2023 Planning Guide for Analytics and Artificial Intelligence

Published 10 October 2022 - ID G00775123 - 51 min read

By Analyst(s): Sumit Agarwal, Joe Antelmi

Initiatives: Analytics and Artificial Intelligence for Technical Professionals; Evolve Technology and Process Capabilities to Support D&A

Innovation through analytics and AI will be vital for enterprise success in 2023. To enable innovation at scale in a distributed environment, data and analytics technical professionals must develop trustworthy solutions and converge processes, models and technology for organizational growth.

Overview

Key Findings

- Enterprise data is proliferating over a very diverse data plane spanning on-premises, public cloud and edge. This diversity is complicating development of consistent analytics and representative AI models.

- Disparate analytics domains are converging to deliver more insights more efficiently. By combining augmented data preparation, DSML and A&BI solutions, technical professionals are harnessing D&A investments, practices and processes to deliver effective decision support while rationalizing service sprawl.

- Obstacles to enabling AI trust, scalability and operationalization are preventing organizations from harnessing AI's business transformation potential. To succeed, AI initiatives require adequate data, transparency within prebuilt ML models, and an XOps framework integrating DataOps and MLOps for operationalization of the right infrastructure.

- Democratization of data opens up an opportunity to engage both technical and subject matter experts in the development of differentiated insights. However, successful adoption of this new organizational model requires upskilling, training investment and clear expectations for the new roles.

Recommendations

Data and analytics technical professionals working to deliver analytics and artificial intelligence (A&AI) initiatives must:

- Strive to build a composable D&A architecture by bringing together A&BI and DSML capabilities that support the continuum of descriptive, diagnostic, predictive and prescriptive analytics on a wide variety of data sources.
- Minimize service disruption and support a governed self-service environment by adopting a federated model. Federation is especially relevant to enabling reuse, diversity and choice with cloud-based analytics services.
- Operationalize analytics architectures and AI models by applying DataOps, MLOps and ModelOps practices and tools to provide repeatability, reusability and resilience.
- Combine responsible AI practices across the entire advanced analytics and ML development cycle. Leverage explainability techniques and frameworks to interpret model outputs, and ensure that they can collectively address fairness, bias mitigation, ethics, risk management, privacy and regulatory compliance.

Analytics and Artificial Intelligence Trends

Download All Graphics in This Material

The 2022 Gartner CIO and Technology Executive Survey highlighted increasing investments in data and analytics (D&A) and artificial intelligence (AI). [1] It asked respondents which technology areas would be receiving the largest amount of new or additional funding in 2022 compared with 2021:

- 51% selected business intelligence (BI) and data analytics
- 30% selected AI and machine learning (ML)

These findings underscore the continued organizational focus on harnessing data to identify business insights.

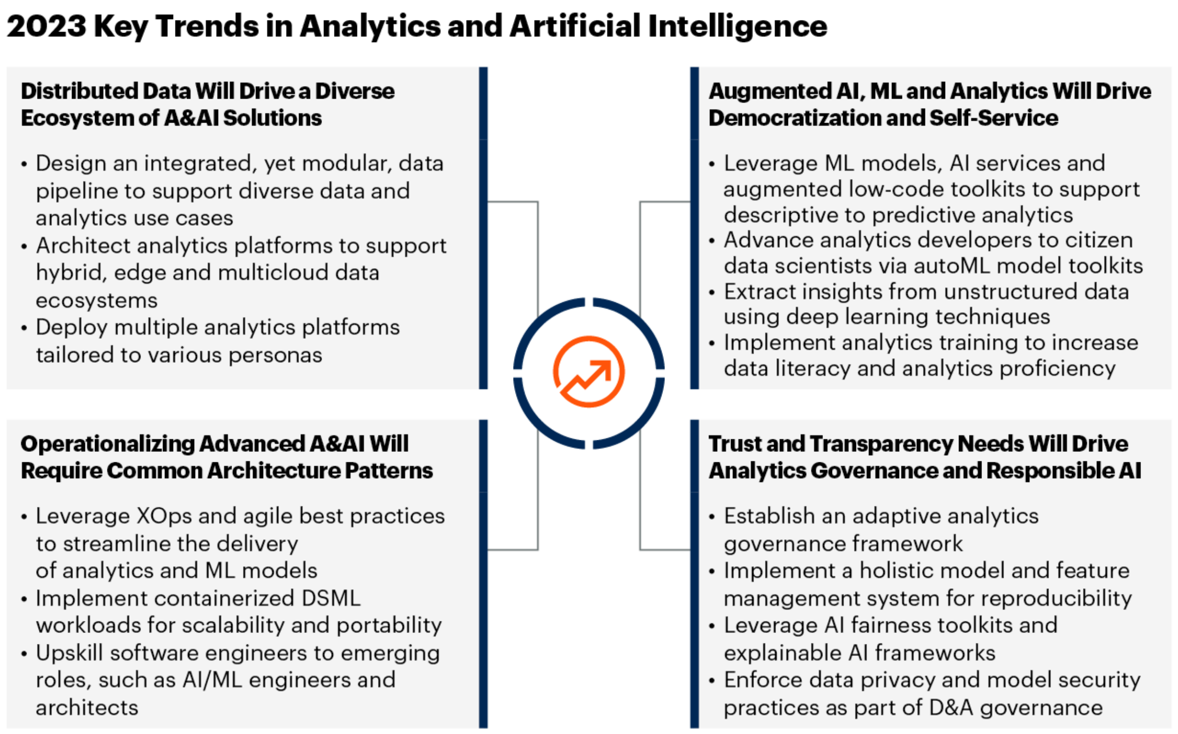
Analytics and AI provide a mix of well-defined methods and transformative technologies. This solution mix provides a massive opportunity for enterprises to use data for competitive advantage at a time of economic and geopolitical shifts and changing employee and customer preferences. Now, it is more important than ever before to:

- Empower employees to analyze data
- Increase agility through cloud adoption
- Develop and scale AI models in a trustworthy way
- Improve productivity and offer differentiated user experiences through AI features

These objectives align with the key trends that organizations must adopt during 2023 (see Figure 1):

- Distributed data will drive a diverse ecosystem of A&AI solutions.
- Augmented AI, ML and analytics will drive democratization and self-service.
- Operationalizing advanced analytics and AI pipelines will require implementation of common architecture patterns.
- Trust and transparency requirements will drive analytics governance and responsible AI practices.

Figure 1: 2023 Key Trends in Analytics and Artificial Intelligence



2023 Key Trends in Analytics and Artificial Intelligence

Distributed Data Will Drive a Diverse Ecosystem of A&AI Solutions

- Design an integrated, yet modular, data pipeline to support diverse data and analytics use cases
- Architect analytics platforms to support hybrid, edge and multicloud data ecosystems
- Deploy multiple analytics platforms tailored to various personas

Augmented AI, ML and Analytics Will Drive Democratization and Self-Service

- Leverage ML models, AI services and augmented low-code toolkits to support descriptive to predictive analytics
- Advance analytics developers to citizen data scientists via autoML model toolkits
- Extract insights from unstructured data using deep learning techniques
- Implement analytics training to increase data literacy and analytics proficiency

Operationalizing Advanced A&AI Will Require Common Architecture Patterns

- Leverage XOps and agile best practices to streamline the delivery of analytics and ML models
- Implement containerized DSML workloads for scalability and portability
- Upskill software engineers to emerging roles, such as AI/ML engineers and architects

Trust and Transparency Needs Will Drive Analytics Governance and Responsible AI

- Establish an adaptive analytics governance framework
- Implement a holistic model and feature management system for reproducibility
- Leverage AI fairness toolkits and explainable AI frameworks
- Enforce data privacy and model security practices as part of D&A governance

Source: Gartner
775123_C

Distributed Data Will Drive a Diverse Ecosystem of Analytics and AI Solutions

Most organizations recognize the need to use data for decision making. To achieve this end, organizations must address the challenge of raising data and analytics maturity. However, data is being generated across the organizational ecosystem — on-premises, in the public cloud and at the edge — making this task even more complex. According to the 2020 Gartner Cloud End-User Buying Behavior Survey, 76% of organizations leverage multiple cloud providers. [2]

> By 2025, more than 50% of enterprise-critical data will be created and processed outside the data center or cloud.[3]

Data and analytics professionals are engaged in several initiatives to effectively manage the increasing volumes and types of data. Organizations have traditionally worked to bring all data into centralized data platforms. These centralized platforms provided a consistent structure to data and strived to deliver a single source of truth. However, they struggled with data latency, rigid structures and lack of agility to accommodate new data types, including text, images, videos and audio.

In the modern world, organizations are deploying data platforms with enhanced support for these required features, as well as support for modularity and interconnectivity. Moreover, distributed data is replacing centralized data. Some decisions are too time-sensitive to rely on data that takes minutes to analyze in the data center or cloud provider. Analytics must occur earlier while the data is in motion, or closer to edge devices, in order to deliver the necessary time to insight. Business semantics must simultaneously become more consistent and more accessible. To meet these requirements, data and analytics technical professionals must:

- Design an integrated, yet modular, data pipeline to support diverse data and analytics use cases
- Architect analytics platforms to support hybrid, edge and multicloud data ecosystems
- Deploy multiple analytics platforms tailored to various personas

Planning Considerations

Design an Integrated, Yet Modular, Data Pipeline to Support Diverse Data and Analytics Use Cases

Organizations need to support increasingly specialized analytics use cases at the edge, in the cloud and on-premises, by bringing in and managing external data, machine data, streaming data and graph data.

Using a single best-of-breed vendor to handle this combination of analytics specialization and data diversity is not currently feasible, as the best products and platforms vary widely for different use cases and clouds. Complicating matters further, a lack of standards, along with intentional vendor lock-in strategies in the analytics market, makes it difficult to assemble a set of analytics services that integrate well together. Additionally, technologists must contend with friction caused by data gravity, cloud-to-cloud data egress charges, and differing standards and integration options between different clouds, software solutions and vendor stacks.

Technical professionals, especially architects, have several implementation choices. Solutions integrated within the platform ecosystem have fewer integration requirements and a lower learning curve. These ease-of-use benefits stem from the platform vendor's continued investments and support to make the various components work together in a more seamless way. However, integrated offerings may not be sufficiently mature in all critical areas. Thus, solutions need the flexibility to fill gaps by incorporating modular, best-of-breed elements.

For example, many organizations struggle to deploy report bursting and paginated reporting in cloud analytics platforms, because of a lack of vendor investment in those features. This scenario provides an advantage to cloud analytics services that can easily integrate with third-party reporting services. An ideal approach balances modularity and integration, taking into account the unique nature of cloud services. Cloud service providers are not market-leading in every service they offer, but they do offer the potential for powerful, cost-optimized, future-oriented architectures. Best-of-breed products offer maturity and support for multiple clouds, on-premises use cases, and domain-specific scenarios that can shore up gaps in the native cloud ecosystem offerings.

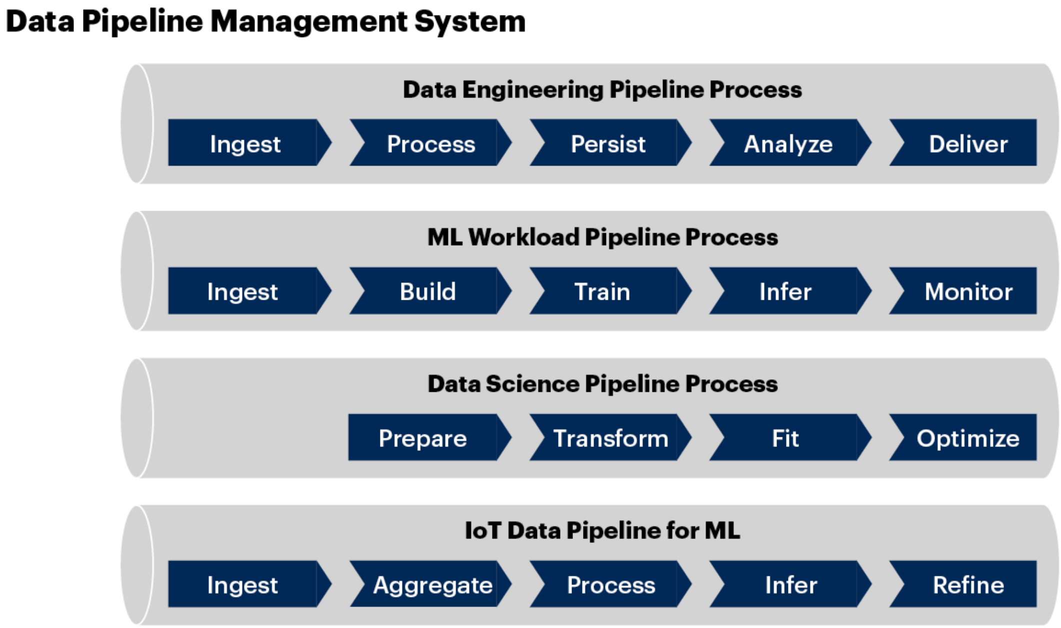
As a result, many organizations are building cloud architectures that both:

- Leverage cloud provider infrastructure services, such as Amazon Simple Storage Service (Amazon S3), Microsoft Azure Data Lake Storage Gen2 or Google Cloud Storage
- Overlay data warehousing services (such as Snowflake), data integration services (such as Informatica) and/or third-party data lake providers

Implementing these architectural practices requires the creation of several different data pipelines on top of the data management systems. A data pipeline is a set of sequential processing elements that transport data from source to target. As shown in Figure 2, four types of data pipelines are commonly used to support analytics and AI:

- Data engineering pipelines
- ML workload pipelines
- Data science pipelines
- Internet of Things (IoT) data flow pipelines

Figure 2: Data Pipeline Management System



Source: Gartner
775123_C

For more information, see:

- Solution Path for Building Modern Analytics and BI Architectures
- Reference Architecture to Enable Self-Service Analytics
- Create a Data Strategy to Ensure Success in Machine Learning Initiatives

Architect Analytics Platforms to Support Hybrid, Edge and Multicloud Data Ecosystems

Traditional D&A platforms are not equipped to meet growing business demands. In addition, the total cost of ownership (TCO) of on-premises D&A solutions continues to grow due to service sprawl and service complexity. These factors are increasing the resources required to support and maintain the D&A environment. At the same time, cloud data and analytics are offering more value and capabilities through new services, simplicity and the agility to handle data modernization. Moreover, the increasing breadth of cloud analytics services is enabling tooling to satisfy new types of analytics demands, such as streaming analytics, specialized data stores and more self-service-friendly tools to support end-to-end deployment.

However, the decision of cloud or on-premises solutions is not binary. In some scenarios, a hybrid cloud solution may be more appropriate, with the data plane extending across on-premises and public cloud environments. In addition, data is increasingly being generated by edge/IoT devices and being collected at the edge server. Transferring the entire video stream from each of the cameras to the cloud would be expensive and would increase latency. An edge server provides the opportunity to apply analytics and related computations closer to the creation of data. It reduces latency and minimizes data privacy risks by applying federated ML techniques.

A data and analytics strategy involves several components across analytics, BI and data science platforms. Each component will require individual assessment and rationalization to determine the implementation approach that meets both short-term and long-term goals. The decision to migrate should be driven by your understanding of the components, characteristics and architecture that best fit the development and operational skills required to build, deploy and operationalize the analytics platform. Hence, you need to:

- **Establish a preliminary architecture and roadmap based on the migration approach:** Defining an approach is an important decision that will drive the strategy needed to deploy the D&A platform and its associated applications in an edge, on-premises, hybrid, cloud or multicloud environment. The migration strategy helps determine the level of effort, based on complexity and the long-term benefit that will be achieved with a new implementation model. Some of the key tasks involved in this first step are:

  - Identifying the cloud deployment model that is best-suited for your organization and the factors that led to that decision

  - Assessing cloud migration and transient architectures

  - Evaluating the decision paths for the migration of various D&A components — data sources, integration, governance, virtualization and analytics consumption

  - Creating a deployment plan and a roadmap for the future-state architecture

- **Evaluate and assess cloud service characteristics:** It is vital to assess the cloud's multiple service characteristics. These include, event streams, data storage, performance, scalability, service and pricing models, metadata management, and compliance capabilities, among others. D&A applications are likely to run more efficiently if you use cloud-based performance optimization tools and architectures designed with proven methodologies. Cost optimization and monitoring tools can help you navigate through the complex vendor pricing models.

- **Identify implementation steps:** Based on the assessment of the various D&A components and cloud service characteristics, it is vital to define steps as part of the migration strategy to a new implementation model. These decisions are usually driven by technical resource availability, the organization's understanding of a distributed analytics workload management system, and the cost to migrate each of the data management and analytics components. The implementation steps include:

  - Establishing a data management framework
  - Identifying data ingestion and integration methods
  - Building data discovery and analytics sandboxes
  - Effectively utilizing tools, such as data catalogs, to assist with governance and metadata management
  - Providing a robust augmented analytical platform as part of the consumption layer

- **Launch, and support postlaunch**: The final step is to operationalize the overall D&A pipeline for oversight and administrative control within the new implementation model. This step comprises platform administration and support; responsible, accountable, consulted and informed (RACI) designation; and change management support. Some common themes that you can adopt to manage your new implementation model effectively include:
  - Roll out analytics services to users in waves via pilots, before moving to full-scale production.
  - Use tools that allow monitoring of analytical workloads, resources and compute costs, and that can adjust resource allocations as necessary.
  - Automate workflows at the platform level. For example, manage hybrid infrastructure, cloud instances, software code migration, workloads and other artifacts at each component level to minimize or eliminate human intervention.
  - Monitor and track individual cloud resources, workloads, service-level engagement, performance and overall user experiences in multicloud deployments.

For more information, see:

- A Guidance Framework for Deploying Data and Analytics in the Cloud
- Identifying Data Access Patterns for Multicloud and Intercloud
- Decision Point for Selecting Cloud Analytics Solution Architecture
- Solution Path for Building a Holistic Data Management and Analytics Architecture

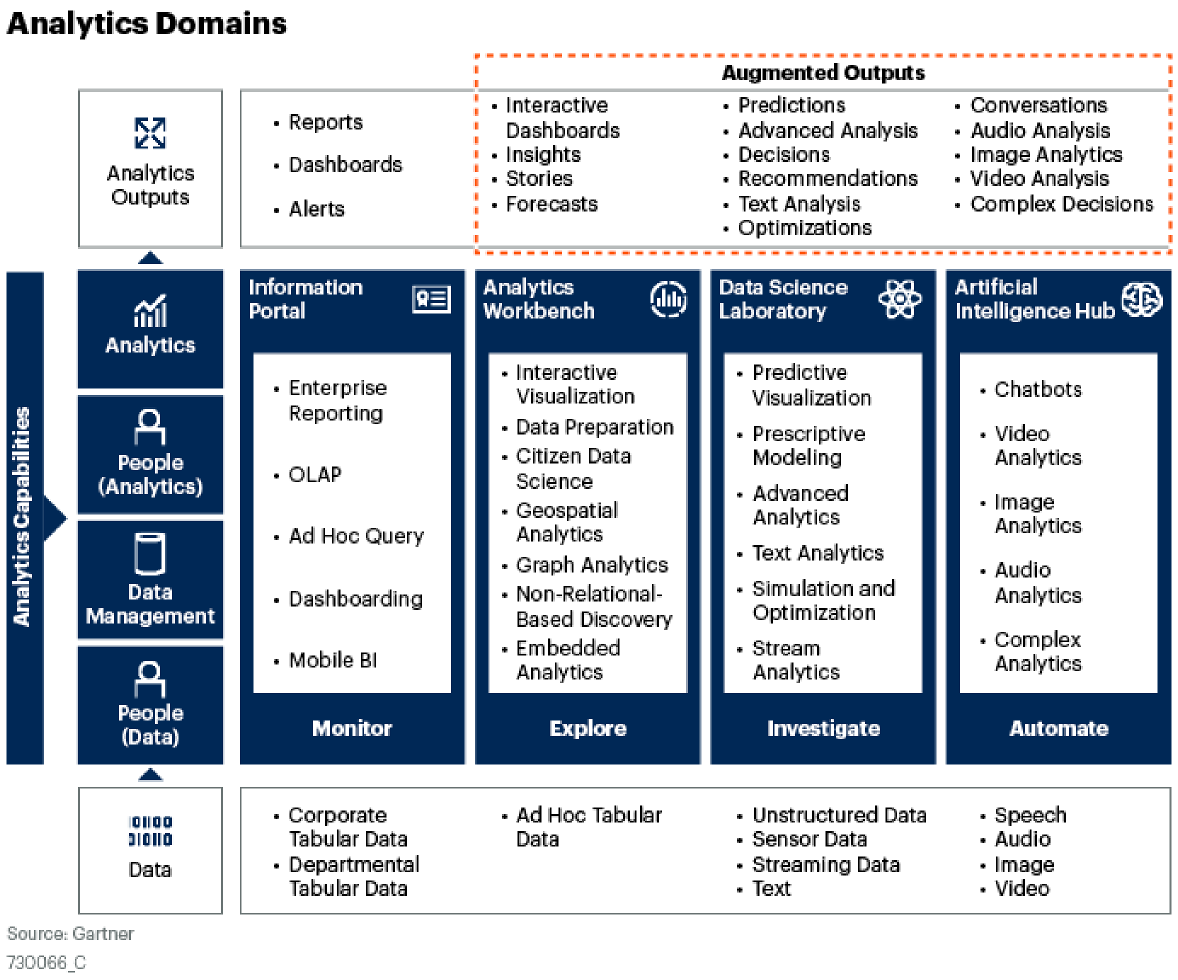
Deploy Multiple Analytics Platforms Tailored to Various Personas

Standardizing all analytics based on a single product or tool sounds simple and ideal. However, it is not realistic due to variances in use cases, user expertise and tool capabilities. Analytics tool selection should be based primarily on use cases, user skills and data needs, although you have to accept that these decisions are subjective and predicated on reasons that are not always technical. Therefore, it is very easy to get into battles over which solution is better than another. One thing is certain, though: No single vendor or platform offers the most comprehensive analytics solution for every use case. Thus, we recommend categorizing the platforms based on their capabilities and features, and excluding vertical-specific analytics solutions (with some exceptions).

As shown in Figure 3, analytics platforms fall under four primary domains:

- **Information portal:** Information portals can provide curated, governed and trusted information to business users in the form of the paginated reports or dashboards offered via traditional (BI) tools. Analytics capabilities from this domain should be selected when reliable metrics are required, and when scalable mass distribution of information is the goal.
- **Analytics workbench:** Tools within this domain empower business users to independently produce and publish insights, mainly through self-service data preparation and interactive visualization. These analytics capabilities should be selected when the objective is to support user independence, agility and flexibility, without requiring advanced analytics skills. This domain requires data governance, so that user-built content can be managed in a process that enhances trust.
- **Data science laboratory:** Tools should support the delivery of advanced analytics outputs, using predictive modeling, prescriptive analytics, machine learning and other sophisticated analytics capabilities. These tools require specialized users — citizen data scientists and data scientists. The analytics capabilities in this domain should be selected when there are clear requirements that justify their potential complexity, and when the resources with the right advanced skills are available.
- **Artificial intelligence hub:** The AI hub should offer technology that emulates human performance. Applications could include natural language query (NLQ), low-code conversational interfaces, text mining, intelligent document processing, and detection of novel concepts and abstractions to work with semistructured and unstructured data. IT will need to play a key role in integrating the analytics solutions within the AI hub with enterprise business and analytics applications.

Figure 3: Analytics Domains



Source: Gartner
730066_C

Technical professionals responsible for supporting multiple analytics and BI (A&BI) tools within their enterprise should:

- Design data architectures with access, virtualization, federation and semantic-layer features to not only deliver reusable data for multiple analytics tools, but also encourage business users to leverage the enterprise-sanctioned analytics portfolio and abandon shadow analytics
- Reduce overlap within analytics domains by retiring slow, expensive and unwieldy products, and by investing in new capabilities where more speed and agility are key to analytics success

For more information, see:

- Demystifying the Analytics and BI Space
- Solution Comparison for SaaS Analytics and Business Intelligence Platforms

Augmented AI, ML and Analytics Will Drive Democratization and Self-Service

Modern enterprises looking to leverage data for decision making need it to answer more than just "what happened" and "why." They also need it to reveal what will happen in the future and what next steps they should take. The increasing requirement for predictive and prescriptive capabilities is driving the convergence of analytics, AI and ML capabilities to generate more proactive and actionable insights. This proliferation of AI and ML capabilities within the analytics ecosystem is also resulting in enhancements to analytics, AI, and ML tools and platforms with the same capabilities.

> **By 2023, 40% of application development teams will be using augmented data science and machine learning (DSML) in their data science initiatives.[4]**

This augmentation of tools is already bringing transformative analytics technologies to the hands of technical and business users. For technical users, it promises to increase productivity. Augmented analytics assists with self-service analytics by lowering the barrier to entry for less experienced technical professionals. Augmented data labeling, data preparation and algorithm selection have enabled data scientists to run quick experiments and, in some cases, even deploy the models to production. For business analysts and business users, the augmentation abstracts away the technical complexity of code and provides AI-based interfaces, insights and workflows. These interfaces, insights and workflows are more user-friendly than code, but still provide the benefits of AI and ML technologies.

Data, analytics and AI architects should use the following planning considerations when seeking to implement or scale self-service BI and data science initiatives:

- Leverage ML models, AI services and augmented low-code toolkits to support descriptive to predictive analytics

- Advance self-service analytics developers to citizen data scientists by embracing autoML model development toolkits
- Extract insights from unstructured data using deep learning techniques
- Implement analytics training to increase data literacy and analytics proficiency

Planning Considerations

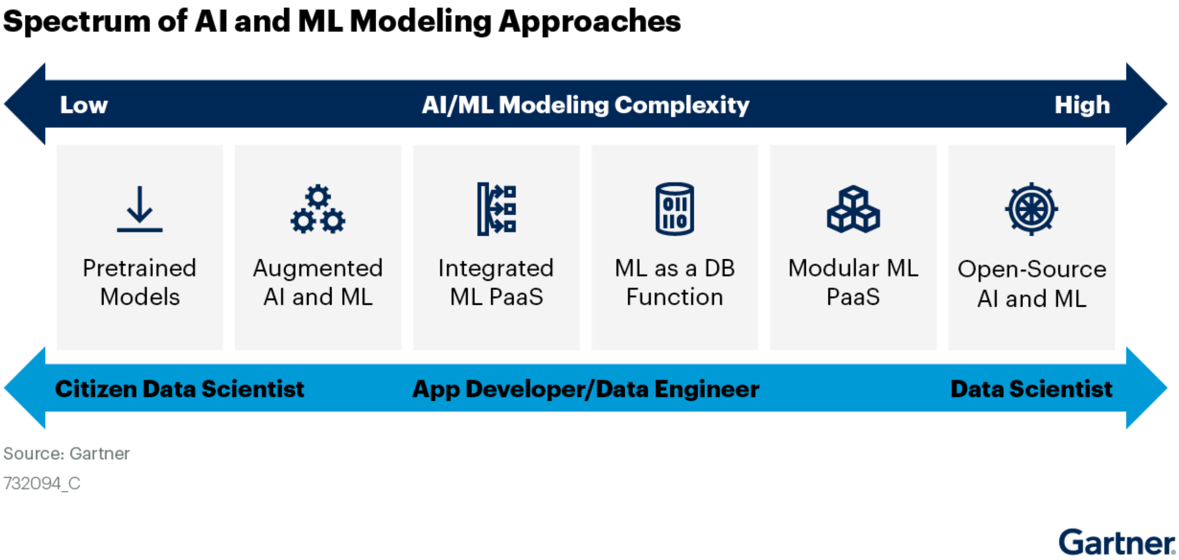
Leverage ML Models, AI Services and Augmented Low-Code Toolkits to Support Descriptive to Predictive Analytics

Augmented analytics can impact the entire analytics cycle, from facilitating ingestion, preparation and analytics on historical data to providing prescriptive insights based on intelligent algorithms and models driven by good-quality data. Such capabilities are being increasingly embedded within analytics products. As a technical professional, you will need to contribute in the following ways to ensure the success of such initiatives:

- **Enable greater accessibility for historical datasets:** All augmented analytics tools require access to data to enable their functionality. One of the key benefits of these toolsets is the ability to provide and enhance access to historical data. Leverage augmented analytics capabilities to assist with data profiling and descriptive analytics. The more you understand what happened in the past, the more you can contextualize predictions for the future. This step is key to ensuring that the predictions and prescriptive insights are backed by sufficient data. Hence, the tools should help domain experts and citizen data scientists analyze data trends and patterns, starting with historical data.
- **Drive consistency with delivering data:** Ensure augmented analytics tools enable end users to support automatic data profiling, data quality checks and data harmonization. Poor quality data is a big problem for organizations, but augmented analytics is improving the experience of data wrangling by delivering AI-assisted data modeling, transformation and enrichment.
- **Assist business analysis:** Ensure modern augmented analytics tools support automated identification, visualization and narration of relevant finds based on pristine-quality data and a common semantic layer. Allow users without a data science background to explore data via NLQ, chatbots and voice interfaces, so that they can identify correlations, clusters, exceptions and predictions.

It is also important to consider the various ways in which augmented analytics capabilities can be introduced to existing enterprise analytics platforms. Figure 4 lists the spectrum of AI and ML modeling approaches.

Figure 4: Spectrum of AI and ML Modeling Approaches



Source: Gartner
732094_C

The following approaches in the figure above enable analytics developers to extend descriptive and diagnostic analytics with predictive components, through embedded AI and ML models:

- **Pretrained models:** Pretrained models provide out-of-the-box functionality with a technology or business focus. Technology-focused pretrained models include generalized language translation, image classification and speech processing. Some examples of industry-focused pretrained models are medical diagnosis, cybersecurity threat detection and anti-money-laundering detection. In addition, there are academic open-image models, such as InceptionV3 and ResNet-50, which can be used as a baseline for more advanced model development using transfer learning. Although end-user organizations may not be able to fully customize the business objectives of third-party pretrained models, they can customize the data used to train those models. For example, end-user organizations can customize image models by using transfer learning to incorporate their proprietary data.

- **Augmented AI and ML:** Augmented AI and ML tools aim to improve the construction of applications powered by AI and machine learning by reducing the burden on humans and reducing the need for expert knowledge. Augmented object detection, document classification and machine learning enable the development of models with minimal knowledge of data science and AI frameworks. The tools give citizen data scientists and application developers the ability to use datasets and develop models without requiring knowledge to use and select algorithms.

- **Integrated ML PaaS:** An AI and ML solution based on Python or R creates a barrier to entry for application developers or citizen data scientists. These users may be working to understand machine learning concepts, but may not be confident about programming a data science solution. A low-code solution, such as integrated ML PaaS, provides an application studio with a palette of tasks. These interfaces are service wrappers that abstract the complexity of data services and deployment tasks. The platform would require connecting through a series of tasks, including data acquisition, feature engineering, algorithm selection, model training, metrics validation and related configuration, to create a workflow for model creation.

For more information, see:

- Evolving Capabilities of Analytics and Business Intelligence Platforms

- Comparing Platforms and Capabilities for Data Science and AI

Advance Self-Service Analytics Developers to Citizen Data Scientists by Embracing AutoML Model Development Toolkits

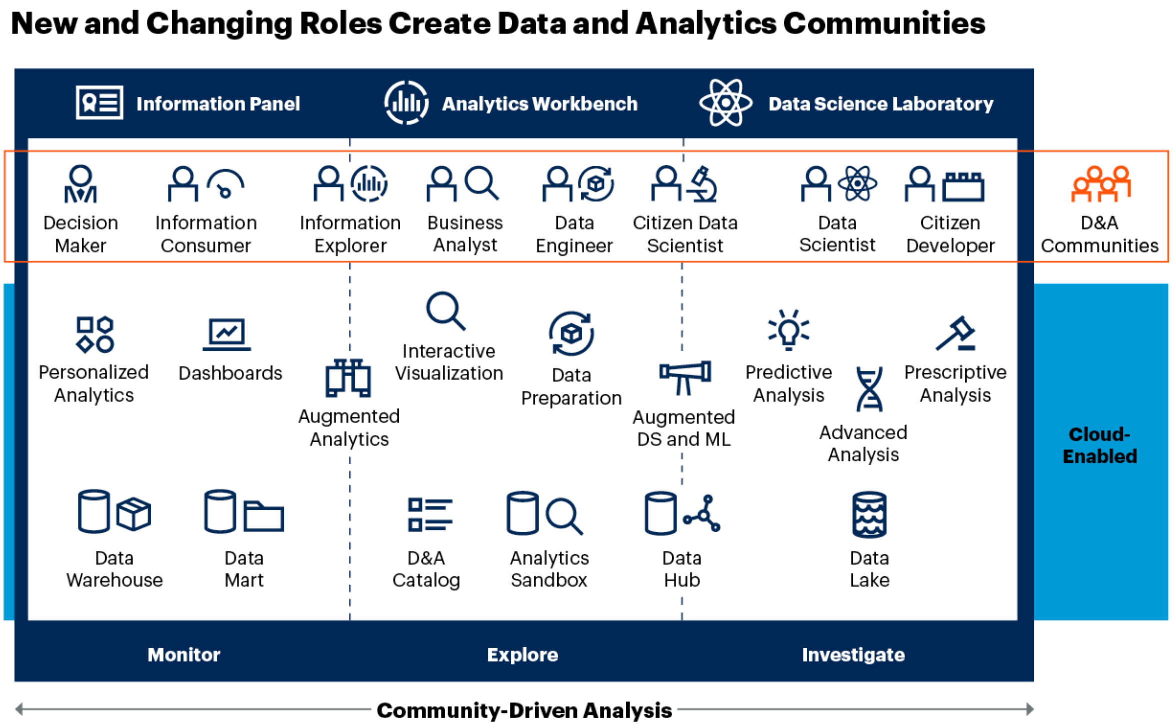
> **Organizations that are successfully adopting augmented analytics are delivering intuitive analytics interfaces and a collaborative experience across the user community.**

Analytics adoption has come in waves of centralization and decentralization. For every push for centralization, such as the data warehouse, there is a counterpull for decentralization, such as self-service analytics. Augmented analytics is unique in that it can be leveraged for both objectives.

Analytics adoption can support decentralized use cases by providing a personalized experience for analytics workbench users who can prepare, analyze and visualize AI-driven data. Time spent in predefined dashboards will be complemented — and likely displaced to a degree — by automated, conversational and dynamically generated insights that are customized to users' contexts and delivered to their points of consumption. By being embedded within existing business applications and workflows, autogenerated and personalized data stories will offer multiple experiences that leverage a variety of devices and consumption methods. Augmented analytics can also support centralization by offering the ability to run machine learning models inside the centralized data warehouse. In 2023, features like Snowpark within Snowflake will bolster interest among data science teams in taking advantage of data warehouses for AI/ML use cases.

Figure 5 shows three separate analytics environments, with augmented analytics bridging the gap between the information panel and the analytics workbench, and augmented DSML bridging the gap between the analytics workbench and the data science laboratory. As shown in the figure, augmented analytics capabilities upskill users across the spectrum.

Figure 5: New and Changing Roles Create Data and Analytics Communities



Source: Gartner
732258_C

For more information, see:

- Democratize Data Science Initiatives With Augmented DSML

Extract Insights From Unstructured Data Using Deep Learning Techniques

Although datasets are growing at a rapid rate, the majority of them are not curated and leveraged toward gaining insights. As a result, organizations have little awareness of the volume, composition, risk and business value of their "dark data," which sits mostly in unstructured formats.

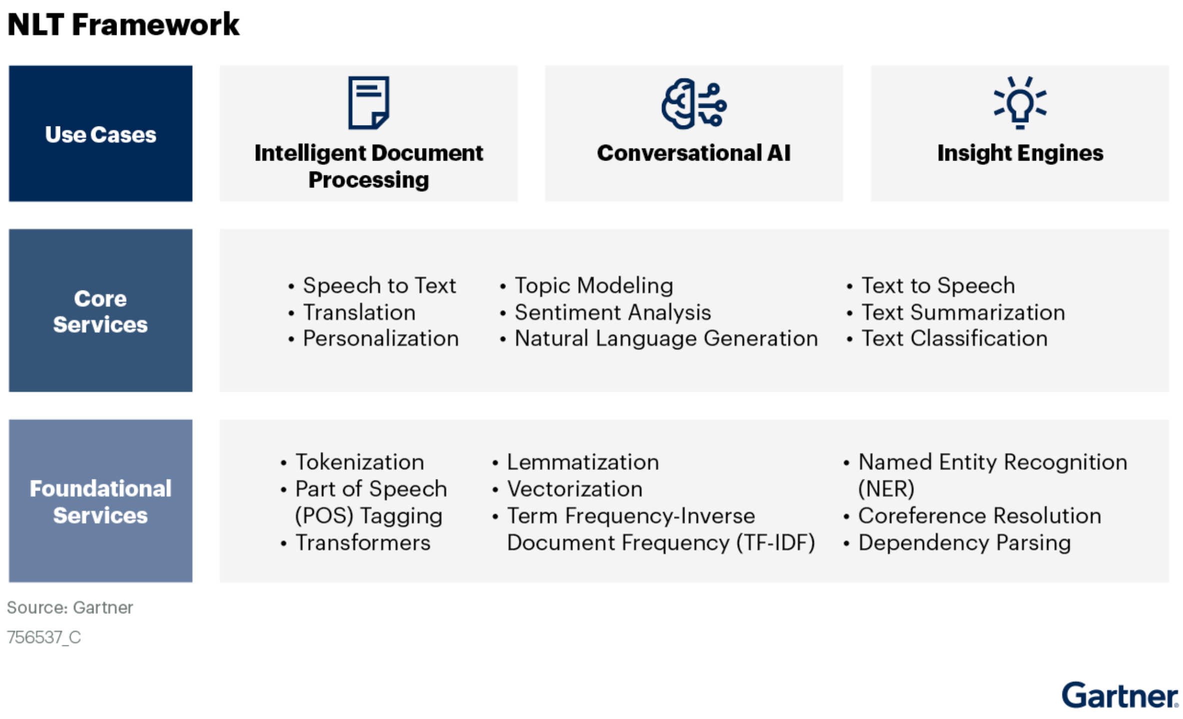
Natural language technology (NLT) offers significant opportunities to improve operations and services. The most immediate use cases focus on improving customer service (impacting cost, service levels, customer satisfaction and upselling) and employee support (augmenting AI with business processes to increase efficiency). Intelligent document processing, conversational platforms and insight engines are the emerging endpoint solutions helping to unlock insights from semistructured and unstructured data.

The NLT framework in Figure 6 consists of foundational services, core services and use cases. Foundational services are the building blocks used to create core NLT services, such as speech to text, text to speech, translation, text classification, topic modeling, text summarization, personalization, sentiment analysis and natural language generation (NLG). These core NLT services make up product solutions:

- **Intelligent document processing** helps ingest data from a wide variety of sources and formats (e.g., PDFs and images), extract information from that data, and consolidate documents across multiple applications into one component. These capabilities support document classification, metadata extraction, knowledge graphs, search, and natural language query and answer.
- **Conversational AI** is being employed by developers to build conversational user interfaces, chatbots and virtual assistants for a variety of use cases. Developers are layering these interfaces on top of CRM applications and analytical platforms across customer service, IT service desk and HR for tasks like recruitment, debt collection, appointment booking and more.

- **Insight engines** use AI to reinvent enterprise search. They enable enterprises to shift gears to semantic search, thereby unlocking patterns held within both unstructured and structured data, sourced internally and externally. Content and data are presented in context, as information, to deliver the insight needed for purposeful action. Insight engines provide the platform to gather content/data from myriad sources, enrich it for indexing, and query it via multiple touchpoints and integrations.

Figure 6: NLT Framework



NLT Framework

Use Cases	**Intelligent Document Processing**	**Conversational AI**	**Insight Engines**
Core Services	• Speech to Text • Translation • Personalization	• Topic Modeling • Sentiment Analysis • Natural Language Generation	• Text to Speech • Text Summarization • Text Classification
Foundational Services	• Tokenization • Part of Speech (POS) Tagging • Transformers	• Lemmatization • Vectorization • Term Frequency-Inverse Document Frequency (TF-IDF)	• Named Entity Recognition (NER) • Coreference Resolution • Dependency Parsing

Source: Gartner
756537_C

For more information, see:

- Working With Semistructured and Unstructured Datasets
- Emerging Use Cases for Natural Language Technology
- Selecting Conversational AI Solutions for Chatbot and Virtual Assistant Initiatives

Additionally, computer vision (CV) techniques are being used to analyze images and videos, allowing machines to extract meaningful and contextual information from the physical world. Deep learning techniques are the underpinning of CV. The following techniques are used in ML models for image and video analytics:

- **Image classification** scans through a set of input images and classifies an object or set of objects within the image. It requires labeled data to process new images and place them into specific categories.
- **Object detection** identifies and defines objects and labels them with bounding boxes. It also applies classification and localization to the objects.
- **Semantic segmentation** divides the entire image into groups of pixels. It classifies every pixel to identify what objects they contain.
- **Instance segmentation** advances semantic segmentation by not only classifying objects in the image at a pixel level, but also differentiating those similar objects.
- **Panoptic segmentation** combines both semantic segmentation and instance segmentation to assign a class label to each pixel, and then detect and segment each object instance.
- **Object tracking** is the process of following a particular object or multiple objects in a video. It uses a generative method to describe characteristics of the objects and a discriminative method to separate the objects from the background.

For more information, see:

- Improve Computer Vision Use Cases With Standardized Implementation Patterns

Implement Analytics Training to Increase Data Literacy and Analytics Proficiency

Gartner defines data literacy as the ability to read, write and communicate data in context. Data literacy includes an understanding of data sources, data constructs, and analytical methods and techniques applied, as well as the ability to describe the use case, application and resulting value.

> By 2023, data literacy will become an explicit and necessary driver of business value, demonstrated by its formal inclusion in over 80% of D&A strategies and change management programs.[5]

A data-literate workforce is central to an effective D&A program and critical to driving measurable business outcomes. With the expansion of self-service analytics initiatives, business teams are empowered to analyze data and generate insights. However, low data literacy can doom self-service initiatives, as it results in lower analytics adoption (due to higher barriers to entry) and less impactful analytics. By contrast, data and analytics professionals are generally considered to be the most content-neutral data-literate employees within an organization. Additionally, data stewards (either business subject matter experts or data stewards working in IT) are an organization's primary source of domain-specific data literacy. Working together, IT professionals and data stewards can develop training and educational materials to help increase the overall data literacy within the organization.

The data literacy training program must have three key tenets:

- Data concepts
- Analytics approaches
- Business value

There are several methods for delivering training to the organization. D&A technical professionals designing upskilling paths for analytics can choose one or multiple modes of learning, such as:

- Classroom-based learning (either in-person or virtual)
- Workshops
- Self-paced online learning
- Game-based training

Successful training programs are flexible and provide a variety of options to suit user preferences and learning styles. When these techniques are used in combination, learning and retention are greatly enhanced. For example, classroom learning may be needed to introduce topics and provide a foundation for users. However, this alone is not sufficient to build proficiency in any topic. Working through generic examples in a classroom setting is unlikely to prepare users for the real challenges they may encounter when attempting to build their own analytics content.

Instead, training approaches should contextualize these skills by getting users to apply them to a specific problem that is relevant to their role within the organization. The organization may encourage users to tackle such problems on their own time, or it may guide them through the process by bringing a specific problem to a workshop. Alternatively, the organization could assign users to a working group to develop a solution as a team. Completing a project can even be made into a game or competition. There are many options.

For more information, see:

- Case Study: Practical Data Literacy (Kraft Heinz)
- Improve Data Literacy and Governance by Using Accurate, Meaningful Names for Data Artifacts

Operationalizing Advanced Analytics and AI Pipelines Will Require Implementation of Common Architecture Patterns

According to Gartner's 2021 AI in Organizations Survey, [6] it takes organizations, on average, 7.3 months to progress an AI initiative from prototype to production. Forty-seven percent of organizations take more than six months on average. This is certainly a substantial improvement from the 2020 survey, where the average was almost 8.6 months. That said, a lot more work is still needed as enterprises look to scale their AI implementations.

By 2025, 50% of enterprises will have devised AI orchestration platforms to operationalize AI, up from fewer than 10% in 2020.[7]

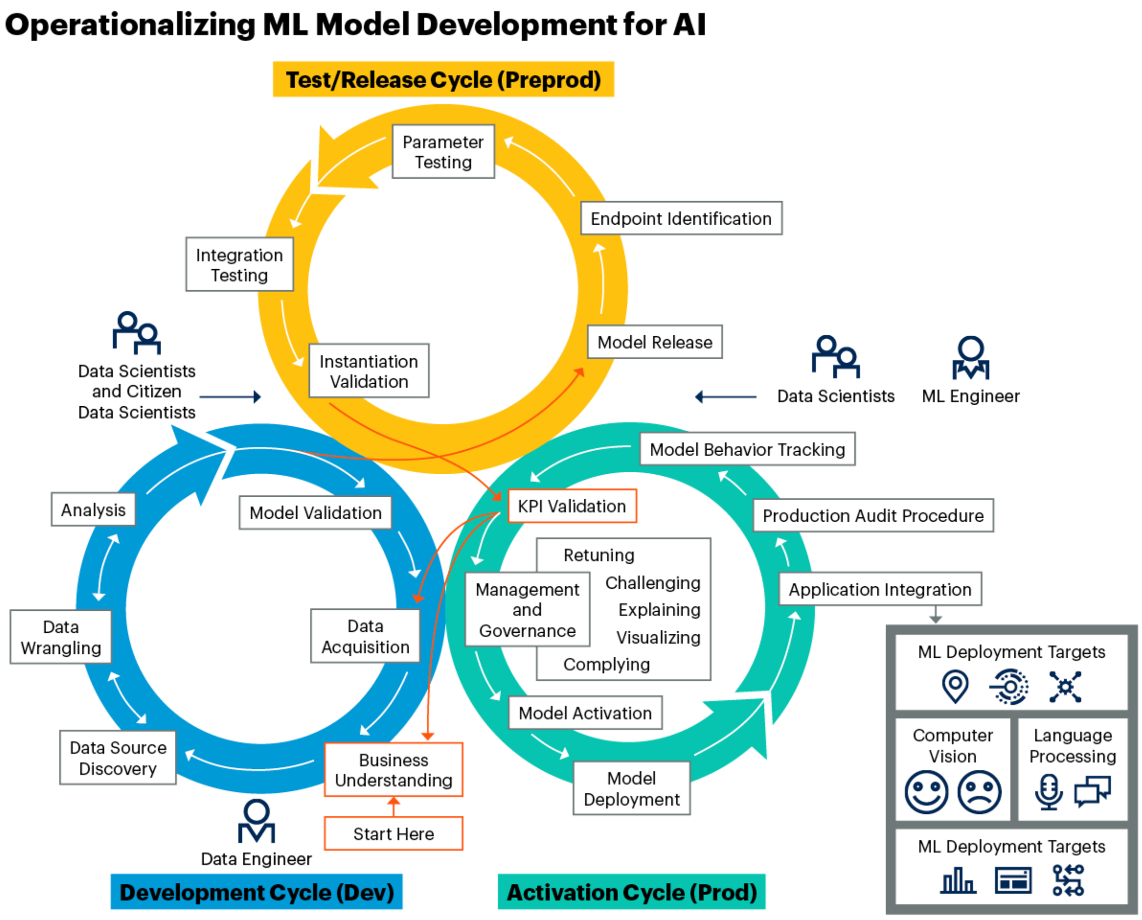
In order to successfully move AI pilots to production — and scale not just tens of analytical models, but thousands — organizations require the following:

- An AI framework to bring the people and processes together
- An orchestration platform to support the continuous integration and delivery of cognitive artifacts within business workflows

The goal is to create an architecture that brings together products, processes and people. Additionally, organizations should operationalize the end-to-end architecture using the Cross-Industry Standard Process for Data Mining (CRISP-DM) methodology, which has been leveraged by data science practitioners for decades. As shown in Figure 7, this simple yet powerful methodology imposes a discipline that promotes integrity and robustness. It results in an ML model development life cycle composed of three main subcycles:

- Development
- Test/release
- Activation (or production)

Figure 7: Operationalizing Advanced Analytics and ML Model Development for AI



Source: Gartner
732258_C

Operationalizing AI requires building a consistent architecture pattern, instituting an XOps (DevOps, DataOps and ModelOps/MLOps) framework and establishing a consistent delivery pipeline for AI-based systems. Enterprises should take the following steps to implement AI models at scale:

- Leverage XOps and agile best practices to streamline the delivery of analytics and ML models
- Implement containerized DSML workloads for scalability and portability
- Upskill software engineers to emerging roles, such as AI/ML engineers and architects

Planning Considerations

Leverage XOps and Agile Best Practices to Streamline the Delivery of Analytics and ML Models

Transitioning analytical and ML models to real-world production systems requires operationalizing the ML development life cycle (MLDLC) by adopting the XOps framework and agile best practices. The goal of XOps (DataOps, MLOps/ModelOps and DevOps) is threefold:

- Achieve efficiencies and economies of scale by applying DevOps principles to the data and analytics domain
- Ensure reliability, reusability and repeatability while reducing duplication of technology
- Enable automation

Tools that support operationalization of analytics and AI pipelines usually have overlapping capabilities, making it even more difficult to identify the right product and framework for implementation. If AI teams don’t leverage XOps for the individual subcomponents of the architecture, they will be overwhelmed by the technical debt, implementation complexity and infrastructure complexity associated with production A&AI pipelines. Hence, technical professionals responsible for delivering AI solutions should:

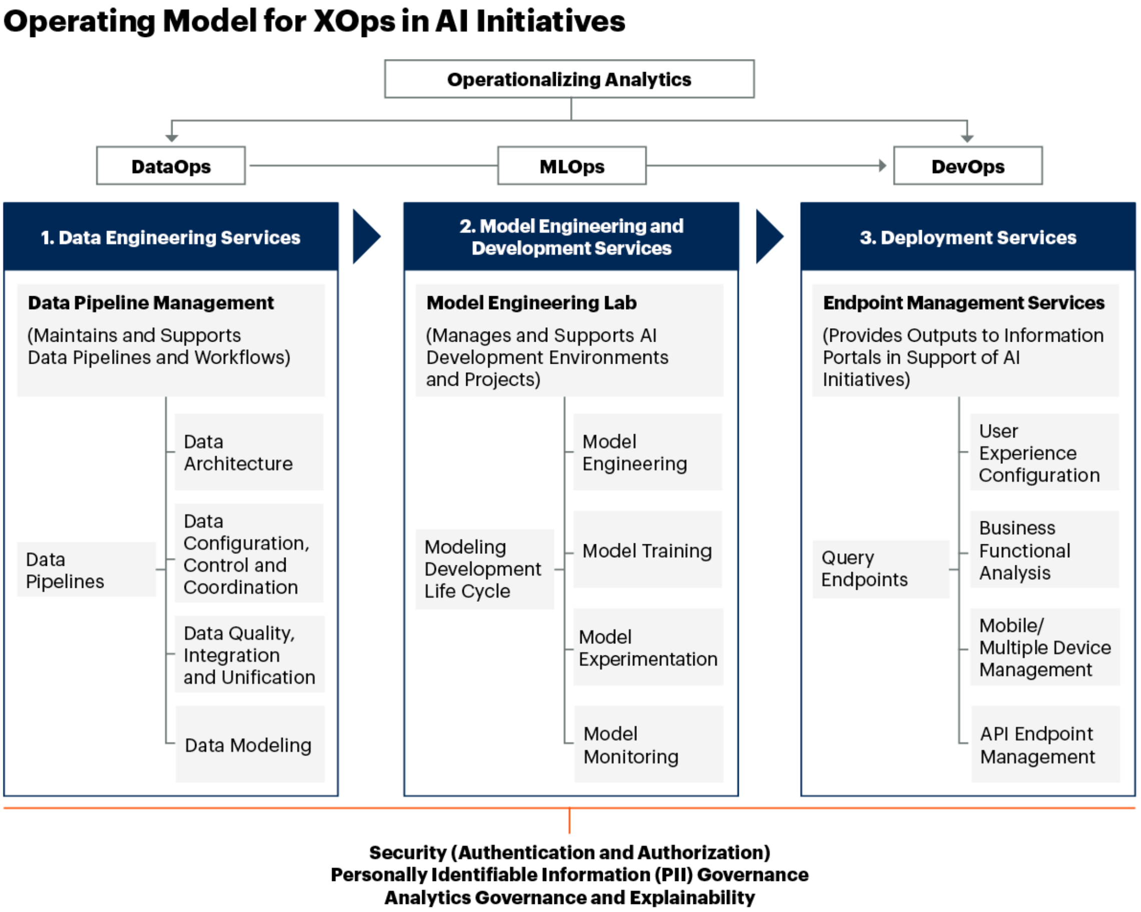
- Operationalize A&AI architectures by applying DevOps principles across the various stages of the analytics delivery pipeline.

- Create an integrated XOps practice that blends disparate functions, teams and processes to support data processing, model training, model management and model monitoring. This will allow for continuous delivery of AI-based systems.
- Leverage DataOps and MLOps/ModelOps practices and expertise to build a composable technical architecture that provides resilience, modularity and autonomy to data and AI services, making them ready for consumption by business functions. Orchestrate the delivery of analytics and AI solutions.
- Adopt the agile method that best meets the needs of the data science team and the organization. Adapt this agile process through regular retrospectives to make it fit for purpose.
- Deliver data science solutions iteratively and incrementally with feedback at each stage. Deliver early and often with knowledge realized first, followed closely by incremental business value.

Figure 8 illustrates how XOps best practices contribute to each of the following three operating modules:

- Data engineering services (DataOps),
- ML engineering and development services (MLOps/ModelOps)
- Deployment services (DevOps)

Figure 8: Operating Model for XOps in AI Initiatives


Operating Model for XOps in AI Initiatives
Operationalizing Analytics
DataOps
MLOps
DevOps
1. Data Engineering Services
Data Pipeline Management
(Maintains and Supports Data Pipelines and Workflows)
Data Pipelines
Data Architecture
Data Configuration, Control and Coordination
Data Quality, Integration and Unification
Data Modeling
2. Model Engineering and Development Services
Model Engineering Lab
(Manages and Supports AI Development Environments and Projects)
Modeling Development Life Cycle
Model Engineering
Model Training
Model Experimentation
Model Monitoring
3. Deployment Services
Endpoint Management Services
(Provides Outputs to Information Portals in Support of AI Initiatives)
Query Endpoints
User Experience Configuration
Business Functional Analysis
Mobile/ Multiple Device Management
API Endpoint Management
Security (Authentication and Authorization)
Personally Identifiable Information (PII) Governance
Analytics Governance and Explainability

Source: Gartner
775123_C

For more information, see:

- Demystifying XOps: DataOps, MLOps, ModelOps, AIOps and Platform Ops for AI
- A Guidance Framework for Operationalizing Machine Learning
- Five Fundamental Truths of Agile Methods in Data Science

Implement Containerized DSML Workloads for Scalability and Portability

Advanced A&AI workloads require varying storage and compute infrastructure. Combined with data gravity challenges, these disparate infrastructure needs complicate selection of the ideal implementation strategy. Hence, to best leverage a distributed compute environment that includes hybrid, multicloud and edge enterprise architectures, organizations may need to containerize their DSML workloads.

Organizations are also looking at ways to reduce data movement (latency) and get more deterministic performance on-premises. However, the cloud is emerging as a preferred implementation model, due to the scalability of compute infrastructure and the availability of pretrained models and cognitive services via simple API integrations. Hence, abstracting the ML pipeline into three stages — build, train and infer — allows organizations to examine and design the underlying compute infrastructure according to the needs of the individual stages. This approach brings flexibility, offering 27 different ways in which to implement the end-to-end pipeline, thereby providing a modular and scalable architecture for delivering AI-based systems.

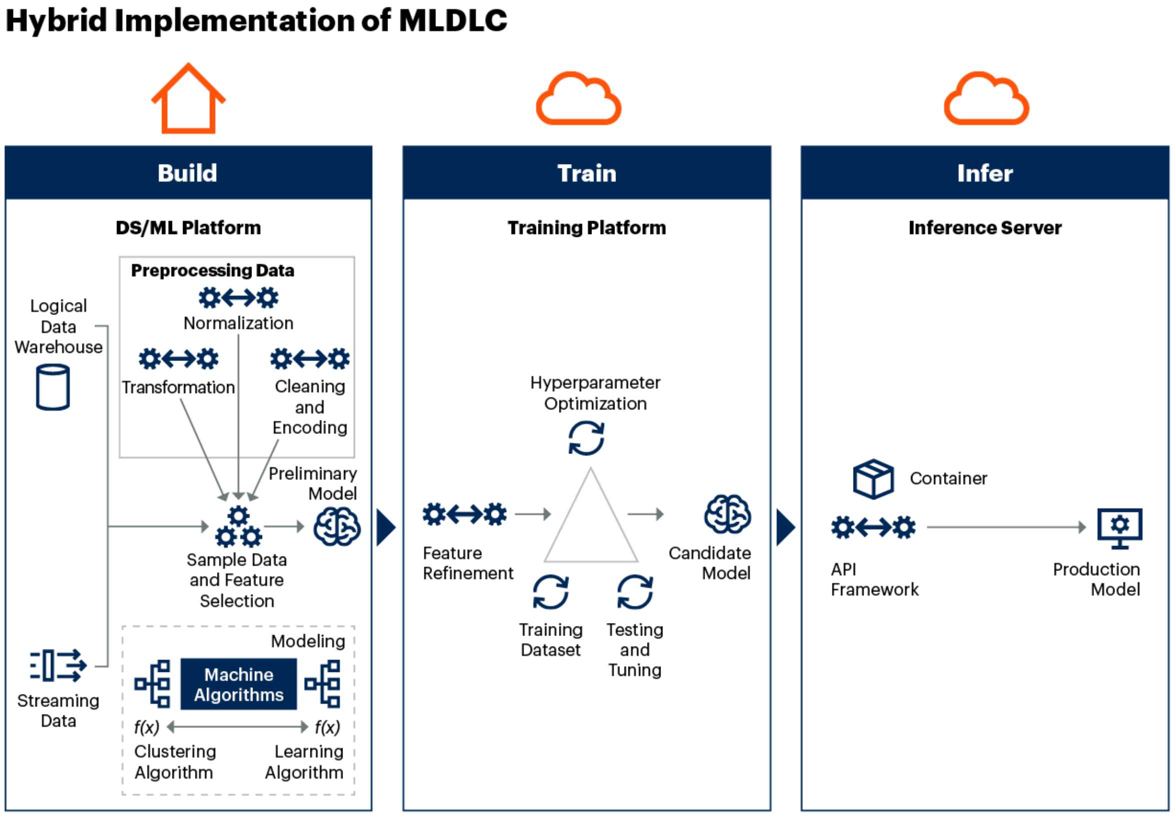
Containers are emerging as a popular abstraction technology for ML workloads because they enable implementation of a rapid continuous integration/continuous delivery (CI/CD) pipeline for ML models. They result in a perfect delivery mechanism to move the model across each of the build, train and infer stages of the MLDLC. Containers could pave the way to providing "data science in a box" within the organization, when it must support multiple analytics and DSML tools with overlapping capabilities.

Containers provide two fundamental benefits for DSML that technical professionals can benefit from:

- **Consistent packaging of DSML libraries and frameworks:** Containerized ML pipelines can be moved quickly and easily across heterogeneous infrastructure, and between private and public cloud environments from multiple providers. Containers therefore provide a smooth path for ML model deployment, in addition to model training and retraining, allowing data scientists and ML engineers to run ML workloads across identical environments at every stage of the MLDLC. This level of mobility with ML models can significantly reduce friction between development, operations and business teams, especially with regard to model management and MLOps/ModelOps.
- **Streamlined configuration management:** The contents of containers and the manifests for orchestrating containers on Kubernetes are defined with text-based artifacts (e.g., Dockerfiles, docker-compose files and YAML Ain't Markup Language [YAML] files). These artifacts can be treated the same way as application source code — that is, integrated into CI/CD pipelines and transactionally updated with a model management system. Model code changes can be selectively applied using canary or blue-green testing methods. Containerization not only provides immutability for the model code, but also enables state changes to be introduced more safely and efficiently.

Figure 9 depicts an example of a hybrid implementation model using containers.

Figure 9: Hybrid Implementation of MLDLC



Source: Gartner
732258_C

To effectively scale an ML and AI implementation, technical professionals need to account for the following:

- **Compute-intensive ML and deep neural networks (DNNs) require large training datasets:** During the train stage, model training, testing and tuning can be performed in batch mode with historical data or in real time with market data. These heuristics are best executed within the data center via compute systems accelerated with technologies such as graphics processing units (GPUs) and application-specific integrated circuits (ASICs), and complemented by high-bandwidth, low-latency networking infrastructures. In addition to specialized hardware, these techniques require optimally integrated middleware and system software. These are important considerations when organizations are deciding whether to train the model on-premises or in the cloud.

- **For inference, the paramount performance goal is low latency:** Automated services are leveraged to minimize the network's end-to-end response time and provide a near-real-time response across multiple inputs. Hence, these services can drive a significant amount of the compute costs. The models deployed within the AI-based systems might be placed within a dynamic training mode instead of a static mode, thus requiring low-latency, high-throughput workloads. Governing the model during the infer stage is imperative to ensure it is providing responses within acceptable thresholds. Therefore, specialized hardware may be needed to efficiently manage models in production, dictating the implementation strategy for this stage of the MLDLC.

- **Cloud comes with trade-offs:** Organizations must consider the trade-offs of selecting cloud as an option for moving the workloads. Such trade-offs may include data migration costs, interim transition architecture and vendor lock-in. Cloud vendors are also getting creative with compute processing costs. For example, "pay as you go" is becoming popular, but it makes estimating the TCO a challenge.

For more information, see:

- Assessing Containerization for Advanced Analytics Initiatives
- Solution Criteria for Data Science and Machine Learning Platforms
- Implementing an Enterprise Open-Source Machine Learning Stack

Upskill Software Engineers to Emerging Roles, Such as AI/ML Engineers and Architects

Although data scientists are responsible for ML model development, new roles have emerged to focus on the entire, end-to-end AI solution and to deploy and integrate the AI-based system within enterprise applications:

- **AI and ML architect:** This individual acts as a bridge between data scientists and application/data analytics developers, filling the gap between science and engineering. ML architects primarily design and build ML solutions as part of data science initiatives. However, they spend a growing portion of their time in operations — deploying and managing ML artifacts in production environments in collaboration with data science teams.

- **AI and ML engineer:** AI and ML engineers bring unique and valuable skill sets spanning multiple domains, such as advanced analytics, software engineering, architecture design and operations. All of these domains are required to effectively deliver comprehensive AI solutions. An important responsibility of the AI and ML engineer is to optimize implementation of AI-based systems that are built using advanced analytics technologies — such as NLT, CV and deep learning — and deployed across distributed compute environments.

- **Analytics engineer:** In some organizations, this role has emerged as a combination of data engineer and AI engineer skills. The analytics engineer is similar to the data engineer, but focuses on different pipelines. The traditional data engineer focuses on developing data pipelines from enterprise operational/transactional sources to the enterprise data repository. By contrast, the analytics engineer focuses on developing data pipelines primarily for AI/ML models and data science teams. Analytics engineers prepare data pipelines that create datasets for analytical purposes, including datasets for end users to import into desktop tools and datasets for BI developers to import into A&BI tools. In addition to regional and local data sources, the analytics engineer uses the “curated” area of the enterprise data repository as the source for these types of data products.

Data scientists should not be burdened with the software and system engineering needed to operationalize the delivery of ML models — and thereby the delivery of AI-based systems within the organization. Engineering is what differentiates AI/ML architects, AI/ML engineers and analytics engineers from data scientists. Technical professionals must be prepared to employ engineering capabilities to accelerate the delivery of ML models. They must support AI initiatives by establishing, hiring and upskilling existing technical groups toward AI/ML engineer and architect roles, or analytics engineer roles. Doing so will allow them to take on the engineering, deployment and delivery responsibilities of AI-based systems.

For more information, see:

- Roles and Skills to Support Advanced Analytics and AI Initiatives

Trust and Transparency Requirements Will Drive Analytics Governance and Responsible AI Practices

Data governance helps ensure access to trustworthy data. While trustworthy data is foundational, analytics and AI are data consumers and support enterprises with decision making. Enterprise leaders and users must be able to trust the metrics presented in an analytics dashboard or in a prediction served by an AI model, since those metrics will drive business actions and outcomes. Analytics governance and responsible AI practices bring in the structure, policies and best practices necessary for the reliability and quality of the outcomes. These practices should apply to the range of steps that make up an analytics or AI implementation, including:

- Data wrangling and preparation
- Data source and data type integration
- Feature engineering and metrics calculation
- Model development and deployment

Recently, many jurisdictions (e.g., the EU, China, the U.S. and the U.K.) globally introduced new and pending AI regulations that challenge data and analytics leaders to respond in meaningful ways. These pending regulations may conflict and lead to difficult business decisions. Responsible AI brings together many aspects of ethical business best practices and ethical decision making that organizations need to address independently. These include business and societal value, risk, trust, transparency, fairness, bias mitigation, explainability, accountability, safety, privacy, and regulatory compliance. Responsible AI operationalizes an organization’s obligation to ensure positive and accountable AI development through interpretability and explainable AI techniques.

Although traditional data governance practices are still important and useful, they must be supplemented with new approaches to accommodate the distributed nature of analytics development and the complexities of AI models. Technical professionals looking to solve these challenges should follow these planning considerations:

- Establish an adaptive analytics governance framework to support a federated self-service analytics implementation
- Implement a holistic model and feature management system to ensure reproducibility, and embed governance into the model life cycle

- Leverage AI fairness toolkits and explainable AI frameworks for compliance and interpretability
- Enforce data privacy and analytics model security practices as part of wider D&A governance

Planning Considerations

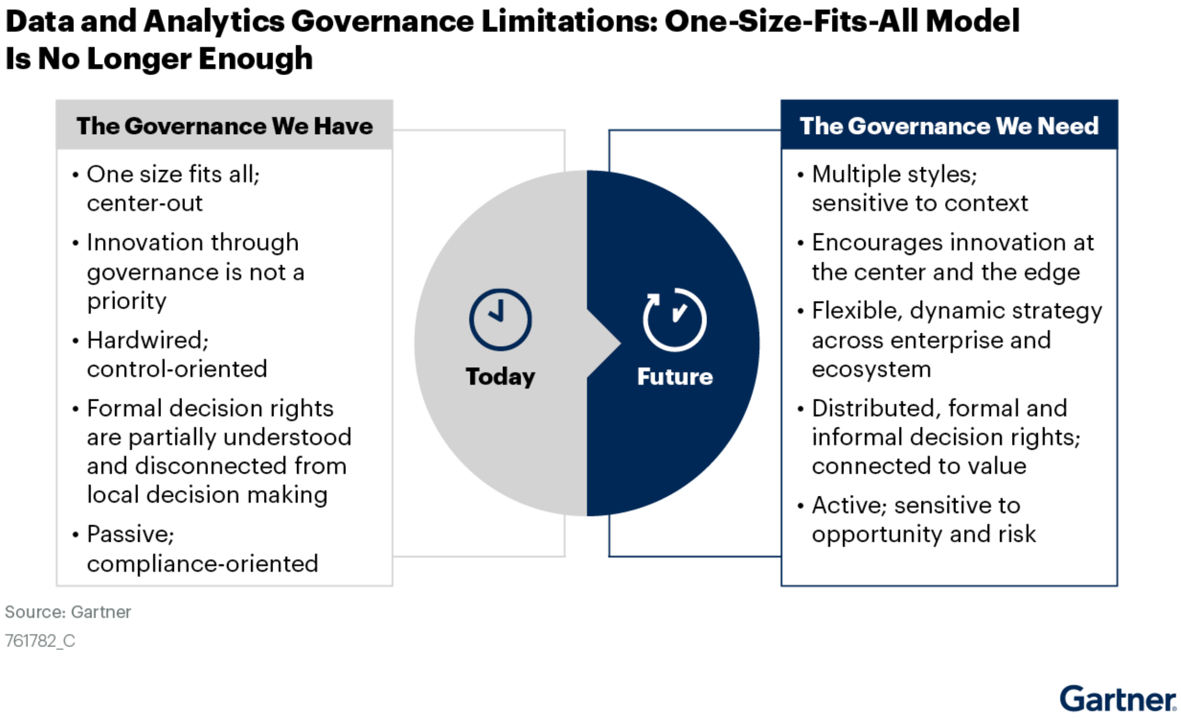
Establish an Adaptive Analytics Governance Framework to Support a Federated Self-Service Analytics Implementation

Self-service analytics creates a distributed analytics development ecosystem, where several groups within an enterprise develop and deploy their dashboards, metrics, measures and KPIs. Self-service enables business teams to develop dashboards that meet their needs in a much shorter cycle time. However, it also creates the potential for inconsistent use of data. A governance model becomes necessary to ensure trust, establish transparency and improve return on investment. *Analytics governance*, aka analytics stewardship, pertains to the enforcement of governance policy along the analytics pipeline — from data discovery, to analytics model deployment, to access of the analysis and insight.

> **By 2023, 90% of the world's top 500 companies will have converged analytics governance into broader data and analytics governance initiatives.[8]**

Analytics governance should not follow a one-size-fits-all approach either (see Figure 10). It must be able to adapt to changing organization needs, people and analytics use cases. At the same time, it must mitigate risk by setting guardrails and providing the tools and education necessary for analysts and decision makers to model, view and act on data. The modern approach to D&A governance is adaptable to different contexts and focuses on driving enablement, adoption and trust in information assets.

Figure 10: Adaptive Governance



Data and Analytics Governance Limitations: One-Size-Fits-All Model Is No Longer Enough

The Governance We Have

- One size fits all; center-out
- Innovation through governance is not a priority
- Hardwired; control-oriented
- Formal decision rights are partially understood and disconnected from local decision making
- Passive; compliance-oriented

The Governance We Need

- Multiple styles; sensitive to context
- Encourages innovation at the center and the edge
- Flexible, dynamic strategy across enterprise and ecosystem
- Distributed, formal and informal decision rights; connected to value
- Active; sensitive to opportunity and risk

Source: Gartner
761782_C

To ensure that the analytics developers generate trustworthy insights, the analytics center of excellence or the analytics leaders within the organization should provide:

- Clear, upfront communication of governance policies and best practices
- Adequate training to improve data literacy, as well as guidance and resources to help end users choose analyses appropriate for answering the question at hand
- Documentation and other resources, including a repository of sample, prototype and template reports, to share best practices
- A support system that users can leverage to get help, seek advice and receive feedback on developed content

For more information, see:

- Data and Analytics Governance Approaches for the Technical Professional
- Self-Service Analytics Governance With Microsoft Power BI
- Deploying Effective Data and Analytics Governance: Three Companies That Got It Right

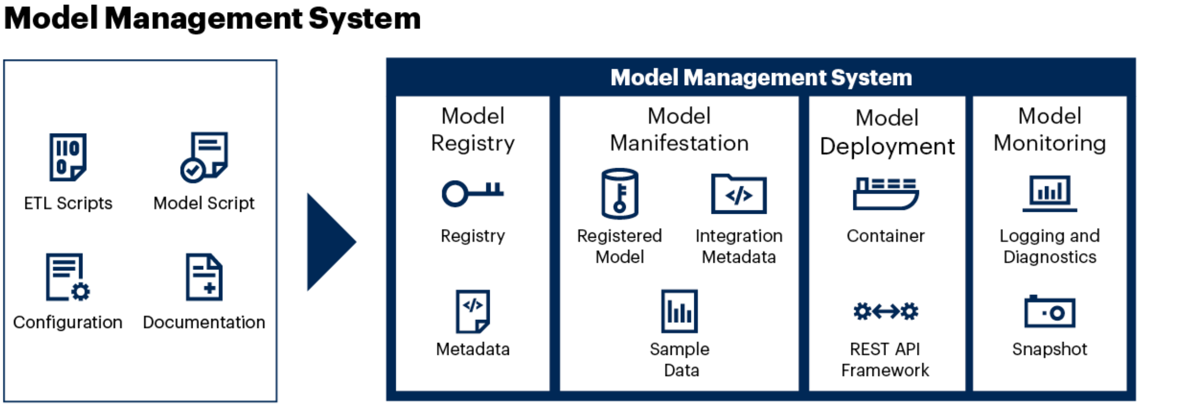
- Improve Data Literacy and Governance by Using Accurate, Meaningful Names for Data Artifacts

Implement a Holistic Model and Feature Management System to Ensure Reproducibility, and Embed Governance Into the Model Life Cycle

ML models are being used as cognitive elements within applications. They help with decision making, business process automation and user experience improvement efforts. It is critical to track various assets, artifacts and metrics associated with the models throughout the entire model life cycle.

A model management system is essential for managing and governing the models efficiently (see Figure 11). Reproducible model inferences and reusable models and workflows help establish a baseline for audit, compliance, risk assessment and adherence to industry regulatory guidelines. They also form the basis for MLOps.

Figure 11: Model Management System



Source: Gartner
732258_C

Technical professionals should implement a model management system supporting the following capabilities:

- **Model registry:** A model registry is necessary for versioning models and managing them across multiple environments — development, preproduction and production.
- **Model manifestation:** Manifests are necessary for tracking the parameters and configurations of models in production.

- **Model serving:** Model serving abstracts models into common industry formats, such as Predictive Model Markup Language (PMML) and Open Neural Network Exchange (ONNX). It also deploys models for inferencing via REST APIs or containerization frameworks, such as Kubernetes.
- **Model monitoring:** Model monitoring, aka ML model monitoring, aims to capture multiple dimensions across the following three tiers:
    - Data (e.g., upstream data, feature validation and training-serving "skew")
    - Model infrastructure (e.g., KPI score, concept drift and model metrics)
    - Infrastructure (e.g., challenger/champion, canary releasing and logging tracing)

ML model monitoring is particularly important, as it will help interpret model inference, thereby establishing trust through explainability. The focus of ML model monitoring must be on maximizing business-oriented KPI gain. Failure to demonstrate a measurable business benefit is why so many ML projects fail. The following provide general guidance on developing an effective ML model monitoring system:

- **Define meaningful and measurable KPIs:** Establish a cross-domain team and combine their knowledge to design the KPIs. The team should include a data scientist, a subject matter expert (SME), an application architect and operations personnel. The ML model monitoring system should help measure and optimize KPI gain, thus facilitating the decision process.
- **Detect anomalies, and suggest the next best action:** ML model monitoring must detect when there is a data pattern change that affects KPI gain (e.g., feature or concept drift), and must suggest what to do next. Most of the time, the next action will be to retrain an existing model or promote a challenger model.
- **Design a monitoring system that supports ongoing model deployment**: Rarely will an organization train and deploy a model that is never retrained or replaced. Therefore, the design must be able to simultaneously support many models that are in various stages of operation, as well as collect and update data for model retraining.

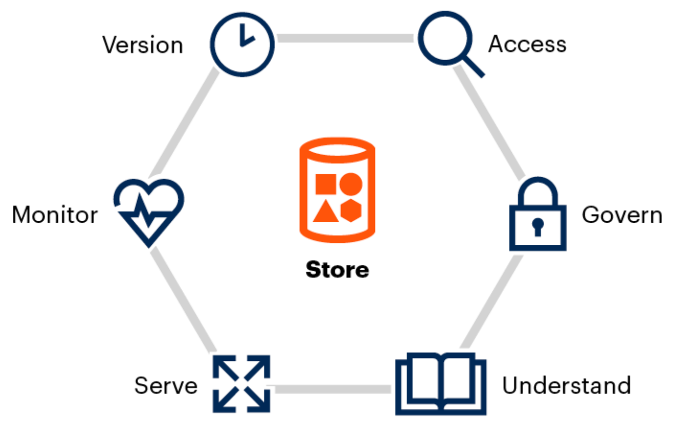
A logical feature store is another key component to implement, alongside the model management system. Data scientists spend a disproportionate amount of time engineering features by finding and cleaning data, writing transformation code, and so on. This work is typically done on local machines, thus duplicating feature-engineering efforts, preventing reuse and making governance near impossible. Data scientists pass ML models and feature-engineering code to data engineers to replicate in the production environment, where subtle differences between frameworks and programming languages create the risk of training-serving skew.

> Feature stores aim to increase the reusability and reliability of features for ML models by providing a repository of curated features from which training and test datasets can be created.

Regardless of the tools and technologies employed to manage ML features within the organization, technical professionals must consider the capabilities that are critical to a feature management solution (see Figure 12). Feature stores play a key role in tracing the lineage of data elements used in the training and testing dataset. Hence, they provide visibility into any bias accidentally introduced as part of the model building process.

Figure 12: Critical Capabilities of a Feature Management Solution

Critical Capabilities of Feature Management Solutions



Source: Gartner

741118_C

Gartner

For more information, see:

- A Guidance Framework for Operationalizing Machine Learning
- Getting Started With Machine Learning Monitoring in Production
- Feature Stores for Machine Learning (Part 1): The Promise of Feature Stores
- Feature Stores for Machine Learning (Part 2): Current State and Future Directions

Leverage AI Fairness Toolkits and Explainable AI Frameworks for Compliance and Interpretability

Many organizations are expecting AI to provide strategic differentiation from their competitors. The more organizations invest in AI, the more they need to understand how the AI models work and whether those models are biased. This knowledge will safeguard organizations from reputation and regulatory risks. At the same time, organizations need to ensure that solutions do not take on the bias of the designers and developers. Bias mitigation is necessary to both comply with regulations and protect the business's reputation. Explainable AI and fairness algorithms promise to resolve some of these challenges.

However, developing a fair model and explaining the rationale behind a prediction must not be an afterthought or a stand-alone activity. In order to make tasks repeatable, organizations must embed the following actions throughout the entire ML development and deployment process:

- Apply fairness metrics during data selection
- Use explainability to improve model training
- Leverage feature importance for validation, and explicitly test for bias
- Monitor the prediction, store the explainability factors and present the explanations in an interpretable way for the user

The following are key points for technical professionals to consider as they define and evaluate explainability and fairness components within the architecture:

- **"Fairness" definitions vary by use case:** Data scientists are often tasked with developing a fair model. However, defining fairness requires an understanding of the business processes and regulatory landscape. Business or product owners need to take the lead and collaborate with the compliance/risk expert and the data scientist to define what fairness means for the specific use case. They should also understand the ramifications of biased predictions and determine the guardrails required once the ML model is implemented in production.

- **Explainability tools and frameworks should align with use cases and impacted personas:** Multiple open-source and proprietary tools and frameworks provide both model-agnostic and model-specific views into the prediction model. However, there is no single solution to the problem. The process includes several decision points — algorithm, accuracy, interpretability, transparency, compute requirements, execution runtime and ease of implementation, to list a few. These decisions must be appropriate to the industry, use case and persona. For example:

  - Insurance actuaries or credit analysts would prefer accuracy and transparency over execution runtime and interpretability. They would use multiple tools — such as SHAP, partial dependence plots and feature importance — and analyze both local and global explanations to identify any model discrepancies.

  - Loan officers or claims processors, on the other hand, would prioritize interpretability and faster availability of results when looking to explain a decision to their customer. LIME, combined with feature importance for a specific prediction, would be more relevant for such use cases.

- **Not all business problems require explainability:** Organizations should not look to use ML explainability for all use cases. Explainability frameworks add complexity and overhead. For applications that create little or no adverse impact, the overhead associated with explainability may not be justified. ML models within games, emails, postal code sorting and movie recommendation systems are some examples.

- **The dynamic solution landscape needs proper assessment:** The landscape for explainability frameworks has been evolving rapidly. A combination of academic research, industry product extensions and open-source community contributions has provided several tools to understand ML predictions. Organizations should monitor academic research, conference presentations and publications from thought leaders to gain an understanding of the forthcoming solutions and capabilities. Most leading product vendors recognize the significance of ML explainability and are investing substantially in product development. This continuing product development, supported by research, should result in improved and more-user-friendly capabilities.

For more information, see:

- Incorporate Explainability and Fairness Within the AI Platform

Enforce Data Privacy and Analytics Model Security Practices as Part of Wider D&A Governance

Organizations are increasingly concerned about data security in several scenarios. These include collecting and retaining sensitive personal information; processing personal information in external environments, such as the cloud; and sharing information internally and externally. Because ML models contain no data, they are usually considered safe from a privacy standpoint. However, the latest research has showcased model inversion techniques that allow adversarial attacks to infer the data used to train the models. [9] To address this potential data security risk, consider several privacy/preservation and enhancement techniques:

- **Synthetic data** is artificially generated, rather than obtained from direct observations of the real world. Data can be generated using different methods, such as statistical sampling, semantic approaches, generative adversarial networks or simulations (where models and processes interact to create completely new datasets of events). Organizations can use synthetic data to comply with data privacy regulations and to overcome data residency restrictions that prevent ML training on real data from different geographies. A synthetic dataset can be created that reflects the trends and dynamics within each individual set. Then, the synthetic dataset can be shared freely. This technique is particularly relevant in healthcare with patient data, as well as in other areas like analysis of customer data across multiple countries or multiple business units.

- **Differential privacy** is a type of data-level transformation that uses data perturbation, often in the form of adding statistical noise, to hide individual data values. Differential privacy is usually available as part of data-masking products or in the form of software components. The primary use case is analysis of "behavioral" data, where aggregating the results is sufficient for building ML models. The value of differential privacy is that it enables organizations to analyze this data without potentially exposing private information, such as specific user travel patterns, locations or web searches. Differential privacy may be applied on synthetic data for additional levels of obfuscation.

- **Secure multiparty computation (SMPC) and homomorphic encryption** are approaches that combine data transformation with specialized software. These approaches are used for computation in untrusted environments, data sharing and shared analysis with multiple parties. With homomorphic encryption, computations can be performed on data while it's in an encrypted state. With SMPC, computations can be performed on data contributed by multiple parties, without any one individual party being able to see more than the portion of data it contributed. This isolation enables secure computation to be performed without the need for a trusted third party. Participants who collaborate on the computation know only the results of that computation, and not the specific data others have contributed.

- **Trusted execution environments and secure hardware systems** implement security mechanisms in hardware. Not all use cases for advanced data protection can handle the software-based approaches discussed above. To address this gap, several emerging techniques protect privacy by securing the hardware environment instead. In such an environment, data processing is protected by virtue of the hardware being resistant to intrusion and tampering. The hardware security mechanisms can be implemented at the microprocessor level, a hardware subsystem level or the hardware system level.

To ensure data privacy and model security while providing a self-service mode of analysis, technical professionals should follow these best practices:

- **Evaluate the organizational sensitivity** of the features within the context of the organization's data governance policies. This assessment will help you apply privacy protection techniques and tools to the ML models. TensorFlow Privacy, an open-source library, provides tools for training ML models with privacy. [10] Synthetic data containing patterns from real data may also be used to protect customer privacy.

- **Continue to use encryption and data masking** as common controls. Select emerging techniques for cases where the need outweighs the challenges (including limited solution availability). Use field-level data transformation for use cases where real data is not required to perform business functions, or when data in its original format has to be shared with a third party. Be aware that sharing detailed, rather than aggregated, data creates additional privacy risk.

- **Employ software-based secure computation** for use cases where a third party must process data and derive conclusions or insights (such as analytics that produce aggregated results).

Data security and privacy are increasingly seen not as distinct and stand-alone efforts, but as elements of a more coordinated D&A governance program. As such, all these efforts should be tied back to the D&A governance program, priorities and outcomes.

For more information, see:

- Three Critical Use Cases for Privacy-Enhancing Computation Techniques

- Machine Learning Training Essentials and Best Practices

Evidence

[1] **The 2022 Gartner CIO and Technology Executive Survey:** This survey was conducted to help CIOs and technology executives adopt business composability as a means to thrive during periods of volatility and uncertainty. It was conducted online from 3 May 2021 through 19 July 2021 among Gartner Executive Programs members and other technology executives. Qualified respondents are each the most senior IT leader (CIO) for their overall organization or a part of their organization (for example, a business unit or region). The total sample is 2,387, with representation from all geographies and industry sectors (public and private). The survey was developed collaboratively by a team of Gartner analysts and Gartner’s Research Data, Analytics and Tools team. *Disclaimer: Results do not represent global findings or the market as a whole, but reflect sentiment of the respondents and companies surveyed.*

[2] **The 2020 Gartner Cloud End-User Buying Behavior Survey:** This survey was conducted to understand how technology leaders approach buying, renewing and using cloud technology. The research was conducted online from July through August 2020, among 850 respondents from midsize and larger ($100 million or more in revenue) organizations in the U.S., Canada, the U.K., Germany, Australia and India. Industries surveyed include energy, financial services, government, healthcare, insurance, manufacturing, retail and utilities. All organizations were required to have cloud deployed. Respondents are involved, either as a decision maker or decision advisor, in new purchases, contract renewals or contract reviews for one of the following cloud types in the past three years: public cloud infrastructure (IaaS), public cloud platform (PaaS), public cloud software (SaaS), private cloud infrastructure, hybrid cloud infrastructure or multicloud infrastructure. Respondents were also required to work in IT-focused roles, with a small subset of procurement respondents. *Disclaimer: Results of this study do not represent global findings or the market as a whole, but do reflect the sentiments of the respondents and companies surveyed*.

[3] Predicts 2022: The Distributed Enterprise Drives Computing to the Edge.

[4] Democratize Data Science Initiatives With Augmented DSML.

[5] Address Both 'Skill' and 'Will' to Deliver Data-Driven Business Change.

[6] **2021 Gartner AI in Organizations Survey:** This survey was conducted to understand the keys to successful AI implementations and the barriers to the operationalization of AI. The research was conducted online from October through December 2021 among 699 respondents from organizations in the U.S., Germany and the U.K. Quotas were established for company size and for industries to ensure a good representation across the sample. Organizations were required to have developed AI or intended to deploy AI within the next three years. Respondents were required to be part of the organization's corporate leadership or report into corporate leadership roles, and have a high level of involvement with at least one AI initiative. Respondents were also required to have one of the following roles when related to AI in their organizations: determine AI business objectives, measure the value derived from AI initiatives, or manage AI initiatives' development and implementation. Disclaimer: Results of this survey do not represent global findings or the market as a whole, but reflect the sentiments of the respondents and companies surveyed.

[7] Cool Vendors in Enterprise AI Operationalization and Engineering.

[8] Self-Service Analytics Governance With Microsoft Power BI.

[9] For example, see:

- Are Your Sensitive Attributes Private? Novel Model Inversion Attribute Inference Attacks on Classification Models, arXiv.
- Algorithms that Remember: Model Inversion Attacks and Data Protection Law, ResearchGate.

[10] Introducing TensorFlow Privacy: Learning With Differential Privacy for Training Data, TensorFlow Blog.

Document Revision History

2022 Planning Guide for Analytics and Artificial Intelligence - 11 October 2021

2021 Planning Guide for Data Analytics and Artificial Intelligence - 9 October 2020

Recommended by the Authors

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

Incorporate Explainability and Fairness Within the AI Platform

A Guidance Framework for Operationalizing Machine Learning

Feature Stores for Machine Learning (Part 1): The Promise of Feature Stores

Feature Stores for Machine Learning (Part 2): Current State and Future Directions

Demystifying the Analytics and BI Space

Solution Comparison for SaaS Analytics and Business Intelligence Platforms

Data and Analytics Governance Approaches for the Technical Professional

Self-Service Analytics Governance With Microsoft Power BI