



CIO

InsideNow

How the cloud can be a force multiplier for AI and analytics solutions

Cloud-enabled AI brings various advantages for organizations throughout the ML pipeline, from development, training, management until deployment.

Authors

Jean-Pierre Maissin - Partner - Strategy, Analytics & M&A Leader - Deloitte

Nicolas Griedlich - Director - Artificial Intelligence & Data - Deloitte

Jordane Le Ponner - Manager - Artificial Intelligence & Data - Deloitte

Anke Joubert - Manager - Artificial Intelligence & Data - Deloitte

Hajar Khizou - Analyst - Artificial Intelligence & Data - Deloitte

Published on 6 May 2021

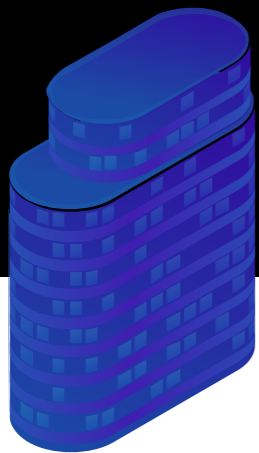


The pace of innovation is moving at a speed never seen before. With this ever-growing evolution, companies need to enhance their activity with new technologies and concepts such as cloud, DevOps, machine learning (ML), artificial intelligence (AI), internet of things (IoT), and many others. At the heart of this evolution comes cloud ML. It minimizes businesses' need for a large team of data scientists to develop ML programs and it enables data engineers to build innovative products and services, while giving flexibility and cost and scale benefits.

AI and ML are enabling organizations across multiple industries to navigate various business challenges, including cost optimization, revenue/margin gain, risk management, operational efficiency, and more. While ML is powerful on its own, when combined with cloud technologies it allows businesses to innovate at speed and scale.

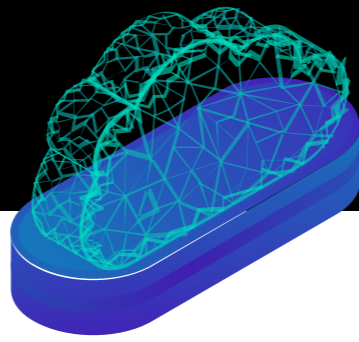
TO CLOUD OR NOT TO CLOUD?

Considering the cloud for analytics and AI solutions is a decision that most organizations need to make. In Deloitte's 2020 *State of AI in the Enterprise* survey, 93% of over 2,500 executive respondents said they were using cloud-based AI capabilities ^[1]. Using cloud technologies is enabling AI innovation at its fullest and accelerating its adoption in various ways including scalability, agility, and increased access.



ENABLING ML AT SCALE ACROSS THE ENTERPRISE

ML models require variable cloud service levels to scale up and down depending on workload.



DRIVING AGILITY AND INNOVATION

Cloud ML offers additional ML features such as pre-trained models, frameworks, and general-purpose algorithms as accelerators for data teams.



ENABLING ACCESS TO HIGH VOLUMES OF DATA

Cloud offers large-scale data stores and the necessary compute resources that can ingest, process, and store high-velocity real-time streaming data, as well as high-volume batch data, needed for artificial intelligence pipelines.

Additionally, using cloud computing for your AI solutions enables you to reach the highest price-service ratio as it comes with different pricing models that can result in significant cost savings. Three typical cost models include ^[2]:

- **On-demand costing**
Users pay when they are utilizing cloud resources. It is ideal for short-term or seasonal workloads.
- **Reservation-based costing**
Users can reserve capacity on their cloud vendors' infrastructure. This model is most suitable for long-term, steady workloads.
- **Preemptible/spot costing**
Offers the user excess capacity within the cloud provider datacenter at massive discount, with the drawback of varying capacity availability. It offers high potential savings and is ideal for modernized and fault-tolerant applications and big data processing.

Given these opportunities, we foresee that many technology and talent barriers will be eliminated by Cloud ML. It is expected that the cloud AI market will reach \$13.1 billion by 2026, registering a compound annual growth rate (CAGR) of 20.3% during the period of 2021-2026 ^[3].

HOW TO GET STARTED

Cloud-enabled business models enable organizations to embrace new channels for a high-level customer experience by differentiating their strategy and solutions from their current competitors. However, they need to ensure that they don't replicate what they have on-premises in terms of storage and developments to get the most value from their transformation. Ultimately, cloud and on-premise infrastructures should complement rather than compete.

While choosing your next cloud platform, there are some essential features to consider:

- **Compliance with standards and quality frameworks**

Choose a cloud platform that can help you adhere to the compliance standards of your industry and organization. Providers that align with approved standards and quality frameworks show adherence to industry best practices.

- **Alignment with current environment**

Understand the commonalities between the cloud platform features, your existing technology stack, and your technical requirements including any complex dependencies.

- **Support for your cloud objectives**

Define your strategic short-, medium-, and long-term objectives. Have a view on the provider's roadmap of service development and ambitions to continue innovation in the future.

- **Ensure data privacy and security**

Business data can include sensitive personal, financial, health, and other identifiable information that falls at the core of organizations' operations. It is important to examine your cloud provider's capability to create a complex and secure data center and robust data ethics framework.

Given the spike in data and usage of various ML and AI services, cloud providers continuously enhance their application programming interface's (API's) and services' quality. Many tech giants provide cloud platforms that come with a wide range of features. Once you have selected your preferred provider, the next step would be to know how to most efficiently use the different capabilities. Two simple questions can help guide organizations to the right cloud services to support their use cases.

START

Define analytics use case
Choose analytics use-cases that would drive value for your business

Ask the right questions
These questions can help to use the cloud to answer to your analytics needs

"Are there suitable existing APIs?"

"Is control over hardware/ software layers a prerequisite?"

Existing APIs
Example: Use conversational AI to develop an intelligent assistant

ML as a service
Example: ML at the service of digital transformation

Infrastructure as a service
Example: Visual intelligence applied to car damage recognition and to face mask detection

The following Deloitte case studies illustrate the different, main approaches to leverage these cloud platforms in implementing your next AI solution:

LEVERAGING HIGH-LEVEL APIs OFFERED BY THE CLOUD PLATFORM

Applied use case on conversational AI

Using cloud APIs for text analytics has been applied to develop a conversational AI solution that can assist and guide users on a search portal. Cloud platforms offer APIs with a high-level of abstraction that allows you to perform multiple tasks (such as computer vision and natural language processing (NLP)) without delving deep into the technical details and API structure^[4]. The project's focus was to increase accessibility to sensitive users, improve document access, and guide users in their search journey, with the goal being to have more relevant search results. The customer's current cloud platform offered text analytics and other ML APIs that could be used to accelerate the development of the search portal chatbot.

As the consumption of these APIs rises, the quality of cognitive services increases, allowing organizations to deliver data and analytics products with enhanced ML models. It takes various steps to develop a solution using ML and deep learning (DL).

Each step can be labor-intensive and require data scientists to make multiple decisions. A speech API was used to integrate speech understanding into the assistant that was being developed, which allowed the project team to build more models with fewer steps. By leveraging further cognitive APIs related to language understanding, the intelligent assistant could understand both English and French text and voice inputs and respond accordingly to guide users in their search journey. Apart from these text analytics APIs that allow NLP-related tasks off the shelf, many cloud providers also offer other pre-trained AI services through APIs. This can cover computer vision, recommendation systems, forecasting, and other functionalities.

It is clear that the cloud enables AI for real-life problem solving through the use of APIs. This use case shows that applying cloud APIs for AI projects can be impactful. The project benefited from the client's cloud provider existing APIs without introducing additional complexities into the client's IT landscape. Additionally, the use of cloud-based AI APIs helped to answer identified pain-points, deliver all required features, and speed up delivery without re-inventing the wheel.

AI INFRASTRUCTURE AS A SERVICE (AI IAAS)

Image recognition for car damage and face mask detection leveraging AI IaaS

One practical use case is recognizing car parts and minor damages such as scratches, dents, and cracks with only a picture as input. An AI model can be trained to automatically output this valuable information that can be advantageous for multiple players in the automotive and insurance industry. However, image recognition models are trained on complex neural networks that require large computing power. AI IaaS provides the computational resources that make it possible to train these complex models. For such complex AI projects that require niche toolkits and libraries, it is best to opt for raw Cloud VMs backed by high-end CPUs and accelerators for an enhanced level of customization and control over the hardware and software layers. A growing interest and maturity in the market of applying visual intelligence an AI automation is evident within the automotive sector. An increasing number of players

is investigating and testing how these solutions can improve their current business processes, showing a growing appetite for cloud-supported AI solutions.

An AI IaaS model also allows the necessary performance for executing calculations at high speeds to solve large, complex business problems. A relevant use case benefiting from this feature arose within the COVID-19 pandemic and focusses on the detection of properly worn face masks in public places. As this requires real-time processing, AI IaaS has significantly facilitated building these computer vision techniques and algorithms by yielding scalability and on-demand service without any internal infrastructure requirement.

Shifting to AI IaaS comes with a range of advantages, and when coupled with the new computing trends such as containers, it allows organizations to harness the power of scalable ML^[5]. These use cases show that leveraging AI IaaS can be a performance powerhouse. AI IaaS enabled transforming the claim value chain and offers on-demand scalability and real-time processing without any requirement for internal infrastructure or maintenance cost from our clients.

MACHINE LEARNING PLATFORM AS A SERVICE (MLAAS)

ML at the service of digital transformation

If you are looking to remove the friction involved in setting up and configuring data science environments while still being able to build customized ML models, MLaaS is the way to go. Major cloud providers offer pre-configured environments to ease some of the complexity involved in AI projects pipeline.

MLaaS offers a wide range of capabilities to enable organizations to develop their AI projects at scale. To illustrate these powerful capabilities, a digital transformation project to support the client in becoming a data-driven organization showcases how to harness the power of MLaaS capabilities. In this use case, a large cloud provider's ML models could help predict future outcomes based on the existing patterns in the historical data, allowing the client to segment its customers based on real-

time insights and providing them more personalized communication.

Using such MLaaS platforms allows data teams to work with their own data and code without the effort of configuring compute, storage, and networking environments to run complex ML jobs. MLaaS provides environments that eases workflows by presenting tools for fast model development and deployment [6]. They join most tools in one place, making it simple to go from creating models to scalable deployment from their interface. These platforms ease data exploration and analysis without the hassle of server management.

MLaaS delivers efficient lifecycle management of ML models and helps users to convey agility into ML model development and deployment by combining the technique of continuous integration and continuous delivery, or deployment (CI/CD) with ML model management. With these features, MLaaS can be a promising option to build your cloud-based AI solution efficiently.

References:


- [1] Deloitte [Thriving in the era of pervasive AI, Deloitte's State of AI in the Enterprise, 3rd Edition.](#)
- [2] Turbonomic Blog [Mastering Cloud Cost Optimization: Cloud Cost Models & Discounts Overview.](#)
- [3] MordorIntelligence [CLOUD AI MARKET - GROWTH, TRENDS, COVID-19 IMPACT, AND FORECASTS \(2021 - 2026\)](#)
- [4] AWS, [Easily add intelligence to applications, No machine learning skills required](#)
- [4] Azure Cognitive Services, [A comprehensive family of AI services and cognitive APIs to help you build intelligent apps](#)
- [4] Google Cloud Platform, [Cloud AI building blocks](#)
- [5] HPE, [Containers as an enabler of AI](#)
- [6] Neptune AI, [Machine Learning as a Service: What It Is, When to Use It and What Are the Best Tools Out There](#)



Conclusion

The cloud enables companies to innovate new solutions and scale their existing ones together while driving efficiency and extending their data team's talents. Cloud-enabled AI brings various advantages for organizations throughout the ML pipeline, from development, to training and management, through to deployment. By leveraging high-level APIs, utilizing AI IaaS, or using machine learning as a service, decision makers can decide on the right cloud ML service to suit their situation.

Share #DeloitteInsideNow



Artificial Intelligence fi& fiData

The world is changing, fast. Customers are demanding a more personalised experience. Competitors are appearing out of nowhere.

© 2021. See [Terms of Use](#) for more information.

Deloitte refers to one or more of Deloitte Touche Tohmatsu Limited, a UK private company limited by guarantee ("DTTL"), its network of member firms, and their related entities. DTTL and each of its member firms are legally separate and independent entities. DTTL (also referred to as "Deloitte Global") does not provide services to clients. Please see www.deloitte.com/about to learn more about our global network of member firms.

The Luxembourg member firm of Deloitte Touche Tohmatsu Limited Privacy Statement notice may be found at www.deloitte.com/lu/privacy.