



ACOT102

Innovate on enterprise data with Amazon Q to improve employee productivity

Vijai Gandikota

Principal Product Manager,
AWS

Agenda

- The opportunity
- Generative AI challenges
- What is Amazon Q?
- Key use cases
- Key features
- Demos
- Customer stories



Employee productivity

Marketing Manager

Sales Manager

R&D Engineer

Financial Analyst

Developer

Help Desk Employee (IT, HR etc.)

Product Manager...



Opportunity with generative AI



Generative AI's ability to understand natural language enables automation for work activities that account for

25% of total
work time



**Imagine a generative AI powered assistant that saves
you**

every
day

2 hours

[McKinsey Digital June 2023](#)

Challenges adopting generative AI



Accuracy



Security & control



Time to value

Introducing

Amazon Q

**Your generative AI assistant designed
for work that can be tailored to your
business, data, code, and operations**



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved. Amazon Confidential and Trademark.

Amazon Q areas of expertise

Your
business



Building
on AWS



Amazon
QuickSight



Amazon
Connect



AWS
Supply Chain



Amazon Q areas of expertise

Your
business



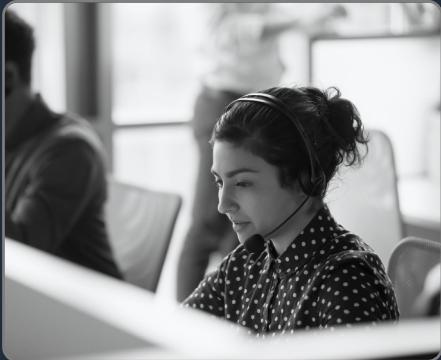
Building
on AWS



Amazon
QuickSight



Amazon
Connect



AWS
Supply Chain



Amazon Q addresses generative AI challenges



Accuracy

- Accurate answers from enterprise data sources using RAG
- Citations for fact-checking
- Broad set of capabilities



Security & control

- Support for enterprise permissions
- Ability to configure response guardrails



Time to value

- Connectors to 40+ popular data sources
- In-built vector index with managed ingestion
- In-built application with SSO

Top use cases



Streamlined search experience

For a unified, conversational Q&A experience, connect internal wikis, SharePoint sites, Confluence, Salesforce, and 40+ other applications



Accelerate content creation

Empower employees to create content faster and smarter across departments—from marketing to sales to engineering



Generate summaries

Quickly understand the essence of documents by creating summaries of uploaded files or existing enterprise documents



Extract key insights

Perform comparative analysis on your documents and unlock powerful insights with the Amazon Q application so that your employees can make faster, better data-driven decisions



Key features – Conversational Q&A

LET'S GET TO WORK



Trusted answers generated from enterprise data



In-context conversations



Source references for fact-checking



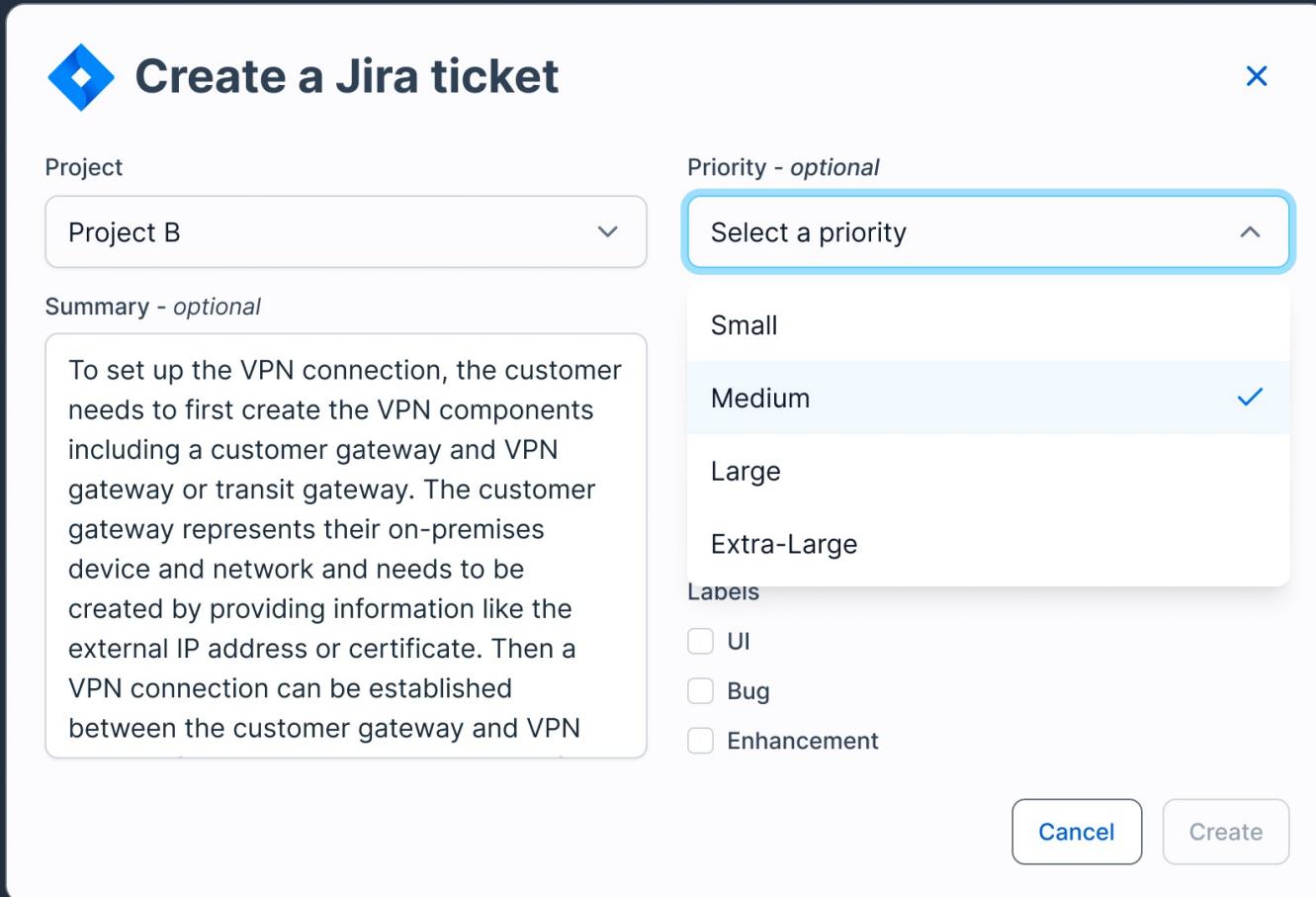
Conversation history

The screenshot shows a conversational interface with four entries:

- What is the reliability of S3?**
Amazon S3 provides 99.99999999% durability for stored objects, which means the reliability of S3 is extremely high. [1] [2] S3 achieves this level of reliability by redundantly storing objects across multiple facilities and devices in each facility. [1] In addition, S3 is designed to reliably store 10,000,000 data assets for 10,000 years. [2] S3 delivers high reliability through its infrastructure of Availability Zones which consist of independent data centers with redundant power and networking to protect stored data. [2]
- Tell me more about availability zones**
Availability zones are composed of independent infrastructure that are physically separated within a region to protect services and data from failure of a single location. Each availability zone runs on its own physically distinct, independent infrastructure, and is engineered to be highly reliable. [1] Availability zones consist of independent data centers with redundant power and networking to protect stored data. [2] They demonstrate compartmentalization within a region and act to isolate failures, but availability zones are also connected through low-latency links. [3] [4] Failure of a single availability zone will not affect resources or services located in another availability zone. [4]
- How many availability zones should my data be stored across?**
According to your source data, it is recommended to store data across multiple availability zones to achieve higher availability and fault tolerance. Some key reasons are:
 - Availability zones are isolated locations within a region so failure of a single availability zone will not affect resources located in other availability zones.
 - Deploying across multiple availability zones helps maximize availability of resources and reduces impact of failure of single availability zone.
 - Critical data should be stored in multiple availability zones in case of failure of single availability zone.
 - Read replicas of databases should be placed in different availability zones to account for loss of single availability zone. Therefore, based on the information provided, to achieve higher availability and fault tolerance of data, it is recommended to store data across multiple availability zones rather than a single availability zone. The number of availability zones to use may depend on specific system requirements and fault tolerance needs. [1] [2] [3]
- Info**
Amazon Q - Preview uses generative AI and it may give inaccurate or inappropriate responses. For more information: [AWS Responsible AI Policy](#)



Key features - Execute actions using plugins



- Enable end-users to perform actions on SaaS applications
- *"summarize conversation and create ticket in Jira"*



Key features – Admin console

LET'S CREATE AND DEPLOY GENERATIVE AI APPLICATIONS



In-built vector index with managed ingestion



In-built application with SSO



3 click setup: Settings, retriever, and data sources



Accuracy of Retrieval Augmented Generation (RAG)

The screenshot shows the 'Create application' wizard in the AWS ExpertQ Admin Console. It is Step 1 of 3:

- Create application** (radio button selected)
- Select retriever
- Connect data sources

Application settings

Application name: AWSBlogExpert

Service access: ExpertQ requires permissions to use other services on your behalf.

Choose a method to authorize ExpertQ:

- Create and use a new service role
- Use an existing service role

Service role name: AmazonExpertQ-Application-AWSBlogExpert

Customize web experience

Title: AWS Blog Expert

Subtitle - optional: Ask me about AWS!

Application tags (selected tab) Web experience tags

Tags - optional (0)

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.

Cancel **Create**



Key features – Safety and security

ADD GUARDRAILS TO THE EXPERIENCE

Update global controls Info

Global controls Info
Application guardrails will apply to all messages returned by Enterprise Q.

Response settings Info
You can limit Enterprise Q from using its own knowledge to generate answers when it cannot find relevant content in your enterprise corpus.

Only produce responses from Retrieval Augmented Generation (RAG)
Responses will be limited to ingested documents in your enterprise corpus.

Blocked words Info
Define blocked words for the application. The application will not respond to questions that contain these words or mention them in any responses.

Enter blocked words

You can block 18 more words.
Account vulnerabilities Project X

Messaging shown for blocked words
I cannot complete this request as the response contains content that is blocked by your Admin. Please contact your Admin for help.

This response can have up to 150 characters. Valid characters are a-z, A-Z, 0-9, _ (underscore) and - (hyphen).

Feature settings Info
Configure features end users have access to in the web experience.

Allow end users to upload files in chat context
This feature enables end users to upload files directly to chat in order to ask questions specific to the document.

Use pre-built guardrails for toxicity

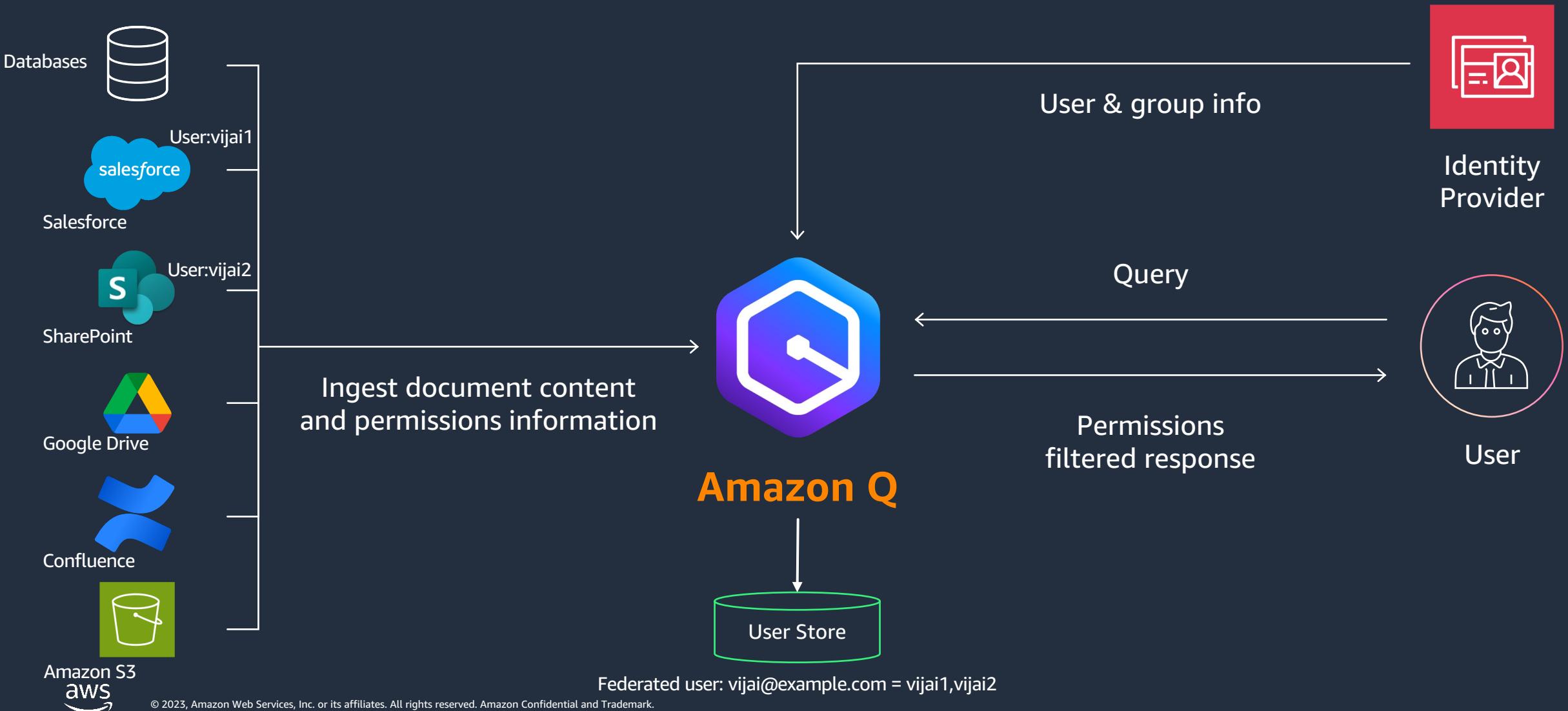
Restrict responses to enterprise content only

Specify blocked words or phrases that never appear in responses



Key features – Safety and security

AWARE OF ENTERPRISE USER PERMISSIONS



Get started quickly with 40+ connectors

Bring content from all your
enterprise sources together
in a few clicks!

Most popular

All Cloud On-prem

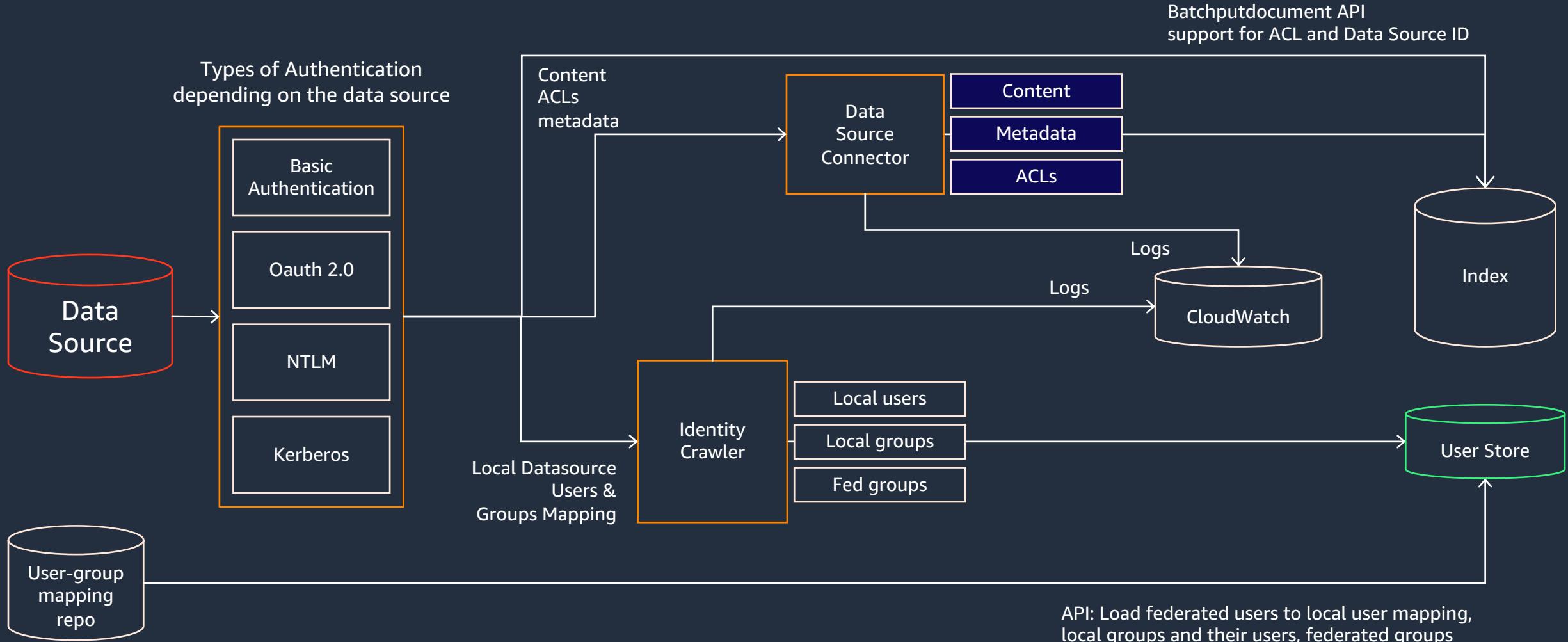
Search

Amazon S3 Cloud	1 added ✓ +	Web crawler	1 added ✓ +	Upload docs	+ (disabled)
Adobe Experience Manager Cloud On-prem	+ (disabled)	Alfresco Cloud On-prem	+ (disabled)	Amazon FSx Cloud	+ (disabled)
Amazon RDS (MySQL) On-prem	+ (disabled)	Amazon RDS (Microsoft SQL Server) On-prem	+ (disabled)	Amazon RDS (Oracle) On-prem	+ (disabled)
Amazon RDS (PostgreSQL) On-prem	+ (disabled)	Aurora (MySQL) On-prem	+ (disabled)	Aurora (PostgreSQL) On-prem	+ (disabled)
Box Cloud	+ (disabled)	Confluence Cloud On-prem	+ (disabled)	Custom data source	+ (disabled)
Dropbox Cloud	+ (disabled)	Drupal Cloud	+ (disabled)	Github Cloud On-prem	+ (disabled)
Gmail Cloud	+ (disabled)	Google Drive Cloud	+ (disabled)	IBM DB2 On-prem	+ (disabled)
Jira Cloud	+ (disabled)	MySQL On-prem	+ (disabled)	Microsoft Exchange Cloud	+ (disabled)
Microsoft SQL Server On-prem	+ (disabled)	Microsoft Teams Cloud	+ (disabled)	Microsoft Yammer Cloud	+ (disabled)
ORACLE DATABASE On-prem	+ (disabled)	OneDrive Cloud	+ (disabled)	PostgreSQL On-prem	+ (disabled)
Quip Cloud	+ (disabled)	Salesforce Online Cloud	+ (disabled)	ServiceNow Online Cloud	+ (disabled)
SharePoint Cloud On-prem	+ (disabled)	Slack Cloud	+ (disabled)	WorkDocs Cloud	+ (disabled)
Zendesk On-prem	+ (disabled)				

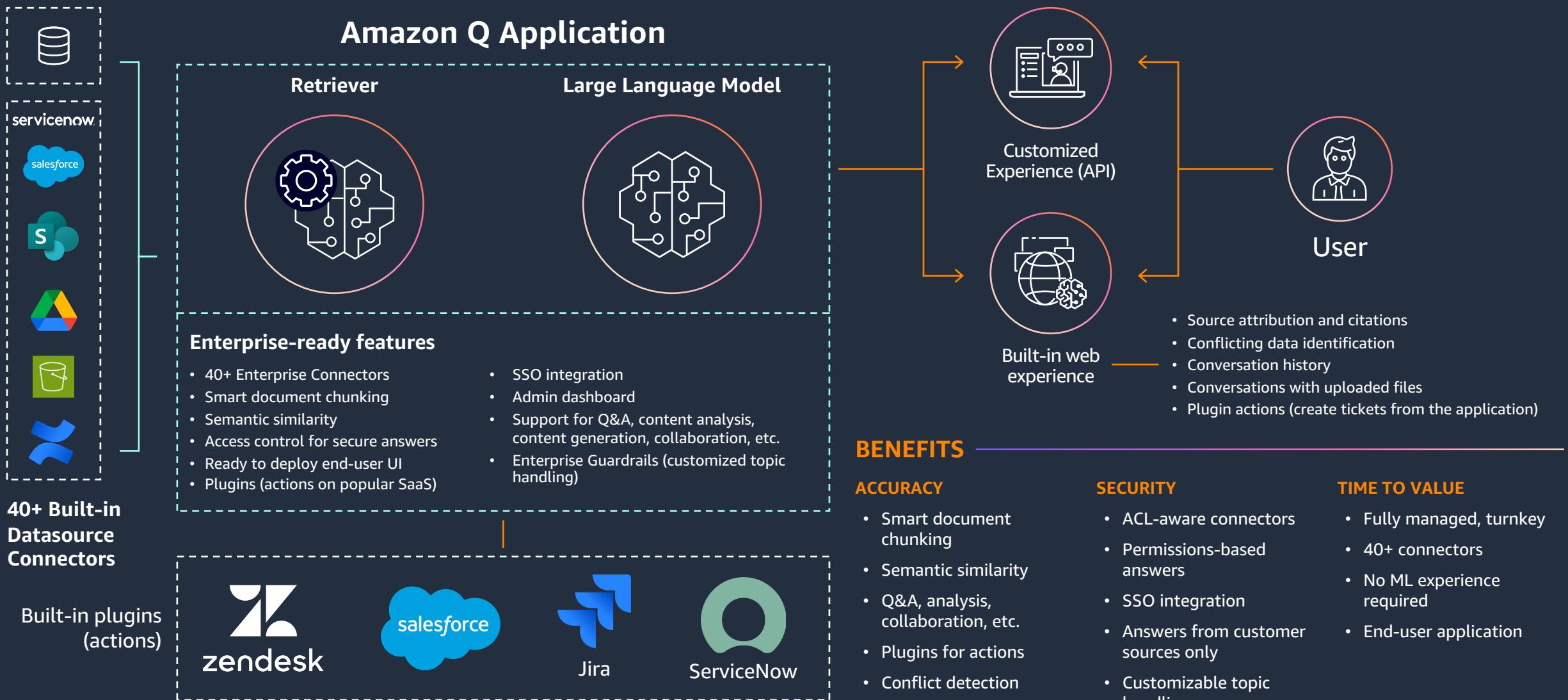


Key features – Connectors

INDEXING CONTENT FROM THE ENTERPRISE



In summary



Demo

Web experience

QUESTION ANSWERING



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved. Amazon Confidential and Trademark.

Q New Conversation

X

Q tell me about Example Corp. what do they do?

November 23, 2023

Q Tell me about Example Corp. What does the company do?

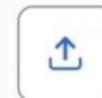
November 23, 2023

Amazon Q

Your AI assistant for work



👉 Hello, I am your AI assistant. Enter a prompt or choose a suggestion to start a conversation. I'll respond using data from within your organization.



i Info

Amazon Q - Preview uses generative AI and it may give inaccurate or inappropriate responses. For more information: [AWS Responsible AI Policy](#) ↗



Demo

Custom Amazon Q applications: Slack integration

EMBEDDING GENERATIVE AI IN MESSAGING APPS



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved. Amazon Confidential and Trademark.

See how useful it is in a multi-person thread!

Try the Amazon Q Slack bot – Source code

GitHub repo: <https://github.com/aws-samples/amazon-q-slack-gateway>



Demo

Amazon Q application setup with S3 connector



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved. Amazon Confidential and Trademark.

Create bucket Info

Buckets are containers for data stored in S3. [Learn more](#)

General configuration

Bucket name

aws-wp-doc

Bucket name must be unique within the global namespace and follow the bucket naming rules. [See rules for bucket naming](#)

AWS Region

US East (N. Virginia) us-east-1



Copy settings from existing bucket - *optional*

Only the bucket settings in the following configuration are copied.

[Choose bucket](#)

Object Ownership Info

Control ownership of objects written to this bucket from other AWS accounts and the use of access control lists (ACLs). Object ownership determines who can specify access to objects.

ACLs disabled (recommended)

All objects in this bucket are owned by this account.
Access to this bucket and its objects is specified using
only policies.

ACLs enabled

Objects in this bucket can be owned by other AWS
accounts. Access to this bucket and its objects can be
specified using ACLs.

Customer stories



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved. Amazon Confidential and Trademark.

Accelerating generative AI application development



“ Gilead's use of generative AI on AWS has led to faster innovation and productivity gains. By leveraging Amazon Q, we will be able to generate insights and accelerate analysis of large amounts of Gilead data across our enterprise. Based on our evaluation, **Amazon Q provides a faster way to create generative AI solutions** by streamlining connections to our data sources, automating complex tasks, such as managing vector stores, and quickly surfacing relevant insights on demand. For life sciences organizations like Gilead, the **productivity benefits unlocked** by generative AI solutions on AWS like Amazon Q are exciting. ”

Kevin Cox

Chief Cloud Architect, Gilead



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved. Amazon Confidential and Trademark.

Reducing content discovery and marketing content creation time

Wunderkind

“ Adding Amazon Q as a topline layer over our various content and data repositories brings a whole new level of efficiency to our customer success and marketing teams. Based on initial estimates, we expect the **time spent on content discovery alone to be reduced by over 30%**, which empowers our success team to service clients faster, and with better accuracy. It also jump-starts the creation of sales and marketing content, such as email drips, whitepapers, and ad copy. With Amazon Q, we anticipate the ability to **accelerate the content creation process by nearly 50%**, allowing us to shift our attention to scaling the personalization of content instead of spending time on the laborious task of creating materials from scratch.”

Richard Jones

Chief Revenue Officer at Wunderkind



Get started

RUN A FLASH POC

1

Pick easily
accessible and
well defined
set of
documents

2

Index your
content
with
Amazon Q
using
built-in
connectors

3

Perform
Q&A,
Summarize,
interact
with your
content in
the console
UI





Thank you!

Vijai Gandikota
Principal Product
Manager,
AWS



Link to console



Webpage



Developer guide