G00311517

2017 Planning Guide for Data and Analytics

Published: 13 October 2016

Analyst(s): John Hagerty

In 2017, analytics will go viral within and outside the enterprise. Technical professionals will need to holistically manage their data and analytics architecture from end to end and leverage cloud wherever appropriate to meet the requirement for "analytics everywhere."

Key Findings

- Data and analytics must drive modern business operations, not just reflect them. Technical professionals must holistically manage an end-to-end data and analytics architecture to acquire, organize, analyze and deliver insights to support that goal.
- Analytics are now infused in places where they never existed before. Demand for delivery of data and analytics at the optimal point of impact will drive innovative machine learning and predictive and prescriptive analytics integration from the core to the edge of the enterprise.
- Data gravity is rapidly shifting to the cloud, with IoT, data providers and cloud-native applications leading the way. It is no longer a question of "if" for using cloud for data and analytics; it's "how."
- Executives will seek strategies to better manage and monetize data for internal and external business ecosystems. This presents data and analytics professionals with the opportunity to assume new roles and enhance skillsets to make these executive dreams a reality.

Recommendations

- Fuse data, analysis and action into a cohesive plan of attack. Design and build a flexible, componentized end-to-end data and analytics architecture to scale to meet needs of a competitive, growing digital business.
- Shift your focus from getting data in and hoping someone uses it to determining how best to get information out to the people and processes that will gain value from it.
- Enable analytics to truly go viral, within and outside the enterprise. Empower more business users to perform analytics by fostering a pragmatic approach to self-service and by embedding analytic capabilities at the point of data ingestion within interactions and processes.

- Incorporate the cloud as a core element of current and future data and analytics architecture. Develop a cloud-first strategy for data and analytics, but be prepared to mix and match multiple services and to blend cloud and on-premises elements in a "hybrid" approach.
- Embrace new roles driven by rising business demand for analytics. Develop both technical and professional effectiveness skills to support the end-to-end architecture vision.

Table of Contents

List of Tables

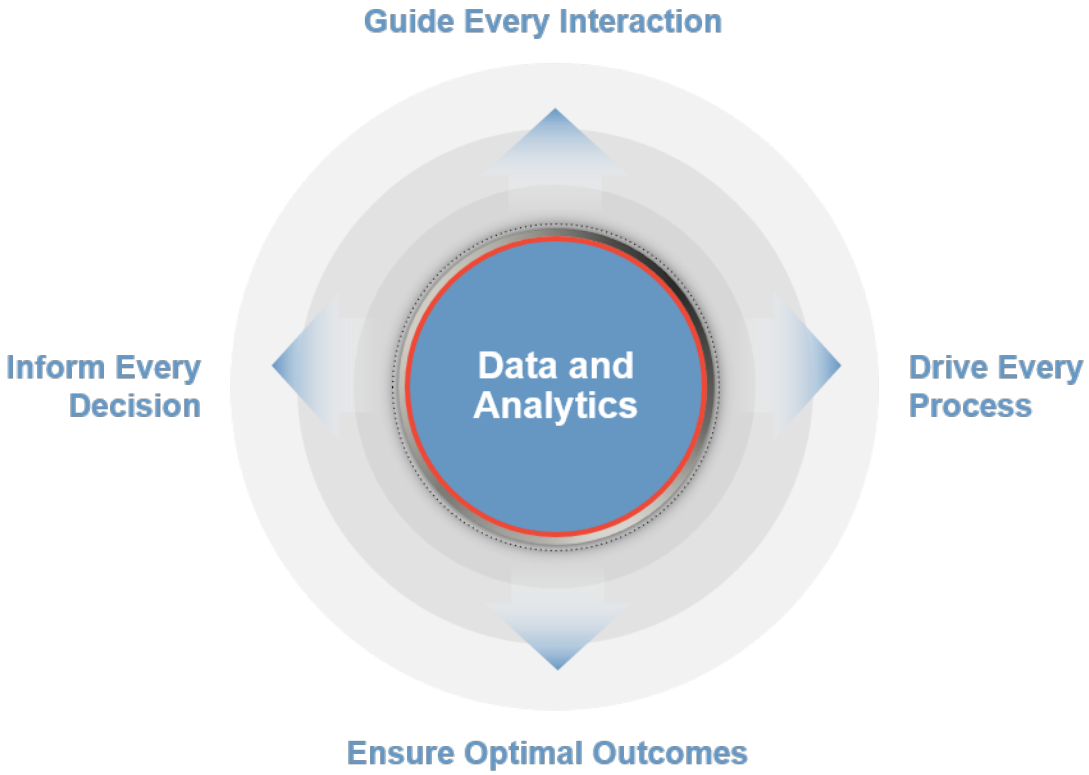
List of Figures

Data and Analytics Trends

Many organizations claim that their business decisions are data-driven. But they often use the term "data-driven" to mean reporting key performance metrics based on historical data — and using analysis of these metrics to support and justify business decisions that will, hopefully, lead to desired business outcomes. While this a good start, it is no longer enough.

Data is the raw material for any decision, and that data comes from both within and outside the enterprise. It exists everywhere: at rest, in motion, on-premises and in the cloud. Data volume, variety and velocity is ever-increasing. To capitalize on opportunities that can be identified, data and analytics are taking on a more active and dynamic role in powering the activities of the entire organization, not just reflecting where it's been (see Figure 1).

Figure 1. Data and Analytics Enable Everything in the Enterprise



Source: Gartner (October 2016)

Beyond data and analytics' traditional role in supporting decision making, they are increasingly being infused in places they haven't existed before. Today, data and analytics are:

- Shaping and molding external and internal customer experiences, based on predicted preferences for how each individual and group wants to interact with the organization.

- Driving business processes, not only by recommending the next best action but also by triggering those actions automatically.

In short, data and analytics are the brain of the enterprise — becoming proactive as well as reactive, and coordinating a host of decisions, interactions and processes in support of business and IT outcomes.

> Data and analytics are at the center of every competitive business.

To enable their organizations to achieve those optimal outcomes, technical professionals must manage the end-to-end data and analytics process holistically. In Gartner's "2016 Planning Guide for Data Management and Analytics," we recommended that organizations deploy a logical data warehouse (LDW) to dynamically connect relevant data across heterogeneous platforms, rather than collecting all data in a monolithic warehouse. We also stressed the business benefits that could be achieved by applying advanced analytics to these vast sources of data — and by providing business users with more self-service data access and analysis capabilities. In 2017, we expect these trends to progress to the next level:

- Distributed data and analytics will demand a comprehensive end-to-end architecture.
- Analytics will go viral, both within and outside the enterprise.

In addition, we expect two other trends in 2017 to fundamentally change the IT architectures supporting data and analytics — and to impact the skillsets and roles of the technical professionals who support these architectures:

- The cloud will be an indispensable platform for data and analytics workloads.
- Executive demands to share data across business ecosystems will drive new roles and skills for technical professionals.

In 2015, technical professionals focused on understanding the data management and analytics options available to them, and defining what their new world could look like. In 2016, for many of these professionals, the focus shifted toward making it all become real. A Gartner survey of nearly 950 IT professionals conducted in early 2016 indicated that 45% of data and analytics projects were already in the "design" and "select" phases — far surpassing the amount that were still in the "plan" and "assess" phases.[1] Now that technical professionals understand the "what" and the "why," it is time to address the "how."

In 2017, forward-thinking IT organizations will realize that rigid data architectures will not scale to meet the needs of a competitive, growing digital business. Technical professionals must utilize the lower-cost, flexible, componentized data management and analytics platforms that are emerging to enable their enterprises to be successful into the digital business era.

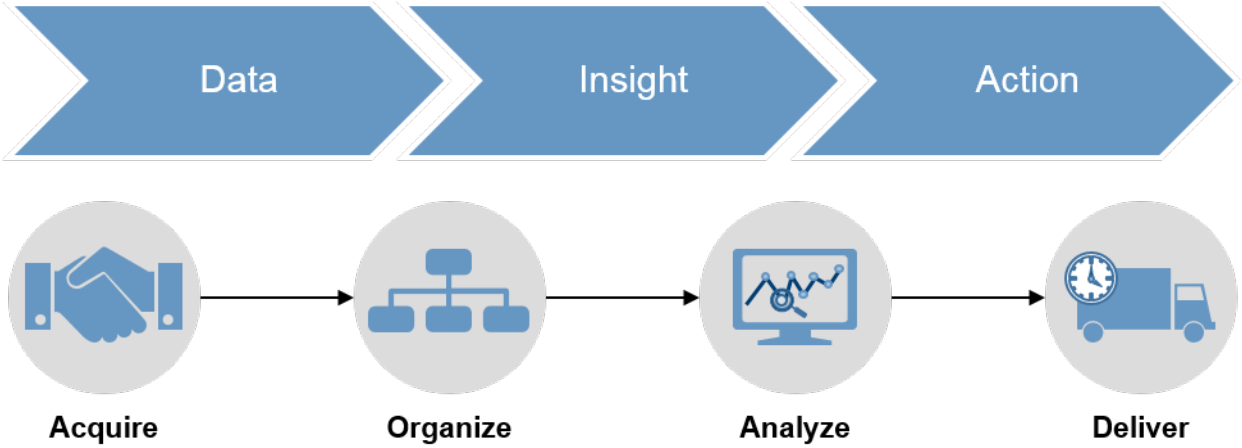
Distributed Data and Analytics Will Demand a Comprehensive End-to-End Architecture

In today's world, the volume, variety and velocity of data is overwhelming. Inexpensive computing at the edge of the enterprise enables a huge amount of information to be captured. Be it video from closed circuit cameras, temperature data from an Internet of Things (IoT) solution or RFID packets indicating product locations in a warehouse, the variety of data is mind-boggling. When you add in the diverse sources of external, often cloud-based data now being used to enrich customer, prospect and partner understanding, a tipping point is quickly reached where the gravity of data skews toward external, rather than internal, data.

These changes will force IT to envision a revitalized data and analytics continuum that incorporates diverse data and can deliver "analytics everywhere" (see Figure 2). While some enterprises are doggedly capturing all data in hopes of uncovering some new insights and spurring possible actions, others are starting with the end goals in mind to streamline the process and holistically manage an end-to-end architecture to support those desired outcomes. Regardless of approach, data, insight and action can no longer represent separate disciplines; they must be fused into one architecture that encompasses:

- **Data acquisition,** regardless of where the information is generated
- **Organization** of that data, using a LDW at the core to connect to data as needed, rather than collect it all in a single source
- **Analysis** of data when and where it makes most sense — including reporting and data visualization, machine learning and everything in between
- **Delivery** of insights and data at the optimal point of impact, whether to support human activities with just-in-time insights, embed analysis into business processes, or feed algorithms that analyze data as it streams into the enterprise and automatically take action on the results

Figure 2. The Revitalized Data and Analytics Continuum



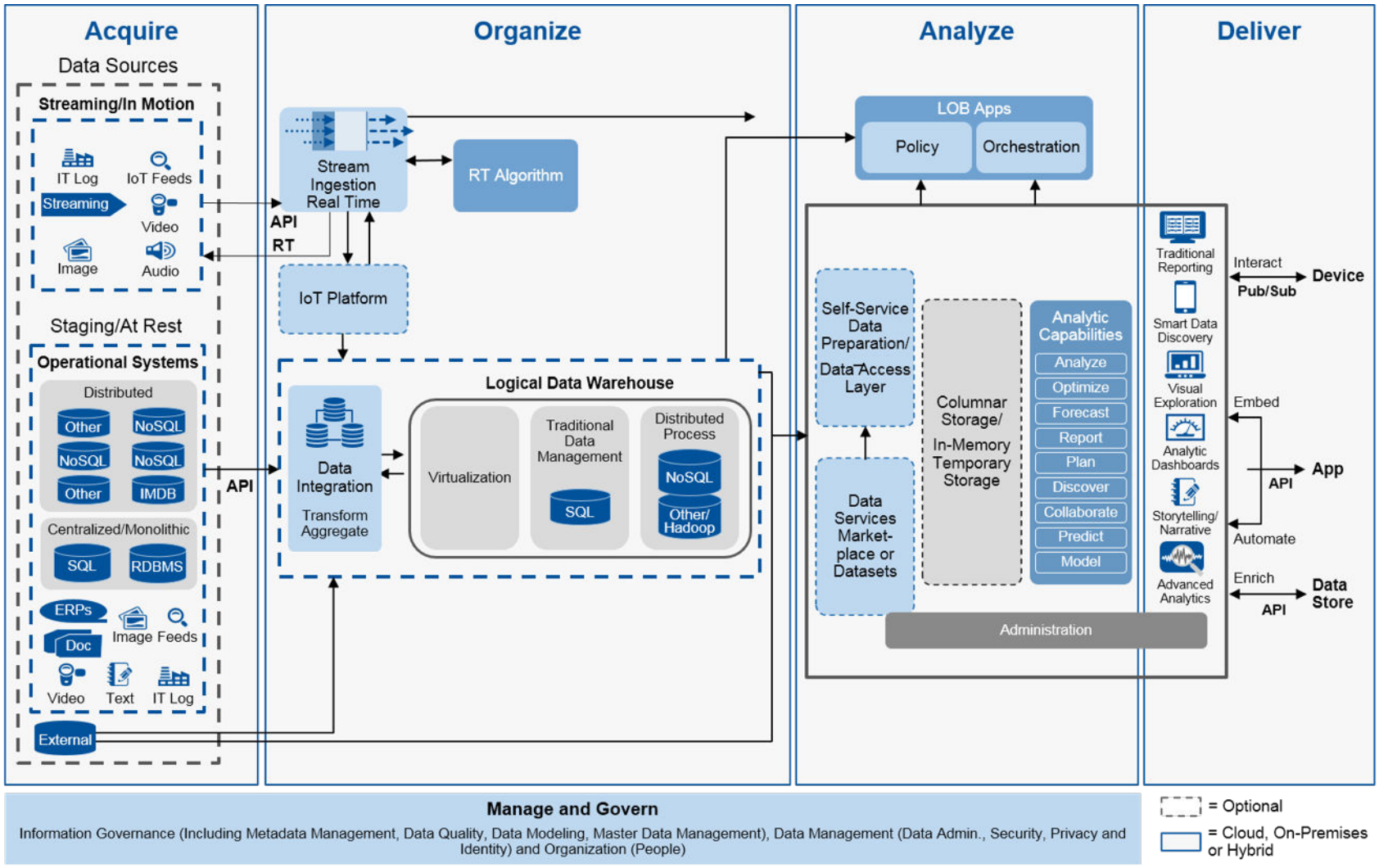
Source: Gartner (October 2016)

This doesn't mean that organizations should immediately discard their traditional data and analytics techniques and approaches and replace them with new ones. The shift will be gradual and incremental — but also inevitable. The key is that data and analytics must drive modern business operations, not just reflect them.

Planning Considerations

In 2017, technical professionals must build a data management and analytics architecture that can support changing and varied data and analysis needs — one that can accommodate not only traditional data analysis, but also newer, advanced analytics techniques. This architecture should be modular by design, to accommodate mix-and-match configuration options as they arise. Figure 3 shows Gartner's four-stage model for an end-to-end data and analytics architecture.

Figure 3. A Comprehensive, End-to-End Data and Analytics Architecture



LOB = line of business; RDBMS = relational database management system; RT = real time

Source: Gartner (October 2016)

Extend the Data Architecture to Acquire Streaming and Cloud-Born External Data

The "Acquire" stage (see Figure 3) embraces all data, regardless of volume, source, speed and type, providing the raw materials needed to enable downstream business processes and analytic activities. For example, the emergence of IoT requires data and analytics professionals to proactively manage, integrate and analyze real-time data. Internal log data often must be inspected in real time to protect against unauthorized intrusion, or to ensure the health of the technology backbone. Strategic IT involvement in sensor and log data management on the technology edge of the organization will bring many benefits, including increased value as such data is used to enhance analytics and improve operations.

In doing so, organizations must shift their traditional focus from getting the data *in* and hoping someone uses it to determining how best to get information *out* to the people and processes that will gain value from it. The sheer volume of data can clog data repositories if technical professionals subscribe to a "store everything" philosophy. For example, machine-learning algorithms can assess incoming streaming data at the edge and decide whether to store, summarize or discard it. When deciding whether and when data will be stored, holistic thinking about *how* the data will be used is another key aspect of the "end-to-end" thinking required.

Above and beyond streaming data, there is so much value-added content available from third parties that organizations are often challenged to find, select and leverage it. Syndicated data comes in a variety of forms, from a variety of sources. Examples include:

- Consumer data from marketing and credit agencies
- Geolocation data for population and traffic information
- Weather data to enhance predictive algorithms that drive diverse use cases from public safety to retail shopping patterns
- Risk management data for insurance

Tapping into this data already enhances analytic and operational activities. Businesses have been leveraging this type of data for decades, often getting a fee-based periodic feed directly from the data provider. Increasingly, vast quantities of this data are available through cloud services — some fee-based, and some free — to be accessed whenever and wherever it's needed. Enabling the data and analytics architecture to embrace these new forms of data in a more dynamic manner is essential to provide contextual information needed to better support data-driven, digital businesses. For more information on the types of data available, see "Understand the Data Brokerage Market Before Choosing a Provider."

Develop a Virtualized Data Organization Layer to Connect to Data, Not Collect It

The different uses, varieties, velocities and volumes of data demand that IT employ multiple data stores across cloud and on-premises environments. But IT cannot allow these multiple data stores to prevent the business from obtaining actionable intelligence. When using an LDW approach, there is no need to create a specialized infrastructure for unique use cases, such as big data. The LDW provides the flexibility to accommodate an infinite number of use cases using a variety of data

stores. "Big data" is the new normal. It is no longer a separate, siloed, tactical use case; it is simply one of many use cases that can be accommodated in the architecture to enable the digital enterprise.

The core of the "Organize" stage of the end-to-end architecture is the LDW. It is *the* data platform for analytics, as defined in Gartner's "Adopt Logical Data Warehouse Architectural Patterns to Mature Your Data Warehouse." Every data warehouse is an LDW initiative waiting to materialize. An LDW:

- Provides modern, scalable data management architecture that is well-positioned to support the data and analytics needs of the digital enterprise
- Supports an incremental development approach that leverages existing enterprise data warehouse architecture and techniques in the organization
- Establishes a shared data access layer that logically relates data, regardless of source

Building the LDW and the end-to-end analytics architecture will require that technical professionals combine technologies and components to provide a complete solution. It requires a significant amount of data integration and an understanding of data inputs and existing data stores. In addition, the numerous technical choices available for building the LDW can be overwhelming. The key is to choose and integrate the technical combination that is most appropriate for the organization's needs. This work needs to be done by technical professionals who specialize in data integration. Hence, 2017 will see the continued rise of the data architect role. (See "Solution Path for Planning and Implementing the Next-Generation Data Warehouse.")

Many clients still directly access various data sources using point-to-point integration. In such cases, any changes in data sources can have a disruptive impact. Although it's often infeasible to completely stop direct access to data, shared data access can minimize the proliferation of one-off direct access methods. This is especially true for use cases that require data from multiple data sources.

To increase the value of shared data access, organizations should:

- Define a business glossary, and enable traceability from data sources to the delivery/ presentation layer.
- Employ various levels of certification for data integration logic, thus creating a healthy ecosystem that enables self-service data integration and analytics.
- Incrementally build this ecosystem as needed to avoid the failures of past "big bang" approaches.

Although technical professionals can custom-code the shared data access layer, commercial data virtualization tools provide many advantages over a custom approach, including the provision of comprehensive connectors, advanced performance techniques and improved sustainability. Gartner recommends that clients deploy these virtualization tools to create the data virtualization layer on

top of the LDW. Cisco Data Virtualization, Denodo, IBM, Informatica, Information Builders, Oracle and Red Hat are examples of stand-alone data virtualization middleware tools.

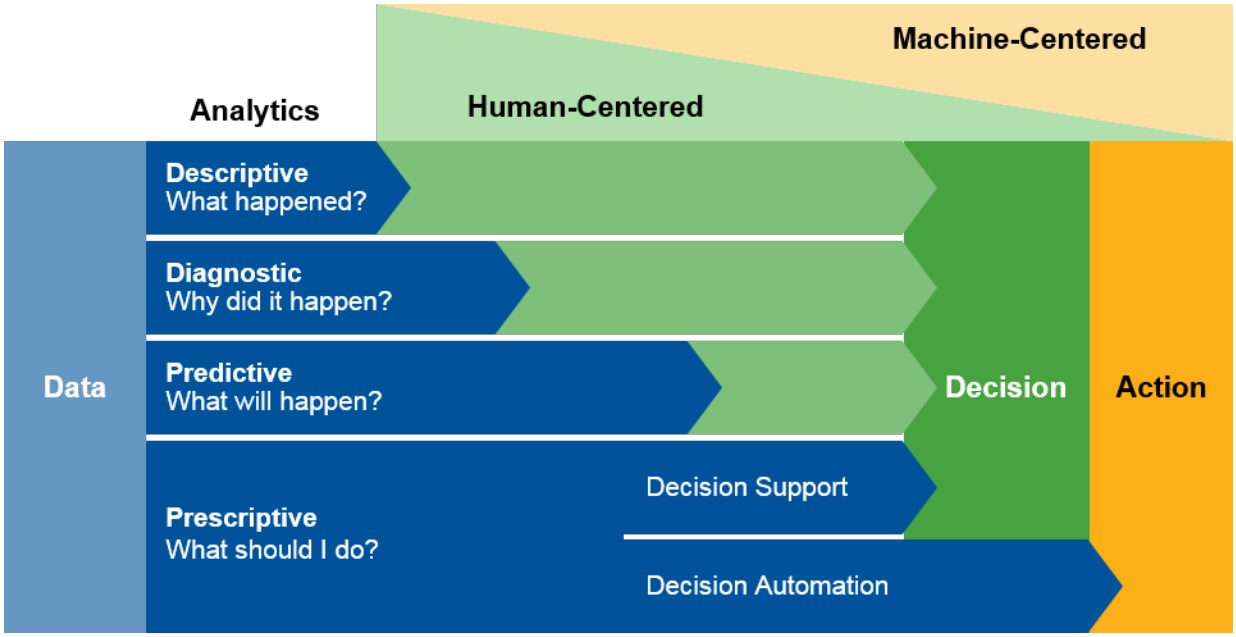
As well, your organization might already have something that can be leveraged; business analytics (BA) tools typically offer embedded functions for data virtualization. However, these are unsuitable as long-term, comprehensive, strategic solutions for providing a data access layer for analytics. They tend to couple the data access layer with specific analytical tools in a way that prevents the integration logic or assets from being leveraged by other tools in the organization.

Develop a Comprehensive Analytics Environment That Spans From Traditional Reporting to Prescriptive Analytics

Comprehensive BA requires more than simply providing tools that support analytics capabilities alone. Multiple components are needed to build out an end-to-end data architecture that encompasses the delivery and presentation of analyses, data ingestion and transformation, data stores, and collaboration on results.

The "Analyze" phase of the end-to-end architecture can be simple for some, but can become increasingly multifaceted as demand for predictions and real-time reactions grows. The range of analytics capabilities available go well beyond traditional data reporting and analysis (see Figure 4). Although Gartner estimates that a vast majority of organizations' analytics efforts (and budgets) are spent on descriptive and diagnostic analytics, a significant chunk of that work is now handled by business users doing their own analysis. This often occurs outside the realm of the sanctioned IT data and analytics architecture. Predictive and prescriptive capabilities, on the other hand, have usually been focused within individual business units and have not been widely leveraged across the organization. That mix must change.

Figure 4. The Four Analytic Capabilities



Source: Gartner (October 2016)

Organizations will need to provide more business and IT institutional support for advanced analytics capabilities. But in the digital business, activities will be interactively guided by data, and processes will be automatically driven by analytics and algorithms. IT organizations must invest in machine learning, data science, artificial intelligence and cognitive computing to automate their business. This will represent a growing percentage of future investment and innovation. But keep in mind that it's not throwing out the old and focusing solely on the new; a fully mature analytic capability set includes a balance of all four types of analytics.

Data and analytics professionals must embrace these advanced capabilities and be prepared to enable and integrate them for maximum impact. Programmatic use of advanced analytics (as opposed to a sandbox approach) is also on the rise, and it must be managed as part of an end-to-end architecture.

Deliver Data and Analytics at the Optimal Point of Impact

The "Deliver" phase of the end-to-end data and analytics architecture (see Figure 3) is often forgotten. For years, this activity has been equated with producing a report, interacting with a visualization or exploring a dataset. But those actions only involve human-to-data interfaces and are managed by BA products and services. Analytics' future will increasingly be partly human-interaction-based, and partly machine-driven. Gartner refers to this mixing of machine-based and human-based capabilities as "augmented intelligence."

Increasingly, key considerations in the delivery of analyzed information will include devices and gateways, applications, processes, or data stores:

- **Devices and gateways:** Users can subscribe to content for delivery to the mobile device of their choice, such as a tablet or a smartphone. Having access to the right information, in the optimal form factor, increases adoption and value. For example, retail district managers may need to access information about store performance and customer demographics while they are in the field, without having to open a laptop, connect to a network and retrieve analysis.
- **Applications:** In-context analytics can be embedded within an application to enrich users' experiences with just-in-time information to support their activities. This could enable a service technician, for example, to view a snapshot of a customer's past service engagements and repairs while diagnosing the cause of a problem. Applications can also be automated using predictions generated by analytics processes running behind the scenes. For example, medical equipment diagnostics can be assessed using IoT in near-real time to determine whether maintenance should be performed on a given machine before it fails.
- **Processes:** The output of an analytic activity — be it in real time or in aggregate — can recommend the next step to take. That result, coupled with rules as to what to do when specific conditions are met, can automate an operational process. For example, if a sensor in a refrigerated storage area of a warehouse indicates that temperature is on the rise, analytics can determine if this is cause for concern, and then dispatch a repairman to the site for immediate inspection and possible repair.

- **Data stores:** Analytics is often used to generate even more data for use in other analytic activities. The output of one activity is input to another. This becomes more critical if the organization wants to monetize its data to external audiences. Insights generated by acquire-organize-analyze activities are output to another data store for eventual access by third parties who need that data to support their decisions and actions. The emergence of connected business ecosystems will drive even more of these changes.

The range of analytic options must now be integrated into the fabric of how you work. We more fully address the "how" of this planning in the next section.

Analytics Will Go Viral, Within and Outside the Enterprise

In many organizations, there is an emerging mandate: Everything must be data-driven. Decisions should no longer be left to gut instinct. Instead, decisions and actions should be based in facts, and those facts also fuel algorithms that predict optimal outcomes. Although it's taken a while to take root, leading enterprises are finally embracing this perspective.

In short, analytics are going viral. More people want to engage with data, and more interactions and processes need analytics to automate and scale. Use cases are exploding in the core of the business, on the edges of the enterprise and beyond. And this trend goes beyond traditional analytics, such as data visualization and reports. Analytics services and algorithms will be activated whenever and wherever they are needed. Whether in support of the next big strategic move or to optimize millions of transactions and interactions a bit at a time, analytics and the data that powers them are showing up in places where they rarely existed before. This is adding a whole new dimension to the concept of "analytics everywhere."

Not long ago, IT systems' main purpose was to automate processes. Data was stored, then analyzed, often as an afterthought, to assess what had already happened. That passive approach has given way to a more proactive, engaged model, where systems are architected and built around analytics, which are rapidly becoming part of every IT system. Today, analytics are:

- Embedded within applications (IoT, mobile and web) to assess data dynamically and/or enrich the application experience
- Just-in-time, personalizing the user experience in the context of what's occurring in the moment
- Running silently behind the scenes and orchestrating processes for efficiency and profitability

Massive quantities of data at rest have fueled innovative use cases for analytics. Add in data in motion —including sensor data streaming within IoT solutions — and you expand scenarios for employing machine learning and artificial intelligence, in real time, to assess, scrub and collect the most useful and meaningful information and insights.

That doesn't mean that traditional analytics activities should be de-emphasized. Business demand for self-service data preparation and analytics continues to accelerate, and IT should enable these capabilities. As data and analytics expand to incorporate ecosystem partners, this demand will also increase from outside the organization.

Planning Considerations

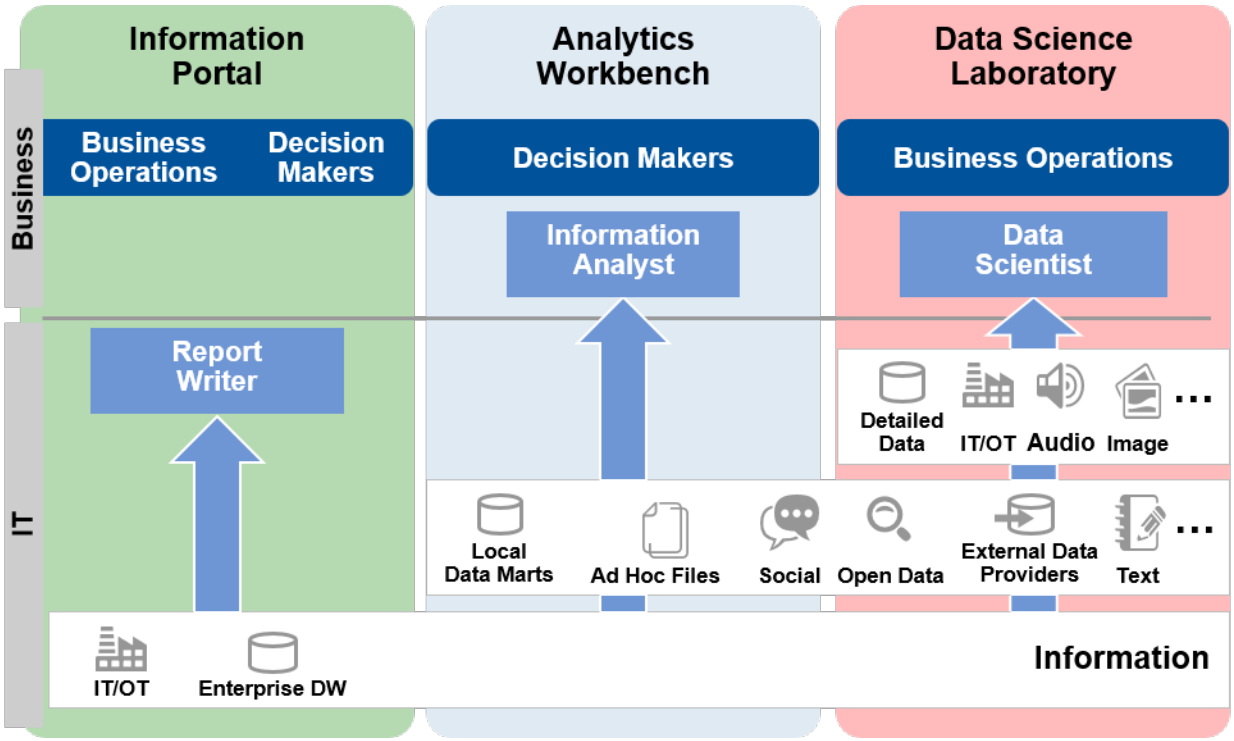
In 2017, technical professionals can expect even more emphasis on analytics as they catch fire throughout the enterprise. The expansion from human-centered interaction to machine-driven automation will have a profound impact on how analytics will be deployed.

Build a Managed Self-Service Environment to Prevent Chaos

Self-service analytics and data preparation enables business users to be self-sufficient, and it gives them the flexibility to iteratively develop their own analytics in a timely fashion. Many businesses organizations have decided that they cannot wait for IT to deliver the data and intelligence they need. They have forged ahead with their own initiatives instead — a situation that has led to "shadow analytics" stacks and to a certain degree of anarchy. To avoid this, technical professionals have a critical role to play. They must establish the infrastructure and environment to drive as much analytical capability as possible into the business, and facilitate a self-service data and analytics approach.

Although most diagnostic and some descriptive analytics are self-service-based, we are still a long way from self-service prediction and prescription. With these increased capabilities comes increased responsibility for users. For these users, technical professionals should establish an environment and processes that facilitate a self-service approach. Gartner recommends building a three-tiered architecture (see Figure 5) to accommodate the four analytical capabilities — descriptive, diagnostic, predictive and prescriptive.

Figure 5. Tiered Business Analytics Environment



DW = data warehouse; IT/OT = information technology/operational technology

Source: Gartner (October 2016)

The three tiers in this model are:

- *The information portal,* an environment similar to a traditional business intelligence (BI) environment. It includes trusted, structured sources for repeatable, relatively slow and expensive descriptive reporting processes.
- The *analytics workbench,* which provides an agile, flexible analytics environment. This environment is easy to use in an exploratory, autonomous way to generate the quick insights required of a diagnostic approach.
- *The data science laboratory,* which caters to advanced analytics (predictive and prescriptive), for heuristic analyses that are often detailed, complex and unique. The process can be somewhat slow and laborious, but can ultimately result in high-impact results.

Incorporate EIM and Governance for Internal and External Use Cases

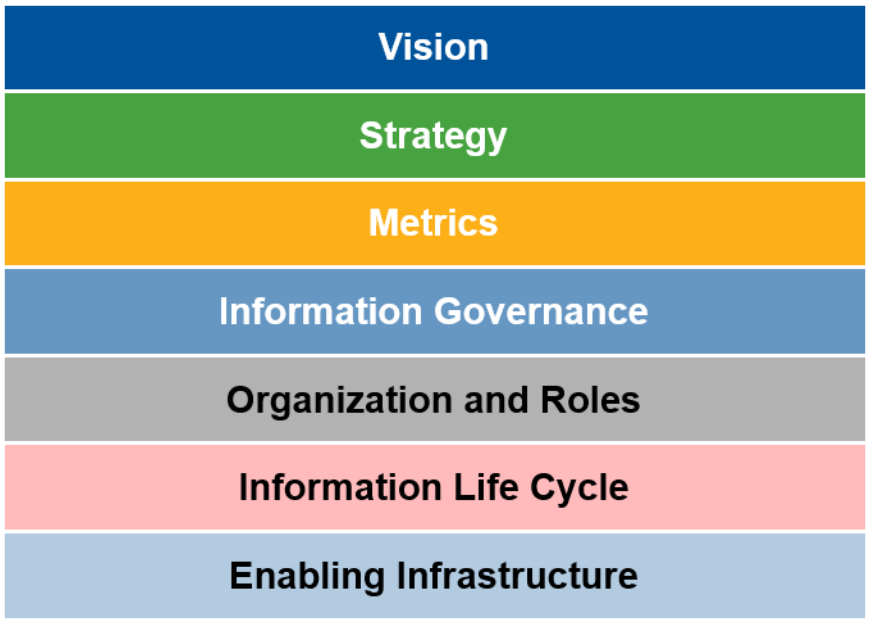
Gartner defines enterprise information management (EIM) as an integrative discipline for structuring, describing and governing information assets — regardless of organizational and technological boundaries — to improve operational efficiency, promote transparency and enable business insight. As more and more data sources for analytics reside outside of the analytics group's or the

organization's control, it becomes even more important to assert just enough governance over all sources of analytic data to enable new and existing use cases.

An EIM program based on sound information governance principles is an effective tool for managing and controlling the ever-increasing volume, velocity and variety of enterprise data to improve business outcomes. In the digital economy, EIM is a necessity, but it remains a struggle to design and implement enterprisewide EIM and information governance programs that yield tangible results. In 2017, a key question for many technical professionals and their business counterparts will be, "How do we successfully set up EIM and information governance?"

Most successful EIM programs start with one or more initial areas of focus, such as master data management (MDM), data quality, data integration or metadata management initiatives. All EIM efforts need to include the same, proven seven components for effective program management depicted in Figure 6. See "EIM 1.0: Setting Up Enterprise Information Management and Governance" for complete guidance on this topic.

Figure 6. The Seven Building Blocks of EIM



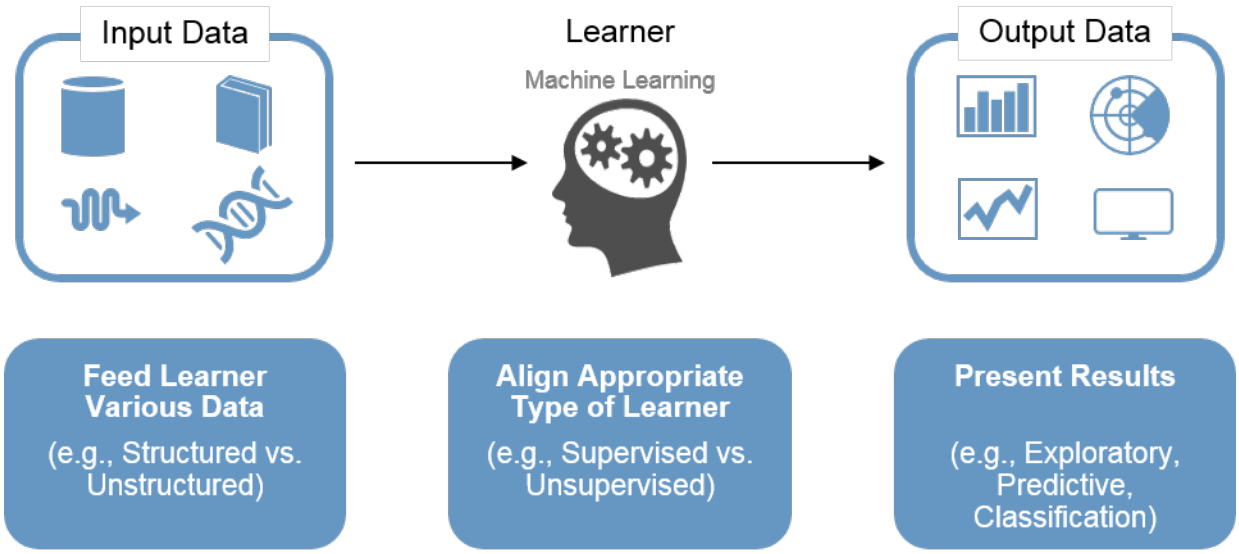
Source: Gartner (October 2016)

Prepare for the Machine-Learning Onslaught

For most data and analytics technical professionals today, advanced analytics and machine-learning techniques are a mystery. But demands from business — fueled by the immense volume, variety and velocity of data now available — mean that machine learning and algorithms must soon become part of the knowledge base of these professionals.

The machine-learning concept is simple: Algorithms learn from data without being explicitly programmed. Machine-learning techniques are based on statistics and mathematics, which are rarely part of traditional data analysis. Any type of data is input, learning occurs and results are output. In supervised learning, known sample outcomes are used for training to achieve desired results. Unsupervised learning relies on machine-learning algorithms to determine the answers (see Figure 7).

Figure 7. The Basics of Machine-Learning Technology



Source: Gartner (October 2016)

To prepare for this exciting and inevitable future, data and analytics technical professionals should start with the machine-learning basics, and learn by doing:

- **Define a business challenge to solve.** This can be *exploratory* (for example, determining what factors contribute to a consumer's default on a bank loan) or *predictive* (for example, predicting when the next natural gas leak will occur and what factors will drive the next failure). Start small, and build in stages. Don't "boil the ocean" on your initial attempts. Evolve your approach over time.
- **Partner with the data science team.** Work with this team to deliver a data and processing environment for the data needed to address the defined business challenge. Enable a platform that will scale to execute the required models and algorithms. This environment might be cloud-based.
- **Get trained now.** Before you act, you must learn. Several online courses offer good basic knowledge on the mechanics of machine learning. Two examples worth reviewing are "Coursera: Machine Learning" and "Udacity: Intro to Machine Learning." Leading consultants such as Deloitte, IBM Global Business Services and Accenture can work with your teams to get these activities off to a strong start.

Enhance Application Integration Skills to Embed Analytics Everywhere

Data and analytics systems are often architected and developed in parallel with systems that capture and process data, and while these systems are logically connected, they are physically separated. In order for data and analytics to be delivered at the optimal point of impact, monolithic analytics systems must be architected and decomposed into callable services so they can be integrated wherever they are needed.

In a mix-and-match world, components must be architected in a more modular way, using features such as:

- Standard data model and transport protocols to locate and retrieve the right data, be it on-premises or in the cloud
- Machine-learning algorithms that can be developed in the "R" environment then executed within a Python program or another analytics tool
- Visualization widgets (for example, components offered by d3.js) that deliver information in the optimal format based on the calling device (web or mobile)
- Data services to deliver raw data to analytic processes via a RESTful APIs

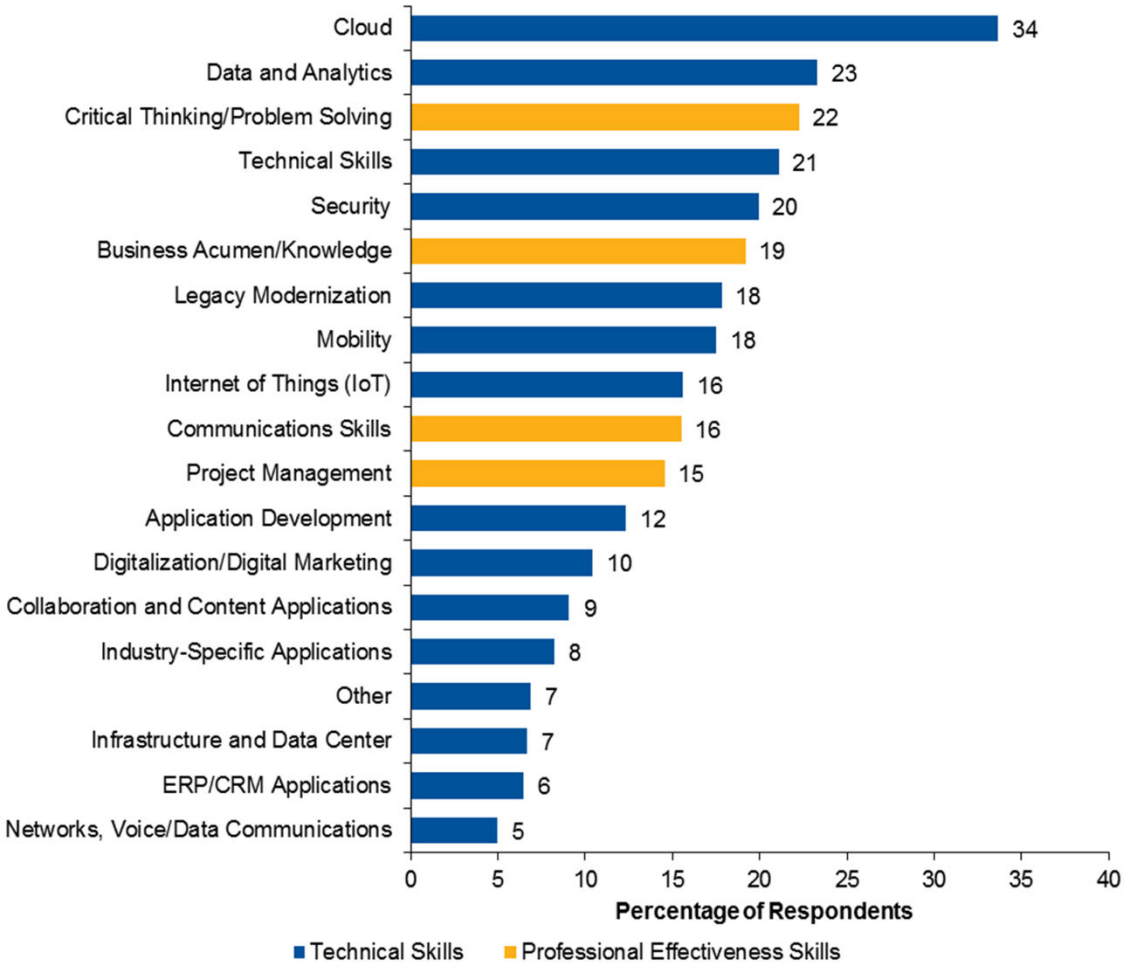
Whether you integrate using a commercial BI and analytics platform or an open-source option, pay particular attention to the provider's API granularity. The finer-grained the services are, the more flexibility you will have.

The Cloud Will Be an Indispensable Platform for Data and Analytics Workloads

Over the past three years, Gartner has seen a steady increase in the adoption of, and inquiries about, cloud computing for data storage — both for operational and analytic data. Much of this interest and adoption can be attributed to cloud-native applications such as salesforce and Workday, emerging IoT platforms, and externally generated data born in the cloud. However, an increasing number of organizations are making a strategic push to incorporate the cloud into *all* aspects of their IT compute and storage infrastructure.

The scale and capacity of the public cloud — coupled with increasing business demand to gather as much data as possible, from as many different sources as possible — is forcing the cloud into the middle of many data and analytics architectures. The data "center of gravity" is rapidly shifting toward the cloud — and as more data moves to the cloud, analytics is sure to follow. Reflecting this trend, both the cloud and analytics are front and center in the minds of architects and technology professionals. In a recent Gartner survey of nearly 950 IT professionals (see Note 1), respondents identified the cloud, followed by data and analytics, as the biggest talent gaps they need to fill (see Figure 8).

Figure 8. Top Skill Gaps Identified by Technical Professionals



Bars of the same value may vary in length due to rounding.
Survey question: What are the three biggest talent gaps related to information, technology or digital business that your organization is trying to fill at the moment?
n = 949 Gartner for Technical Professionals seatholders

Source: Gartner (October 2016)

Cloud is already fundamentally impacting the end-to-end architecture for data and analytics (see Figure 3). Technology related to each stage of the data and analytics continuum — acquire, organize, analyze and deliver — can be deployed in the cloud or on-premises. Data and analytics can also be deployed using "hybrid" combinations of both cloud and on-premises technologies and data stores.

In fact, Gartner expects such hybrid IT approaches and deployments to be a reality of most IT environments in 2017 and beyond. Even with rapid adoption of cloud databases, integration services and analytics tools, enterprises will have to maintain traditional, on-premises databases. The key to success will be to manage all of the integrations and interdependencies, while adopting cloud databases to deliver new capabilities for the business. While this makes for a potentially

complex architecture in the near term, we expect to see data and analytics continue their inexorable march into the cloud.

Planning Considerations

As part of incorporating cloud into every aspect of data and analytics, technical professionals need to focus on long-term objectives, coupled with near-term actions to flesh out the right approach for their organization.

Start Developing a Cloud-First Strategy for Data, Followed by Analytics

Public cloud services, such as Amazon Web Services (AWS), Microsoft Azure and IBM Cloud, are innovation juggernauts that offer highly operating-cost-competitive alternatives to traditional, on-premises hosting environments. Cloud databases are now essential for emerging digital business use cases, next-generation applications and initiatives such as IoT. Gartner recommends that enterprises make cloud databases the preferred deployment model for all new business processes, workloads and applications. As such, architects and tech professionals should start building a cloud-first data strategy now, if they haven't done so already.

This team should also develop a strategy for how the cloud will be used in analytics deployments. Data gravity, latency and governance are the major determinants that will influence when to consider deploying analytics to the cloud, and analytic database services for cloud are numerous. For example, if streaming data is processed in the cloud, it makes sense to deploy analytic capabilities there as well. If application data is resident in the cloud, you should strongly consider deploying BI and analytics as close to the data as possible. Additionally, cloud-born data sources from outside the enterprise will take on an increasingly important role in any data and analytics architecture.

Determine the Right Database Service — or Services — for Your Needs

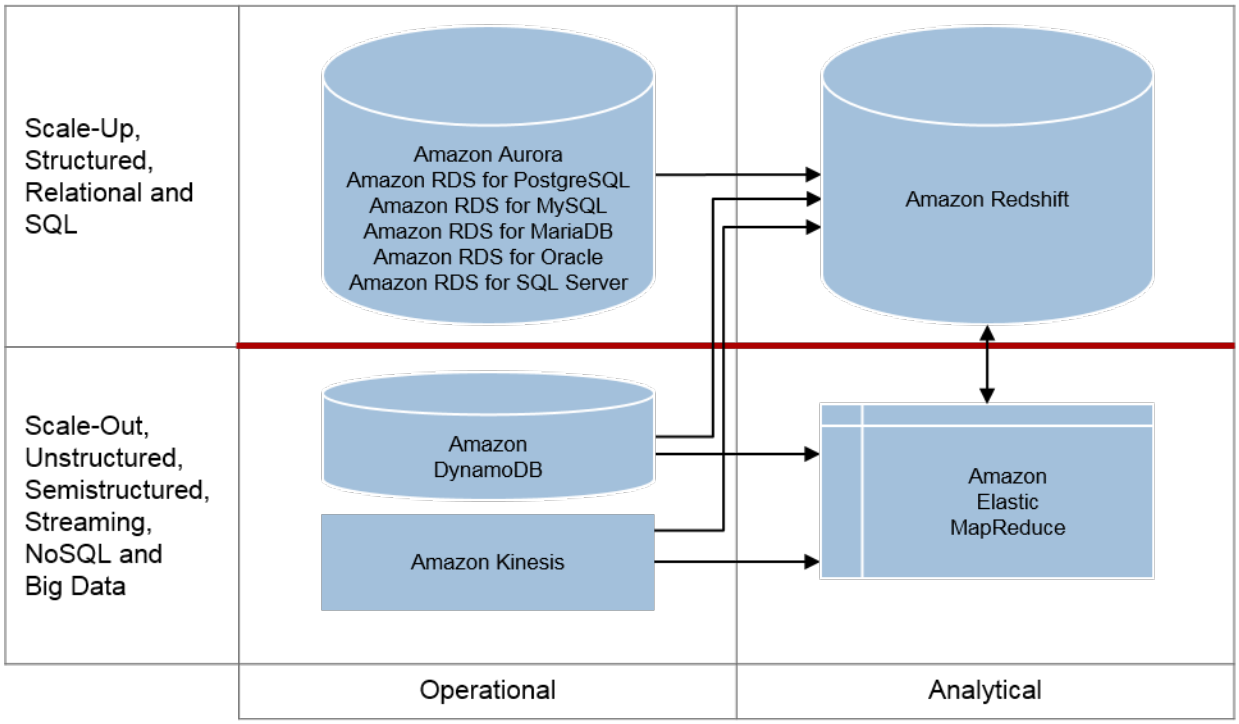
Depending on which cloud service provider you choose, many database options may be available to you. For example, AWS introduced its analytic database service Amazon Redshift in 2014, and Microsoft released Azure SQL Data Warehouse in July 2016. Determining which database service or services to use is a key priority. It is important to understand the ideal usage patterns — as well anti-patterns (i.e., scenarios that aren't recommended) — of each possible option. Matching the right technology to a specific use case is critical to success when using these products. This may lead you to choose different database services for unique workloads.

For example, AWS offers several standard services that are broadly characterized as operational or analytic for structured or unstructured data, as shown in Figure 9. (For more information, see "Evaluating the Cloud Databases from Amazon Web Services.") You may use one service for transaction processing and another for analytics. One service does not have to fit all use cases.

This same model holds true for other cloud providers. For example, Microsoft offers Azure SQL Database for operational needs and Azure SQL Data Warehouse for analytics, among other

offerings. In addition, a database service from an independent vendor can be run in the cloud — either licensed through a marketplace or by bringing your own license.

Figure 9. AWS Databases and Data Flows



RDS = Relational Database Service

Source: Adapted from AWS

Adopt a Use-Case-Driven Approach to Cloud Business Analytics

As mentioned above, data gravity, latency and governance are important factors in determining when and how to deploy BI and analytics in the cloud. But another factor also weighs heavily — the reuse of existing functionality. In fact, this is the No. 1 concern raised in Gartner client inquiries about business analytics.

Historically, many data and analytics technical professionals have been conditioned to "standardize" on as few business analytics tools as possible. Thus, their first inclination is to replicate what's been done on-premises in the cloud. They often seek to "lift and shift" from one computing environment to another, leveraging knowledge and skills in the process.

However, Gartner believes that this is not the right approach for many organizations. Rather, they should take a use-case-driven planning approach to the incremental adoption of cloud analytics — not try to create a singular platform for all BI and analytics out of the gate. This is not an all-or-none approach. Instead, the goal is to gradually transition over time, based on what each organization needs to accomplish. Gartner has identified seven criteria that should be evaluated to help determine whether analytic use cases should be deployed to the cloud (see Table 1).

Table 1. Seven Criteria for Determining a Cloud Analytics Architecture

Criteria	Essential Questions
Data Gravity	Where is the current center of gravity for data?
Data Latency	How fresh does the data need to be?
Governance	How much governance is required based on domains and use cases?
Skills	What skills (tools and platforms) are available in your organization?
Agility	How quickly must new requirements/components be added/updated?
Functionality	Are certain functions only available in the cloud or on-premises?
Reuse	How much existing investment do you want to carry forward from your on-premises analytics platform?

Source: Gartner (October 2016)

Model Cloud Data and Analytics Costs Carefully Based on Anticipated Workloads

The cost model for cloud data and analytics is completely different from on-premises chargeback models. Pricing constructs vary considerably among vendors, with several analytics vendors offering cloud services directly as well as through major marketplaces. Factors such as data volumes, transfer rates, processing power and service uptime will all impact monthly charges. Use-case evaluations should include a goal of avoiding unexpected costs into the future. For more information, see "How to Budget, Track and Reduce Public Cloud Spending."

Executive Demands to Share Data Across Business Ecosystems Will Drive New Roles and Skills for Technical Professionals

Most companies want to wring as much business value as possible from their data. Some organizations have even created roles specifically to fulfill that goal, including:

- The chief data officer (CDO), who can have up to three primary objectives:
  - To manage the organization's information assets
  - To deliver insights to the business to improve decision making
  - To generate incremental business value.[2]
- The chief analytics officer (CAO), who is charged with crafting a business analytics strategy for the enterprise.[3]

Increasingly, these mandates encompass external as well as internal constituencies.

As part of generating incremental business value, CDOs are looking at ways to monetize data — turning internal data assets into external data and analytics products that offer value to partners and suppliers in a connected business ecosystem. For example, an automotive ecosystem might encompass the brand manufacturer, its myriad parts suppliers (and the suppliers' suppliers), financial services partners to provide vehicle financing and insurance companies to tap into driver performance data for variable insurance pricing. The combinations are abundant. These business ecosystems connect virtually all sectors of the global economy — and they are all connected by data and analytics. With CDO and CAO roles leading the data and analytics agenda, architects and technical professionals must work closely with these roles to turn executive vision into a reality that is repeatable, sustainable and effective.

Planning Considerations

In 2017, the development of technical and professional effectiveness skills will be important priorities for technical professionals as they partner with CDOs and CAOs to connect the data and analytics architecture to external and internal product platforms.

Focus on New and Emerging Architecture, Technical and Product Management Roles

New opportunities will open up for technical professionals to play new roles — roles that will help their enterprises exploit data and analytics technologies to improve and transform their businesses. Some may already exist in your organization, such as data architects and analytics architects. These roles will have significant input in designing and developing the end-to-end architecture discussed earlier (see Figure 3), and will become more prominent in 2017.

Data engineers — a role often linked with data science — design, build and integrate data stores from diverse sources to smooth the path for ever-more complex analysis. It is a natural progression from data integration specialist, and will become an essential part of any data science effort that furthers predictive, prescriptive and machine-learning analytics efforts.

We also expect new teams to appear, most likely in the form of transformation teams or centers of excellence. These teams will emphasize refinement, efficiency and ongoing improvement as data and analytics activities work their way into the fabric of the organization's processes and capabilities.

In addition, as more data and analytics services become outward-facing to connect ecosystems and monetize data to external constituents, architects or other technical professional functions may also take on the role of "product manager" — a role that sits at the intersection of business, technology and user experience. Although this is a long-established role in the software vendor and OEM marketplaces, product managers are starting to appear with greater frequency in many other firms. This position occupies a unique role in an organization, with responsibilities that include:

- Researching market needs and customer preferences
- Setting the vision for the product, and selling that vision to the rest of the organization
- Defining and prioritizing the business outcomes required to attain the vision

- Obtaining the resources necessary to build and sustain the product
- Working with development teams to translate the business outcomes into features
- Working across the organization — with stakeholders, users, development teams and operations — to ensure product success
- Working with sales, marketing, ecosystem partners and customers on products aimed at external customers

This last point — working with ecosystem partners and customers — is a key reason why any data products created for external consumption should have a product manager. This role is needed to ensure that the organization delivers the right products to the right markets at the right time. This role is not limited to external data products; the product management discipline is also a great addition for internally facing data and analytics products. See "Moving From Project to Products Requires a Product Manager" for more information.

> Existing roles, such as the project manager or scrum product owner, are neither appropriate nor sufficient for managing a significant product that requires many teams to build.[4]

Dedicate Time to Enhance Technical and Professional Effectiveness Skills

To capitalize on emerging opportunities, it is important to develop a broad range of technical and professional effectiveness skills. Although technical skills are a minimum requirement, effectiveness skills can make or break your success in any project or program you work on. Gartner has long advocated that technical professionals supplement their technical capabilities with additional "soft skills," such as the ability to:

- Better understand business goals and scenarios
- Critically think through problem resolution
- Articulate points of view in the language of the business audience

With the emergence of new, increasingly business-related and customer-facing roles in IT, communication skills and business acumen are even more important than ever.

When Gartner asked nearly 950 technical professionals where they saw skill gaps today, three of the top 10 responses were related to professional effectiveness skills (critical thinking/problem solving, business acumen/knowledge, and communication skills; see Figure 8).[1]

Effectiveness skills without requisite technical prowess are only half a story. The trends outlined in this Planning Guide will require technical professionals to enhance their technical expertise in cloud technology, advanced analytics and machine learning, data virtualization and LDW, streaming ingestion, and integration capabilities to incorporate data and analytics everywhere.

Technical professionals should take the following steps to improve their skillsets:

- Identify the skills you need to improve. It's useful to ask others you work with for their opinions.
- Research whether employee development and technical training programs are in place in your organization. HR often has relevant courses and programs in position. If available, enroll in these courses.
- If no programs are available internally, look to external resources. For communications training, turn to vendors such as Toastmasters International. Explore course work at local universities or online courseware (e.g., Coursera). Depending on cost, determine if your organization will assist you in these efforts.
- Spend time putting what you learn into practice. Participating in training without follow-up reinforcement is not always worth the time, effort and cost. Make this new knowledge part of your new standard operating procedure.
- Take personal responsibility for this improvement. It will not only benefit your company; it will benefit you in any future endeavor.

Setting Priorities

Data and analytics technical professionals must focus on four key areas as they plan and prioritize their activities in 2017:

1. Design and build a comprehensive end-to-end architecture.
2. Enable analytics to truly go viral, within and outside the enterprise.
3. Incorporate the cloud as a core element of both current and future data and analytics architecture.
4. Expand roles and skillsets to deliver data services products for internal and external business ecosystems.

Technical professionals must begin by taking inventory of their existing environments. All of the planning considerations discussed in this report should be approached as part of an evolution to a strategic end state, not as a rip-and-replace strategy. Some actions will move faster than others.

- IT and business must work jointly to design an end-to-end architecture for data and analytics. Technical professionals must start with the business goals in mind and holistically manage an architecture to support those outcomes. The phases — acquire, organize, analyze and deliver — must be planned together, with each feeding off the others. Data, analysis and action can no longer represent separate disciplines; they must be fused into a cohesive plan of attack.
- With more people wanting to engage with data — and more interactions and processes needing analytics to automate and scale — demand for analytics will continue to expand. Use cases are exploding in the core of the business, on the edges of the enterprise and beyond. It's critical to be prepared for more business user enablement by fostering a pragmatic approach to better

self-service coupled with processes to prioritize, facilitate and manage the onslaught. Machine learning is rising quickly, and technical professionals need to understand the concepts, experiment with the technologies and integrate analytics wherever they are needed for optimal impact.

- The cloud needs to become part — or even the centerpiece — of the organization's data and analytics architecture. Developing a cloud-first strategy for data, followed by analytics, is an essential first step. Choosing the right cloud service providers and technologies should follow. With many possible services available, technical professionals may select a mix-and-match approach for data and analytics as they gradually migrate data storage and computing capabilities to the cloud. As well, technical professionals should exploit as much data and analytics services in the cloud as are available through data marketplaces.
- With CDOs striving to increase business value, data and analytics products are being evaluated and designed for internal and external business ecosystem consumption. As internal projects turn into external products, new roles for technical professionals will emerge. Because solid technical and professional effectiveness skills are important components of these architect, engineer and product manager roles, it's important to devote time and effort to improving these capabilities.

Finally, it's critical that data and analytics technical professionals keep the end goal in mind. It is easy to become enamored of new technology choices, but business value must be front and center in every decision. It is important to maintain open channels of communication with constituents — both internal and external — and to explain any technical actions or concepts in terms they can understand, support and champion. This is an exciting time for data and analytics professionals, in which they can play an increasingly critical role in helping the organization achieve business success.

Gartner Recommended Reading

Some documents may not be available as part of your current Gartner subscription.

"Solution Path for Evolving Your Business Analytics Program"

"Solution Path for Planning and Implementing the Next-Generation Data Warehouse"

"Solution Path: Implementing Big Data for Analytics"

"Adopt Logical Data Warehouse Architectural Patterns to Mature Your Data Warehouse"

"Doing 'Just Enough' Master Data Management for Business Analytics"

"Top Skills for IT's Future: Cloud, Analytics, Mobility and Security"

"EIM 1.0: Setting Up Enterprise Information Management and Governance"

"Three Architecture Styles for a Useful Data Lake"

"Comparing Four iPaaS-Based Architectures for Data and App Integration in Public Cloud"

"Comparing Four Hadoop Integration Architectures"

"Comparing Three Self-Service Integration Architectures"

"Evaluating the Cloud Databases From Amazon Web Services"

"Understand the Data Brokerage Market Before Choosing a Provider"

Evidence

[1] "Top Skills for IT's Future: Cloud, Analytics, Mobility and Security"

[2] "Staffing the Office of the CDO"

[3] "The Chief Analytics Officer's Vision Sets the Narrative for the Business Analytics Strategy"

[4] "Moving From Project to Products Requires a Product Manager"

Note 1 2016 Gartner Technical Professionals Survey

Results presented are based on a Gartner study conducted to provide an overview of how technical professionals are dealing with changes related to digital business. The research was conducted online from 7 to 29 January 2016 among 949 respondents in every region of the world. Gartner for Technical Professionals seatholders were invited to participate.

Respondents were required to be a member of their organization's IT staff or department (or serve in an IT function). They could not serve as a member of the board, president, or in an executive-level or IT leadership position. The survey was developed collaboratively by a team of Gartner analysts who follow technical professionals and was reviewed, tested and administered by Gartner's Research Data Analytics team. For more information on the survey demographics and methodology, see "Top Skills for IT's Future: Cloud, Analytics, Mobility and Security."

More on This Topic

This is part of an in-depth collection of research. See the collection:

- 2017 Planning Guide Overview: Architecting a Digital Business With Sensing, Adapting and Scaling

GARTNER HEADQUARTERS

Corporate Headquarters
56 Top Gallant Road
Stamford, CT 06902-7700
USA
+1 203 964 0096

Regional Headquarters
AUSTRALIA
BRAZIL
JAPAN
UNITED KINGDOM

For a complete list of worldwide locations,
visit http://www.gartner.com/technology/about.jsp