

Tema 1. Computación en la nube en la era del big data y la inteligencia artificial

construcción de aplicaciones empresariales modernas y la implementación de soluciones de inteligencia artificial.

Azure se distingue por su integración con las **tecnologías y herramientas de Microsoft**, lo que les permite a los clientes aprovechar sus inversiones existentes en productos como Windows Server, SQL Server, Active Directory y Visual Studio. Además, Azure es compatible con una amplia variedad de sistemas operativos, lenguajes de programación, *frameworks* y herramientas de desarrollo, lo que brinda a los clientes flexibilidad y opciones para construir y desplegar aplicaciones en su entorno preferido.

Además, ofrece opciones de **precios flexibles**, que incluyen modelos de pago por uso y descuentos por compromisos a largo plazo, lo que les permite a las organizaciones optimizar sus costos y maximizar el valor de su inversión en la nube.

¿Cuáles son las ventajas competitivas de Azure?

- ▶ **Escalabilidad y flexibilidad.** Escala automáticamente según las necesidades de tu negocio y paga solo por lo que utilizas.
- ▶ **Integración profunda.** Integración con herramientas y tecnologías de Microsoft, como Windows Server, SQL Server, Visual Studio y más.
- ▶ **Amplia gama de servicios.** Más de doscientos servicios para cubrir todas las necesidades de infraestructura y desarrollo.
- ▶ **Seguridad y confiabilidad.** Cumplimiento de los más altos estándares de seguridad y disponibilidad.
- ▶ **Innovación continua.** Constante lanzamiento de nuevos servicios y características para impulsar la innovación en la nube.

Tema 1. Computación en la nube en la era del big data y la inteligencia artificial

Principales servicios de Azure

Entre los principales servicios que la nube de Microsoft ofrece, se pueden destacar de manera categorizada los siguientes:

- ▶ **Cómputo:** Azure Virtual Machines, Azure Kubernetes Service y Azure Functions.
- ▶ **Almacenamiento:** Azure Blob Storage, Azure Files y Azure Disk Storage.
- ▶ **Bases de datos:** Azure SQL Database, Azure Cosmos DB y Azure Synapse Analytics.
- ▶ **Redes:** Azure Virtual Network, Azure Load Balancer y Azure VPN Gateway.
- ▶ **IA y ML:** Azure Machine Learning, Azure Cognitive Services y Azure Databricks.
- ▶ **IoT:** Azure IoT Hub, Azure IoT Central y Azure IoT Edge.
- ▶ **Seguridad:** Azure Active Directory, Azure Security Center y Azure Sentinel.
- ▶ **Desarrollo de aplicaciones:** Azure DevOps, Azure App Service y Azure Functions.

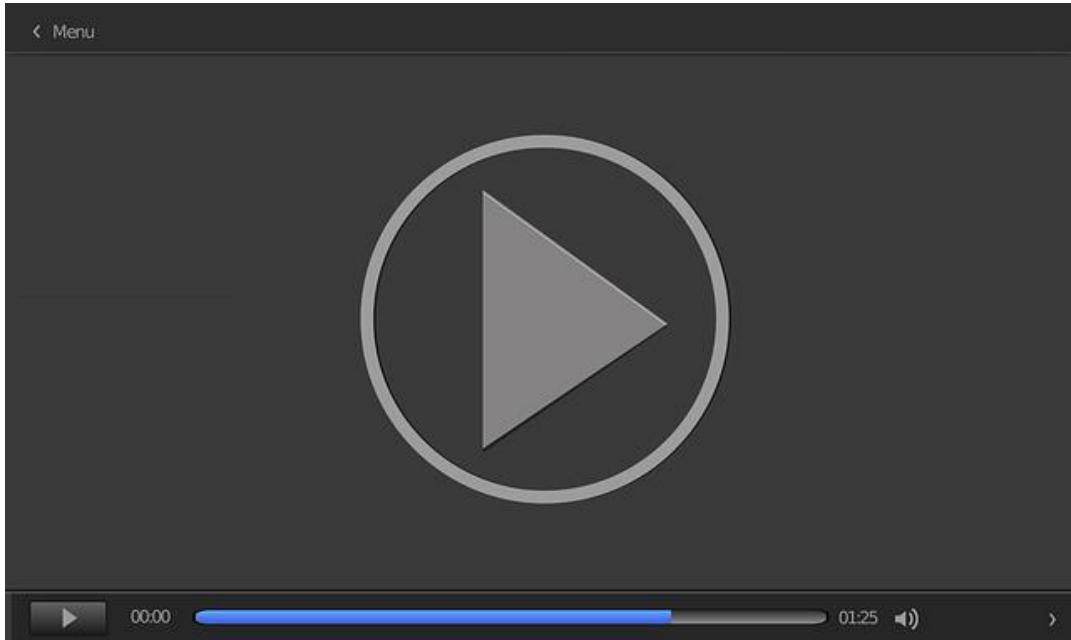
Casos de éxito de Azure

A continuación, se muestra una lista de casos de éxito de Azure en el ámbito del *big data* y de la IA.

- ▶ **Johnson Controls.** Esta empresa global de tecnología para edificios utilizó Azure para desarrollar soluciones de *big data* e IoT para optimizar la eficiencia energética y mejorar la gestión de sus sistemas de calefacción, ventilación y aire acondicionado en edificios comerciales e industriales. Con Azure, Johnson Controls pudo analizar grandes volúmenes de datos en tiempo real y proporcionar recomendaciones inteligentes para mejorar el rendimiento y reducir los costos operativos.

Tema 1. Computación en la nube en la era del big data y la inteligencia artificial

Si quieras saber más sobre el proyecto Johnson Controls, te recomendamos ver el vídeo *Johnson Controls wins 2023 Microsoft Global ISV Partner of the Year for OpenBlue Building Solutions* (Johnson Controls, 2023).



Johnson Controls wins 2023 Microsoft Global ISV Partner of the Year for OpenBlue Building Solutions.

Accede al vídeo:

<https://www.youtube.com/embed/8qHLAZUIMXk>

- ▶ **ASOS.** Como una de las principales empresas de moda *online*, ASOS aprovechó Azure para impulsar su estrategia de personalización y recomendación. Utilizando soluciones de inteligencia artificial y análisis de *big data* en Azure, ASOS pudo analizar el comportamiento de compra de sus clientes y ofrecer recomendaciones de productos altamente relevantes y personalizadas, lo que resultó en un aumento significativo en las conversiones y la retención de clientes.

Tema 1. Computación en la nube en la era del big data y la inteligencia artificial

Microsoft. (2017, octubre 25). *Online retailer uses cloud database to deliver world-class shopping experiences.* <https://customers.microsoft.com/en-us/story/asos-retail-and-consumer-goods-azure>

- ▶ **Shell.** La multinacional energética Shell utilizó Azure para optimizar sus operaciones de exploración y producción de petróleo y gas. Con soluciones de *big data* e IA en Azure, Shell pudo analizar datos sísmicos, de perforación y de producción en tiempo real para identificar oportunidades de optimización, mejorar la seguridad y aumentar la eficiencia en sus operaciones en todo el mundo.

Microsoft. 2018, septiembre 24). *Shell invests in safety with Azure, AI, and machine vision to better protect customers and service champions.* <https://customers.microsoft.com/en-us/story/shell-mining-oil-gas-azure-databricks>

- ▶ **Maersk.** Como una de las mayores empresas de transporte marítimo del mundo, Maersk utilizó Azure para desarrollar soluciones de *big data* e IA que mejoraron la eficiencia operativa y redujeron los costos en su cadena de suministro global. Con Azure, Maersk pudo analizar grandes volúmenes de datos de seguimiento de contenedores, rutas marítimas, condiciones meteorológicas y más, lo que le permitió optimizar la planificación de rutas, minimizar los tiempos de inactividad de los buques y mejorar la visibilidad de la cadena de suministro.

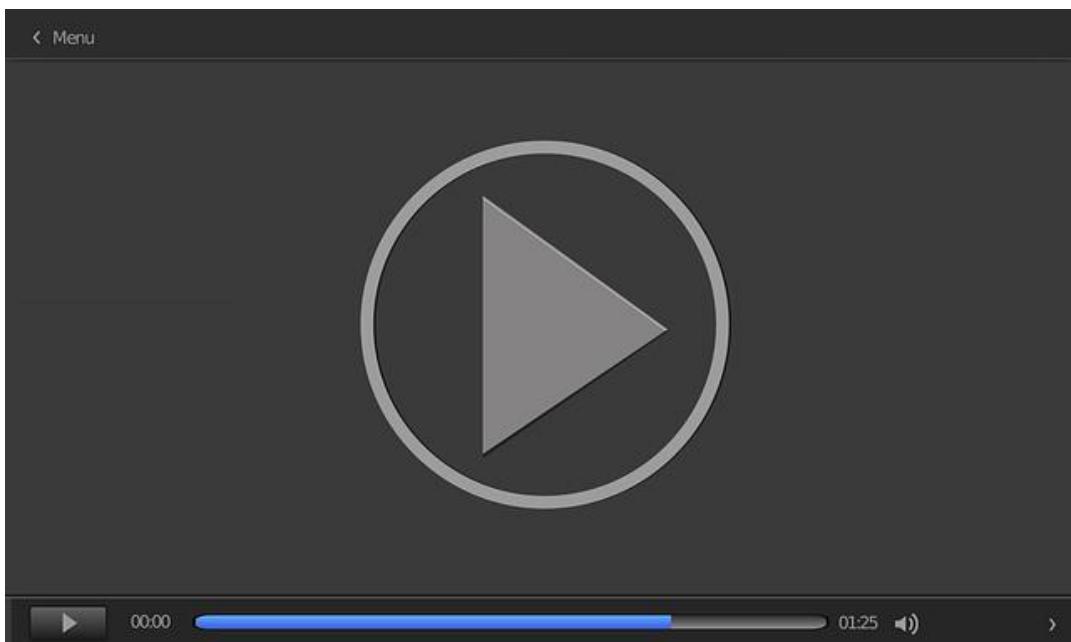
Microsoft. (2017, diciembre 21). *Global transport and logistics company goes digital to transform its operations.* <https://customers.microsoft.com/en-us/story/maersk-travel-transportation-microsoft-services>

- ▶ **BMW.** El fabricante de automóviles BMW utilizó Azure para desarrollar soluciones de *big data* e IA que mejoraron la experiencia del cliente y optimizaron la producción.

Tema 1. Computación en la nube en la era del big data y la inteligencia artificial

Con Azure, BMW pudo analizar datos de sensores en tiempo real de sus vehículos conectados para identificar patrones de uso, predecir fallos mecánicos y ofrecer servicios de mantenimiento predictivo a los propietarios de vehículos. Además, Azure ayudó a BMW a optimizar sus procesos de fabricación mediante el análisis de datos de sensores en las líneas de producción para identificar áreas de mejora y aumentar la eficiencia.

Si quieras saber más sobre cómo BMW utilizó Azure, te recomendamos ver el vídeo *Microsoft Intelligent Manufacturing Award 2021 – BMW* (Roland Berger, 2022).



Microsoft Intelligent Manufacturing Award 2021 – BMW.

Accede al vídeo:

<https://www.youtube.com/embed/bACcdP7kz5A>

Estos son solo algunos ejemplos de cómo las empresas líderes en diversos sectores están utilizando Azure para impulsar la innovación, mejorar la eficiencia operativa y ofrecer experiencias personalizadas a los clientes mediante el aprovechamiento de soluciones de *big data* e IA.

Tema 1. Computación en la nube en la era del big data y la inteligencia artificial

Servicios *big data* e IA de Azure

Azure Big Data, la plataforma en la nube de Microsoft, ofrece una amplia gama de **servicios especializados** en *big data* e IA, que permiten a las organizaciones construir, desplegar y escalar soluciones avanzadas de análisis de datos y aprendizaje automático.

A continuación, se expone una descripción general de algunos de los principales servicios que Azure ofrece en estas áreas.

- ▶ **Big data.** En la Figura 3 se muestra la amplia gama de servicios de *big data* e IA ofrecidos por esta plataforma.

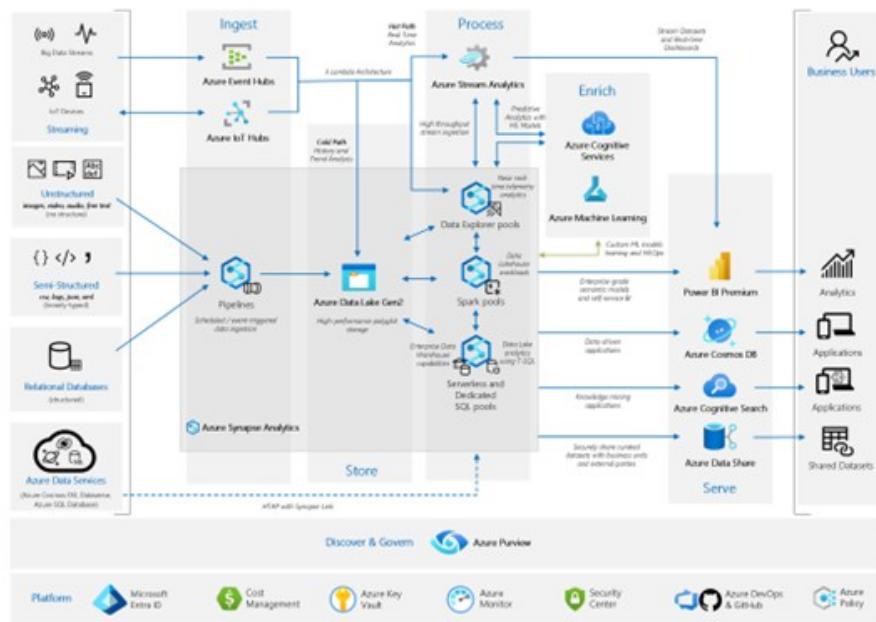


Figura 3. servicios de *big data* e IA. Fuente: Microsoft Learn, 2023.

Profundizando en los principales **servicios** de *big data* que aparecen en la Figura 3, tenemos los siguientes:

- ▶ **Azure Event Hubs.** Es un servicio de ingestión de datos en tiempo real y escalable que puede procesar millones de eventos por segundo. Permite a las organizaciones

Tema 1. Computación en la nube en la era del big data y la inteligencia artificial

recopilar y procesar datos de dispositivos IoT, aplicaciones móviles, registros de servidores, redes sociales y otros sistemas en tiempo real. Event Hubs integra fácilmente con otros servicios de Azure, como Azure Stream Analytics y Azure Functions, para realizar análisis en tiempo real y activar acciones automatizadas.

- ▶ **Azure IoT Hub.** Es un servicio gestionado de IoT que les permite a las organizaciones conectar, supervisar y administrar dispositivos IoT a escala. Proporciona capacidades de ingestión de datos para dispositivos IoT, así como funciones de seguridad, control de acceso y administración de dispositivos. IoT Hub permite la integración con otros servicios de Azure, como Azure Stream Analytics y Azure Functions, para procesar y analizar datos de dispositivos IoT en tiempo real.
- ▶ **Azure Data Factory.** Es un servicio de orquestación de datos en la nube que les permite a las organizaciones crear, programar y administrar flujos de datos entre diversos orígenes y destinos. Permite la ingestión de datos en lotes y en tiempo real, así como la transformación y el enriquecimiento de datos durante el proceso de ingestión. Data Factory integra con una amplia gama de servicios de Azure y fuentes de datos externas.
- ▶ **Azure Data Lake Storage.** Es un servicio de almacenamiento de datos escalable y seguro que les permite a las organizaciones almacenar grandes volúmenes de datos estructurados y no estructurados. Proporciona capacidades de ingestión de datos para datos en lotes y en tiempo real, así como funciones avanzadas de análisis y procesamiento de datos utilizando herramientas como Azure Data Lake Analytics y Azure Databricks.
- ▶ **Azure Blob Storage.** Es un servicio de almacenamiento en la nube escalable y económico diseñado para almacenar grandes volúmenes de datos no estructurados, como archivos de registro, imágenes, vídeos y archivos de audio. Se puede utilizar como destino de datos para la ingestión de datos en lotes, así como para la ingestión de datos en tiempo real utilizando Azure Event Hubs, Azure IoT Hub y otros servicios.

Tema 1. Computación en la nube en la era del big data y la inteligencia artificial

- ▶ **Azure Data Lake Storage.** Es un servicio de almacenamiento escalable y seguro diseñado para almacenar grandes volúmenes de datos sin estructurar. Permite el análisis de datos en tiempo real y el procesamiento distribuido de conjuntos de datos masivos.
- ▶ **Azure HDInsight.** Es un servicio completamente gestionado de Apache Hadoop, Spark, HBase y otros *frameworks* de *big data*. Proporciona capacidades de procesamiento distribuido y análisis de datos en clústeres escalables y seguros.
- ▶ **Azure Databricks.** Es una plataforma de análisis de *big data* basada en Apache Spark que simplifica el desarrollo, la implementación y la administración de *pipelines* de datos y modelos de ML.
- ▶ **Azure Synapse Analytics.** Es un servicio de análisis de datos que combina el almacenamiento de datos empresariales con la integración de *big data* y el análisis de datos en tiempo real. Permite realizar consultas SQL en datos estructurados y no estructurados a gran escala.

En el área de IA, la Figura 4 proporciona una visión de la oferta de esta plataforma.

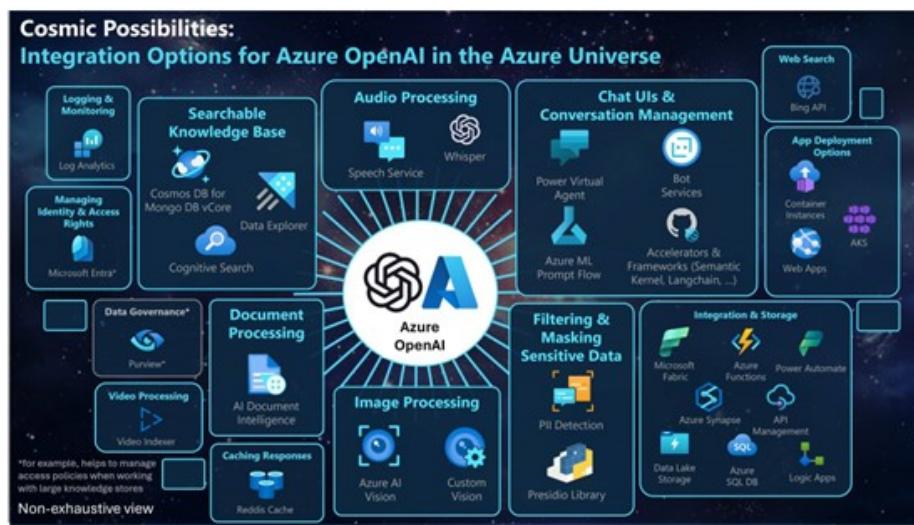


Figura 4. Catálogo de Azure. Fuente: Rautenstrauch, 2023.

Tema 1. Computación en la nube en la era del big data y la inteligencia artificial

Entre su amplia oferta de IA, se destacan los siguientes **servicios**:

- ▶ **Azure Machine Learning.** Es un servicio de ML que les permite a los científicos de datos y desarrolladores construir, entrenar y desplegar modelos de ML en la nube y en el borde de la red.
- ▶ **Azure Cognitive Services.** Es una colección de servicios de IA preconstruidos que ofrecen capacidades de visión artificial, procesamiento de lenguaje natural, reconocimiento de voz, entre otros. Estos servicios permiten agregar inteligencia a las aplicaciones sin necesidad de experiencia en ML.
- ▶ **Azure Bot Service.** Es un servicio que les permite a los desarrolladores construir, conectar y desplegar chatbots inteligentes que interactúan de manera natural con los usuarios a través de múltiples canales.
- ▶ **Azure Computer Vision.** Es un servicio que proporciona capacidades avanzadas de reconocimiento visual, incluyendo etiquetado de imágenes, detección de objetos, reconocimiento facial y más.
- ▶ **Azure Speech Services.** Es un conjunto de servicios de IA que permiten la conversión de voz a texto, la traducción de voz en tiempo real y la generación de voz natural.

Estos son solo algunos ejemplos de la amplia gama de servicios de *big data* e IA que Azure ofrece. Con estas herramientas, las organizaciones pueden aprovechar la nube de Microsoft para obtener *insights* valiosos de sus datos y construir soluciones innovadoras impulsadas por la inteligencia artificial.

Google Cloud Platform (GCP)

El lema de la **plataforma GCP** es «transformando negocios con la nube», con este lema quieren transmitir el carácter innovador de esta nube. Pero ¿qué es GCP? GCP

Tema 1. Computación en la nube en la era del big data y la inteligencia artificial

es la plataforma en la nube de Google, que ofrece una amplia gama de servicios de infraestructura, análisis, ML, almacenamiento, bases de datos, desarrollo de aplicaciones, etc. Google Cloud Platform ofrece una solución completa y escalable para las necesidades de infraestructura y desarrollo de cualquier organización.

Con su enfoque en la innovación, escalabilidad, seguridad y precio competitivo, GCP es la **opción preferida** para empresas de todos los tamaños y sectores que buscan transformar sus negocios con la nube.

¿Cuáles son las principales ventajas competitivas de GCP?

Esta plataforma proporciona herramientas y servicios para ayudar a las organizaciones a construir, escalar y administrar aplicaciones y servicios de manera eficiente en la nube. Entre sus **características** fundamentales, se pueden enumerar las siguientes:

- ▶ **Innovación y tecnología avanzada.** Google es conocido por su enfoque en la innovación y la tecnología de vanguardia, lo que se refleja en GCP con servicios líderes en la industria.
- ▶ **Escalabilidad y rendimiento.** GCP ofrece una infraestructura altamente escalable y un rendimiento confiable, lo que les permite a las organizaciones escalar según la demanda y mantener una alta disponibilidad.
- ▶ **Integración con herramientas de Google,** como TensorFlow para ML, Kubernetes para contenedores y BigQuery para análisis de datos.
- ▶ **Seguridad y cumplimiento.** Compromiso con la seguridad y el cumplimiento de los más altos estándares de la industria para proteger los datos y la privacidad del cliente.
- ▶ **Precio competitivo.** Precios competitivos y opciones flexibles de pago por uso que ayudan a optimizar los costos y maximizar el valor.

Tema 1. Computación en la nube en la era del big data y la inteligencia artificial

Principales servicios de GCP

Entre los principales servicios que la nube de Google ofrece, se pueden destacar de manera categorizada los siguientes:

- ▶ **Cómputo:** Compute Engine, Kubernetes Engine y App Engine.
- ▶ **Almacenamiento:** Cloud Storage, Cloud SQL y Cloud Bigtable.
- ▶ **Bases de datos:** Cloud Firestore, Cloud Spanner y BigQuery.
- ▶ **Redes:** Virtual Private Cloud (VPC), Cloud Load Balancing y Cloud CDN.
- ▶ **IA y ML:** AI Platform, TensorFlow y AutoML.
- ▶ **IoT:** Cloud IoT Core, Edge TPU y IoT Edge.
- ▶ **Seguridad:** Identity and Access Management (IAM), Cloud Identity y Security Command Center.
- ▶ **Desarrollo de aplicaciones:** Cloud Functions, Cloud Run y Firebase.

Casos de éxito de GCP

A continuación, se muestra una lista de casos de éxito de GCP en *big data* e IA.

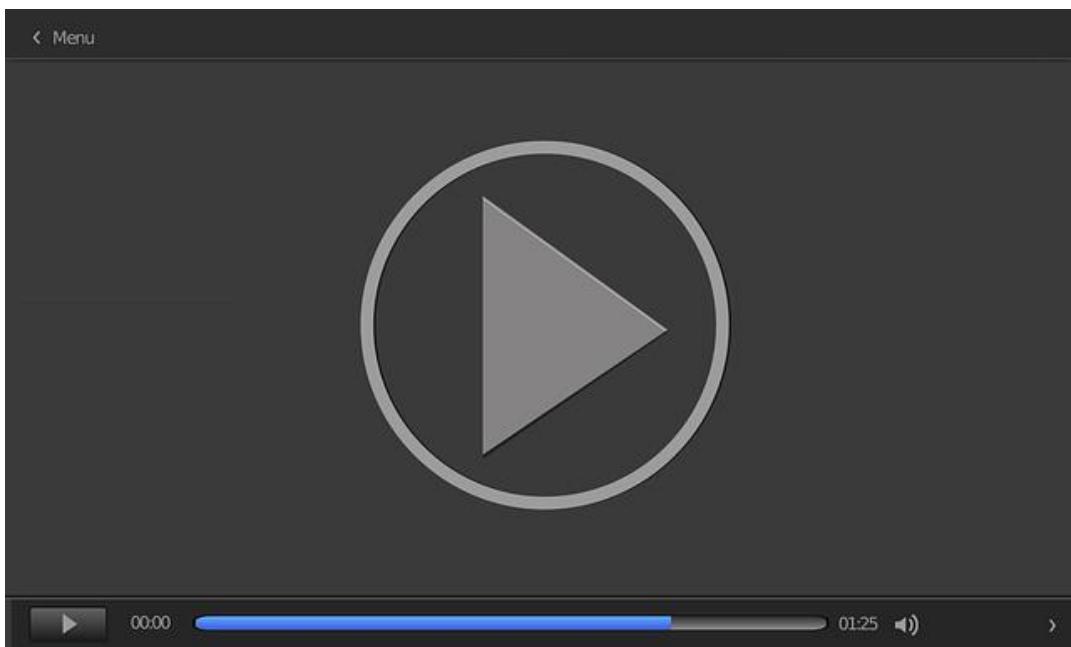
- ▶ **Snap Inc.** La empresa detrás de Snapchat utiliza GCP para manejar grandes volúmenes de datos generados por sus usuarios. GCP proporciona a Snap herramientas escalables de *big data* para procesar y analizar datos en tiempo real, lo que permite a la empresa ofrecer características innovadoras y personalizadas a sus usuarios.

Tema 1. Computación en la nube en la era del big data y la inteligencia artificial

Google Cloud. (2022, mayo 31). *Snap Inc. adopts Google Cloud TPU for deep learning recommendation models.* <https://cloud.google.com/blog/products/ai-machine-learning/snap-inc-uses-google-cloud-tpu-for-deep-learning-recommendation-models>

- ▶ **Spotify.** La popular plataforma de *streaming* de música aprovecha GCP para almacenar y procesar grandes cantidades de datos de audio. GCP proporciona a Spotify la capacidad de analizar los gustos musicales de sus usuarios y ofrecer recomendaciones personalizadas, lo que mejora la experiencia del usuario y aumenta la retención.

Si quieras saber más sobre cómo Spotify utiliza la plataforma GCP, te recomendamos ver el vídeo *Spotify uses Google Cloud to unlock infinite capacity and faster innovation* (Google Cloud, 2020).



Spotify uses Google Cloud to unlock infinite capacity and faster innovation.

Accede al vídeo:

https://www.youtube.com/embed/55xgR_o4PGs

Tema 1. Computación en la nube en la era del big data y la inteligencia artificial

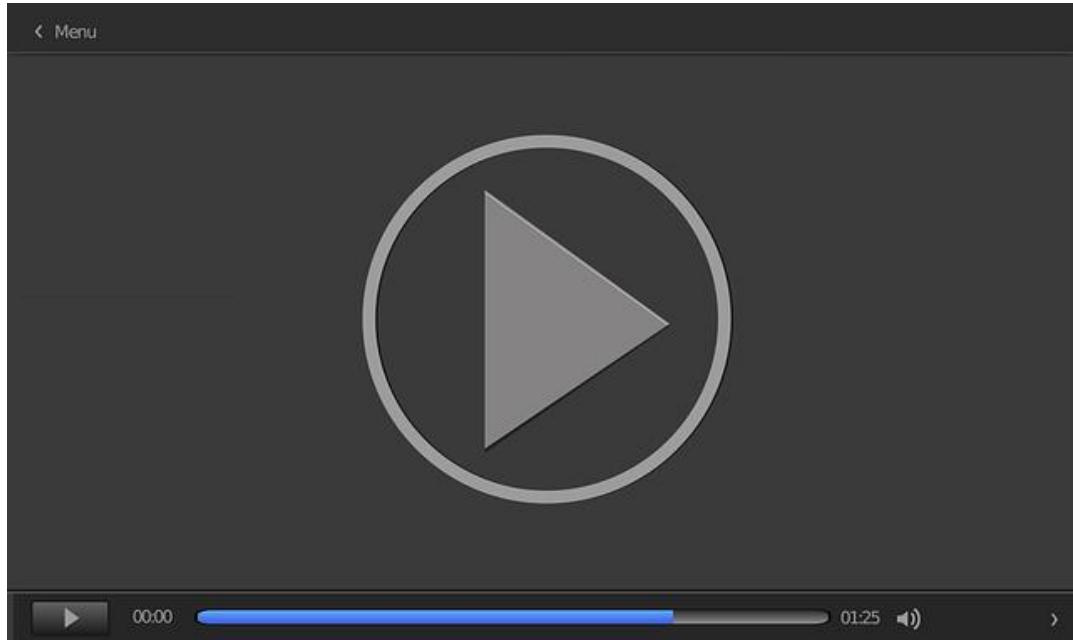
- ▶ **Niantic.** La empresa detrás del juego de realidad aumentada Pokémon GO utiliza GCP para gestionar la gran cantidad de datos generados por millones de jugadores en todo el mundo. GCP proporciona a Niantic la escalabilidad necesaria para manejar picos de tráfico repentinos y analizar datos en tiempo real para mejorar la jugabilidad y la experiencia del usuario.

Google Cloud. (2021, octubre 27). *How Pokémon GO scales to millions of requests?*
<https://cloud.google.com/blog/topics/developers-practitioners/how-pok%C3%A9mon-go-scales-millions-requests>

- ▶ **PayPal.** La plataforma de pagos en línea PayPal utiliza GCP para mejorar la seguridad y prevenir el fraude. GCP proporciona a PayPal herramientas avanzadas de análisis de *big data* e IA para detectar patrones de fraude y proteger las transacciones de los usuarios en tiempo real.

Tema 1. Computación en la nube en la era del big data y la inteligencia artificial

Si quieres saber más sobre cómo PayPal utiliza la plataforma GCP, te recomendamos ver el vídeo *PayPal partners with Google Cloud* (Google Cloud, 2018).



PayPal partners with Google Cloud.

Accede al vídeo:

<https://www.youtube.com/embed/9jJ6xLOSS3c>

- ▶ **X (antes Twitter).** La red social X utiliza GCP para almacenar y analizar grandes volúmenes de datos generados por millones de tweets diarios. GCP proporciona a X la capacidad de procesar y analizar datos en tiempo real para identificar tendencias, mejorar la segmentación de anuncios y proporcionar recomendaciones personalizadas a los usuarios.

Tema 1. Computación en la nube en la era del big data y la inteligencia artificial

Google Cloud. (2022, mayo 17). *Twitter takes data activation to new heights with Google Cloud.* <https://cloud.google.com/blog/products/data-analytics/how-twitter-modernized-its-data-processing-with-google-cloud>

- ▶ **HSBC.** El banco multinacional HSBC utiliza GCP para desarrollar soluciones de *big data* e IA que mejoran la experiencia del cliente y optimizan las operaciones bancarias. GCP proporciona a HSBC la capacidad de analizar grandes volúmenes de datos de transacciones financieras para detectar fraudes, identificar oportunidades de *upselling* y ofrecer servicios personalizados a sus clientes.

Google Cloud. (s. f.). *HSBC: Using Google Cloud to improve customer service at call centers.* <https://cloud.google.com/customers/hsbc>

Estos son solo algunos ejemplos de cómo empresas líderes están utilizando GCP para impulsar la innovación, mejorar la experiencia del cliente y optimizar sus operaciones mediante el aprovechamiento de soluciones de *big data* e IA.

Servicios de GCP de *big data* e IA

GCP ofrece una amplia gama de servicios y productos en el ámbito de *big data* e IA que permiten a las organizaciones gestionar, procesar y analizar grandes volúmenes de datos de manera eficiente, así como construir y desplegar modelos avanzados de IA.

A continuación, se expone una descripción general de algunos de los principales **servicios** que GCP ofrece en estas áreas:

- ▶ **Big data.** En la Figura 5 se muestra un resumen de los servicios de *big data* e IA ofrecidos por esta plataforma.

Tema 1. Computación en la nube en la era del big data y la inteligencia artificial

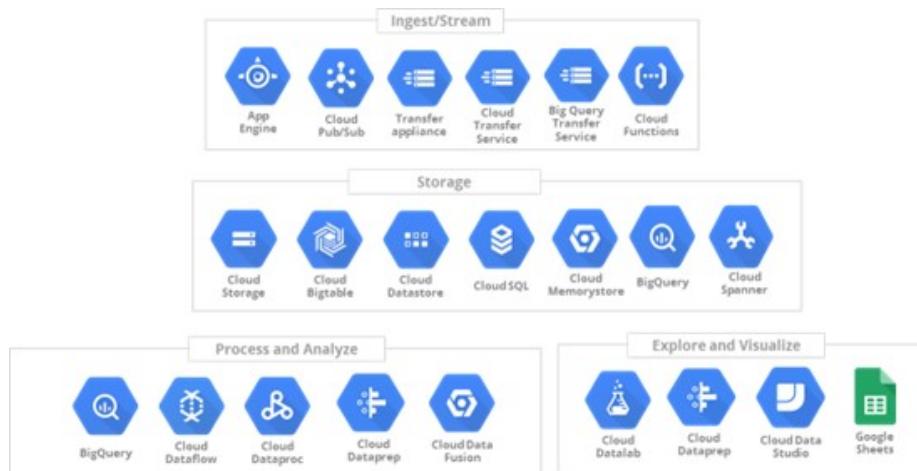


Figura 5. Servicios de *big data* e IA de GCP. Fuente: Srinivasan, 2019.

Si ahondamos en los principales **servicios** de *big data* que aparecen en la Figura 5, tendremos los siguientes:

- ▶ **BigQuery.** Es un almacén de datos completamente gestionado y altamente escalable que permite analizar conjuntos de datos grandes y complejos de manera rápida y eficiente mediante SQL estándar. Ofrece capacidades de procesamiento en paralelo y permite ejecutar consultas en tiempo real sobre petabytes de datos. BigQuery también ofrece integración con otras herramientas de Google Cloud, como Dataflow y TensorFlow.
- ▶ **Dataflow.** Es un servicio de procesamiento de datos en tiempo real y por lotes que permite crear, ejecutar y escalar *pipelines* de datos de manera sencilla. Utiliza un modelo de programación basado en Apache Beam para procesar datos de manera paralela y distribuida en entornos de nube y local. Dataflow es compatible con una amplia variedad de fuentes y destinos de datos, incluidos BigQuery, Cloud Storage y Pub/Sub.
- ▶ **Dataproc.** Es un servicio de procesamiento de datos gestionado que permite ejecutar clústeres de Hadoop y Spark en la nube de manera rápida y sencilla. Proporciona capacidades de escalado automático y administración de clústeres, lo

Tema 1. Computación en la nube en la era del big data y la inteligencia artificial

que les permite a las organizaciones procesar grandes volúmenes de datos de manera eficiente y rentable. Dataproc también ofrece integración con otros servicios de GCP, como BigQuery y Dataflow.

- ▶ **Pub/Sub.** Es un servicio de mensajería en tiempo real que les permite a las organizaciones enviar y recibir mensajes entre aplicaciones y servicios de manera segura y escalable. Utiliza un modelo de publicación/suscripción para transmitir datos de manera confiable y duradera. Pub/Sub es ideal para casos de uso de ingestión de datos en tiempo real, como el procesamiento de registros de eventos, telemetría de dispositivos IoT y más.

En el área de la IA, la plataforma GCP ofrece la plataforma **Google Cloud Vertex AI**, que es un servicio completamente gestionado que les permite a las organizaciones construir, entrenar y desplegar modelos de aprendizaje automático de manera rápida y sencilla en la nube de Google.

Vertex AI proporciona una **plataforma unificada** que integra una variedad de herramientas y servicios de Google Cloud para facilitar el desarrollo y la implementación de modelos de inteligencia artificial a escala.

La Figura 6 proporciona una visión de la oferta de la plataforma Google Vertex AI.



Figura 6. Catálogo de Google Vertex AI. Fuente: Inderpreet Singh, 2024.

Tema 1. Computación en la nube en la era del big data y la inteligencia artificial

Para el área de **unificación de servicios**, tenemos los siguiente:

- ▶ **Integración de servicios de aprendizaje automático.** Vertex AI integra varios servicios de aprendizaje automático de Google Cloud, incluidos AutoML, TensorFlow, AI Platform y más, en una plataforma unificada. Esto les permite a los usuarios acceder a una amplia gama de herramientas y capacidades para construir y desplegar modelos de inteligencia artificial de manera eficiente.

Dentro del **desarrollo de modelos** podemos destacar:

- ▶ **AutoML.** Vertex AI ofrece capacidades de AutoML para permitir a los usuarios entrenar modelos de aprendizaje automático personalizados sin necesidad de experiencia en ciencia de datos o aprendizaje automático. Los usuarios pueden utilizar AutoML Vision, AutoML Natural Language, AutoML Tables y otros servicios para construir modelos de clasificación, detección de objetos, análisis de sentimientos y más.
- ▶ **TensorFlow.** Vertex AI proporciona soporte para TensorFlow, la biblioteca de aprendizaje automático de código abierto de Google. Los usuarios pueden aprovechar TensorFlow para construir modelos de aprendizaje profundo personalizados utilizando una variedad de herramientas y API, incluyendo TensorFlow Extended (TFX) para el desarrollo de *pipelines* de datos y TensorFlow Serving para el despliegue de modelos en producción.

En cuanto a los servicios dedicados al **entrenamiento y evaluación de modelos**:

- ▶ **AI Platform Training.** Vertex AI ofrece capacidades de entrenamiento de modelos escalables y distribuidas a través de AI Platform Training. Los usuarios pueden entrenar modelos utilizando recursos de procesamiento distribuido y GPU en la nube de Google, lo que permite acelerar el proceso de entrenamiento y mejorar el rendimiento de los modelos.

Tema 1. Computación en la nube en la era del big data y la inteligencia artificial

- ▶ **Evaluación de modelos.** Vertex AI proporciona herramientas para evaluar y validar modelos de aprendizaje automático, incluidas las métricas de evaluación, la visualización de resultados y la comparación de modelos. Esto les permite a los usuarios medir el rendimiento de los modelos en datos de prueba y ajustar los hiperparámetros para mejorar la precisión y la generalización del modelo.

Respecto a la **implementación y despliegue de modelos**, destacan los siguientes:

- ▶ **AI Platform Prediction.** Vertex AI ofrece capacidades de despliegue de modelos escalables y seguras a través de AI Platform Prediction. Los usuarios pueden implementar modelos entrenados en la nube de Google y exponerlos como servicios web RESTful para realizar inferencias en tiempo real. AI Platform Prediction ofrece integración con otros servicios de Google Cloud, como Cloud Functions y Cloud Run, para permitir la automatización y la escalabilidad de las inferencias.
- ▶ **Model Monitoring.** Vertex AI proporciona herramientas para monitorear el rendimiento y la salud de los modelos desplegados en producción. Los usuarios pueden supervisar métricas de rendimiento, detectar anomalías y recibir alertas en tiempo real para garantizar la fiabilidad y la disponibilidad de los modelos en producción.

Finalmente, para la **automatización y optimización** se ofrece lo siguiente:

- ▶ **Vertex AI MLOps.** Vertex AI ofrece capacidades de MLOps para permitir la automatización y la optimización del ciclo de vida de los modelos de aprendizaje automático. Esto incluye la automatización de flujos de trabajo de entrenamiento, evaluación y despliegue de modelos, así como la gestión de versiones y el control de calidad de los modelos en producción.

Tema 1. Computación en la nube en la era del big data y la inteligencia artificial

- ▶ **Dialogflow.** Dialogflow es un servicio de Google Cloud que les permite a los desarrolladores crear interfaces de conversación para aplicaciones y servicios utilizando inteligencia artificial y procesamiento de lenguaje natural. Proporciona capacidades avanzadas de reconocimiento de voz, comprensión del lenguaje natural y generación de respuestas para crear experiencias de usuario conversacionales y personalizadas. Dialogflow se utiliza comúnmente para construir chatbots, asistentes virtuales y aplicaciones de voz.

En resumen, **Google Cloud Vertex AI** es un servicio integral que les permite a las organizaciones construir, entrenar y desplegar modelos de aprendizaje automático de manera eficiente y escalable en la nube de Google. Con su enfoque en la unificación de servicios, desarrollo de modelos, entrenamiento y evaluación, implementación y despliegue, así como automatización y optimización, Vertex AI facilita el proceso de desarrollo y operación de modelos de inteligencia artificial en producción. Estos son solo algunos ejemplos de los principales servicios y productos de Google Cloud Platform en *big data* e IA.

GCP ofrece una amplia variedad de herramientas y servicios que permiten a las organizaciones aprovechar la potencia de la nube para gestionar y analizar datos a gran escala, así como construir y desplegar aplicaciones y servicios inteligentes impulsados por la IA.

Tema 1. Computación en la nube en la era del big data y la inteligencia artificial

Amazon Web Services (AWS)

AWS es la **plataforma líder** en servicios en la nube, ofrecida por Amazon.com. Proporciona una amplia gama de servicios de infraestructura informática, almacenamiento, bases de datos, análisis, inteligencia artificial, IoT, seguridad, desarrollo de aplicaciones y más. Ayuda a las organizaciones a escalar y crecer, eliminando la necesidad de invertir en infraestructura física costosa y ofreciendo flexibilidad, agilidad y escalabilidad.

AWS es, actualmente, la plataforma líder de servicios en la nube, la cual presta sus servicios a grandes empresas (Expedia, Shell), *startups* (Airbnb, Lyft), sector público (Coursera, FDA) y particulares. En la Figura 7 se puede observar la posición de cada **proveedor de servicios cloud**, según el cuadrante del año 2023 de la consultora Gartner; en él se aprecia cómo AWS tiene un papel dominante, seguido de cerca por Microsoft y Google Cloud Services.

Tema 1. Computación en la nube en la era del big data y la inteligencia artificial

Figure 1: Magic Quadrant for Strategic Cloud Platform Services



Figura 7. Cuadrante mágico de Gartner para servicios cloud de infraestructura (IaaS) y plataforma (PaaS), a fecha de agosto de 2020. Fuente: Amazon Web Service, s. f.-a.

AWS ofrece un gran abanico de servicios en la nube mediante una infraestructura global que proporciona acceso desde casi cualquier parte del mundo, alta disponibilidad y seguridad, y un esquema de pago por uso (*pay-as-you-go*). Entre los **servicios** que ofrece, podemos encontrar los comunes a todas las plataformas de *cloud computing*, como son el almacenamiento, el procesamiento, las redes y la seguridad; pero también otros más específicos, relacionados con sectores concretos, como la analítica de datos, el DevOps, las bases de datos, la inteligencia artificial, el IoT, el desarrollo móvil, entre muchos otros.

Tema 1. Computación en la nube en la era del big data y la inteligencia artificial

Además, al igual que otras plataformas de *cloud computing*, permite desplegar **modelos** tanto completamente en la nube como de forma híbrida, es decir, ofreciendo algunos servicios en sus propios servidores y otros en la nube, y dando posibilidad de migrar cuando el usuario decida hacerlo.

Además de los servicios que ofrece la propia plataforma AWS, cuenta con el denominado AWS Marketplace, un **catálogo digital** con numerosos recursos adicionales proporcionados por empresas independientes de Amazon, pero basadas en los servicios en la nube de AWS, como imágenes de máquinas virtuales, herramientas de BI, bases de datos, etc.

¿Cuáles son las ventajas competitivas de AWS?

- ▶ **Escalabilidad y flexibilidad.** Escala a medida que tu negocio crece y paga solo por lo que usas.
- ▶ **Amplia gama de servicios.** Más de doscientos servicios y productos para cubrir todas las necesidades de infraestructura y desarrollo.
- ▶ **Seguridad y confiabilidad.** Cumplimiento de los más altos estándares de seguridad y disponibilidad.
- ▶ **Facilidad de uso.** Interfaz intuitiva y herramientas de gestión para facilitar la administración de tus recursos.
- ▶ **Innovación continua.** Líder en innovación, con lanzamientos regulares de nuevos servicios y características.

Principales servicios de AWS

Entre los principales servicios que la nube de Amazon ofrece, se pueden destacar de manera categorizada los siguientes:

- ▶ **Cómputo:** Amazon EC2, Amazon ECS y AWS Lambda.

Tema 1. Computación en la nube en la era del big data y la inteligencia artificial

- ▶ **Almacenamiento:** Amazon S3, Amazon EBS y Amazon Glacier.
- ▶ **Bases de datos:** Amazon RDS, Amazon DynamoDB y Amazon Redshift.
- ▶ **Redes:** Amazon VPC, Amazon Route 53 y AWS Direct Connect.
- ▶ **IA y ML:** Amazon SageMaker, AWS Deep Learning AMIs y Amazon Rekognition.
- ▶ **IoT:** AWS IoT Core, AWS IoT Greengrass y Amazon FreeRTOS.
- ▶ **Seguridad:** AWS IAM, AWS Key Management Service (KMS) y AWS WAF.
- ▶ **Desarrollo de aplicaciones:** AWS SDKs, AWS Amplify y AWS CloudFormation.

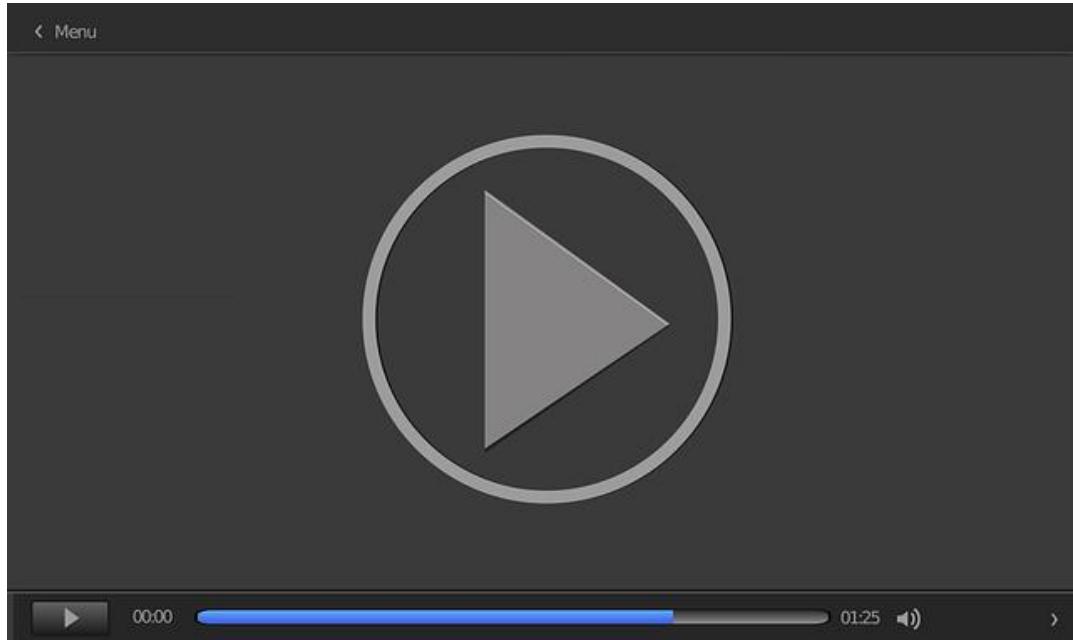
Casos de éxito de AWS

A continuación, se muestra una lista de casos de éxito de AWS en *big data* e IA:

- ▶ **Netflix.** La plataforma de *streaming* de vídeo más grande del mundo, Netflix, utiliza AWS para almacenar, procesar y analizar grandes volúmenes de datos de sus usuarios y contenido. AWS proporciona a Netflix la escalabilidad y las herramientas necesarias para recomendar contenido personalizado a sus usuarios, mejorar la experiencia del usuario y optimizar la entrega de contenido en tiempo real.

Tema 1. Computación en la nube en la era del big data y la inteligencia artificial

Si quieres saber más sobre cómo Netflix utiliza la plataforma AWS, te recomendamos ver el vídeo *Netflix Delivers Billions of Hours of Content Globally by Running on AWS* (Amazon Web Services, 2015).



Netflix Delivers Billions of Hours of Content Globally by Running on AWS.

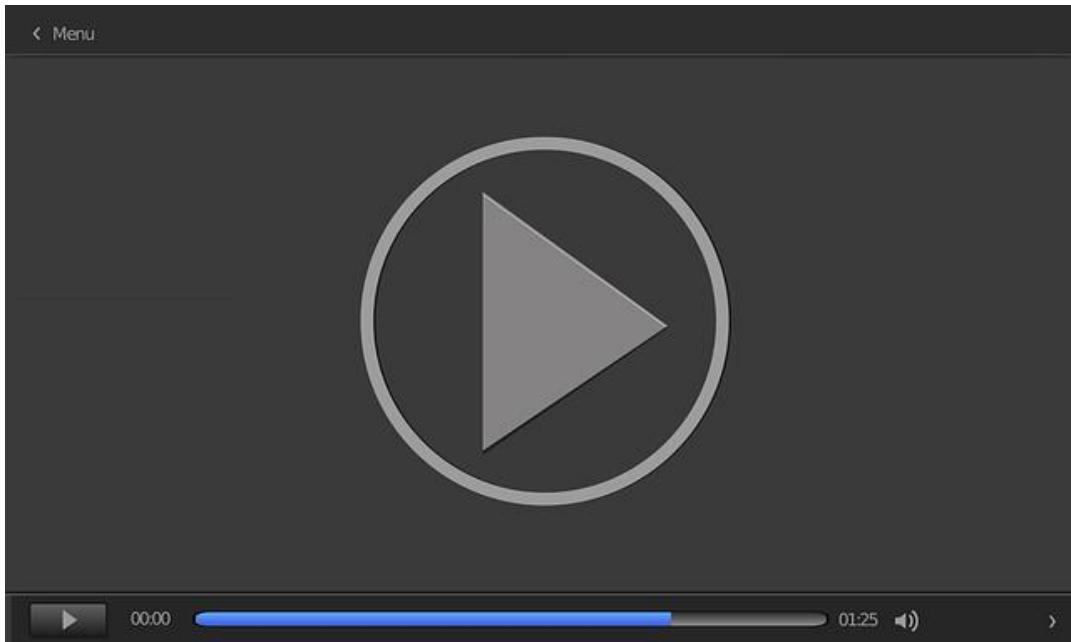
Accede al vídeo:

<https://www.youtube.com/embed/IQGHsBOZJBw>

- ▶ **Airbnb.** El popular servicio de alquiler de alojamientos Airbnb utiliza AWS para gestionar grandes volúmenes de datos generados por sus usuarios y anfitriones. AWS proporciona a Airbnb la capacidad de analizar datos en tiempo real para mejorar la segmentación de usuarios, ofrecer recomendaciones personalizadas y optimizar los precios de los alojamientos.

Tema 1. Computación en la nube en la era del big data y la inteligencia artificial

Si quieres saber más sobre cómo Airbnb utiliza la plataforma AWS, te recomendamos ver el vídeo *AWS re:Invent 2018: Airbnb's Journey from Self-Managed Redis to ElastiCache for Redis (DAT319)* (Amazon Web Services, 2018a).



AWS re:Invent 2018: Airbnb's Journey from Self-Managed Redis to ElastiCache for Redis (DAT319).

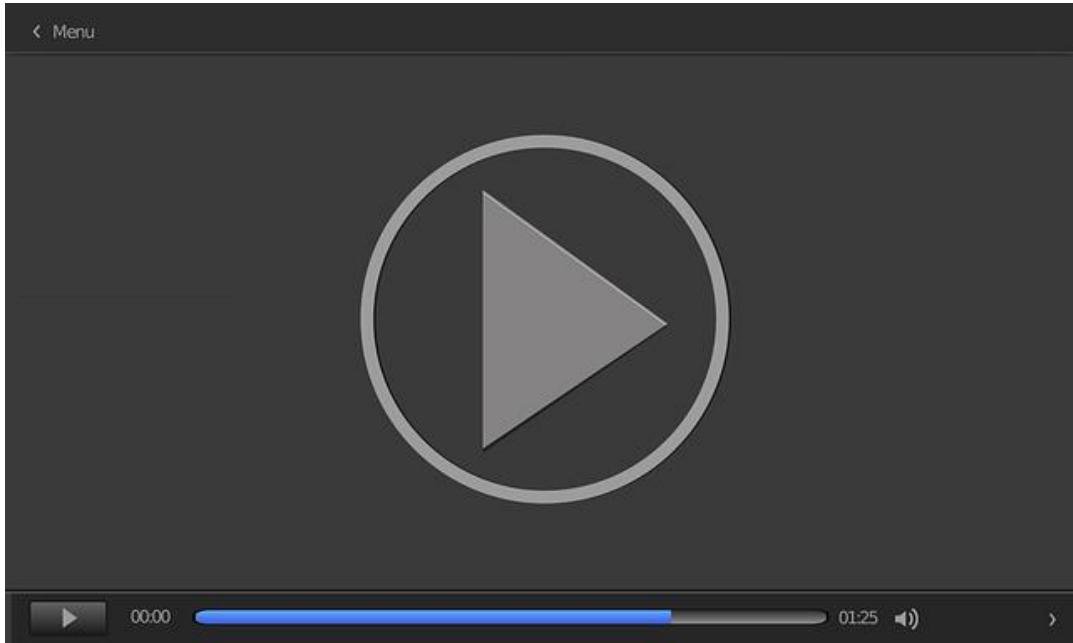
Accede al vídeo:

https://www.youtube.com/embed/eyd_8efUCwM

- ▶ **NASA.** La NASA utiliza AWS para almacenar y analizar grandes cantidades de datos generados por sus misiones espaciales y proyectos de investigación. AWS proporciona a la NASA la capacidad de procesar y analizar datos de telemetría, imágenes satelitales y simulaciones numéricas para obtener información valiosa sobre el espacio y la Tierra.

Tema 1. Computación en la nube en la era del big data y la inteligencia artificial

Si quieres saber más sobre cómo Airbnb utiliza la plataforma AWS, te recomendamos ver el vídeo *AWS re:Invent 2017: #EarthOnAWS: How NASA Is Using AWS (STG205)* (Amazon Web Services, 2017).



AWS re:Invent 2017: #EarthOnAWS: How NASA Is Using AWS (STG205).

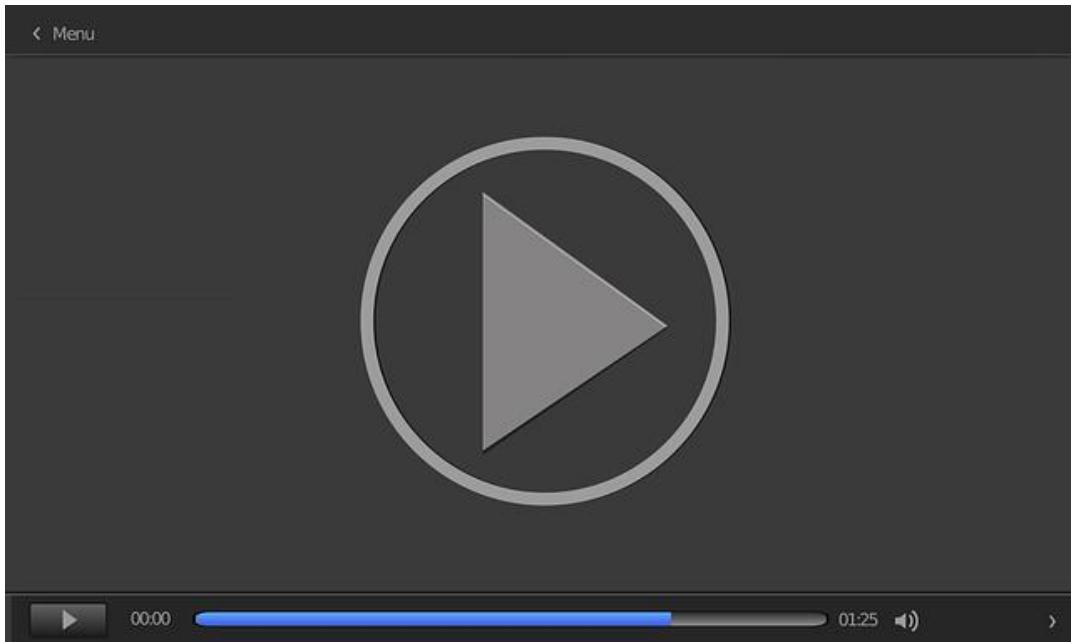
Accede al vídeo:

<https://www.youtube.com/embed/Sh7FB-tkYXM>

- ▶ **Pinterest.** La plataforma de descubrimiento visual Pinterest utiliza AWS para almacenar, procesar y analizar grandes volúmenes de datos generados por sus usuarios y contenido. AWS proporciona a Pinterest la capacidad de ofrecer recomendaciones personalizadas, mejorar la segmentación de anuncios y optimizar la experiencia del usuario en tiempo real.

Tema 1. Computación en la nube en la era del big data y la inteligencia artificial

Si quieres saber más sobre cómo Pinterest utiliza la plataforma AWS, te recomendamos ver el vídeo *AWS re:Invent 2023 - Pinterest extends existing data lake with generative AI / AWS Events* (Amazon Web Services, 2024).



WS re:Invent 2023 - Pinterest extends existing data lake with generative AI | AWS Events.

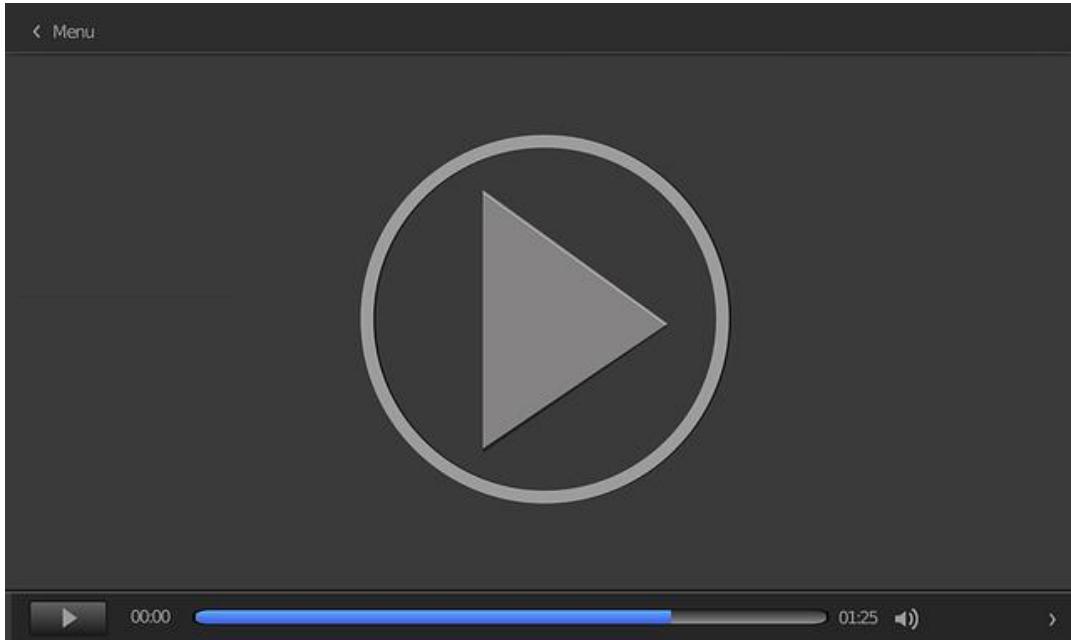
Accede al vídeo:

<https://www.youtube.com/embed/p6B7QX2osRU>

- ▶ **Capital One.** La empresa de servicios financieros Capital One utiliza AWS para desarrollar soluciones de *big data* e IA que mejoran la experiencia del cliente y optimizan las operaciones bancarias. AWS proporciona a Capital One la capacidad de analizar grandes volúmenes de datos de transacciones financieras para detectar fraudes, identificar oportunidades de *upselling* y ofrecer servicios personalizados a sus clientes.

Tema 1. Computación en la nube en la era del big data y la inteligencia artificial

Si quieres saber más sobre cómo Capital One utiliza la plataforma AWS, te recomendamos ver el vídeo *Capital One Reimagines Banking Using AWS* (Amazon Web Services, 2018b).



Capital One Reimagines Banking Using AWS.

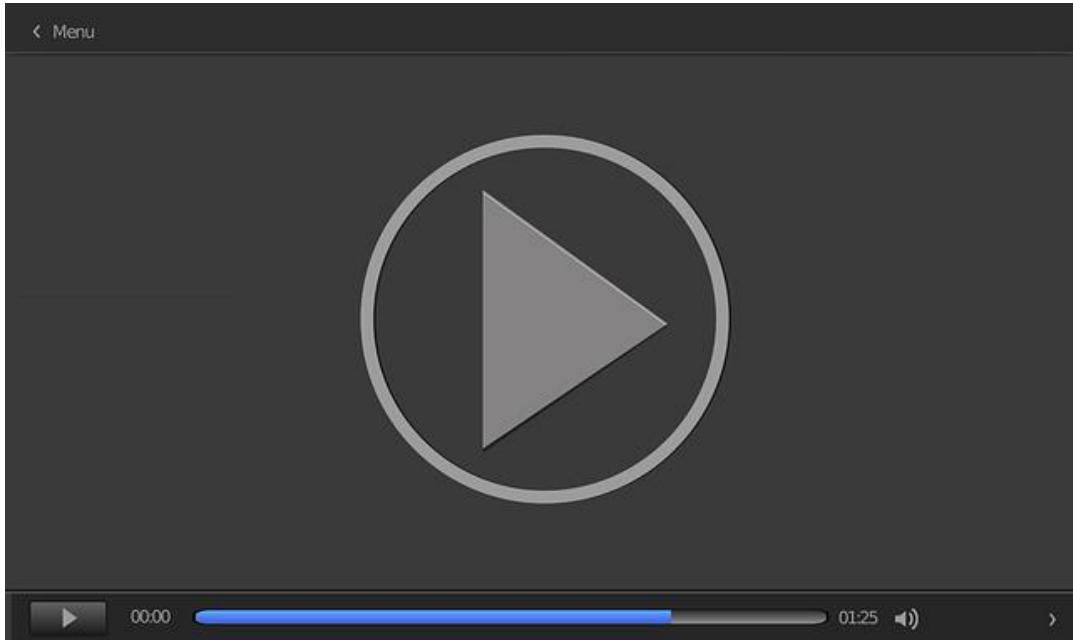
Accede al vídeo:

<https://www.youtube.com/embed/qU0HuWtzDC4>

- ▶ **General Electric (GE).** La empresa multinacional de tecnología e industria GE utiliza AWS para desarrollar soluciones de *big data* e IoT que mejoran la eficiencia operativa y optimizan el mantenimiento predictivo en sus productos y servicios. AWS proporciona a GE la capacidad de analizar grandes volúmenes de datos de sensores en tiempo real para identificar patrones de uso, predecir fallos mecánicos y mejorar la fiabilidad de sus productos y servicios.

Tema 1. Computación en la nube en la era del big data y la inteligencia artificial

Si quieres saber más sobre cómo General Electric utiliza la plataforma AWS, te recomendamos ver el vídeo *Accelerating Enterprise-Wide AI/ML Innovation: GE Healthcare and Autodesk* (Amazon Web Services, 2021).



Accelerating Enterprise-Wide AI/ML Innovation: GE Healthcare and Autodesk.

Accede al vídeo:

<https://www.youtube.com/embed/0vWIAOYkjRk>

Estos son solo algunos ejemplos de cómo empresas líderes están utilizando AWS para impulsar la innovación, mejorar la experiencia del cliente y optimizar sus operaciones mediante el aprovechamiento de soluciones de *big data* e IA.