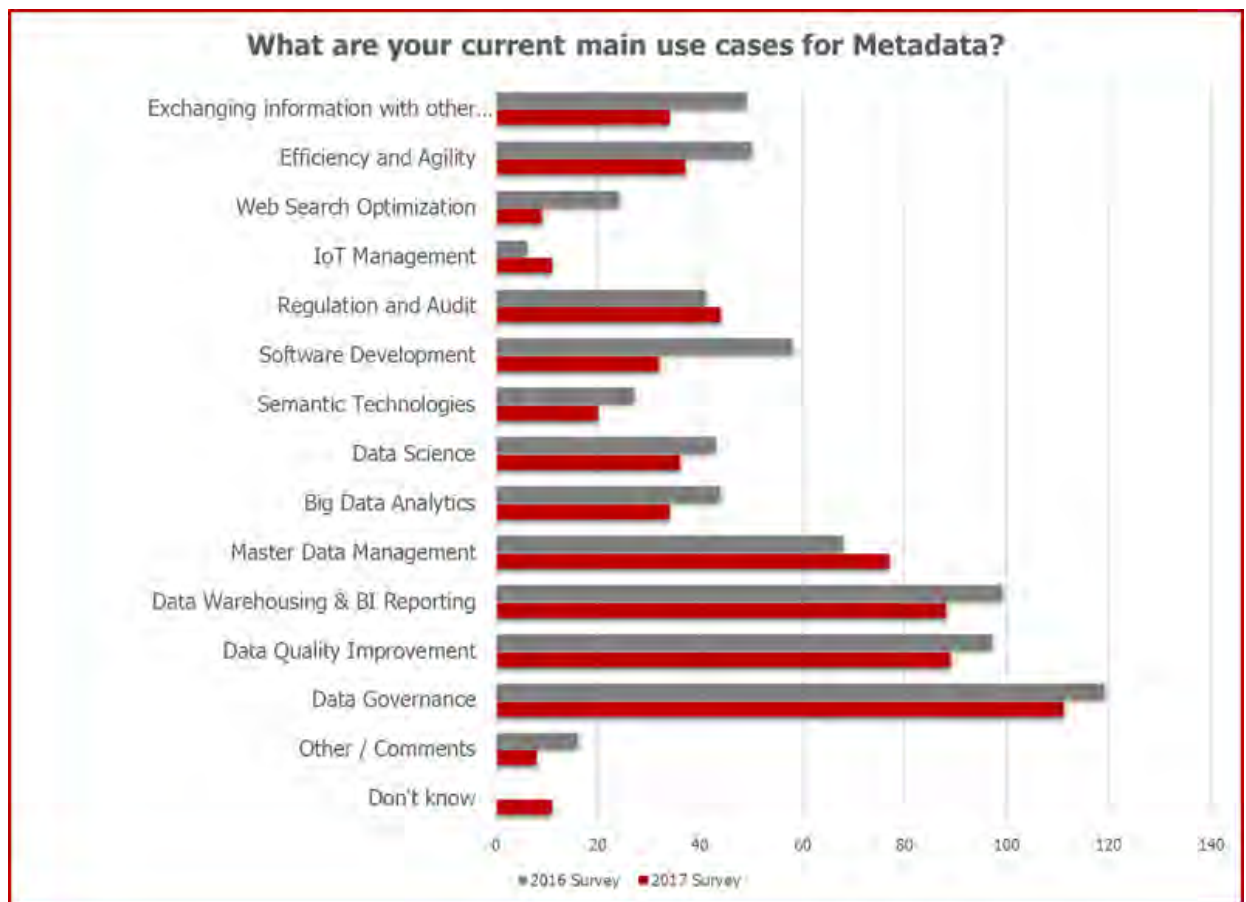


Figure 27

# METADATA

## Conclusion

With the vast array of new technologies that are currently available or on the horizon, it is an exciting time for organizations looking to better leverage their data assets. More than ever, business and technical personnel can work together with data-centric technologies to create new business models, increase efficiencies, and gain strategic advantages.

At the same time, the core foundations of Data Architecture still apply in Metadata, Data Modeling, and Data Governance. Organizations realize that with opportunity comes responsibility and that, if data is seen as an asset, it needs to be managed as one with proper governance, quality, and a foundational architecture behind it.

Collaboration is critical to success, particularly with the abundance of technologies and diverse sets of both business and technical stakeholders involved with data across today's enterprise. Notably, many of the efforts that are most commonly in place (Data Governance, MDM, Metadata Management, Data Modeling, etc.) seek to integrate multiple systems and stakeholders in order to get the most value from data, while at the same time reducing risk. While this is a difficult challenge, it is a very profitable and high-yielding one if done correctly, which highlights the importance of solid Data Architecture in today's rapidly-changing data landscape.

“

***Collaboration is critical to success, particularly with the abundance of technologies and diverse sets of both business and technical stakeholders involved with data across today's enterprise.***

”

## Sponsor



[www.idera.com](http://www.idera.com)

IDERA understands that IT doesn't run on the network – it runs on the data and databases that power your business. That's why we design our products with the database as the nucleus of your IT universe.

Our database lifecycle management solutions allow database and IT professionals to design, monitor and manage data systems with complete confidence, whether in the cloud or on-premises.

We offer a diverse portfolio of free tools and educational resources to help you do more with less while giving you the knowledge to deliver even more than you did yesterday.

Whatever your need, IDERA has a solution.



# DATA ARCHITECTURE SUMMIT

**November 13 – 16, 2017, Chicago, Illinois**

**Use Coupon Code: “ARCHRP” to  
receive 10% off your registration**

**[www.dataarchitecturesummit.com](http://www.dataarchitecturesummit.com)**

## About the Authors



**DONNA BURBANK** is a recognized industry expert in information management with over 20 years of experience helping organizations enrich their business opportunities through data and information. She currently is the Managing Director of [Global Data Strategy Ltd](#), where she assists organizations around the globe in driving value from their data. She has worked with dozens of Fortune 500 companies worldwide in the Americas, Europe, Asia, and Africa and speaks regularly at industry conferences. She has co-authored several books on data management and is a regular contributor

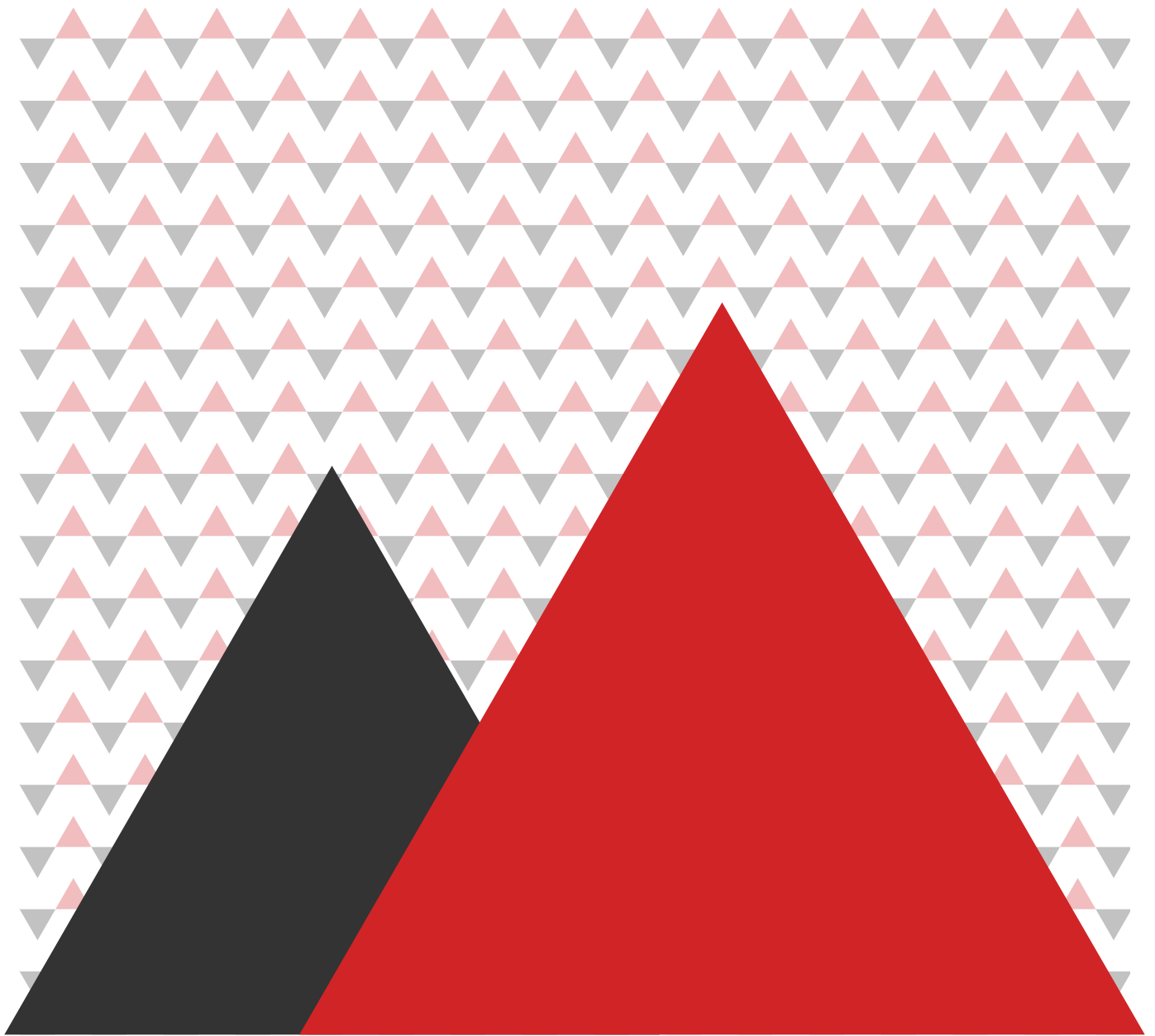
to industry publications. She can be reached at [donna.burbank@globaldatastrategy.com](mailto:donna.burbank@globaldatastrategy.com) and you can follow her on Twitter @donnaburbank.



**CHARLES ROE** has been a professional freelance writer and copy editor for more 15 years, and has been writing for the Data Management industry since 2009. He is the founder of CRWrites.com, his own writing/editing business. Charles has written on a range of industry topics in numerous articles, white papers, and research reports including Data Governance, Big Data, NoSQL technologies, Data Science, Cognitive Computing, Business Intelligence & Analytics, Information Architecture, Data Modeling, Executive Management, Metadata Management, and a host of others. Charles is backed with advanced degrees

in English, History, and a Cambridge degree in Language Instruction. He has worked for almost 20 years as an instructor of English, History, Culture, and Writing at the college level in the USA, Europe, and Turkey. He writes creatively in his spare time.





I D E R A

 DATAVERSITY®