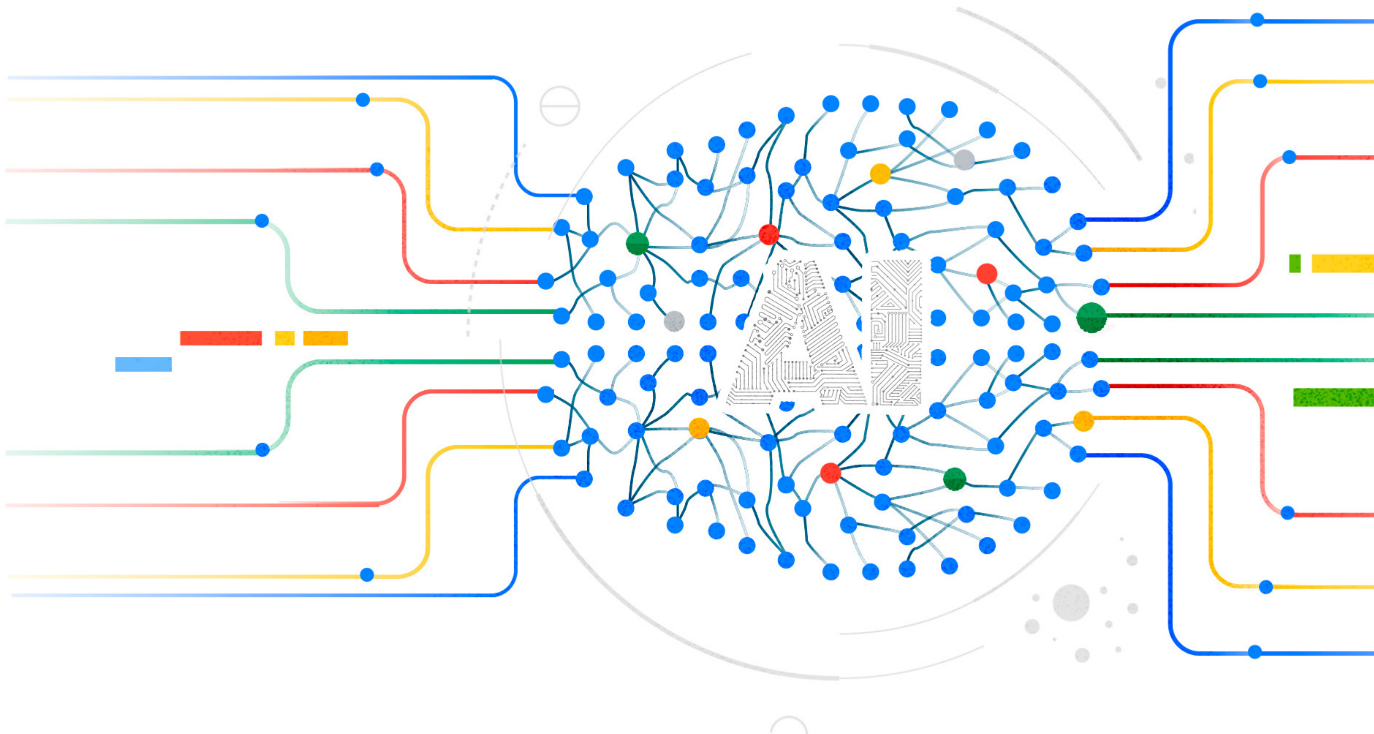
Google Cloud's AI Adoption Framework



Google Cloud

Google Cloud

Contents

Creating value through AI, every step of the way

Companies everywhere are seeking to leverage the power of AI. And rightly so. The smart applications of AI enable organizations to improve, to scale, and to accelerate the decision-making process across most business functions, so as to work both more efficiently and more effectively. It can also open up new avenues and new revenue streams, providing the organization with an additional competitive edge.

Recent advances in technology are making AI more versatile — and all but indispensable

In short, many believe (as we do) that the enterprises that invest in building industry-specific AI solutions today are positioning themselves to be the global economic leaders of tomorrow.

But the path to building an effective AI capability is not an easy one. There are many challenges to overcome. Challenges with the technology to develop platforms and solutions. With the people who will implement and manage that technology. With the data that fuels the technology. And with the processes that govern the whole of it. How do you harness the power inherent in AI, while avoiding any potential missteps?

That's where Google Cloud comes in. Our framework for AI adoption provides a guide to technology leaders who want to build an effective AI capability, one that enables them to leverage the power of AI to enhance and streamline their business, smoothly and smartly. The framework is informed by Google's own evolution, innovation, and leadership in AI, including experience deploying AI in production through products such as Gmail and Google Photos. It is also inspired by many years of experience helping cloud customers[1] — from startups to enterprises, in various industries — to solve complex challenges.[2]

1 Google Cloud Customer Voices Digital Book 2019.
2 Google Cloud named a leader in The Forrester New Wave™: Computer Vision Platforms, Q4 2019.

With Google Cloud's AI Adoption Framework, you'll be able to create and evolve your own transformative AI capability. You'll have a map for assessing where you are in the journey and where, at the end of it, you'd like to be. You'll have a structure for building scalable AI capabilities to create better insights from big data with powerful algorithms across the entire business.

With Google Cloud as your guide, the path to AI is considerably smoother.

Part 1:

Executive summary

Google Cloud

The power of AI

New ideas are brought to market daily — some from established companies equipped with industry experience and capital, some from new companies armed with new technologies and a desire to disrupt. What often bridges the gap between plans and outcomes is a company's ability to effectively make data-driven decisions and execute at scale.

This is precisely what artificial intelligence with machine learning (ML)[3] can do for the companies that know how to take that information and use it well. Machine learning is particularly adept at finding patterns in complex datasets to solve complex problems, including perceptual tasks, such as visual perception and speech recognition. The use cases are both wide reaching and dynamic. Manufacturers, for example, are streamlining their capital expenses by implementing predictive maintenance. Financial institutions are enhancing their risk analysis. Retailers and media providers are personalizing their customer experience. And the travel industry is offering their customers dynamic pricing predictions.

At the same time, academic and industrial advances in AI have resulted in better tooling, smarter algorithms, and more effective implementation techniques covering a wide range of use cases and datasets. These advances — combined with exponential gains in the cost of data storage, compute power, AI-centric hardware, and cloud computing — have democratized AI for industry in an unprecedented way.

AI and ML are increasingly implicated in companies gaining a competitive edge, with direct and attributable business value. In a research study we conducted in partnership with MIT Technology Review,[4] we found that the adoption of ML results in 2x more data-driven decisions, 5x faster decision-making, and 3x faster execution. Enterprises that invest in building industry-specific AI solutions are proven to be better positioned as future global economic leaders. By 2030, companies that fully absorb AI could double their cash flow.[5]

Machine learning truly is reshaping the marketplace.

3 Artificial intelligence is the theory and development of systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, and decision-making. Machine learning is an effective way for building AI systems through automatically discovering useful patterns from data, rather than feeding human-written rules to the system.
4 Machine Learning: The New Proving Ground for Competitive Advantage.
5 Notes from the AI frontier: Modeling the impact of AI on the world economy, McKinsey & Company, September 2018.

Leveraging the power of AI

How do you structure your teams for success? How can you create, discover, share, and manage data assets? How can you leverage native cloud technologies to scale AI? How do you streamline the process of updating and monitoring your ML models in production?

We have a solution for that.

The AI maturity themes

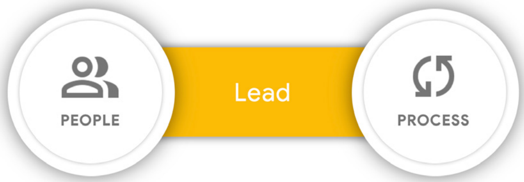
Google Cloud's AI Adoption Framework is anchored in the familiar rubric of people, process, technology, and data. The interplay between these four key areas gives rise to six themes: Learn, Lead, Access, Scale, Automate, and Secure. These themes are foundational to the AI Adoption Framework.

Figure 1: The AI maturity themes

Each of the themes draws its character from the two areas it bridges:

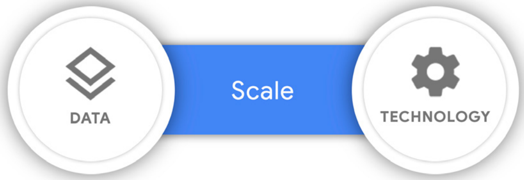
Learn concerns the quality and scale of learning programs to upskill your staff, hire external talent, and augment your data science and ML engineering staff with experienced partners. What data and ML skill sets are required in the organization? What data science and engineering roles should be hired? To what extent do learning plans reflect business needs? What is the nature of the partnership with AI third parties?

Lead concerns the extent to which your data scientists are supported by a mandate from leadership to apply ML to business use cases, and the degree to which the data scientists are cross-functional, collaborative, and self-motivated. How are the teams structured? Do they have executive sponsorship and empowerment? How are AI projects budgeted, governed, assessed?



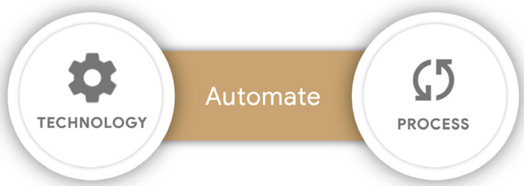
Access concerns the extent to which your organization recognizes data management as a key element to enable AI and the degree to which data scientists can share, discover, and reuse data and other ML artifacts. How is the dataset created, curated, and annotated? Who owns the dataset? Is it discoverable and reusable? Can you share, reuse, and expand trained models, notebooks, and other ML components and solutions?

Scale concerns the extent to which you use cloud-native ML services that scale with large amounts of data and large numbers of data processing and ML jobs, with reduced operational overhead. How are cloud-based services provisioned? Are they on demand or long-living? How is capacity for workloads allocated?



Secure concerns the extent to which you understand and protect your data and ML services from unauthorized and inappropriate access, in addition to ensuring responsible and explainable AI. What controls are in place? What strategies govern the whole? How does an organization establish trust in its AI capabilities so that it is leveraged to drive business value?

Automate concerns the extent to which you are able to deploy, execute, and operate technology for data processing and ML pipelines in production efficiently, frequently, and reliably. What triggers a process? How do you track data lineage? Are your pipelines fault tolerant and resumable? How do you manage logging, monitoring, and notifications?



The AI maturity phases

Your readiness for success in adopting AI in your business is determined by your current business practices in each of these six themes. For each theme, those practices will fall into one of the following phases:

Tactical

Tactical: Simple use cases for AI are in place, but they are typically short-term and narrow. There may be no coherent plan with a strategy for building out to the future.

The focus is on easy adoption, minimal disruption, and quick wins.

Strategic

Strategic: AI is now used to deliver sustainable business value for the organization, with several ML systems deployed and maintained in production, leveraging both ready-to-use and custom models.

A broader vision governs AI adoption. ML is no longer seen as the domain of a special few. Perception starts to move beyond the hype, becoming a pivotal accelerator for the business.

Transformational

Transformational: AI plays a key role in the organization: stimulating innovation, supporting agility, and helping to cultivate a culture where experimentation and learning is continuous and encouraged.

ML expertise has diffused across lines of business. There is a mechanism in place for scaling and promoting ML capabilities across the organization.

The AI Maturity Scale

When you evaluate the AI maturity themes in light of the three phases, the result is the AI Maturity Scale.

	Tactical	Strategic	Transformational
Learn	Self-motivated, isolated learning using online resources Third parties cover the skills gap in the organization No hiring for ML skills	Hiring data science and ML skills Organizing structured and continuous training programs Strategic partner selected to provide consulting and specialized knowledge	Learning by embedding data scientists to the business function teams Hiring data science and ML talent for innovation with industry expertise Partnering to innovate, co-create, and augment technical resources
Lead	AI adoption driven by individual contributors "Heroic" project manager with team budget AI/ML link to business goals not always clear	Creating a centralized cross-functional advanced analytics team to establish common ML patterns and practices Senior executive sponsorship and dedicated budget by C-level for innovative projects Aligning AI efforts with business objectives and priorities	Endorsement and dedicated budget within each line of business Function-specific data science teams with domain expertise, in addition to the centralized advanced analytics team Innovation and research teams
Access	No asset sharing Isolated data islands Building a data lake	Managing an enterprise data warehouse Defining and sharing a unified data model Centralized data and ML asset management	Discovering, sharing, and reusing datasets and AI assets Standardized ML feature stores and datasets Encouraging contributions from across the organization

	Tactical	Strategic	Transformational
Scale	Dedicated hardware for cost control Working with a limited number of small datasets	Using a fully managed serverless data warehouse for ad hoc querying and data exploration Using fully managed serverless data services for ingestion and processing Using fully managed serverless ML services for training and prediction serving	Operating an integrated ML experimentation and production platform Using specialized ML accelerators (GPUs, TPUs) on demand Orchestrating end-to-end data and ML pipelines
Secure	Implementing private networks with primitive IAM accessed and managed by a dedicated team Ensuring privacy through sensitive data classification and obfuscation Enabling data protection through encryption	Implementing principle of least privilege Exploring explainable AI techniques Investing in establishing AI ethics	IAM continuously monitored and improved Considering AI safety and robustness Developing fair ML systems
Automate	Ad hoc, manual data processing and ML model training and serving High-risk changes reviewed and deployed infrequently and manually	Automating (scheduled and event-driven) data pipelines Automating ML training and batch-prediction pipelines Managing logging, monitoring, and notifications	Implementing ML training pipelines with continuous integration and delivery Implementing ML prediction services with continuous integration and delivery Managing ML model registry, ML metadata, and ML artifacts

In each of the themes, you can see what happens when you move from adopting AI approaches ad hoc, to working with them more and more comprehensively across the organization — which means deeper and more consistent training for your people, which in turn means streamlined and updated processes, which in turn drives collaboration and, in time, innovation. The organization transforms.

When AI has been integrated into all parts of your organization, then you are fully harnessing the power it offers to transform your position in the industry. But at every step along the way, adding in effective AI capabilities brings benefits.

Putting it all together

And that's the essence of the framework that Google Cloud uses to guide customers successfully through the process of adopting AI in their organization.

With the framework, you can assess your organization's AI maturity and determine what you'll need to bridge the gap to where you'd like to be. While we touch on the Google Cloud products, you can use this information however you would like: the framework is technology agnostic. We're here to offer further guidance, if that alignment dovetails with your vision. We've worked hard to make AI accessible to all, not only ML researchers and engineers, but to a vast array of customers across industries as well. And our ongoing work in tooling, frameworks, datasets, and models is well documented in the open source community. AI and ML are central to who we are.

Whether or not we accompany you on the journey, however, our framework can help you find your way, from your initial changes all the way to becoming fully AI-powered.

Next steps

Find out more

To dive more deeply into the details of Google Cloud's AI Adoption Framework, see Part 2 of this paper.

To learn about transforming your organization with the cloud, see the Google Cloud Adoption Framework whitepaper. Cloud computing enables organizations to build scalable AI capabilities to create better insights from big data with powerful algorithms.

To learn about setting up a Cloud COE to drive and manage change, see Building a Cloud Center of Excellence.

To expand your ability to formulate ML solutions to solve real-world problems, to discover relevant ML use cases, and to start an ML project, consider the Machine Learning for Business Professionals course.

Work with Google experts

Google Cloud offers a range of professional consulting services and workshops, which enable you to work directly with Google's experts to discover, assess, deploy, and upskill in ML. If you need help thinking through some of the questions explored in this paper, get in touch with your Google Cloud representative or contact us.

Part 2:

Technical deep-dive

Google Cloud

The AI maturity phases

There are three natural phases to AI maturity: tactical, strategic, and transformational. Each organization's current approaches to AI will fall within one of these three phases. Each phase offers opportunities for further exploration and development.

Tactical

Characteristics

Organizations at the tactical phase are exploring the potential of AI to deliver in the short term. Use cases tend to be narrow, and developers are typically leveraging exploratory data analysis (EDA) tools and ready-to-use AI and ML services for proofs of concept and prototyping, for example, using a prebuilt computer vision service to detect printed and handwritten text, or using descriptive analytics to create a customer segmentation model.

At this phase, organizations are aware of the promise of advanced analytics,[6] but ML can be seen as unattainable, and for this reason, complex problems are outsourced. Since the ML projects that exist do so through individual efforts, those projects might not be aligned with the organization's business goals, and so even the most successfully deployed ML projects may have only limited business impact.

At this phase, too, there may be no process to scale solutions consistency, nor the skill set to solve complex analytics problems.

Opportunities for growth and advancement

At this phase, organizations can benefit substantially from better access to data. Integrated, cleaner, and fresher data leads, in turn, to actionable insights better tailored to your needs, which translates to sharper and more informed decision-making. Such swift improvement is immediately validating, and can help to paint a picture for stakeholders of the power of ML.

6 Broadly, we think of advanced analytics as the ability to use data science and machine learning to create insight for more effective decision-making.

Organizations at this phase should look to develop the foundational skill set for core data wrangling and descriptive analytics. To drive effective collaboration across an organization and to innovate with data from many sources, a key next move is to start bringing together siloed data from the many lines of business into a central, unified data lake, but with decentralized access. With a central data lake, it's easier to derive insights from unstructured data and easier also to perform batch integration for reporting. Data is available through centralized tooling, but teams do not lose autonomy of access management or data ownership.

Organizations at this phase can benefit, too, from the move toward use cases that are feasible and designed to deliver business value, but that begin to tackle increasingly complex problems.

And the quality of the analytics solutions improves when the organization focuses on developing core standards:

- A set of unified standards and technical practices that ensures the security of all data access and protection
- A set of common principles and procedures that prevents building any ML that may harm your brand (such as, for example, inadvertently building socially biased models)

Strategic

Characteristics

Organizations at the strategic phase are focused on delivering sustainable business value, with several ML systems deployed and maintained in production that leverage both ready-to-use and custom models. ML is no longer seen as the domain of a select few, but is instead in the process of becoming a pivotal accelerator for the business.

At this phase, organizations typically require a degree of centralized coordination and so they will often create an advanced analytics[7] team with the right skill set to build solutions for various ML use cases across business functions. This team can be part of a broader Cloud

7 A centralized team, spun out from the core Cloud COE. This team should have engineering skills in data, analytics, and ML; and they should operate as the hub in a hub-and-spoke pattern, working closely with other ML and data science teams in the various lines of business.

Center of Excellence (COE),[8] where engineers and data scientists with domain expertise work side by side, drawing on business subject matter experts, as required by the use case. The centralized team aims to achieve several strategic goals:

- To ensure consistent standards and governance
- To assess the feasibility of new use cases
- To weigh in on new data collection
- To remain current on which products can be bought rather than built
- To help build ML solutions for different product teams
- To prioritize scarce ML resources around key business problems

As the surrounding organizational fabric matures (for example, increased awareness of AI capabilities, increasing demand for utilizing those capabilities, increased consistency of tooling and approach), it is more feasible for ML to be used independently in different products and business areas.

At this phase, teams have skills in data wrangling and descriptive and predictive analytics; they use existing frameworks, methods, and techniques to solve a variety of use cases; and they often deploy custom ML models in production. Teams retrieve their information from a single source: an enterprise data warehouse (EDW), with complete, consistent, correct, and concurrent data. Extract, transform, load (ETL) and extract, load, transform (ELT) routines are automated and scalable.

Opportunities for growth and advancement

At this phase, organizations can benefit substantially from developing an AI capability that is more tailored to their business model and their distinct business needs. And from adding automated processes to their customized models.

Technically, the next key moves are about protecting data quality, preventing ML models from going stale, and enabling valuable solutions to be production ready. In addition, common practices and guidelines are established for building secure, ethics-complaint AI solutions.

8 No two Cloud COEs are quite the same. Data science skills are an important part of the picture. However, as the level of maturity deepens and the focus of the Cloud COE evolves, it can make increasing sense to establish a distinct advanced analytics capability. You can read more about what makes an effective Cloud COE in this whitepaper.

Organizationally, the centralized advanced analytics team with technical ML skills starts to work smoothly in conjunction with various functional teams with specialized domain expertise.

Transformational

Characteristics

Organizations at the transformational phase are actively using AI to stimulate innovation, to support agility, and to cultivate a culture where experimentation and learning is ongoing. Models are built and deployed from a unified ML platform, making ML accessible to everyone in the organization.

At this phase, organizations typically model a hybrid approach to AI, with functional or product-specific AI teams embedded into the broader product teams — supported by the advanced analytics team, which might become its own hub or center of excellence. The central team enables a mechanism for scaling and cultivating these capabilities across the organization: for example, implementing a mechanism for secondment or talent rotation, whereby business or product experts are immersed in ML techniques to learn hands-on. Over time, ML expertise diffuses across lines of business. The role of the centralized advanced analytics team becomes more confined to establishing common patterns and best practices and to providing standard tools and libraries for accelerating ML projects. By contrast, the data science teams embedded in product groups or lines of business are responsible for building their function-specific ML models. This division of responsibility drives consistency in building high-quality, technical solutions with real business impact.

All teams are empowered through a platform that enables access to useful datasets, prepared features, reusable components, and trained models. That platform is supported by scalable and serverless compute for batch and online data ingestion and processing, distributed ML training and serving, access to databases with specialized storage and querying capabilities, and hardware accelerators.

Opportunities for growth and advancement

At this phase, organizations can benefit substantially from focusing on best practices, ensuring that AI practices are responsible, that they are based on sound principles, and that AI systems are safe and robust.

The ML platform is supported with tools for continuous integration, continuous training, and continuous model serving and monitoring. Building and maintaining such a platform is a shared responsibility between ML and software engineers with skills in infrastructure, DevOps, and SRE.

Organizations in this phase often assemble teams to conduct cutting-edge research and present at academic conferences and AI events. To do this, they focus their research and innovation in areas where they have unique capabilities, either in terms of domain understanding or data availability, so that they can build a sustainable advantage over time.

Besides the competitive edge that such a sustainable advantage provides, the in-house research capabilities can add to an organization's employee value proposition, becoming a point of differentiation to attract the best talent.

Google Cloud smart analytics and AI

Google Cloud offers a wide set of products that enable organizations to accelerate their AI journey, ranging from ready-to-use AI services to an integrated data science development environment for building custom solutions.

Prebuilt APIs. Prebuilt AI and ML APIs offer ready access to easy-to-use building blocks that require no in-house ML expertise. These APIs address various perceptual tasks: Vision API, Video API, Natural Language API, Speech-to-Text API, Text-to-Speech API, and Translation API.

Cloud AutoML. Cloud AutoML services allow developers, with limited ML expertise, to build high-quality custom models specific to your business needs. For language, you can use AutoML Natural Language and AutoML Translation. For sight, you can use AutoML Vision and AutoML Video. For structured data, you can automatically build and deploy models using AutoML Tables.

AI solutions. These ready-to-use AI solutions enable you to run your business faster and smoother. Contact Center AI positions you to deliver exceptional customer service, while with Document AI you can easily extract insights, information, and knowledge from enterprise-wide text sources.

AI Platform. Through serverless, scalable training and serving capabilities for custom ML models, AI Platform makes it easy for data scientists to take their ML projects from initial concept to production and deployment, quickly and cost-effectively. With AI Platform, you can use Tensorflow Enterprise, which offers enterprise-grade support, performance, and managed services for your AI workloads, and Explainable AI, which lets you interpret models with confidence.

Data management. Cloud Storage and BigQuery provide you with a powerful foundation for evolving from siloed data to a centralized data lake, driving enterprise-grade data management. BigQuery ML also enables you to build powerful ML models without moving any data out of your data warehouse. In addition, Google Cloud provides a variety of data services for various operational and analytical workloads, including Cloud SQL, Cloud Spanner, Cloud Bigtable, Firestore, and Memorystore. With Cloud Pub/Sub, Dataflow, Cloud Data Fusion, and Dataproc, you can implement batch and real-time data ingestion and processing at scale.

Automation and instrumentation. Google Cloud provides tools to manage your ML systems in production at scale. With AI Platform Pipelines and Cloud Composer, you can orchestrate and automate data and ML pipelines. Cloud Build and Container Registry enable you to build and deploy custom ML systems.

The AI Maturity Scale

When you evaluate the six AI maturity themes in terms of the three AI maturity phases — where each phase is descriptive of how a given organization is currently functioning in that theme — you get the AI Maturity Scale.

In each of the themes, you can see what happens when an organization moves from experimenting with ML tools and technologies, to working with them more strategically, to building a transformational AI capability.

Learn

Bridging People and Technology, "Learn" concerns the quality and scale of learning programs to upskill your staff, hire outside talent, and augment your data science and ML engineering staff with experienced partners.

An organization's maturity in the Learn theme reflects that organization's ability both to keep up with the latest advances in ML and to evolve AI capabilities toward solving ever more complex business problems.

Tactical maturity

Learning about data science and ML is self-motivated by a few members of your IT and data teams, using publicy available online resources. While this approach is useful for developing technical skills around AI and ML in general, the training courses don't follow a planned learning path that is aligned with your organization's current and future needs. Meanwhile, a few members may be working on ML prototypes, but when faced with a compelling use case where ML is needed, your organization tends to turn to third-party consultants, rather than hiring to fill the skill set gaps in-house.

Strategic maturity

You're starting to build greater AI capability in-house by hiring the required data science and ML engineering skills. As there are plenty of job titles and descriptions in the field of ML,[9] your focus is on the skills and seniority levels that the organization needs.

9 For a breakdown of the top 10 role profiles needed in a data science team, see this article from Google's Head of Decision Intelligence.

To ensure that your team is well qualified to use the necessary tools and technology, you're developing clear learning paths with certifications and aligning those paths with your organization's priority use cases. You've recognized that the field of AI is constantly changing, and you keep up with those changes by following relevant events, conferences, and experts in the field. You strive to do this systematically, providing employees with training opportunities and requirements to fulfill.

In addition, you choose a strategic partner for advisory, consultation, and program management, someone who also provides specialized knowledge in specific AI and ML use cases (for example, chatbots and conversational apps). You use partnerships to help accelerate your AI adoption by taking advantage of best practices and other organizational lessons that experienced AI and ML subject matter experts bring to the table.

Transformational maturity

You aim to head-hunt well-known AI and ML talent to lead increasing research and innovation efforts — while also gaining a greater reputation for effective talent development. Both industry expertise (like finance, healthcare, telco) and domain expertise (churn prediction, credit risk assessment, medical diagnostics) become increasingly important for hiring great candidates. This talent and experience is essential for solving groundbreaking domain-specific problems with AI.

For learning and development, you encourage peer-to-peer and community learning, creating an internal knowledge base, wikis, and tailored courses and learning tracks. You also organize internal conferences to showcase AI applications and promote knowledge sharing. Enabling rotations across business functions ensures not only tighter collaboration between members of different teams, but also an exchange of experience and learning on the job. This mechanism for diffusing and developing talent across your organization can become part of your employee value proposition, helping to attract the best talent. This can, in turn, help start a virtuous talent cycle.

At this maturity phase, a partnership can evolve to become a co-creation relationship, where both you and your partners' ML researchers and engineers work together to break new ground and solve cutting-edge problems in the field. Such research also becomes part of your value proposition for employees, and a competitive advantage in the market.

Lead

Bridging People and Process, "Lead" concerns the extent to which your data scientists are supported by a mandate from leadership to apply ML to business use cases, and the degree to which the data scientists are cross-functional, collaborative, and self-motivated.

An organization's maturity in the Lead theme reflects the effectiveness with which that organization will adopt AI in line with business priorities.

Tactical maturity

AI adoption is driven across your organization by individual contributors or a manager within one project team, with little or no executive sponsorship. There is typically limited collaboration between data scientists in different teams, and there is often an unclear line of sight between ML initiatives and the organization's strategic goals. You're identifying the right use cases for applying ML and also pursuing executive sponsorship. The focus is to explore the power and prove the value of ML to stakeholders, while the funding is part of the project team's budget.

The first successful use case

The first step to realizing value from AI is to identify the right business problem and a sponsor committed to using AI to solve that problem. To ensure alignment, start with your organization's business strategy and key priorities. Identify the right business priority use cases to address with AI, and find the senior executive to own it. Work with their team to get their buy-in and sponsorship. AI projects are more likely to be successful when they have a senior executive sponsor who will champion them with other leaders in your organization.

Strategic maturity

Your senior executives support AI capabilities and projects to deliver value to several business functions, an approach that amounts to a competitive advantage. You've begun to establish a mechanism for standardizing practices and guidelines and sharing accumulated knowledge. To accomplish this, you may have put into place a centralized advanced analytics team, with a dedicated budget established by a set of C-level executives with a technology agenda.

The advanced analytics team works on delivering prioritized projects or short-term consulting to other teams. In addition, they proactively evangelize and advocate their AI capabilities to other lines of business, clarifying how AI can address the various use cases. What members of this team all share is a forward-thinking and self-motivated interest in using AI to deliver business value. Equally important, they share an understanding of the importance of scaling this impact by growing and embedding data science capabilities in functional teams across the business.

Transformational maturity

Each functional team has its own data scientists with the right domain-specific expertise and support from a technical project manager with the right organizational influence. This ensures that models can be successfully deployed and scaled in a timely manner. This team is supported by a centralized advanced analytics team, which provides standards and best practices, as well as the tooling and platform for implementing ML projects. Having both a centralized team and embedded data scientists in the business functions ensures technical standards and business alignment, respectively.

With standards firmly in place, and the structure to support further investigation, research and innovation activities flourish. This work is sponsored by senior executives both to solve cutting-edge business problems and to establish the organization's name as a market leader. In addition, to stimulate innovation, you facilitate competitions that reward employees for sharing ideas on how ML can solve key problems.

Your C-level executives act passionately, continuously demonstrating active sponsorship for AI projects and encouraging contributions from across the organization. The leadership functions effectively for an AI-driven environment, fostering a culture of blamelessness and open communication channels, where sharing failures openly is encouraged and mistakes are treated as opportunities for improvement. Experimentation drives innovation and market success. Teams are free to try out many different ideas with the goal of failing faster, failing better, and learning from the experience to improve and innovate.

Access

Bridging People and Data, "Access" concerns the extent to which your organization recognizes data management as a key element to enable AI and the degree to which data scientists can share, discover, and reuse data and other ML assets.

An organization's maturity in the Access theme reflects that organization's ability both to accelerate an ML project and to improve the adoption of the AI capabilities across different teams and functions.

Tactical maturity

Each team usually manages its own data island and ML assets,[10] with no sharing among other teams. Even within the same team, it might not be possible to reuse assets, as the technology and processes have not been standardized yet.

In a bid to update your data management and access strategy, you are looking into building a unified data lake, containing the raw (structured and unstructured) data feeds, for example, using Cloud Storage. This data lake can then form the foundation for collecting, curating, and sharing data across the organization, that is, for building managed information stores.

Strategic maturity

You recognize data as a vital enterprise asset that fuels AI, and so you've improved both the utilization and the management of data, working to keep it safe and useful. You likely have invested in an EDW, which provides a unified data model, with an integrated, consistent information store for various business functions. You have also made data quality management a business priority to ensure the right foundation for analytics and AI. The EDW, which can be implemented using BigQuery, empowers data analysts in their reporting, data scientists in their exploratory data analysis and ML experimentation tasks, and other business intelligence (BI) activities.

10 *ML assets* refers to notebooks, trained models, reusable components, and code snippets. From the strategic phase onwards, solution templates are another asset used to create consistency.

In addition, your organization is starting now to put together a centralized repository for ML assets and to develop tools for annotating and categorizing data, for example, using Data Catalog. Such assets are created and curated by a centralized advanced analytics team, which creates and provides access to such assets on demand. Although each team maintains their own assets, there is typically limited discoverability and shareability outside this team.

To facilitate AI adoption, you're looking now at ways of better handling sensitive data and of better protecting data overall.

Transformational maturity

You are actively working to develop the technology and processes to allow data scientists to create, share, discover, and reuse data and ML assets across different teams and functions, which accelerates the launch of new ML-based services. And your teams are using specialized data storage tools (such as Cloud SQL, Cloud Spanner, Cloud Bigtable, Firestore, and Memorystore) based on the requirements for data volume, structure, and consistency and on read/write workloads.

To help standardize the definition, storage, and access of features for training and serving ML models, you have a feature store as a unified repository, which provides functionality for registering new features to be discovered and used by data scientists. This approach prevents inconsistency between the data used in training the model and the data subsequently used in serving the model. Data scientists from across the organization are encouraged to use and contribute to the feature store.

Your AI efforts are typically driven and maintained by a centralized team that acts as a COE, with the goal of driving consistency by defining the standards, the procedures, and the templates to create different assets, rather than actively creating them.

Scale

Bridging Data and Technology, “Scale” concerns the extent to which you use cloud-native ML services that scale with large amounts of data and large numbers of data processing and ML jobs, with reduced operational overhead.

An organization’s maturity in the Scale theme reflects that organization’s ability to scale data processing and ML workloads.

Tactical maturity

To control the cost of exploring and experimenting with ML and predictive analytics, you’ve got dedicated hardware or cloud compute instances for a few data scientists. Both the size and the lifespan of these compute instances are dictated by the IT operations team. In addition, your team of data scientists work with small, offline datasets for exploratory data analysis and processing, using simple data wrangling tools and ready-to-use AI services.

Data processing and ML training are done locally, using the dedicated data science VMs, which means limited compute power. You are actively exploring the potential of AI by finding simple, yet interesting, use cases that demonstrate proof of value (POV), so as to increase the awareness across the organization of the promise of advanced analytics.

Strategic maturity

You’ve enabled an enterprise-wide, fully managed cloud data warehouse for ad hoc analytical queries, for example, using BigQuery. Data is now ingested into your cloud data warehouse from various sources spread across the organization in different systems. Such a scalable data warehouse enables the data scientists across your organization to perform complex analytics and information retrieval on a large amount of data in a timely fashion.

Data ingestion and transformation are performed using a serverless, fault-tolerant, autoscaling, parallel processing service for handling large batch and streaming ETL/ELT workloads to populate the data warehouse and prepare ML datasets. You use tools such as Dataflow, Dataproc, and Cloud Data Fusion to combine and process both structured and unstructured data at scale so as to gain actionable insights.

Your data scientists build custom ML models tailored to your data and business needs, using advanced ML tools and frameworks such as TensorFlow Enterprise. Long-running ML training jobs are performed using cloud serverless platforms for distributed training and automatic hyperparameter tuning — with no dedicated infrastructure — for example, with AI Platform Training.

Using cloud-native technologies helps your data scientists and engineers focus on data processing and ML modeling activities, enabling them to run numerous jobs at scale on large datasets with no infrastructure operations overhead.

Transformational maturity

You've invested in building online complex event processing and stream analytics pipelines, powered by ML, to achieve (near) real-time operation optimization and decision-making. And you're engaged in architecting an integrated ML platform, for experimentation and productionalization, to serve all the data scientists and ML engineers in different teams. This approach provides access to standardized tools, services, compute, processes, and best practices, both from internal and external sources. The goal is to give data science teams a solid starting point when approaching new ML projects and to provide them with the support to carry them through to completion of those projects.

You use ML accelerators (like GPUs or Cloud TPUs[11]) at scale to train complex ML models, using a large amount of data in a short time. In addition, your data engineering team creates metadata-driven data processing templates to configure and deploy new workflows without coding. End-to-end data and ML pipelines are orchestrated and automated with the required instrumentation.

11 A TPU is a Tensor Processing Unit, a chip specifically designed to be faster and more power-efficient than GPUs for certain machine learning tasks.

Secure

Bridging Data and Process, "Secure" concerns the extent to which you understand and protect your data and ML services from unauthorized and inappropriate access, in addition to ensuring responsible and explainable AI.

An organization's maturity in the Secure theme reflects that organization's ability to ensure that their data is appropriately protected, catalogued, encrypted, and guarded from exfiltration, in accordance with ethical AI principles and practices.

Tactical maturity

Your Cloud Identity and Access Management (IAM) policies predominantly rely on the convenience of project-level primitive roles (Owner, Editor, Viewer) rather than following the principle of least privilege. Default permissions allow for any user to create projects and billing accounts. Cloud IAM permissions are not continuously monitored, and the admin activity and data access logs are not systematically audited. Service accounts can be created freely, and private keys for service accounts are not automatically rotated.

You are constantly identifying sensitive data, defined as data containing personally identifiable information (PII), in order to govern its protection, and the user's privacy.

Data classification and protection

Poor data protection and sensitive data handling are big blockers for AI adoption, as they can result in breaches or other issues and, consequently, reputational damage or regulatory sanctions. Data classification and loss prevention, for example, using Cloud Data Loss Prevention, can help you identify and handle sensitive data through encryption, removing, masking, or coarsening. In addition, establishing a governance policy and adhering to regulatory compliance are important considerations at this phase. Another key aspect is to track data lineage and its chain of custody, to ensure that the ML models informing business decisions are based on untampered data.

Strategic maturity

Your Cloud IAM policies reference a much more granular set of predefined roles, rather than the coarse primitive roles. In addition, your AI capabilities are supported by clear governance and decision-making responsibilities. Building on this foundation, you've also provided a verification route for decisions guided by the principles of AI,[12] ensuring that the desired level of trust in AI is maintained throughout the organization. This impacts how confident your organization is in relying on AI to influence decision-making, while avoiding potential biases in human-centric[13] use cases.

Explainable AI

Explainable AI methods and techniques render AI solutions and outputs intelligible to human experts. This approach mitigates the concept of "blackboxing" in ML, where it is hard to explain specific decisions from an ML model. Your AI-enabled business may impact, or even redefine, many areas of society. The usefulness and fairness of these AI systems will be gated both by their transparency and by your ability to understand, explain, and control them. Activating the right Google tools and capabilities, such as What-If tool, Fairness Indicators, and Explainable AI, will not only speed up and secure the AI journey, it will also enable your organization to stay compliant with current regulations, and to react quickly when they change.

Transformational maturity

You aim to have a comprehensive understanding of the contents of all your data stores, so as to obtain the threat profiles necessary for designing more effective security and data governance models, models that consider scenarios of both unauthorized and inappropriate access. All Cloud Admin activity and data access logs are regularly audited, while automatic alerts have been configured to watch for patterns that match your threat profiles. Cloud IAM permissions and firewall rules are continuously monitored and corrected.

12 We recognize that technologies that solve important problems also raise important challenges that we need to address clearly, thoughtfully, and affirmatively. Artificial Intelligence at Google: Our Principles sets out our commitment to develop technology responsibly and establishes specific application areas we will not pursue.
13 People + AI Research (PAIR) is a multidisciplinary research and development team at Google that explores the human side of AI by working with diverse communities. PAIR released a guidebook to help user experience (UX) professionals and product managers follow a human-centered approach to AI.

Your data governance is streamlined, for example, by using automated workflows to rapidly validate new use cases against your AI principles — allowing greater focus and discussion time on the edge cases. A team specialized in AI safety and robustness works to improve the reliability and generalizability of ML models, recognizing the importance of well-calibrated uncertainty and protecting against adversarial attacks.[14]

14 *Adversarial attacks* refers to how ML models can be vulnerable to inputs maliciously constructed by adversaries to force misclassification.

Automate

Bridging Technology and Process, "Automate" concerns the extent to which you are able to deploy, execute, and operate technology for data processing and ML pipelines in production efficiently, frequently, and reliably.

An organization's maturity in the Automate theme reflects the ability of its AI systems to adapt to changes in data and the environment, which provides the means for making timely data-driven decisions.

Tactical maturity

The technologies adopted across your organization are easy to use, such as data wrangling tools, sheet-based data visualization, and ready-to-use AI services. The process of using such technologies to build AI systems is usually manual in every step — from data analysis and preparation, to model training and validation. The process is also driven by experimental code that is written and executed interactively (for example, using Jupyter Notebooks) by data scientists until a workable model is produced.

Data analytics and ML are ad hoc, with no automation, meaning that insights are not delivered regularly in a timely manner. As is appropriate when directed by this manual, data scientist–driven process, your models are rarely changed or retrained. However, such a manual process prevents you from working with a large number of models or with frequent updates to these models.

Strategic maturity

You automate data processing and analytics pipelines (for example, using Cloud Composer) either on a recurrent schedule (for example, using Cloud Scheduler) or, to tackle critical events, on a trigger (like data anomaly detection). You've started to standardize processes and to consolidate technology to govern the flow of data from one step of the data lifecycle to the next, that is, from data ingestion and transformation to data analysis and reporting. This strategy has led to increased agility and decreased cycle times, while reducing data defects, giving developers and business users greater confidence in the insights of analytics for timely decisions.

You are using ML models in production. However, the models often break when deployed in the real world because they fail to adapt to changes in the dynamics of the environment or to changes in the data that describes the environment. For this reason, to regularly update your models with new data, you deploy end-to-end continuous training[15] ML pipelines.

In addition, your trained models are integrated with your automated ETL/ELT routines to perform batch scoring. For online use cases, your models are deployed as services to be used in apps and stream-processing pipelines for serving real-time predictions.

Continuous training

For reliable continuous training (CT) automation, the validation and quality control aspects of your ML training need to be backed into your pipelines. That is why you need data validation and model validation steps in the beginning and the end of the pipeline, respectively. Such validation steps act as the "gatekeepers" to your model training. Data validation makes sure that the schema and data types of the new data for retraining the model are as expected. Model validation makes sure that the produced model meets the required predictive performance for deployment, for example, that it outperforms the current model in production and that it meets the fairness measures — if any have been specified.

Quality control in machine learning

When a deployed ML model produces bad predictions, the poor ML quality may imply a wide range of problems, including the presence of the bugs typical of any program, but also data skews and anomalies, and the absence of proper procedures for evaluating models after training and for validating models before deployment. Testing is required not only for development, but for deployment and production as well. Instrumentations like logging, monitoring, and notifications are critical to maintain the system's health and operate it reliably. See Testing and Debugging in Machine Learning and What's your ML test score? A rubric for ML production systems, by Google's ML experts.

15 *Continuous training pipelines* are automatically executed through schedules or event-based triggers that you can set up with tools like Cloud Scheduler or Cloud Functions. You can orchestrate your training workflows using AI Platform Pipelines.

Transformational maturity

Your goal is to develop an ML engineering culture and practice that unifies ML system development (ML) and ML system operations (Ops). MLOps strongly advocates automation and monitoring at all steps of ML system construction — from integration, testing, and releasing to deployment, model serving, and infrastructure management. With this approach, you are working towards shorter development cycles, increased deployment velocity, and more dependable releases that are also in close alignment with your business objectives. This is crucial because you have a large number of models in production that require frequent updating both with new data to capture the emerging patterns and with the new implementation of state-of-the-art ML ideas.

Automatic detection of skews and anomalies in the data through regular data validation jobs help data scientists to monitor the performance of their model. In addition, you have an automated A/B testing system to evaluate the effectiveness of a newly released model service.

Continuous integration and delivery

In ML, continuous integration (CI) covers testing and validating code and components, as well as testing and validating data, data schemas, and models. Continuous delivery (CD) covers two aspects: 1) deploying the CT pipeline, which produces a newly trained (and validated) model every time it is executed, and 2) deploying the trained model as a prediction service to the serving infrastructure. Services like Cloud Build, Container Registry, Model Registry, and ML Metadata are required to streamline the CI/CD/CT of production ML systems, while maintaining reproducibility, resumability, and reliability.

Additional technical resources

To get started with ML on Google Cloud, take a look at the following guides, courses, articles, and product documentation.

- Rules of Machine Learning offers an overview of best practices for ML engineering.
- Machine Learning Crash Course provides a fast-paced, practical introduction to ML.
- Data Engineering with Google Cloud Professional Certificate provides the skills you need to advance your career in data engineering and recommends training to support your preparation for the industry-recognized Google Cloud Professional Data Engineer certification.
- Machine Learning with TensorFlow on Google Cloud teaches you how to use Cloud AI and big data products to build production-grade ML systems.
- TensorFlow in Practice Specialization helps you to discover the tools that software developers use to build scalable AI-powered algorithms in TensorFlow, a popular open-source machine learning framework.
- MLOps: Continuous delivery and automation in ML discusses techniques for implementing and automating continuous integration (CI), continuous delivery (CD), and continuous training (CT) for ML systems.
- Google Cloud Smart Analytics describes Google Cloud's fully managed serverless analytics platform, which can be leveraged to empower the business while eliminating constraints of scale, performance, and cost.
- Google Cloud AI solutions presents and enables organizations to use high-quality, scalable, continuously improving, and fully managed AI solutions.