

# Guía Práctica: AWS Bedrock con RAG

## Introducción

Esta guía te llevará paso a paso en la implementación de **AWS Bedrock** con **Retrieval-Augmented Generation (RAG)**, un enfoque que combina modelos de lenguaje con bases de conocimiento personalizadas para generar respuestas más precisas y contextuales.

### ¿Qué es RAG?

RAG (Retrieval-Augmented Generation) es una técnica que:

- **Recupera** información relevante de una base de conocimiento
- **Aumenta** el prompt del modelo con esta información
- **Genera** respuestas más precisas y fundamentadas

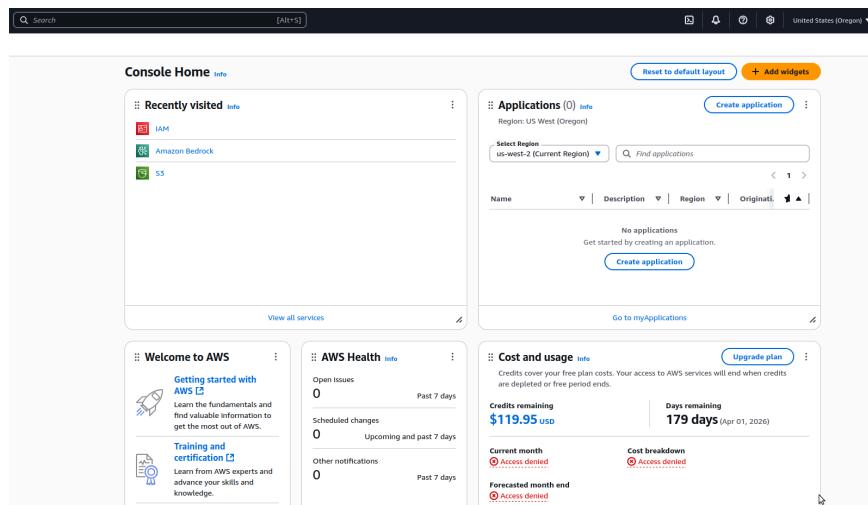
## Paso 1: Configuración Inicial de AWS

### 1.1 Crear Cuenta AWS

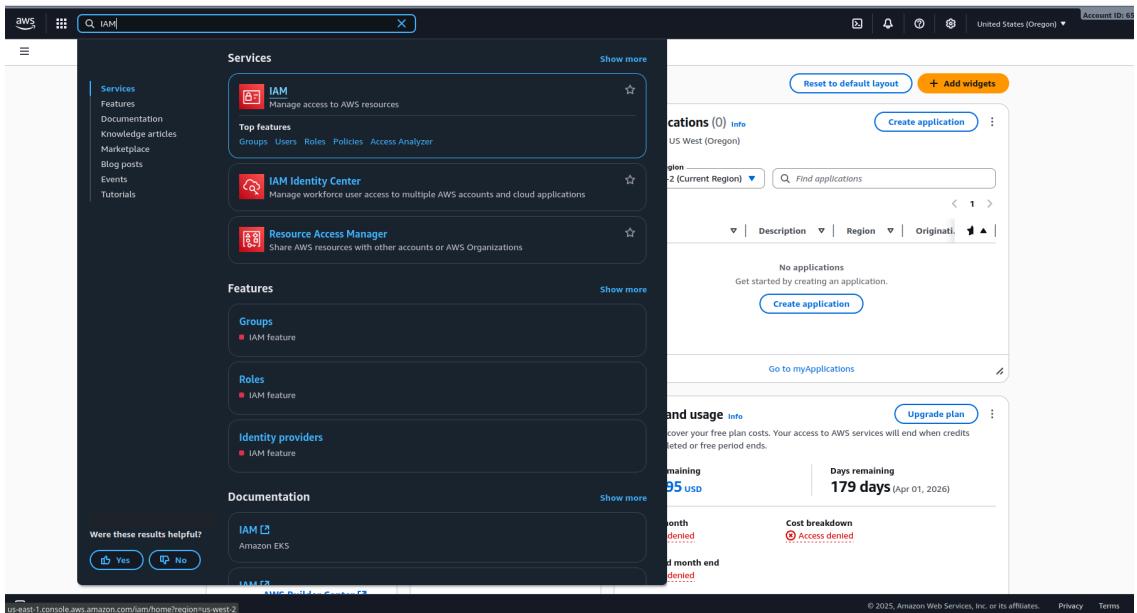
- Ve a <https://aws.amazon.com>
- Haz clic en "Create an AWS Account"
- Completa el proceso de registro (requiere tarjeta de crédito, pero hay tier gratuito)

### 1.2 Configurar IAM User

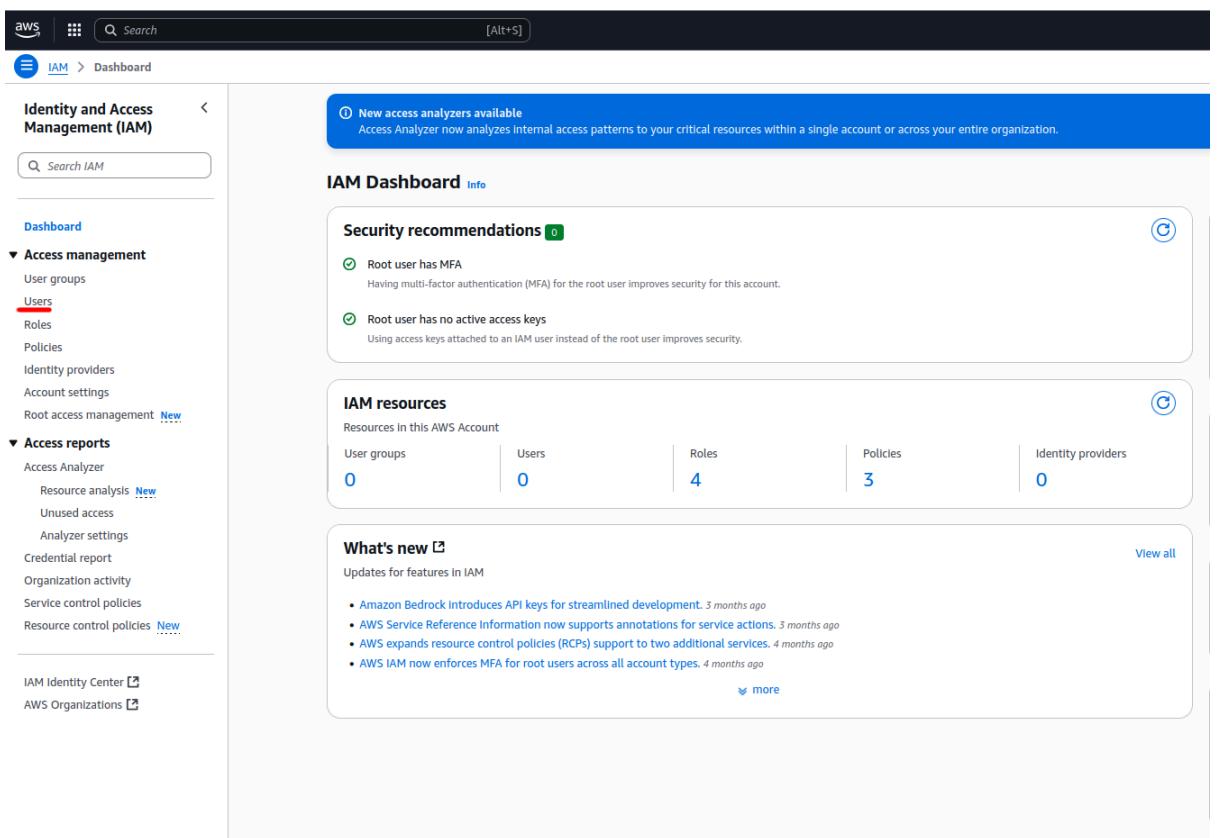
Una vez completado el anterior paso, seremos redirigidos aquí:



## 1. Buscar IAM en la barra de navegación superior



una vez hecho click serás redirigido a:



Ir a "users" en el menú lateral

## 2. Crear usuario:

- User name

### 3. Permisos:

- Attach existing policies directly
- Seleccionar: por simpleza le asignaremos el **AdministratorAccess** y ya

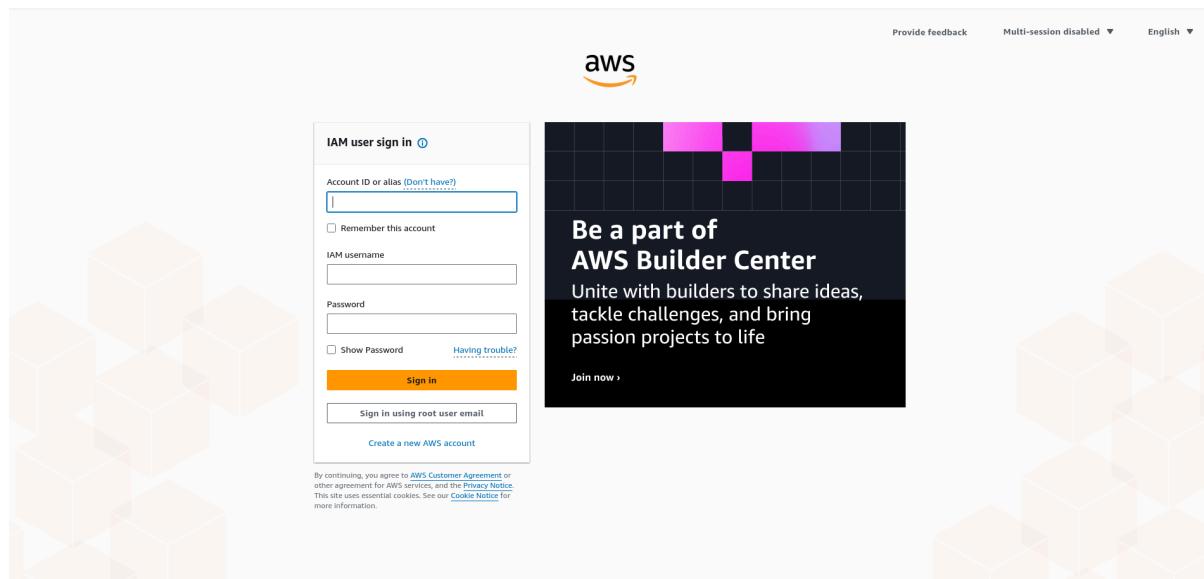
### 4. Crear usuario y guardar las credenciales

- Hacer click en crear usuario y ser redirigido a:

The screenshot shows the AWS IAM Users page. At the top, a green banner says "User created successfully". Below it, a table lists one user: "pablox1". The table has columns for User name, Path, Groups, Last activity, MFA, Password age, Console last sign-in, Access key ID, Active key age, Access key last use, and ARN. On the right side of the table, there are "Delete" and "Create user" buttons. The left sidebar shows navigation options like Dashboard, Access management (selected), User groups, Users, Roles, Policies, Identity providers, Account settings, Root access management, Access reports, Access Analyzer, and IAM Identity Center.

5. Con el usuario creado, copiamos nuestro account-id que sale en la parte superior derecha que se puede observar en la anterior imagen, para posteriormente hacer sign out.

6. Iniciar sesión con el usuario creado. Se realiza el inicio de sesión



**Nota:** Inicialmente la password es la misma que la cuenta principal, pero cuando entres con el usuario que acabas de crear te pedirá cambiarla para ese usuario en específico.

## Paso 2: Preparar Datos en Amazon S3

## 2.1 ¿Por qué usar S3?

Amazon S3 nos permite almacenar los documentos que nuestra base de conocimiento utilizará para el RAG.

### 1. Ubicarse en el home

The screenshot shows the AWS Home page. On the left, there's a sidebar with links like Amazon Bedrock, Billing and Cost Management, and Amazon Polly. The main area has three sections: "Welcome to AWS" (with Getting started with AWS, Training and certification, and AWS Builder Center), "AWS Health" (with Open issues, Scheduled changes, and Other notifications), and "Cost and usage" (showing credits remaining of \$99.95 USD, days remaining of 181 days, and a chart of forecasted monthly costs from May 25 to Oct 25).

### 2. Buscar "S3" en la barra de navegación superior

The screenshot shows the search results for "S3". The top navigation bar has a search bar with "S3" typed in. The search results show the "Services" section with cards for S3 (Scalable Storage in the Cloud), S3 Glacier (Archive Storage in the Cloud), and AWS Snow Family (Large Scale Data Transport). Below this are sections for "Features" (Imports from S3, Feature spotlight, S3 Access Grants) and "Resources" (Introducing resource search). To the right, there's a sidebar with "Applications" (no applications listed), "Cost and usage" (showing credits remaining of \$0.05 USD, days remaining of 181 days, and a chart of forecasted monthly costs from May 25 to Oct 25), and "Billing and Cost Management".

### 3. Crear bucket:

- dejamos todos los campos por defecto, en este caso lo único que se cambió fue el nombre, se le colocó "ia2demo"
- Después de crear el bucket, serás redireccionado a la lista de buckets.

The screenshot shows the AWS S3 console interface. On the left, there's a sidebar with navigation links like 'Amazon S3', 'General purpose buckets', 'Storage Lens', and 'AWS Marketplace for S3'. The main area displays a table for 'General purpose buckets'. The table has one row with the following details:

Name	AWS Region	Creation date
ia2demo	US West (Oregon) us-west-2	October 3, 2025, 23:10:59 (UTC-05:00)

On the right side of the main area, there are three informational boxes: 'Account snapshot', 'External access summary - new', and 'Block Public Access settings for this account'.

## 2.2 Subir Documentos

- Seleccionar el bucket recién creado**
- Hacer clic en "Upload" para ser redireccionado aquí:**

The screenshot shows the 'Upload objects' page for the 'ia2demo' bucket. The top navigation bar includes links for 'Buckets' and 'Upload'. The main content area is divided into several sections:

- Upload**: A large input field for dragging and dropping files or choosing them from a file browser.
- Files and folders (0)**: A table showing the selected files and folders. It includes columns for Name, Folder, Type, and Size. A message indicates 'No files or folders' have been chosen.
- Destination**: Set to 's3://ia2demo'. Below it, 'Destination details' and 'Permissions' (granting public access) are shown.
- Properties**: Options for specifying storage class, encryption settings, tags, and more.
- Buttons**: 'Cancel' and 'Upload' buttons at the bottom right.

- Añadir archivos:** Seleccionar documentos PDF, TXT o DOC, en este caso añadiremos el siguiente pdf:  
<https://drive.google.com/file/d/1oktMo1th-JxuKEd5XVvbncGC1qWxciSb/view?usp=ssharing>

#### 4. Hacer clic en "Upload" y debería de mostrarse esto:

The screenshot shows the AWS Lambda console interface. At the top, there's a green success message: "Upload succeeded. For more information, see the Files and folders table." Below it, a banner says "Upload: status" with a "Close" button. A note says "After you navigate away from this page, the following information is no longer available." The main area has a "Summary" section with "Destination" set to "s3://iazedemo". It shows two rows: "Succeeded" (1 file, 210.9 KB (100.0%)) and "Failed" (0 files, 0 B (0%)). Below this is a "Files and folders" table with one entry: "BATMAN.pdf" (application/pdf, 210.9 KB, Succeeded). The table has columns for Name, Folder, Type, Size, Status, and Error.

## Paso 3: Configurar AWS Bedrock

### 3.1 Acceder a Bedrock

- Volver al home
- Buscar "Amazon Bedrock" en la barra de navegación

The screenshot shows the AWS search results for "Amazon bedrock". The search bar at the top contains the query. On the left, a sidebar lists services like Features, Resources, Documentation, and Tutorials. The main content area shows three service cards: "Amazon Bedrock" (selected), "Amazon Bedrock AgentCore", and "AWS Private Certificate Authority". Below these are sections for "Features" (Prompt Management, Imported Models, Marketplace Deployments) and "Resources" (Introducing resource search). On the right, there's a "Billing and usage" section showing a chart of remaining credits (181 days) and a "Create application" button. The top navigation bar includes account information (Account ID: 6584-8522-1516, United States (Oregon), Pablo).

- Seleccionar el servicio

### 3.2 Habilitar Model Access

¿Por qué? AWS requiere que habilites explícitamente los modelos que quieras usar.

## 1. Ir a "Model access" en el menú lateral

The screenshot shows the Amazon Bedrock console interface. On the left, there is a sidebar with several sections: **Image / Video playground**, **Watermark detection**, **Infer** (with sub-options: Cross-region inference, Batch inference, Provisioned Throughput, Custom model on-demand), **Tune** (with sub-options: Custom models, Prompt router models, Imported models, Marketplace model deployment), **Build** (with sub-options: Agents, Flows, Knowledge Bases, Automated Reasoning, Guardrails, Prompt Management, Data Automation, AgentCore), **Assess** (with sub-options: Evaluations), **Configure and learn** (with sub-options: Settings, Model access, User guide, Bedrock Service Terms). The **Model access** option is highlighted with a red border.

The main content area has three main sections: **Overview** (Info), **Model catalog** (Amazon Bedrock supports over 100 foundation models from industry-leading providers and emerging leaders. Select a serverless model or Bedrock Marketplace model that is best suited for achieving your unique goals. [View Model catalog](#)), and **Model spotlight** (DeepSeek V3.1, Hybrid reasoning model. [Open in chat playground](#)). Below these are sections for **Get started by using API Keys** (Create an Amazon Bedrock API key in one click using the console and use it to authenticate API requests to Amazon Bedrock. After you create the key, you can integrate it into your setup to be automatically recognized when making requests. [View API keys](#)), **Test** (Chat / Text playground, Experiment with Bedrock models to generate high-quality text for a variety of language processing tasks. Choose between single-prompt generation or chat-style interactions that incorporate context from previous prompts. [Open playground](#)), **Image / Video playground** (Use Bedrock models to create and edit high-quality images or generate videos using simple text prompts. For videos, choose between single-shot generation or try the storyboarding capability for more control over visual outputs. [Open playground](#)), and **Watermark detection** (Verify whether an image was generated by a Titan Image Generator model by checking for the presence of an embedded watermark. [View watermark detection](#)).

## 2. Hacer clic en "Modify model access"

The screenshot shows the 'Modify model access' page. At the top, there is a yellow banner with the following text: **⚠️ Model access page retiring Oct 8 2025**. It states: Starting Oct 8 2025, Amazon Bedrock will simplify access to all serverless foundation models, and any new models, by automatically enabling them for every AWS account, eliminating the need to manually activate access through the Bedrock console. Account administrators retain full control over model access through IAM policies and Service Control Policies (SCPs) to restrict model access as needed. Note: Selected new models launched before Oct 8 2025 will be auto-enabled, so they won't appear on this page; you can still manage their access via IAM and SCPs. For a complete list of supported models, refer to our documentation.

Below the banner, there is a section titled **What is Model access?** with the following text: To use Bedrock serverless models, account users with the correct IAM Permissions must enable access to available Bedrock foundation models (FMs). View all Bedrock Model Terms for Bedrock FMs. There is a **Modify model access** button.

Further down, there is a section titled **Base models (63)** with the following text: Not seeing a model you're interested in? Check out all supported models by region [here](#). There is a search bar labeled **Find model** and a dropdown menu labeled **Group by provider**.

Models	Access status	Modality	EULA
▼ Amazon (12)	1/12 access granted		
Titan Text G1 - Lite	<span style="color: green;">Access granted</span>	Text	<a href="#">EULA</a>
Titan Text G1 - Express	<span style="color: grey;">Available to request</span>	Text	<a href="#">EULA</a>
Titan Embeddings G1 - Text	<span style="color: grey;">Available to request</span>	Embedding	<a href="#">EULA</a>
Titan Text Embeddings V2	<span style="color: grey;">Available to request</span>	Embedding	<a href="#">EULA</a>
Titan Image Generator G1	<span style="color: grey;">Available to request</span>	Image	<a href="#">EULA</a>
Titan Image Generator G1 v2	<span style="color: grey;">Available to request</span>	Image	<a href="#">EULA</a>

## 3. Seleccionar modelos:

- Seleccionamos toda la familia de modelos de amazon:

The screenshot shows the 'Base models' section of the Amazon Bedrock interface. It lists various AI models grouped by provider. The 'Amazon (12)' group includes Titan Text G1 - Lite, Titan Text G1 - Express, Titan Embeddings G1 - Text, Titan Text Embeddings V2, Titan Image Generator G1, Titan Image Generator G1 v2, Titan Multimodal Embeddings G1, Rerank 1.0, Nova Pro, Nova Lite, Nova Micro, and Nova Premier. The 'Anthropic (1)' group includes Claude Connect A. The 'Modality' column indicates the type of model (Text, Embedding, Image, Text & Vision), and the 'EULA' column shows the license agreement status.

#### 4. Hacer clic en "Next" y posteriormente en "Save changes"

**Nota:** Hay que esperar unos minutos para que nos habiliten el acceso a estos modelos.

## Paso 4: Crear Base de Conocimiento (Knowledge Base)

### 4.1 ¿Qué es una Knowledge Base?

Es el componente central del RAG que:

- **Almacena** tus documentos procesados
- **Crea embeddings** (representaciones vectoriales del texto)
- **Permite búsquedas semánticas** para encontrar información relevante

### 4.2 Proceso de Creación

1. Ir a "Knowledge bases" en el menú lateral
2. Seleccionar "Create knowledge base" → "With vector store"

The screenshot shows the 'Create knowledge base' wizard in the Amazon Bedrock interface. Step 1: 'How it works' explains three methods: Vector store, Structured data store, and Kendra GenAI Index. Step 2: 'Test the Knowledge Base' shows a test window where you can query your knowledge base. Step 3: 'Use the Knowledge Base' shows how to integrate it into an application. The final step shows the 'Knowledge Bases' list with a 'Create' button. A sidebar on the left provides navigation links for Discover, Infer, Tune, and Build.

## Paso 4.2.1: Configure knowledge base details

- dejaremos todos los campos por defecto y hacemos click en “Next”:

The screenshot shows the 'Provide Knowledge Base details' step in the Amazon Bedrock console. On the left, there's a sidebar with navigation links like 'Discover', 'Test', 'Infer', 'Tune', 'Build', and 'Knowledge Bases'. The main area has a title 'Provide Knowledge Base details' with a progress bar showing Step 1 (selected), Step 2, Step 3, and Step 4. Under 'Knowledge Base details', there's a 'Knowledge Base name' input field containing 'knowledge-base-quick-start-f35dv'. Below it is a 'Knowledge Base description - optional' input field with placeholder text. Under 'IAM permissions', there are options for creating a new service role or using an existing one, and a service role name 'AmazonBedrockExecutionRoleForKnowledgeBase\_f35dv'. In the 'Choose data source type' section, 'Amazon S3' is selected, showing its description as an object storage service. Other options include 'Web Crawler - Preview' and 'Custom'.

## Paso 4.2.2: Configure data source

The screenshot shows the 'Configure data source' step in the Amazon Bedrock console. The sidebar shows the previous step was 'Provide Knowledge Base details'. The main area has a title 'Configure data source' with a progress bar showing Step 1 (selected), Step 2 (selected), Step 3, and Step 4. Under 'Data source: knowledge-base-quick-start-fdrp2-data-source', there's a 'Data source name' input field with 'knowledge-base-quick-start-fdrp2-data-source'. The 'Data source location' section shows 'This AWS account' selected. The 'S3 URI' section includes fields for 'Bucket name' (with placeholder 's3://<bucket-name>/<prefix>/<object>'), 'View' (with a 'Browse S3' button), and 'Format'. There's also a checkbox for 'Add customer-managed KMS key for S3 data - optional'. The 'Parsing strategy' section includes 'Amazon Bedrock default parser' (selected), 'Amazon Bedrock Data Automation as parser', and 'Foundation models as a parser'. The 'Chunking strategy' section includes 'Default chunking' (selected) with a note about token limits.

Los campos los dejamos por defecto a excepción del S3 URI, ahí le haremos click en “Browse S3” y seleccionaremos el bucket que creamos anteriormente para posteriormente clickear en “Next”.

## Paso 4.2.3: Configure embeddings and vector store

Step 1  
Provide Knowledge Base details

Step 2  
Configure data source

**Step 3  
Configure data storage and processing**

Step 4  
Review and create

### Configure data storage and processing

Choose an embeddings model to convert the data that you will provide in the next step, and provide details for a vector data store in which Bedrock can store, manage, and update your embeddings. The embeddings model and vector store cannot be changed after creation of Knowledge Base.

#### Embeddings model

Select an embeddings model to convert your data into an embedding. Your selection may limit vector stores and embedding types that are available. Pricing depends on the model. [Learn more](#)

**Select model**

Additional configurations

#### Vector store Info

Let Amazon create a vector store on your behalf or select a previously created store to allow Bedrock to store, update and manage embeddings. You will be billed directly from the vector store provider. [Learn more](#)

#### Vector store creation method

**Quick create a new vector store - Recommended**  
A vector store will be created on your behalf in this AWS account during Knowledge Base creation.

Use an existing vector store  
Connect to an existing vector store to store, update, and manage embeddings.

#### Vector store type - new Info

A vector store holds the vector embeddings representation of your data. This choice cannot be changed later. Not all vector stores support binary vector embeddings.

Select a vector store

Cancel Previous Next

- **Embeddings model:** seleccionamos el modelo "Amazon Titan Embeddings V2"
  - **¿Qué son embeddings?** Son representaciones numéricas del texto que permiten comparar similitudes semánticas
- **Vector store:** "Quick create a new vector store"
- **Vector store type:** Amazon S3 vectores preview

Así quedaría la configuración completa para posteriormente clickear en "Next":

Step 1  
Provide Knowledge Base details

Step 2  
Configure data source

**Step 3  
Configure data storage and processing**

Step 4  
Review and create

### Configure data storage and processing

Choose an embeddings model to convert the data that you will provide in the next step, and provide details for a vector data store in which Bedrock can store, manage, and update your embeddings. The embeddings model and vector store cannot be changed after creation of Knowledge Base.

#### Embeddings model

Select an embeddings model to convert your data into an embedding. Your selection may limit vector stores and embedding types that are available. Pricing depends on the model. [Learn more](#)

**Titan Text Emb...**  On-demand

Additional configurations

#### Vector store Info

Let Amazon create a vector store on your behalf or select a previously created store to allow Bedrock to store, update and manage embeddings. You will be billed directly from the vector store provider. [Learn more](#)

#### Vector store creation method

**Quick create a new vector store - Recommended**  
A vector store will be created on your behalf in this AWS account during Knowledge Base creation.

Use an existing vector store  
Connect to an existing vector store to store, update, and manage embeddings.

#### Vector store type - new Info

A vector store holds the vector embeddings representation of your data. This choice cannot be changed later. Not all vector stores support binary vector embeddings.

**Amazon S3 Vectors - Preview**  
Select to store and index vector embeddings optimized for the cost-effective and durable storage of large, long-term vector data sets while maintaining sub-second query performance.

Additional configurations

Cancel Previous Next

## Paso 4.2.4: Review and create

Bases > Create knowledge base with vector store

S3 URI s3://ia2demo	Chunking strategy Default	S3 bucket for Lambda function -
		Data deletion policy DELETE

**Step 3: Configure data storage and processing**

<b>Embeddings model</b>		
Model Titan Text Embeddings v2	Embedding type Float vector embeddings	Vector dimensions 1024
<b>Vector store</b>		
Quick create vector store - recommended Amazon S3 Vectors	Encryption key ARN -	
<b>Multimodal storage destination</b>		
S3 URI -		

**Create Knowledge Base**

- Revisar el resumen de la configuración
- Hacer clic en "Create knowledge base"
- Si se ha creado correctamente la base de conocimiento serás redireccionado aquí:

Amazon Bedrock > Knowledge Bases > knowledge-base-quick-start-4sg22

Discover	Knowledge Base 'knowledge-base-quick-start-4sg22' created successfully. Sync one or more data sources to index your content for searching. Syncing can take from a few minutes to a few hours.									
Infer	Test Knowledge Base   Delete   Edit									
Tune	Log Deliveries   Configure log deliveries and event logs in the <a href="#">Edit</a> page.									
Build	Retrieval-Augmented Generation (RAG) type   Vector store									
<b>Knowledge Base overview</b> <table border="1"> <tr> <td>Knowledge Base name knowledge-base-quick-start-4sg22</td> <td>Knowledge Base ID Z7OKLR9GGL</td> <td>Log Deliveries</td> </tr> <tr> <td>Knowledge Base description —</td> <td>Status Available</td> <td>Configure log deliveries and event logs in the <a href="#">Edit</a> page.</td> </tr> <tr> <td>Service Role AmazonBedrockExecutionRoleForKnowledgeBase_4sg22</td> <td>Created date October 04, 2025, 16:54 (UTC-05:00)</td> <td>Retrieval-Augmented Generation (RAG) type Vector store</td> </tr> </table>		Knowledge Base name knowledge-base-quick-start-4sg22	Knowledge Base ID Z7OKLR9GGL	Log Deliveries	Knowledge Base description —	Status Available	Configure log deliveries and event logs in the <a href="#">Edit</a> page.	Service Role AmazonBedrockExecutionRoleForKnowledgeBase_4sg22	Created date October 04, 2025, 16:54 (UTC-05:00)	Retrieval-Augmented Generation (RAG) type Vector store
Knowledge Base name knowledge-base-quick-start-4sg22	Knowledge Base ID Z7OKLR9GGL	Log Deliveries								
Knowledge Base description —	Status Available	Configure log deliveries and event logs in the <a href="#">Edit</a> page.								
Service Role AmazonBedrockExecutionRoleForKnowledgeBase_4sg22	Created date October 04, 2025, 16:54 (UTC-05:00)	Retrieval-Augmented Generation (RAG) type Vector store								
<b>Data source (1)</b> Data sources contain information returned when querying a Knowledge Base. <table border="1"> <thead> <tr> <th>Find data source</th> <th>Sync</th> <th>Stop sync</th> <th>Add</th> </tr> </thead> <tbody> <tr> <td><input type="checkbox"/> knowledge... <span>Available</span></td> <td>S3</td> <td>6584522... <a href="#">s3://ia2de...</a></td> <td>Default</td> </tr> </tbody> </table>		Find data source	Sync	Stop sync	Add	<input type="checkbox"/> knowledge... <span>Available</span>	S3	6584522... <a href="#">s3://ia2de...</a>	Default	
Find data source	Sync	Stop sync	Add							
<input type="checkbox"/> knowledge... <span>Available</span>	S3	6584522... <a href="#">s3://ia2de...</a>	Default							
<b>Tags</b> A tag is a label that you assign to an AWS resource. Each tag consists of a key and an optional value. You can use tags to search and filter your resources or track your AWS costs. <table border="1"> <thead> <tr> <th>Key</th> <th>Value</th> </tr> </thead> <tbody> <tr> <td colspan="2">No tags</td> </tr> </tbody> </table>		Key	Value	No tags						
Key	Value									
No tags										

## 4.3 Sincronizar Data Source

¿Por qué? Para procesar los documentos y crear los embeddings.

1. Seleccionar la knowledge base creada
2. Ir a la pestaña "Data sources"
3. Seleccionar el data source

**Data source overview**

- Name:** knowledge-base-quick-start-jk3q5-data-source
- Data source type:** Amazon S3
- Serverside KMS key:** -
- Account ID:** -
- Chunking strategy:** Default
- Lambda function:** -

**Data source ID:** YS8P7DNXDS  
**Created date:** October 04, 2025, 16:54 (UTC-05:00)  
**Status:** Available  
**Data deletion policy:** DELETE  
**Parsing strategy:** Default  
**S3 bucket for Lambda function:** -

**Sync history (0)**

Start time	End time	Status	Source files	Metadata files	Failed files	Added	Deleted	Modified	Metadata fil..
No sync history									

**Documents (0)**

Find document	Delete document	Add documents
< 1 >		

#### 4. Hacer clic en "Sync"

**Nota:** Si aparece error de permisos, verificar que el modelo de embedding que colocaste esté habilitado en "Model access".

## Paso 5: Probar la Knowledge Base

### 5.1 Test en Consola

1. Ir a "Knowledge bases" en el menú lateral
2. Ir a tu knowledge base creada
3. Hacer clic en "Test knowledge base"

**Knowledge Base overview**

- Knowledge Base name:** knowledge-base-quick-start-4sg22
- Knowledge Base description:** -
- Service Role:** AmazonBedrockExecutionRoleForKnowledgeBase\_4sg22

**Knowledge Base ID:** Z7OKL95GGU  
**Status:** Available  
**Created date:** October 04, 2025, 16:54 (UTC-05:00)

**Log Deliveries:** Configure log deliveries and event logs in the [Edit](#) page.  
**Retrieval-Augmented Generation (RAG) type:** Vector store

**Data source (1)**

Find data source	Data source ...	Status	Data source type	Account ID	Source Link	Last sync time	Last sync warn...	Chunking stra...	Parsing strategy	Data deletion ...
	<input type="checkbox"/> knowledge-ba...	Available	S3	65848522151...	s3://la2demo	October 04, 20...	-	Default	Default	Delete

**Tags**

A tag is a label that you assign to an AWS resource. Each tag consists of a key and an optional value. You can use tags to search and filter your resources or track your AWS costs.

Key	Value
No tags	

**Embedding model**

4. **Seleccionar modelo:** Puedes usar cualquiera, sin embargo para esta guía se usará el "Nova Micro 1.0" ya que habilitamos el uso de todos estos modelos de amazon anteriormente.
5. **Hacer preguntas** sobre el contenido de tus documentos

The screenshot shows the AWS Bedrock interface. On the left, under 'Configurations', there's a section for 'Retrieval and response generation' where 'Retrieval and response generation: data sources and model' is selected. A 'Model' section shows 'Nova Micro 1.0' is chosen. Below that are sections for 'Source' (with 'Source chunks' and 'Search Type'), 'Data manipulation' (with 'Filters' and 'Guardrails'), and a 'Preview' window. In the 'Preview' window, a question 'Quién es Batman?' is asked, and the response is: 'Batman es un personaje de ficción que ha trascendido las páginas del cómic para convertirse en un símbolo universal de disciplina, inteligencia y justicia implacable.<sup>[1]</sup> Su origen se remonta a 1939, cuando Bob Kane y Bill Finger crearon un nuevo héroe para Detective Comics.<sup>[2][3]</sup> A diferencia de otros superhéroes, Batman no posee superpoderes. Su fuerza proviene de su voluntad inquebrantable, de una mente brillante, y del dolor de un trauma que transformó la tragedia en propósito.<sup>[4]</sup> El asesinato de sus padres no solo lo marcó emocionalmente, sino que lo convirtió en el reflejo más humano del mito heroico.<sup>[5]</sup> Bruce Wayne nació en el seno de una de las familias más admiradas de Gotham City.<sup>[6]</sup> Durante más de una década, perfeccionó su cuerpo y su mente hasta alcanzar el límite humano.<sup>[7]</sup> Una noche, mientras meditaba en su mansión, un murciélagos rompió el ventanal e intruyó en su estudio. Bruce vio en ello un presagio. El murciélagos sería su emblema.<sup>[8]</sup> Así nació Batman, el vigilante que infunde terror en los corazones de los criminales.<sup>[9]</sup> Gotham no es solo un escenario; es un personaje vivo. Oscura, gótica y corrupta, representa el espejo distorsionado del alma humana y, al mismo tiempo, del propio Bruce Wayne.<sup>[10]</sup>'.

## 5.2 ¿Qué observar?

- **Respuestas contextuales** basadas en tus documentos
- **Citas numeradas** que referencian las fuentes
- **Hacer clic en las citas** para ver el texto original

# Paso 6: Demo Práctica con Python

## 6.1 Código de la Demo

descarga el archivo llamado “`demo_bedrock.py`” del repositorio

## 6.2 Instalación de Librerías Requeridas

Antes de ejecutar el código, instala la librería necesaria:

```
pip install boto3
```

### ¿Qué es boto3?

- Es el **SDK oficial de AWS para Python**
- Permite interactuar con servicios de AWS como Bedrock, S3, IAM, etc.

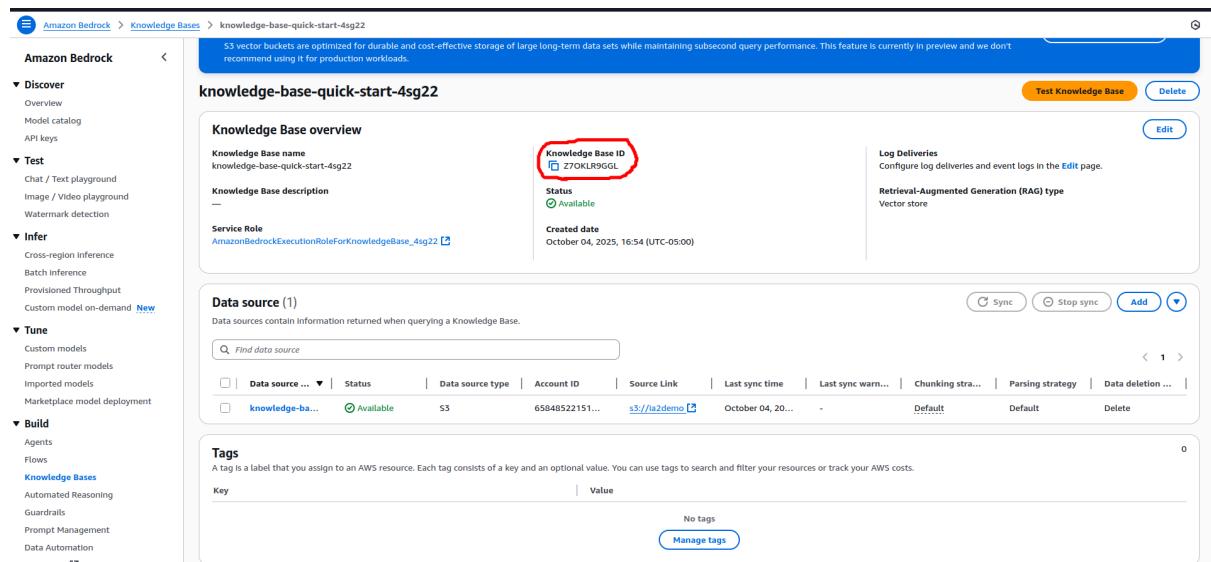
## 6.3 Configuración de Valores

Reemplaza estos 4 valores en el código:

1. **KNOWLEDGE\_BASE\_ID**: Tu Knowledge Base ID de Bedrock
2. **AWS\_ACCESS\_KEY**: Tu AWS Access Key ID
3. **AWS\_SECRET\_KEY**: Tu AWS Secret Access Key
4. **AWS\_REGION**: **us-west-2** (ya configurada, pero debes cambiarla según tu región)
5. **MODEL\_NAME**: **nova-micro-1-0** (ya configurado)

¿Dónde encontrar el **KNOWLEDGE\_BASE\_ID**?

se encuentra dirigiéndose a la base de conocimiento que creamos anteriormente:



The screenshot shows the Amazon Bedrock Knowledge Bases console. On the left, there's a sidebar with navigation links for Discover, Test, Infer, Tune, and Build. The main area displays the 'knowledge-base-quick-start-4sg22' overview. It includes sections for Knowledge Base overview (with a note about vector buckets), Knowledge Base description, Service Role (AmazonBedrockExecutionRoleForKnowledgeBase\_4sg22), and Log Deliveries (Retrieval-Augmented Generation (RAG) type, Vector store). Below this, there's a Data source section with a table showing one entry: knowledge-base-quick-start-4sg22 (Status: Available, Source Link: s3://a2demo, Last sync time: October 04, 2025, 16:54 (UTC-05:00)). At the bottom, there's a Tags section where users can manage tags for the resource.

¿Dónde encontrar el **AWS\_ACCESS\_KEY** y **AWS\_SECRET\_KEY**?

Haremos click al nuestro nombre de usuario que aparece en la parte superior derecha de la pantalla y le haremos click a security credentials

The screenshot shows the AWS Console Home page. At the top right, it displays the account ID (6584-0522-1516), United States (Oregon) region, and a message about free access ending on April 1, 2026. The 'Security credentials' section is highlighted.

**Free plan status**

- Credits remaining: \$119.95 USD
- Days remaining: 179 days

**Applications (0)**

No applications. Get started by creating an application.

**Cost and usage**

Credits cover your free plan costs. Your access to AWS services will end when credits are depleted or free period ends.

**Security credentials**

Account ID: 6584-0522-1516  
 Account color: Unset  
 IAM user: pablox1  
 Account: Organization: Service Quotas: Billing and Cost Management: Security credentials:

Turn on multi-session support | Switch role | Sign out

Serás redirigido a la siguiente página, donde deberás hacer click en “create access key para después seleccionar la opción de CLI”:

The screenshot shows the 'My security credentials' page. The 'AWS IAM credentials' tab is selected.

**Account details**

- User name: pablox1
- AWS account ID: [REDACTED]
- User ARN: [REDACTED]
- Canonical user ID: [REDACTED]

**AWS IAM credentials**

Console sign-in

- Console sign-in link: https://658485221516.sigin.aws.amazon.com/console
- Console password: Updated 20 hours ago (2025-10-04 16:50 GMT-5)
- Last console sign-in: 3 hours ago (2025-10-05 10:11 GMT-5)

**Multi-factor authentication (MFA) (0)**

Use MFA to increase the security of your AWS environment. Signing in with MFA requires an authentication code from an MFA device. Each user can have a maximum of 8 MFA devices assigned. [Learn more](#)

Type	Identifier	Certifications	Created on
No MFA devices. Assign an MFA device to improve the security of your AWS environment			
<a href="#">Assign MFA device</a>			

**Access keys (0)**

Use access keys to send programmatic calls to AWS from the AWS CLI, AWS Tools for PowerShell, AWS SDKs, or direct AWS API calls. You can have a maximum of two access keys (active or inactive) at a time. [Learn more](#)

[Create access key](#)

Después de los pasos anteriores nos arrojará las credenciales las cuales deberemos reemplazar en el código:

The screenshot shows the AWS IAM Access Key creation wizard. The top bar is green with the message: "Access key created. This is the only time that the secret access key can be viewed or downloaded. You cannot recover it later. However, you can create a new access key any time." On the left, a sidebar lists steps: Step 1 (Access key best practices & alternatives), Step 2 (optional Set description tag), Step 3 (Retrieve access keys). The main area is titled "Retrieve access keys" with an "Info" link. It contains a section for "Access key" and "Secret access key". The "Access key" field is filled with a redacted value, and the "Secret access key" field shows a redacted value with a "Show" link. Below this is a "Access key best practices" section with a bulleted list: "Never store your access key in plain text, in a code repository, or in code.", "Disable or delete access key when no longer needed.", "Enable least-privilege permissions.", and "Rotate access keys regularly.". A note at the bottom says "For more details about managing access keys, see the [best practices for managing AWS access keys](#)". At the bottom right are "Download .csv file" and "Done" buttons.

**Nota:** cuando tenga las credenciales inmediatamente pasalas al código, una vez hecho click en “Done” ya no puedes verlas.

## 6.3 Ejecutar la Demo

```
(env) pablo@pablo-VivoBook-ASUSLaptop-X415DA-M415DA:~/Desktop/IA2/practica aws$ python demo_bedrock.py
💡 Pregunta lo que quieras sobre tus documentos:
Tú: quien es batman?
Bot: Batman es un personaje ficticio que ha trascendido las páginas del cómic para convertirse en un símbolo universal de disciplina, inteligencia y justicia implacable. Su origen se remonta a 1939, cuando Bob Kane y Bill Finger crearon un nuevo héroe para Detective Comics. A diferencia de otros superhéroes, Batman no posee superpoderes, sino que su fuerza proviene de su voluntad inquebrantable, de una mente brillante, y del dolor de un trauma que transformó la tragedia en propósito. El asesinato de sus padres no solo lo marcó emocionalmente, sino que lo convirtió en el reflejo más humano del mito heroico. Desde su creación, Batman ha sido reinventado y reinterpretado por escritores, artistas y cineastas. Cada generación ha visto una nueva versión de este arquetipo moderno, una representación de la capacidad humana para enfrentarse a la oscuridad interior y exterior sin perder la cordura ni la ética.
```

## Paso 7: Limpieza de Recursos (Opcional pero Recomendado si ya no harás uso de estas herramientas para que no consuma creditos)

### 7.1 Eliminar Knowledge Base

1. Ve a AWS Bedrock → Knowledge bases
2. Selecciona tu knowledge base
3. Haz clic en "Delete"
4. Confirma la eliminación

### 7.2 Eliminar Bucket S3

1. Ve a Amazon S3 → Buckets
2. Selecciona tu bucket
3. Haz clic en "Empty" (vaciar bucket primero)

**4. Luego haz clic en "Delete"**

## **Conclusiones**

Esta práctica demostró la implementación exitosa de un sistema de IA generativa empresarial usando AWS Bedrock con RAG, donde se creó una base de conocimiento que combina modelos fundacionales como Nova Micro con documentos personalizados almacenados en S3, permitiendo respuestas precisas basadas en información específica. El proceso incluyó la configuración completa desde la creación de usuarios IAM con políticas adecuadas, implementación de buckets S3 para almacenamiento documental, habilitación de modelos en Bedrock, construcción de knowledge bases con vector stores para búsqueda semántica, y desarrollo de una aplicación Python funcional que consume las APIs de Bedrock, validando así la capacidad de AWS para democratizar el acceso a IA avanzada y permitir la creación de asistentes inteligentes contextuales que responden basándose en documentación corporativa específica, todo ello implementable en cuestión de horas sin requerir expertise especializado en machine learning.