Amazon Q Business

Introduction to Amazon Q Business

What does Amazon Q Business do?

Amazon Q Business is a generative AI-powered assistant that can answer questions, generate content, create summaries, and complete tasks, all based on the information in your enterprise. Amazon Q Business is delivered using a built-in web experience or through APIs. This helps business users leverage the power of generative AI without any overhead.

Amazon Q Business can connect to your company data, information, and systems with more than 40 built-in connectors. It has built-in plug-ins for systems such as Salesforce, Jira, ServiceNow, and Zendesk to help complete tasks like creation of tickets, directly within your enterprise systems.

What problems does Amazon Q Business solve?

Amazon Q Business helps solve problems around building and using generative AI-powered digital assistants.

1. **User experience:** provides a built-in web experience that can be deployed for users to interact with the application. Additionally, can be embedded into existing enterprise applications such as Slack and Microsoft Teams to have a integrated user experience and conversation.
2. **Time to value:** allow quickly create a generative AI-powered digital assistant without any coding. It provides a user-friendly console, where an administrator can create an application with simple configurations. Has built-in web experience, generative AI capability, data integrations to enterprise data sources, plug-ins for enterprise applications, and APIs.
3. **Infrastructure overhead:** is a fully managed service that removes all infrastructure overhead from application creation, deployment, or management.
4. **User access controls:** retrieves and uses the existing access controls for users within integrated enterprise applications and data sources. This allows the users to view the data with their existing authorization.
5. **Data source integrations:** provides 40+ built-in integrations to popular enterprise data sources like Amazon S3, Salesforce, Oracle, and so on. It can connect to both cloud-based and on-premise data sources.
6. **Guardrails:** provides straightforward configurations for administrative controls and guardrails. For example, you can apply restrictions such as blocking specific words or topics.

What are the benefits of Amazon Q Business?

- Delivers quick, accurate, and relevant answers to your business questions: quickly connects to your enterprise systems so you can have tailored conversations, solve problems, generate content, and take actions relevant to your business. It generates answers and insights according to the material and knowledge that you provide, backed by references and citations to source documents.
- Connects to over 40 popular enterprise applications and document repositories: has over 40 built-in connectors to popular enterprise applications and document repositories, including Amazon Simple Storage Service (Amazon S3), Salesforce, Google Drive, Microsoft

365, ServiceNow, Gmail, Slack, Atlassian, and Zendesk. This helps with faster integrations to your enterprise systems, providing a tailored response to user queries. The connectors include both cloud-based systems and on-premise systems.

- Respects existing access controls based on user permissions: is built to be secure and private. It can understand and respect your existing identities, roles, and permissions within enterprise data sources. If a user doesn't have permission to access certain data without Amazon Q Business, they can't access it using Amazon Q Business either. This reduces security overhead for administrators while providing relevant responses to individual user queries.
- Helps administrators easily apply guardrails to customize and control responses: provides administrative controls, such as the ability to block entire topics and filter both questions and finalized answers using keywords. This helps ensure that it responds in a way that is consistent with a company's guidelines. You can also choose to limit the response to the knowledge available in the connected data sources or allow Amazon Q Business to use its world knowledge to deliver a response.

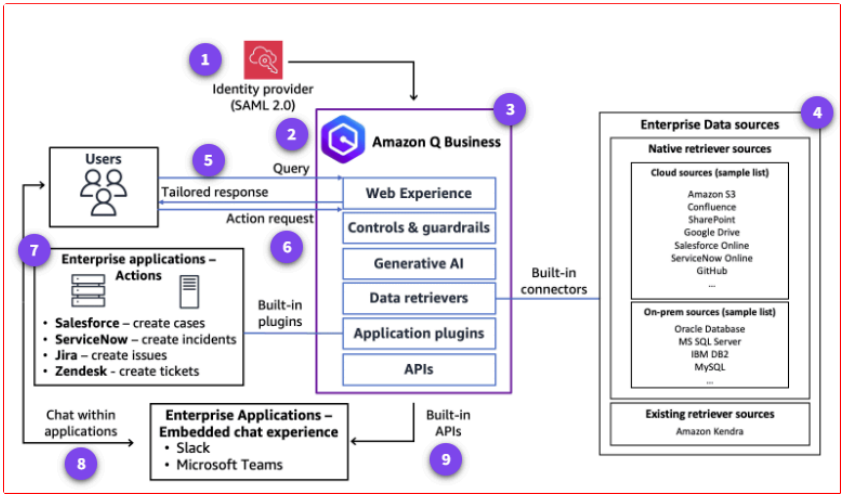
Quiz:

1. Which of the following problems does Amazon Q Business solve?
   Amazon Q Business addresses the need for faster integrations with enterprise systems and data repositories, alignment with existing user access controls for enterprise data, and reduces the time spent on coding to create a generative AI-powered digital assistant.
2. Which of the following options describes capabilities of Amazon Q Business?
   The key concepts for Amazon Q Business include built-in data integrations to enterprise data sources, plugins for enterprise applications, and fully managed RAG capability.

Architecture and Use Cases

How does Amazon Q Business generative AI-powered assistant work?

The following architecture logically illustrates how you can utilize Amazon Q Business to get query responses, take actions, or have conversations within existing applications.



1. **Users:** users are authenticated and authorized using a SAML 2.0 supported identity provider.
2. **User query:** users can provide a natural language query or prompt to the Amazon Q business web experience or chat application.
3. **Amazon Q Business:** has a built-in web experience for conversation. It has administered controls and guardrails. It uses generative AI capabilities, built-in data retrievers, application plug-ins, and APIs to deliver a tailored response.
4. **Enterprise data sources:** has 40+ built-in connectors for data retrieval from enterprise data sources, which includes a list of native data retrievers and existing data retrievers.
5. **Tailored response:** a tailored response is provided back to the user or consumer. This response includes specific data from the enterprise data sources based on users' existing access controls within those enterprise data sources.
6. **User action request:** users can directly create enterprise application actions like creating tickets, cases, incidents, and issues from the Amazon Q Business web experience interface without separately logging into enterprise applications.
7. **Enterprise applications for actions:** actions are created in enterprise applications like Salesforce, ServiceNow, Jira, and Zendesk. These are integrated with Amazon Q Business using built-in plugins.
8. **Chat within applications:** users can chat with Amazon Q Business application directly within existing enterprise applications without having to use a separate user interface.
9. **Embedded chat experience:** Amazon Q Business application can be directly embedded within enterprise applications like Slack and Microsoft Teams using built-in APIs.

What are the basic technical concepts of Amazon Q Business?

- IAM Identity Center: To create and use an application in Amazon Q Business, you need to configure and connect IAM Identity Center. IAM Identity Center is not meant to replace your existing identity provider, rather it serves as an overlay to your identity provider of choice. In case already have an IAM Identity Center configured, you may use that or create a new one for your account to connect to the application. Using IAM Identity Center, you can manage access to the application by assigning existing users or creating new users and groups from your identity center directory. Once the IAM Identity Center is connected to the application, assigned users can use the Amazon Q Business built-in web experience.
- Retrieval Augmented Generation (RAG): Generative AI foundation models are usually trained offline, making the model agnostic to any data that is created after the model was trained. Additionally, foundation models are trained on general domain data, making them less effective for domain-specific response or tasks. RAG is a technique used to retrieve data from outside a foundation model and augment the prompts by adding the relevant retrieved data as context. This helps generative AI based digital assistants to provide tailored response, by augmenting enterprise data sources to a foundation model. Is provided a fully managed RAG approach where administrators and users won't have to manage the underlying data augmentations, prompt engineering, and vector embeddings. Administrators can configure Amazon Q Business to respond strictly from enterprise documents or allow it to use external knowledge to respond to queries when the answer is not available in enterprise documents.
- Enterprise data access control: supports access control for your data so that users have access to the right content based on their permissions. Can integrate the Amazon Q web experience with the external SAML 2.0-supported identity provider (such as Okta, Microsoft Entra ID, and Ping Identity) to manage user authentication and authorization.
- Data integration and updates: has multiple pre-built connectors that can connect to enterprise data sources so that you can implement your generative AI solution with minimal configuration. Connectors offer modes for full synchronization or incremental data synchronization.
- Plugins: provides built-in plugins to interact with popular third-party applications, such as Jira, ServiceNow, Salesforce, and Zendesk. Administrators can enable these plugins to extend the capabilities of their Amazon Q application.

What are typical use cases for Amazon Q Business?

Typical use cases	Definition	Examples
		You can ask Amazon Q Business to write a compelling blog post

Accelerated Content Creation	Using Amazon Q Business, you can accelerate content creation for various business functions like marketing, sales, research, HR, legal, and so on.	and three social media headlines announcing the product described in this uploaded document. Amazon Q Business will generate content based on your enterprise's knowledge and the uploaded file.
Streamlined Enterprise Search Experience	Using Amazon Q Business, you can integrate your enterprise systems and knowledge repositories using pre-built connectors. This provides a streamlined enterprise search experience where users can get relevant response with specific source references.	You can ask Amazon Q Business, "where can I find the latest brand guidelines for logo usage?" Amazon Q Business will find what you need without having to jump between multiple systems. Because Amazon Q Business understands follow-up questions, you can continue asking questions, such as, "where can I find different color sets of our logo?" Amazon Q Business will understand the context of the previous question, surfacing the location of the relevant information.
Generates Summaries	Using Amazon Q Business, you can quickly generate a summary of enterprise documents within repositories, content within systems, and uploaded files. This improves productivity and information consumption.	You can ask Amazon Q Business to create a nicely formatted summary of the customer feedback on the new pricing structure in Slack. Amazon Q Business will generate a summary to help you understand the essence of the feedback, scaling knowledge, and speeding comprehension.
Extracts Key Insights for Business Decisions	Using Amazon Q Business, you can generate and extract insights from multiple documents through natural language queries. You can compare documents, ask insight questions relevant to the documents, and more.	You can ask Amazon Q Business to analyze the two latest customer satisfaction reports from Q1 and Q2 and identify the main differences between them. Amazon Q Business will sift through the reports and provide insights so that you can make faster and more informed decisions.

What else should I keep in mind about Amazon Q Business?

- Use of customer data: customer data is not used to train, improve, or enhance the machine learning models used by Amazon Q business. Additionally, the following data uses AWS Key Management Service (AWS KMS) encryption for protection with service-managed AWS KMS keys.:
  - Indexed customer data

    - Conversation history stored by the service
    - Feedback data (thumbs up or thumbs down)
- Chat history: automatically saves your conversation history for one month. Administrators have the ability to delete conversation history from the application.
- Supported file types: supports many common document types and formats, such as .PDF, .CSV, .DOCX, .HTML, .JSON, and .PPT.
- Choice of large language model (LLM): doesn't provide the ability to choose the underlying LLM. However, AWS uses various foundation models from Amazon Bedrock within Amazon Q Business.
- Custom connectors: allows to use custom connectors. You can add custom connectors and then use the Amazon Q SDK to implement them.

How do I create a data source for Amazon Q Business?

1. To begin, log in to the AWS Management Console and select the appropriate AWS Region from the Region menu.
2. Next, create an S3 bucket to store the files you will use in this demonstration. In the search bar, enter S3 and select S3 from the results. Now, on the S3 landing page, choose **Create bucket**.
3. Then, in the General configuration panel of the Create bucket page, provide a bucket name that is globally unique and lowercase. Your bucket name will be different from the name shown in this demo.
4. Now, scroll to the end of the page. Leave the rest of the configuration options set to their default values. Choose **Create bucket**.
5. Then, a banner is displayed after the bucket is created. The bucket is also listed in the General purpose buckets panel. Select the bucket name link.
6. Now, you can upload the files needed for this demonstration. You will find a .zip file following this video that contains the files you will need. Extract the files to your system. Please note that you should not upload the file named Amazon-Q4_2023_Transcript.pdf in this step. It will be used in a later demonstration. To add the files to your bucket, choose **Upload** or drag the files to this page from your system.
7. Next, your files are listed on the **Upload** page.
8. Now, scroll to the end of the page. Choose **Upload**.
9. Finally, after the files have been successfully uploaded to your bucket, a success banner is displayed.

How Do I Set Up an AWS IAM Identity Center and Create Users?

In this lesson, you will learn how to enable and use the AWS IAM Identity Center to create and manage users for Amazon Q Business.

How do I manage users for Amazon Q Business?

1. You will create an IAM Identity Center instance to manage users for Amazon Q Business. To create an application in Amazon Q Business, you need to configure and use IAM Identity Center to manage users and access for the application user interface. (Note that IAM Identity Center is not meant to replace your existing identity provider. It serves as an overlay to your identity provider of choice). Also, IAM Identity Center and user setup can be done either on the IAM Identity Center service console or within the Amazon Q Business console. In this demo, you will use the IAM Identity Center console to show how to enable the instance and add users.
2. To begin, log in to the AWS Management Console and select the appropriate AWS Region from the Region menu.
3. In the search bar, enter IAM Identity Center and select **IAM Identity Center** from the Services list.
4. Now, in the IAM Identity Center landing page, you can enable IAM Identity Center. Choose **Enable**.
5. You have the option to enable IAM Identity Center either at an organization level or at the account level. For this demo, choose **Enable in only this AWS account**. Then, choose **Continue**. In an enterprise-level setup, the recommendation is to use IAM Identity Center at an organization level. Organization level setup may require necessary administrative privileges.
6. Next, choose **Enable**.
7. After you create the IAM Identity Center instance, it will show on the dashboard. Choose **Users** from the navigation menu on the left.
8. Next, choose **Add user**.

9. Specify and enter user details like **user name**, **email address**, **confirm email address**, **first name**, **last name**, and **display name**. Leave all other settings as the default and scroll down.
10. Then, choose **Next**.
11. Adding a group is optional. Choose **Next**.
12. Now, review the details and choose **Add user**.
13. A banner will display, showing that you have successfully added the user.
14. An automated email is sent to the new user's email address. The email will have a button to access the invitation link. Choose **Accept invitation** within the email.
15. A new web browser window is opened where the user will need to add a new password and confirm the password. After you enter a password, choose **Set new password**.
16. Then, enter the username and choose **Next**.
17. Next, enter the password and choose **Sign in**.
18. For this multi-factor authentication (MFA) step, please use a separate mobile device with an authenticator application like Google Authenticator mobile app. On this page, choose **Show QR code**. Scan the QR code using your mobile device within the authenticator app. Add the **authentication code** displayed on the authenticator app for the user, **demo_user**. Choose **Assign MFA**.
19. After the MFA is set up, it will show that the authenticator application is registered for the specified user. Choose **Done**.
20. Next, you can test the newly added user. Add the username and choose **Next**.
21. Then, add the password and choose **Sign in**.
22. Now, add the MFA code from the authenticator app for the user and choose **Sign in**.
23. The user is taken to the AWS access portal where any application assigned to the user will be shown. Currently, no application is assigned to the user, so the screen is blank.

How do I Create an Amazon Q Business Application?

In this demo, you will learn how to create and customize an Amazon Q Business application, integrate a data source, and assign users.

Setting Up an Amazon Q Business Application

1. To begin, verify that your Region is selected from the Region menu in the AWS Management Console. Then, enter **Amazon Q Business** in the search bar. Select **Amazon Q Business** from the results panel.
2. Now, from the Amazon Q Business landing page, choose **Get Started**.
3. In the Amazon Q Business console, you can create, deploy, and manage generative AI digital assistant applications. You can also create a quick application to experiment before you create or deploy an application. Choose **Create application**.
4. Step 1 is to create the application. In the **Application details** pane, enter an **Application name**. Leave other settings as default. The IAM Identity Center instance that was created earlier as a part of this course is automatically assigned to this application. Choose **Create**.
5. Step 2 is to select a data retriever. You can either use native data retrievers or use existing retrievers. Select **Use native retriever** and choose **Next**.
6. Step 3 is to connect data sources for the application. Amazon Q Business has over 40 built-in connectors to various enterprise systems, like Amazon S3, SharePoint, Google Drive, and more. It has connectors for both cloud based systems and on-premises systems. Choose the **plus sign** by **Amazon S3**.
7. In the connect data sources Amazon S3 page, add a **data source name**. In the IAM role dropdown list, select **Create a new service role**, which is the recommended option. Then, choose **Browse S3** to select the S3 bucket that you created in a previous demonstration. Scroll down.
8. Under **Sync run schedule**, select **Run on demand** from the **Frequency** menu. Then, choose **Add data source**. It can take a few seconds or minutes to add the data source. Then, scroll down. After you successfully add the data source, choose **Next**.
9. Step 4 is to add groups and users to the application. Choose **Add groups and users**. In the **add or assign users and groups** pop-up window, choose **Assign existing users and groups** and choose **Next.**

10. Then, choose **Get started**. In the search bar, search for the username that you created. In this demo, the username is **demo_user**. Choose the user from the dropdown list.
11. Then, choose **Assign**. In the **Add groups and users** page, verify the user is added by choosing the **Users** tab.
12. After the user assignment and the appropriate subscription level is verified, choose **Create application**.
13. Now, the Amazon Q Business application is created and the web experience is deployed. You can verify the web experience status has deployed successfully under the **Web experience status** column. You can also verify the Web Experience URL, which will be used by the assigned users to access the application interface. Choose the demo application name link. In the demo application page, choose the radio button by your data source. Then, choose **Sync now**.
14. After the sync process starts, you will see a notification on the top of the page indicating that the sync started successfully. The sync process can take from a few minutes to a few hours. Sync speeds are limited by factors such as remote repository throughput and throttling, network bandwidth, and the size of the documents in the source system. The sync process continues in the background and doesn't need to be actively monitored, and the user session doesn't need to be active.
15. After the sync process is complete, you will see the last sync status is completed and the last sync time listed. As an admin, you have options to set Admin controls and guardrails for the application. Choose **Admin controls and guardrails** from the menu on the left. In the this page, you can set global controls. You can restrict the application response to only the connected enterprise data sources. You can also augment the response with generative AI large language models (LLMs) within Amazon Q Business to provide a generic response.
16. Now, choose **Edit**. Select the check boxes next to the two options **Allow end users to send queries directly to the LLM** and **Allow Amazon Q to fall back to the LLM knowledge**. Then, choose **Save**.
17. Now, choose **Applications** from the menu on the left.
18. Then, choose the radio button by the demo application you created. Choose **Customize web experience**. As an admin, you can preview the web experience for the application and also customize the title, subtitle, and welcome message. Choose **Save**. As an admin, you won't be able to have conversations with the application on this page. To have a conversation, a user needs to access the web experience URL with their respective credentials.
19. Now, you can access the web experience URL either from this preview page or from the application list. On this page, choose **View web experience**. After you select the web experience URL, you will be redirected to a separate web experience page where the assigned user needs to be authenticated before using the application. You have now successfully created an Amazon Q Business application.

How Do I Chat with an Amazon Q Business Application?

In this lesson, you will learn how to use the Amazon Q Business built-in web experience on a web browser to have a conversation with the Amazon Q Business application.

Chatting with Amazon Q Business

1. After the admin user creates the application in Amazon Q Business, you can choose the web experience URL to have a conversation as a business user. The Web experience URL is shown in the Applications list. Use the built-in web experience URL in a web browser. The assigned business user needs to follow the authentication process. This user was set up in IAM Identity Center and assigned to the Amazon Q Business application. Enter the user name and choose **Next**.
2. Then, enter the password and choose **Sign in**.
3. Then, enter the multi-factor authentication code from the Authenticator app that you set up during the user creation process in IAM Identity Center. Then, choose **Sign in**.
4. This is the personalized web experience page of the Amazon Q Business application, which was created and configured in a previous demonstration. You can start having conversations using the **Enter a prompt** section. You can also see and manage all the chat history under **Conversations** on the upper left panel of the page. For this demonstration, you will use the publicly available Amazon 10K financial annual report documents in a S3 bucket as a data source for this application.
5. You can ask questions, such as "What was Amazon's revenue in 2021." Amazon Q Business will respond with the correct answer and also show the data source references within the answer and in the sources dropdown list. In this case, the response was completed

using one of the data source files you uploaded to your S3 bucket. You have the option of providing feedback to the application by choosing the thumbs up and thumbs down icons. You can also choose the copy icon to copy the response for further use.

6. When you ask a question, like "What was Amazon's revenue in 2015," the application doesn't find any answer within the data source files. Because you allowed the application to fall back to the generative AI large language models (LLMs) of Amazon Q Business, it was able to provide an answer from its own knowledge. Please keep in mind that the answers from the LLMs could be generic and sometimes inaccurate.
7. You can ask the application to write a summary blog of the CEO shareholder letter within the 2022 annual report. You can write a specific prompt to direct the application based on your specification. The application will provide a response along with source references.
8. As a user, you can upload files to the prompt section and have a specific conversation related to the file. You can use the file in the zipped folder provided in this course called **Amazon-Q4_2023_Transcript.pdf**. This is a publicly available earnings call transcript from Q4 2023. Choose the paper clip icon to upload the file. Then, navigate to the file within your computer's directory and choose the file.
9. After you upload the file to the conversation, you can verify that the file successfully uploaded within the prompt section. Now, you are ready to have a specific conversation around this file. In the prompt section, enter "Write a 200 word summary of the attached transcript file." Then, choose the arrow icon to submit the prompt.
10. Amazon Q Business will create a summary from the uploaded file.

How Do I Clean Up Resources?

In this lesson, you will learn how to delete the resources and application you created using AWS management console.

Cleaning Up Resources

1. To begin, return to the Amazon Q Business page for Applications. After the application is no longer needed, select the radio button next to your application. Then, from the Actions menu, choose **Delete**.
2. When you choose Delete, you will be asked for confirmation before the application is deleted permanently. Enter **Delete** in the text box, and then choose **Delete**.
3. Then, after you have deleted the application, you can delete the S3 bucket you created. Return to the S3 landing page. From the menu on the left, choose **Buckets**. Select the radio button next to the bucket you created for this demonstration. Before you can delete the bucket, it must be empty. Choose **Empty**.
4. On the **Empty bucket** page, enter **permanently delete** in the text box to permanently delete all objects in your bucket. Then, choose **Empty**. You will see a banner notification that you successfully emptied the bucket. Choose **Exit**.
5. Now, you will select the radio button next to your bucket. Choose **Delete**. On the **Delete bucket** page, enter the name of the bucket in the text box.
6. Then, choose **Delete bucket**. You will see a banner notification that you successfully deleted the bucket.