

AI in Qlik Cloud®

Copyright © 2018-2026 QlikTech International AB. All rights reserved.
Published: March 2026

1 Qlik Cloud® overview	4
1.1 Architecture overview	4
Qlik Cloud Analytics®	5
Qlik Talend Cloud®	5
1.2 Qlik Cloud® Editions	7
2 AI in Qlik Cloud	8
2.1 Overview	8
Qlik Answers®	8
Qlik Discovery Agent™	8
Qlik MCP Server™	8
2.2 AI in Qlik Cloud	8
Design Principles for AI on Qlik Cloud	8
Guardrails	9
Data Sovereignty	9
Authorization & AI in Qlik Cloud	10
AI Services in Qlik Cloud	10
AI Models at Qlik	11
2.3 Generative AI with Qlik Answers®	11
What is Qlik Answers ?	11
Why Qlik Answers ?	11
How Qlik Answers works	12
Answer formulation in Qlik Answers	14
Access control in Qlik Answers	16
2.4 Qlik Answers® - Security Architecture and Controls	17
Introduction	17
Architecture Overview	17
How Cross-Region Inference Works	19
Common Misconceptions	20
Security Controls	20
Data Storage by Location	21
Amazon Bedrock as a Security Layer	22
Why Cross-Region Inference Is Necessary	22
Security Assessment Guidance	23
2.5 Qlik MCP Server™	23
What is the Model Context Protocol?	23
Business value	24
Use cases and opportunities	25
User personas	26
2.6 Qlik MCP Server™ - Security Architecture and Controls	26
Executive Summary	27
What MCP Is / What MCP Is Not	27
Qlik's Trust Boundary	27
Common Misconceptions vs. Reality	28
Security Controls Deep Dive	28
What Gets Stored Where	30
MCP Risk Profile	30
Summary and Recommendations	30

2.7 Qlik Discovery Agent™	31
Overview	31
Value and Business Impact	32
What Discovery Agent Does	32
Where Discovery Agent Fits in Qlik Cloud	34
Security architecture	34
3 About Qlik Evaluation Guides	35
3.1 Document history	35
3.2 Changelog	35
Changelog – AI in Qlik Cloud	35

1 Qlik Cloud® overview

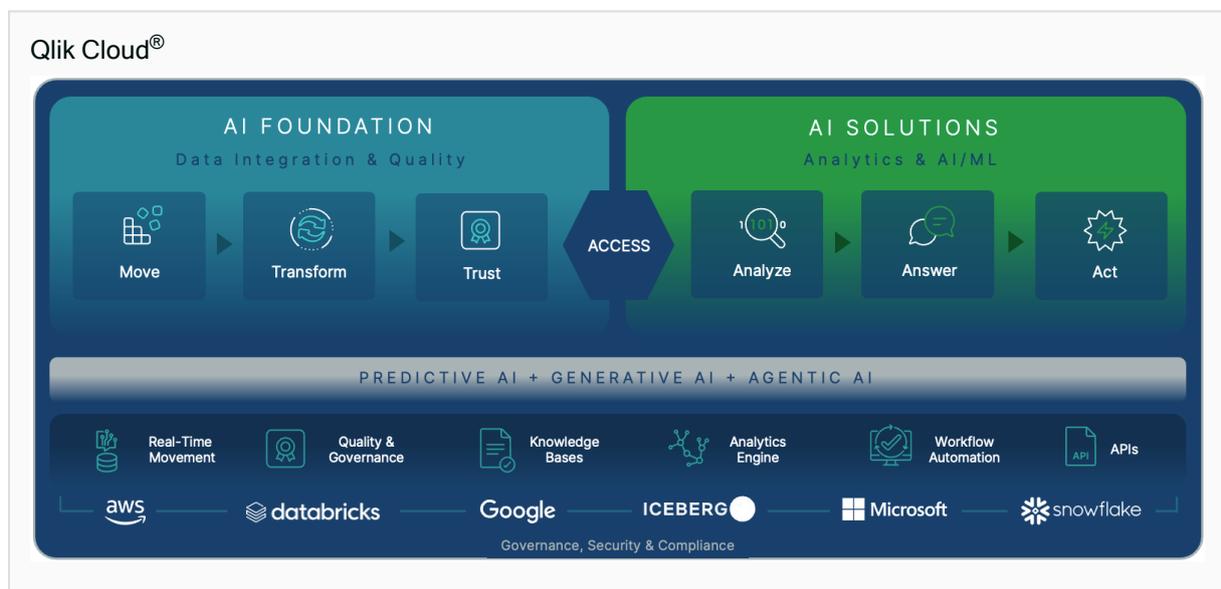
Qlik is a leader in data and analytics with a core mission to provide solutions that ensure organizations can work smarter and use data as a competitive edge. Qlik Cloud is a powerful end-to-end solution for data and analytics services. Our platform empowers curiosity-driven exploration offering everyone - at any skill level - the ability to use data to make transformative change for their organization.

Through several data-focused services, the Qlik Cloud platform supports a full range of users and use-cases across the lifecycle from data integration to insight generation. These services include change data capture, transformation, data cataloging, application automation, self-service analytics and dashboards, conversational analytics, custom and embedded analytics, and alerting.

This document highlights key aspects of the Qlik Cloud platform, including architecture, security, governance, and reliability. It is designed to complement the technical documents for the Qlik solutions that run on the Qlik Cloud platform.

1.1 Architecture overview

All of Qlik's SaaS offerings and services, known collectively as Qlik Cloud®, run on the Qlik Cloud platform. The platform delivers the underlying compute, storage security, and governance features to provide services to our customers. The Qlik Cloud® Platform enables the creation of the analytics data pipeline. Powered by Qlik Cloud and a rich set of foundational services, it provides all the data integration and analytics services you need to transform raw data into informed action.

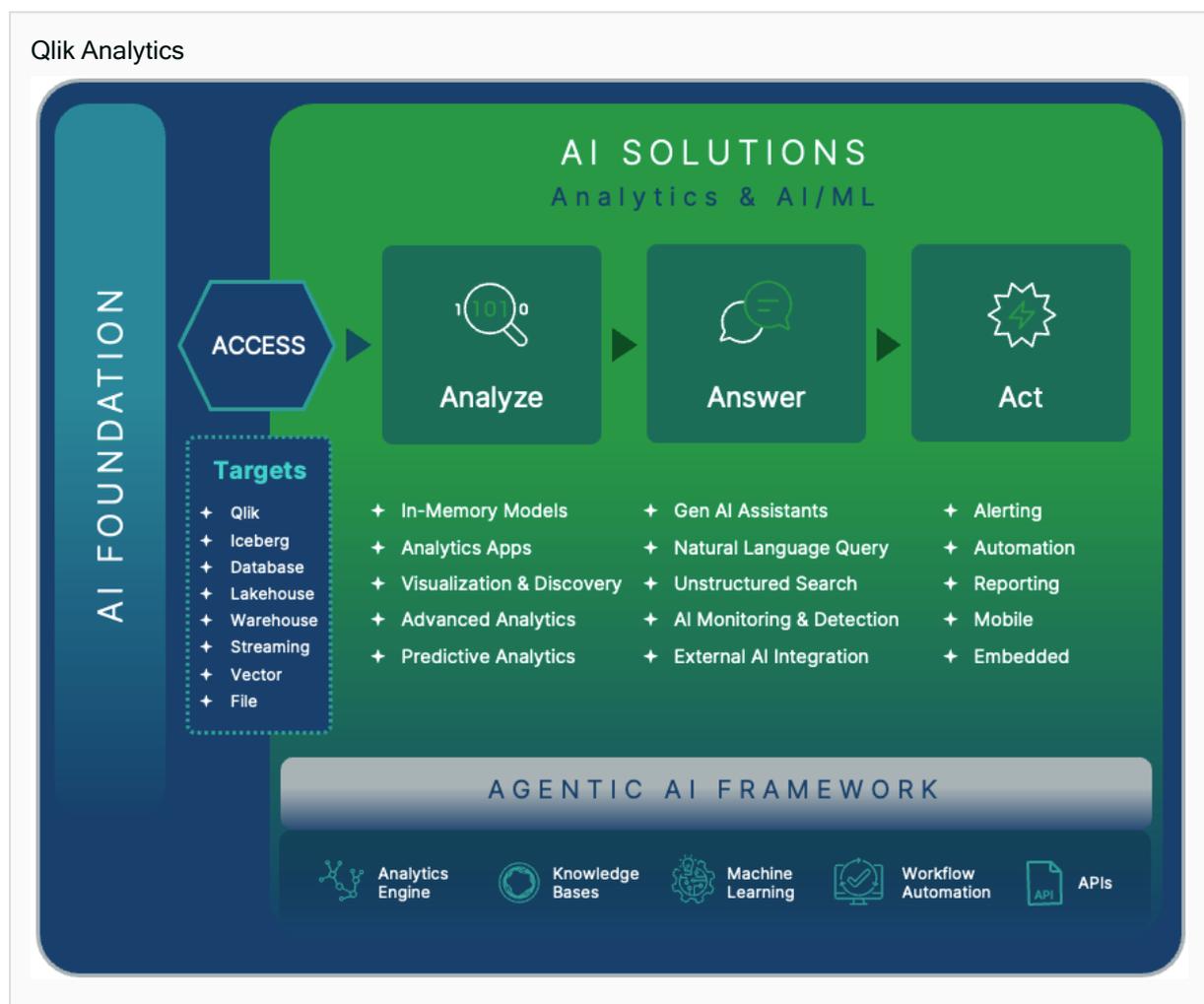


A customer's instance of the Qlik Cloud platform is called a tenant. It is logically separated from other tenants by using unique encryption keys. Access to the platform is controlled by the customer's configured identity provider and any access to functions within the platform is based on the entitlements the customer has assigned across roles and users. A number of services are available on the Qlik Cloud platform:

- Analytics - provides a complete third-generation analytics solution - Qlik Cloud Analytics
- Data Integration - provides the ability to manage your data assets and utilize change data capture to provide real-time access to your data, as well as application automation to automate integrations between cloud applications.

Qlik Cloud Analytics®

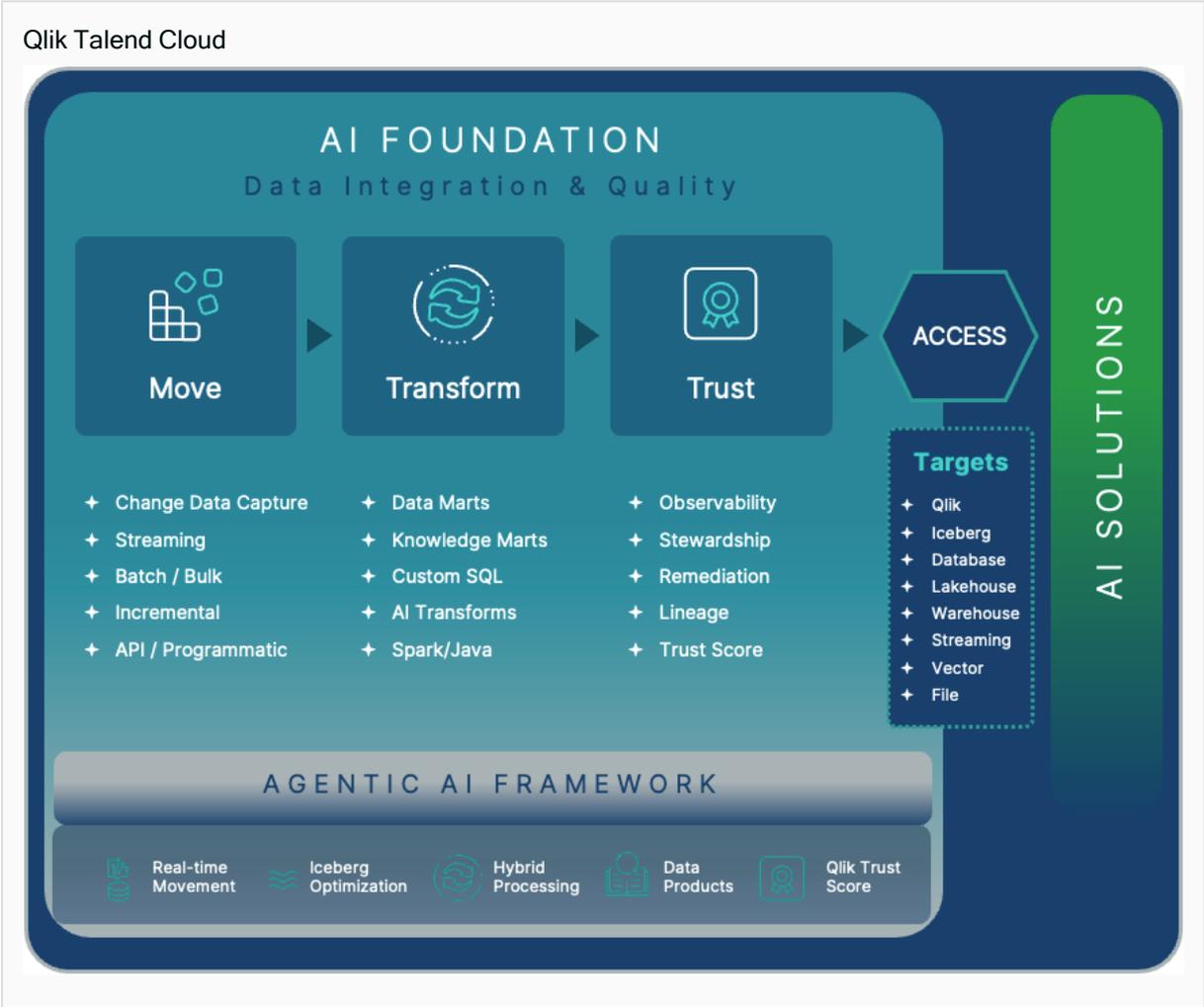
Incorporating our premier offering Qlik Sense, Qlik Cloud Analytics sets the benchmark for third-generation analytics platforms, empowering everyone in your organization to make data-driven decisions. Built on our unique Associative Engine, it supports a full range of users and use-cases across the lifecycle from data to insight: self-service analytics, interactive dashboards, conversational analytics, custom and embedded analytics, mobile analytics, reporting, and alerting. It augments and enhances human intuition with AI-powered insight suggestions, automation, and natural language interaction.



Qlik Talend Cloud®

Qlik Talend Cloud is Qlik's hosted and managed data integration platform as a service (iPaaS). Our vision is to provide a broad variety of data integration services aimed at helping you move from passive to active BI. Qlik Talend Cloud is architected for real-time data capture, transformation, and analytics-ready data

delivery leveraging a variety of methods in our unique change data capture approach.



Real-time data movement

Replicate data from on-premises or cloud sources into Qlik Cloud and other leading cloud data platforms. Automatically and continuously ingest data without the need for job scheduling or scripting. Your data is updated without manual intervention to drive insights and actions of important business moments.

Data transformation

Quickly turn raw transaction records into consumption-ready data via auto-generated, push-down SQL. Our no-code interface helps you create reusable transformation pipelines that intelligently conform data to dimensional models or custom formats.

Qlik Automate™

Qlik Automate is an integration platform to build integrations and automation flows between cloud applications. Closely integrated with the other Qlik Cloud services, Qlik Automate is able to build workflows

between your cloud applications using a no-code approach by connecting data sources, applying conditions, calling webhooks, adding loops, scheduling runs, and setting up triggers. For example, a webhook in your CRM system could initialize a reload of your sales performance Qlik Sense application.

1.2 Qlik Cloud[®] Editions

There are several Editions of Qlik Cloud to meet the varied needs of our customers. The Evaluation Guides are designed to cover all the functionality of Qlik Cloud Analytics, Qlik Talend Data Integration and the Qlik Cloud Platform. There is also a special version of Qlik Cloud for Government organizations in the US, known as Qlik Cloud Government. Information provided in the Evaluation Guides does not imply that those features are available in all Editions.

For information on your edition, see the relevant link below:

- [Qlik Cloud Analytics user-based subscriptions](#)
- [Qlik Cloud Analytics capacity-based subscriptions](#)
- [Qlik Talend Cloud capacity-based subscriptions](#)
- [Qlik Cloud Government](#)

2 AI in Qlik Cloud

2.1 Overview

Qlik leverages AI capabilities within multiple Qlik Cloud services. This includes our Agentic AI Framework, Generative AI & machine learning feature sets. When adding new AI based services, or enhancing existing services with AI, Qlik's priority is to put the protection of customer data first. The full list of AI capabilities are kept on our [Trust & AI](#) page.

AI is not new to Qlik. Qlik introduced the Cognitive engine in 2017, enhancing our associative technology with AI features. At the same time we introduced Server Side Extensions (SSE) to integrate machine learning technologies into Qlik Applications, one of the first Analytics vendors to move in this direction. This was further expanded in 2018 with Insight advisor. Qlik has continued to develop and innovate our products with AI through the years and have a comprehensive set of AI capabilities throughout the Data & Analytics lifecycle.

Qlik Answers[®]

Qlik Answers is Qlik's AI-powered question-and-answer capability built into Qlik Cloud. It allows users to ask natural language questions about their data and receive AI-generated responses grounded in that data, rather than requiring them to build charts or navigate dashboards manually. Qlik Answers supports both structured and unstructured data.

Qlik Discovery Agent[™]

Qlik Discovery Agent automatically detects meaningful changes, anomalies, and trends across your Qlik apps – no rules, thresholds, or manual setup required. Get clear, prioritized insights delivered in plain language so your team can act before issues escalate. All insights surface in the Feed – a central hub that aggregates findings across multiple Qlik apps. Filter by app, insight type, or severity to focus on what matters most, and revisit saved insights whenever you need them.

Qlik MCP Server[™]

Qlik MCP Server is an interface that allows third-party AI clients to query Qlik Cloud data using the user's own credentials and permissions. It implements the Model Context Protocol (MCP) – an open standard for connecting AI assistants to external data sources. This allows users to connect tools like Claude, ChatGPT, GitHub Copilot, Cursor, or VS Code directly to their Qlik Cloud data and ask questions in natural language, with the AI drawing on live Qlik data to generate responses.

2.2 AI in Qlik Cloud

Design Principles for AI on Qlik Cloud

Qlik believes the productivity and efficiency gains AI provide should not require customers to take risks with their data. As a company, Qlik's approach to AI is based key principles. These principles are explained at our [Trust & AI page](#). The design principles drive our development efforts and ensure the solutions we

deliver meet Qlik's ethical standards, giving our customers confidence that leveraging AI with Qlik does not compromise their data security & governance needs.

Based on Qlik's overarching principles, AI principles and governance requirements, the following design principles apply to our AI initiatives at Qlik:

1. Qlik does not use the customer's data to train any AI model external to the customer's tenant. Any training of models internal to the customer's tenant, such as using the machine learning aspects of Qlik Predict, is exclusively to the benefit of that customer and is not shared with others.
2. Qlik Cloud is a highly governed and audited platform, holding many certifications and accreditations (see [Qlik Trust & Security](#)). Any design decisions must align with the requirements of these laws, certifications and accreditations.
3. AI compliance reviews are conducted as part of our product development processes to ensure our AI implementations align with our general AI principles.

We also apply our general cloud design principles to AI, namely:

1. Qlik Cloud is a no view service. This means neither Qlik employees, our cloud providers, nor any parties other than the customer themselves can see the customers data without the customer explicitly providing access. Our AI initiatives must meet this requirement also.
2. Qlik cloud uses multiple layers of encryption (optionally using customer provided encryption keys) for all data stored on the platform including unstructured data used by our AI solutions.
3. All data moving into and out of Qlik Cloud (including AI) is secured with TLS.

Guardrails

When interacting with large language models (LLMs) Qlik applies a number of guardrails to minimize risks for our customers:

- LLM monitoring for harmful content. Qlik scans and rejects questions that contain invisible text or prompt injections.
- Hallucinations. Qlik mitigates hallucinations by detecting contextual relevance in the provided content. Additionally, users can view the content used as a source to verify responses.

Data Sovereignty

When a customer creates their Qlik Cloud tenant, they are able to choose the region their tenant resides in. A customer's data does not leave their chosen region by default. Due to the speed of rollout of some AI services, as well as resource constraints from our providers, certain features may necessitate limited data to be sent securely to another region. Where this is required, customers must opt-in to this. Qlik will not transfer data to other regions without the customers express authorization.



Cross-region inference

Certain AI features in Qlik cloud require customers to allow Cross-region data processing in a very limited capacity in the use of large language models (LLMs). This is known as cross-region inference, or CRIS for short. Enabling cross-region inference allows Qlik Cloud to temporarily send the data required for processing a question to the AWS region where the service is available, process the request, and return the results. **Enabling cross-region inference does not give Qlik or AWS access to your data.** This requirement exists due to the high demand for AI services and limited availability of hardware from AWS. This is not unique to Qlik and is required for most AI services today due to industry-wide resource constraints.

What does this mean for my data?

- All data at rest including tenant event logging information is stored encrypted in the customer's tenant in the customer's chosen region using encryption keys unique to the customer; these can optionally be provided by the customer (See [Tenant encryption](#)).
- Data sent between AWS regions is encrypted with AWS's Encryption keys and travels through the secure AWS private network with end-to-end encryption for data in transit.
- No data is persisted in the other regions.

More information on how this impacts specific AI services is covered in the relevant sections of this document. For a detailed understanding of these processes at AWS, see [Securing Amazon Bedrock cross-Region inference](#)

Authorization & AI in Qlik Cloud

AI services in Qlik cloud fully respect any configured authorization in Qlik Cloud. Access to applications, files and connections rely on permissions granted to the spaces those assets reside in. Within Applications, our AI services are bound by any section access configured for that application. If a user is restricted from data when using the application in a conventional way, they will also be restricted when interacting through AI.

AI Services in Qlik Cloud

AI is used throughout Qlik Cloud in many areas. Broadly speaking, our AI services fit into two key areas:

- Specific AI services and capabilities in Qlik Cloud
- AI Accelerators and helpers for non-AI services and capabilities in Qlik Cloud

AWS change and replace services over time, so it is possible Cross-region inference requirements may change in the future. The following is true at the time of writing.

The following AI services in Qlik Cloud do not require Cross-region inference to function:

- Qlik Answers[®] (legacy assistants only)
- Qlik Predict[™]
- Qlik MCP Server[™]

The following AI services are available in Qlik Cloud that rely on Cross-region inference to function:

- Qlik Answers[®] (excluding legacy assistants)
- Discovery Agent[™]
- Qlik Talend Cloud Pipelines: AI SQL generation and AI model suggestions (see [Generating a SQL transformation from a text prompt](#)).
- Data Products: AI description generation (see [Working with validation rules](#)) and AI DQ Validation rule suggestions (see [Generating an AI-based descriptions](#)) .

AI Models at Qlik

Qlik selects AI models and CRIS inference profiles to best fit with our customers Data Sovereignty, data security and geo-political requirements wherever possible. This is bounded by the limitations enforced on us by AWS - i.e. what infrastructure and CRIS profiles are made available for each Qlik home region by AWS. This is under constant review by AWS and Qlik to best serve the needs of our customers. For the latest breakdown of potential inference processing regions by product feature and home region please see: [Enabling cross-region data processing - Qlik Help](#).. You can also subscribe to the following community page to receive updates to processing regions when changes are announced: [Qlik Cross-Region Data Processing - Qlik Community](#) .

2.3 Generative AI with Qlik Answers[®]

What is Qlik Answers ?

Qlik Answers is a plug-and-play, generative AI-powered knowledge assistant that provides business users with personalized, contextually relevant answers to questions sourced from structured and unstructured content. Unlike traditional search, generative AI delivers personalized answers to questions instead of just lists of content.

Unstructured responses are derived from a number of underlying sources and documents that have been carefully curated into domain-specific knowledge bases. You just ask it a question and get an answer - it's that simple. Answers are reliable and consistent, and with full explainability, you'll always know where things came from and have access to those sources - ensuring consistency, trust and transparency.

Structured responses come from your own Qlik Applications allowing you to leverage your investment in Analytics for much wider audiences. Qlik Answers can leverage your existing content and create new content as it fully understands the business logic in your applications.

Why Qlik Answers ?

The opportunity

Qlik has long been a leader in structured analytics and data integration. However not all data in organizations is structured data. According to Forester, 80% of the worlds data is unstructured and 70% of enterprises agree they are not effectively leveraging the value in this data. Search engines have been seen as a way to leverage this data; however they provide limited benefits in this space as they are really only doing pattern matches and often return irrelevant results.

Gen-AI (generative artificial intelligence) has grown massively over the last year as a solution for unstructured data. Gen-AI's ability to answer natural language questions and create content based on existing knowledge provides solutions such as virtual assistants, code generation, and process optimization. When used in an Enterprise context, Gen-AI helps provide value from that 80% of unstructured data.

The challenges

The barrier to entry for Gen-AI is considerable. While LLMs (large language models) receive a lot of attention, the reality is this is just one of many technologies needed to implement a Gen-AI solution. The cost and complexity to do this well is beyond the resources of all but the largest organizations. This necessitates most organizations using a third-party solution. In most cases this involves storing the organization's data with that third-party and risks questions being answered not just with an organization's own data, but with other publicly available data which may be outdated or inaccurate. Also, many of these solutions are subject to hallucinations, where content is made up by the Gen-AI solution. These solutions typically lack the governance and security that organizations come to expect around their data and raise the risk of legal issues, reputational issues, or both.

The solution

These factors are why we have developed our Gen-AI solution, Qlik Answers. Qlik Answers is an Enterprise-grade, plug-and-play, generative AI-powered knowledge assistant that provides business users with personalized, contextually relevant answers to questions sourced from Structured and unstructured content. Running on the Qlik Cloud platform, Qlik Answers leverages Qlik Cloud's existing security and governance features to provide an Enterprise-ready, plug-and-play, generative AI solution.

See the [Security and Governance](#) section of our Qlik Cloud platform guide for more information.

A mature foundation in AI

Qlik's AI solutions are not new - we have been augmenting our analytics solutions with AI technologies since 2018. In January 2024, Qlik acquired Kyndi, an innovator in natural language processing, search, and generative AI. Kyndi's proprietary technology along with an experienced team of generative AI experts have enabled Qlik to quickly integrate Gen-AI into our existing AI teams and bring a solution to market which provides a plug-and-play solution to our customers.

How Qlik Answers works

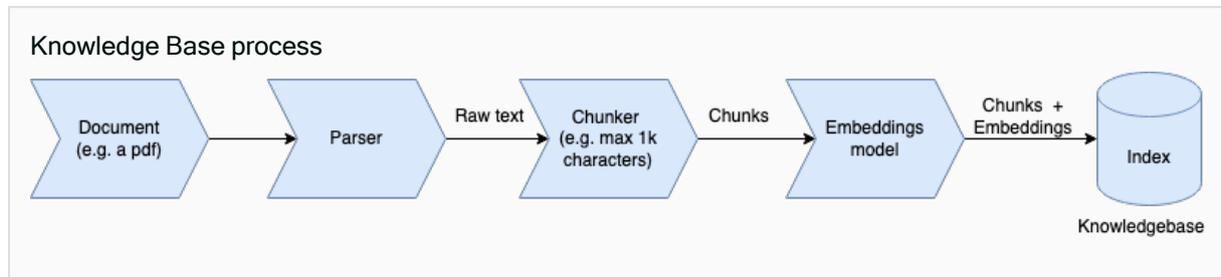
Qlik is building a simple, plug-and-play solution that can be deployed out-of-the-box without the complexity of custom-built solutions. However, simple does not mean simplistic - Qlik Answers uses cutting edge semantic search, vector databases, Gen-AI, LLMs and RAG (retrieval augmented generation) technologies under the hood.

Components of Qlik Answers

Qlik Answers consists of the following components:

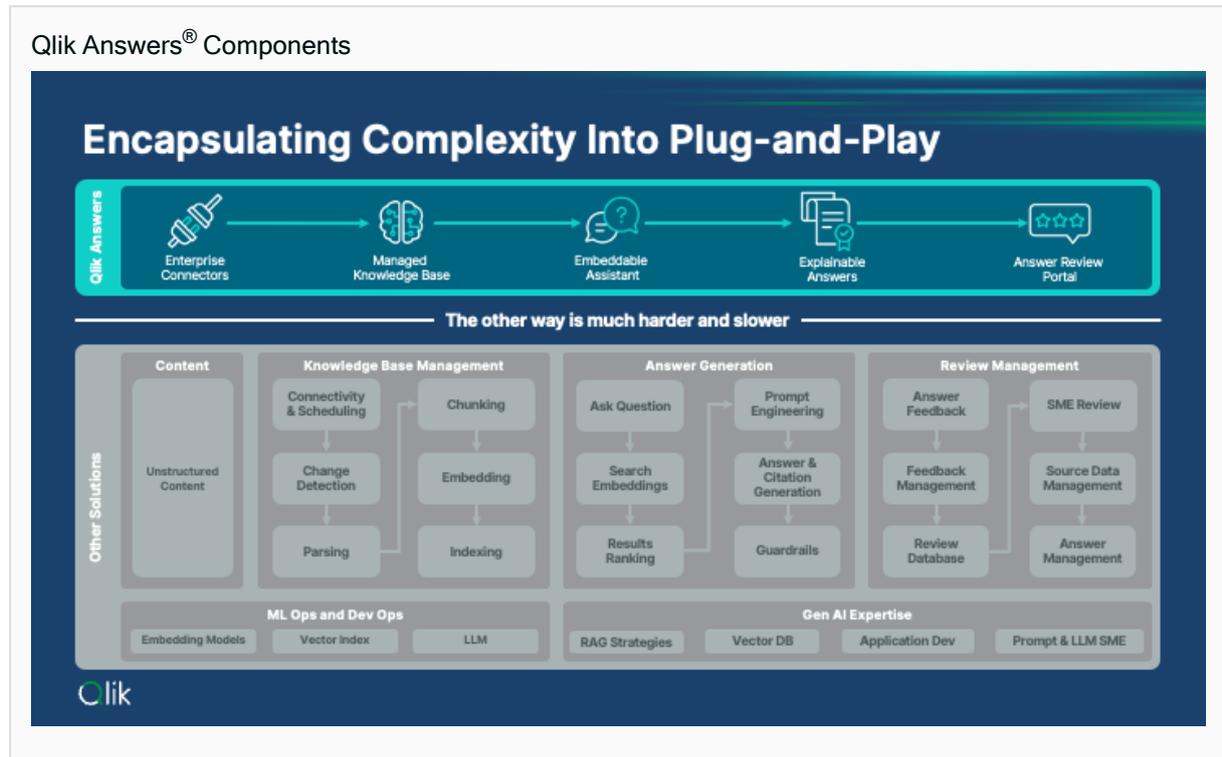
- **Enterprise connectivity** - Using our existing file store connectors, we can access your organization's data in-place.

- **Managed knowledge bases** - These contain references to your organization's data. Knowledge bases are generally based around a particular subject. For example, a customer may have HR, Finance, and Sales knowledge bases. Knowledge bases do not store all your source data in Qlik Cloud (unless you are using a space's DataFiles storage as your source). They contain an index of that data in a numeric format.



- **Qlik Applications** - Qlik Answers can index Qlik applications to allow users to get insights directly from Qlik Applications, either independently or in combination with knowledge bases. Questions can be asked directly from within an app or through assistants configured to use that Application.
- **Embeddable assistants** - These are the part of Qlik Answers that users interact with. An assistant can be based on one or more knowledge bases. This allows us to provide customized assistants to different audiences without having to duplicate content. Assistants are available through Qlik Cloud or can be embedded into your organization's own web portals. Users can provide feedback on the answers received which is then available for administrators to review through Qlik Answer's Review Portal.
- **Explainable answers** - Answers returned from assistants provide references to the documents the answers were sourced from. This allows customers to understand how the answer was derived.
- **Answer review portal** - This allows administrators to see what questions users are asking and any feedback provided. This information can be used to identify areas of interest and potential areas of improvement if users are not satisfied with the answers given.

Technology behind Qlik Answers

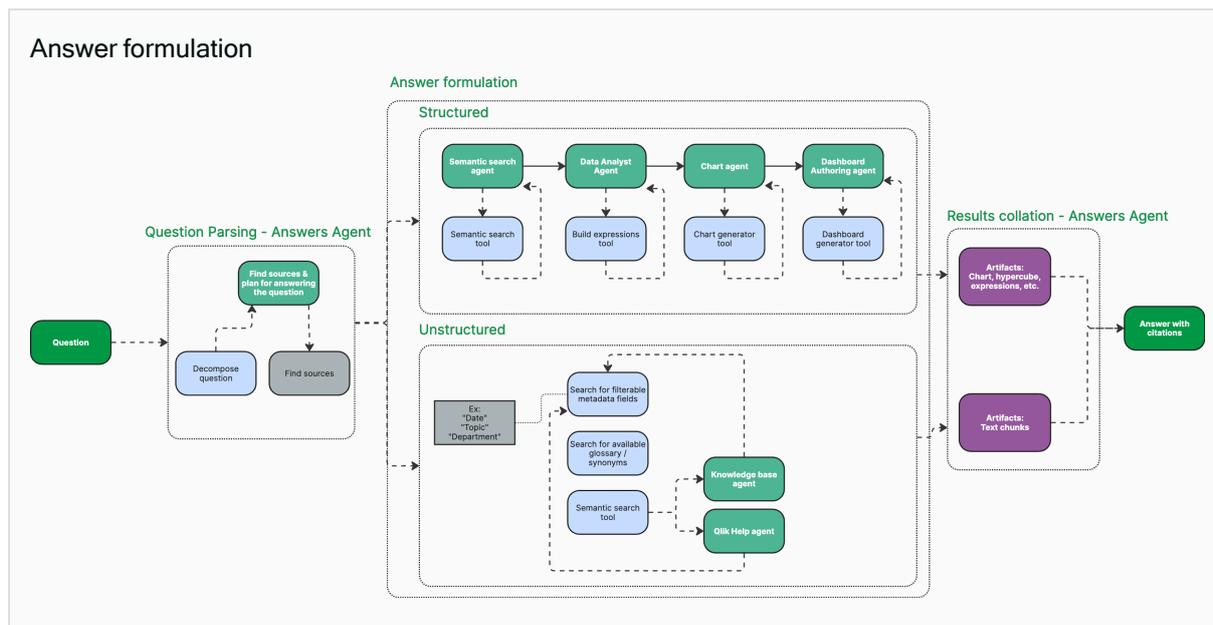


Behind the major components of Qlik Answers are over 25 different Answers-specific processes and numerous other processes of the Qlik Cloud platform. A few of the key technologies used within Qlik Answers are:

- **Large Language Models (LLMs)** - A large language model (LLM) is a computational model notable for its ability to achieve general-purpose language generation and other natural language processing tasks such as classification. LLMs are very large deep learning models that are pre-trained on vast amounts of data. See [What are Large Language Models?](#) for more information.
- **Semantic search** - Semantic search seeks to improve search accuracy by understanding the searcher's intent and the contextual meaning of terms to generate more relevant results. See [Semantic search](#) for more information.
- **Retrieval Augmented Generation (RAG)** - RAG is a technique by which you include context retrieved from some form of search to generate a result. See [What is Retrieval-Augmented Generation?](#) for more information.
- **Qlik Analytics Engine** - We have developed our own enhancements to our existing associative engine to provide the power of Qlik Analytics within Qlik Answers as well as other generative AI solutions.

Answer formulation in Qlik Answers

The following diagram illustrates the process a Qlik Answers assistant uses to answer a question across both structured and unstructured data.



Question parsing - Answers Agent

The first thing Qlik Answers does is parse the question. This is done by the **Answers Agent**. It does this to determine both what the user is asking as well as which resources it has access to that are relevant to asking the question. Once it had determined this it will pass it on the parsed question and resources to the answer formulation process; this may be Structured, Unstructured or both. Specifically, the agent:

- Decomposes the user question into logical sub-task
- Identifies intent and context behind the question
- Selects and triggers the relevant tools to answer each sub-task

Answer formulation - Structured

When formulating answers on structured data, Qlik Answers first uses the **Semantic Search Agent** to identify the relevant master items, measures and dimensions needed to answer the question.

The **Data Analyst Agent** builds an analysis package that specifies

- What data to use
- How to analyze it
- What output best answers the question

It is Designed to always respect master items - Master measures and master dimensions are evaluated before raw fields. Developer defined items is treated as the trusted source - this is designed to ensure answers align with a customer's business outcomes. Business logic defined in an app is beneficial but not essential in all cases. Areas when business logic particularly benefits Qlik Answers are with synonyms and hiding fields you do not want used by the assistant.

The **Chart Agent** receives the analysis package from the Data Analyst Agent and then interprets execution instructions for types of visualization best suited to the data. It applies chart logic based on dimensions, measures, and intent and builds the chart as the final output.

The **Dashboard Authoring Agent** consumes outputs from Qlik Answers Data Analyst and Chart Agents, and translates these analytical outputs into sheet-ready objects. It helps users place, structure, and assemble content on into dashboards.

Answer formulation - Unstructured

The **Knowledge Base Agent** performs semantic searches through your knowledge bases (vector databases) to get relevant unstructured data and then uses RAG to pass that data to the LLM, which then provides an answer. The answer includes references back to the relevant documents that are indexed in your knowledge base.

The **Help Agent** is a special type of knowledge base that is provided by Qlik to customers. It contains Qlik Cloud product documentation and is designed to help you learn Qlik Cloud and take advantage of the power of our platform.

Results Collation - Answers Agent

At the end of this process, the **Answers Agent** takes the results from the other agents and composes their outputs back into a single, coherent answer.

Third-party technologies

The following third-party technologies are used as part of Qlik Answers at the time of writing. Qlik may choose to introduce additional or alternate technologies for Qlik Answers in the future.

- [Temporal](#) - Temporal is used to distribute workflows in the file ingestion workflow (i.e. Knowledge Base indexing).
- [AWS Bedrock](#) - AWS Bedrock is used as a service to connect to LLMs (Large Language Models) such as Anthropic Claude. The actual models we use change over time based on requirements and the evolution of the models.
- [OpenSearch](#) - OpenSearch is the vector store that stores embeddings extracted from documents. OpenSearch is also the semantic search engine that is used to gather relevant text chunks during answer generation.
- [Cohere Rerank](#) - Cohere Rerank is used to rank the most relevant embeddings returned by OpenSearch.
- [Amazon SageMaker](#) - Amazon SageMaker is used to host the Cohere Rerank model.

Access control in Qlik Answers

Qlik Answers leverages the existing access control model in Qlik Cloud. Access is restricted both by space access and by roles. From a user's point of view, it is possible to provide access to the assistant and only some of the knowledge bases connected to the assistant, so certain privileged data can be restricted to only certain users. It is also possible to restrict access to the source data, so users may be able to receive an answer, but not see the document the answer was derived from. When interacting with Qlik Applications, Qlik Answers fully supports section access and users will only receive responses based on what they are allowed to see in the application.

From an administrative point of view, the same approach is used. It is possible to allow users to create an assistant against a knowledge base, without having access to change or create knowledge bases. It is possible to allow the indexing of existing knowledge bases without allowing the creation of new knowledge bases.

These controls can be assigned to users or groups and are fully governed and audit logs are available either through the platform UI or via API integration into your external systems.

2.4 Qlik Answers® - Security Architecture and Controls

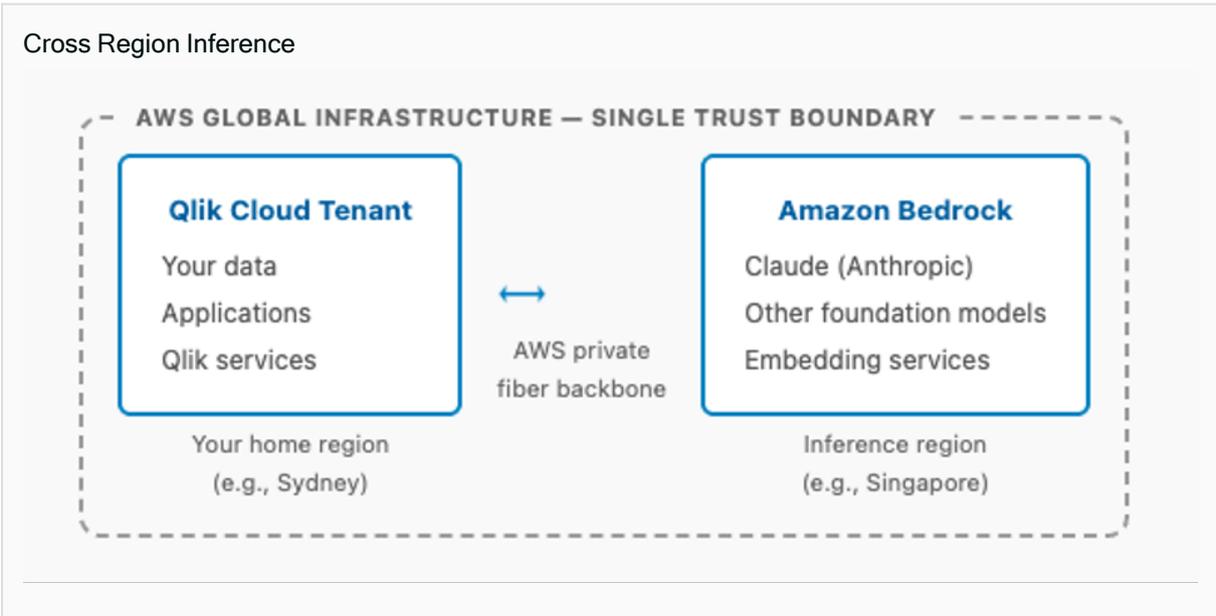
Introduction

This topic explains how Qlik Cloud handles data when AI capabilities such as Qlik Answers, semantic search, and natural language insights are used. It is intended to help security and compliance teams understand the technical controls that apply when inference requests are routed across AWS regions - known as *Cross-Region Inference*.

Qlik Cloud runs on AWS, and AI inference is provided by Amazon Bedrock – also on AWS. All processing occurs within a single AWS security boundary. **No third-party cloud providers or public internet routes are involved.**

Architecture Overview

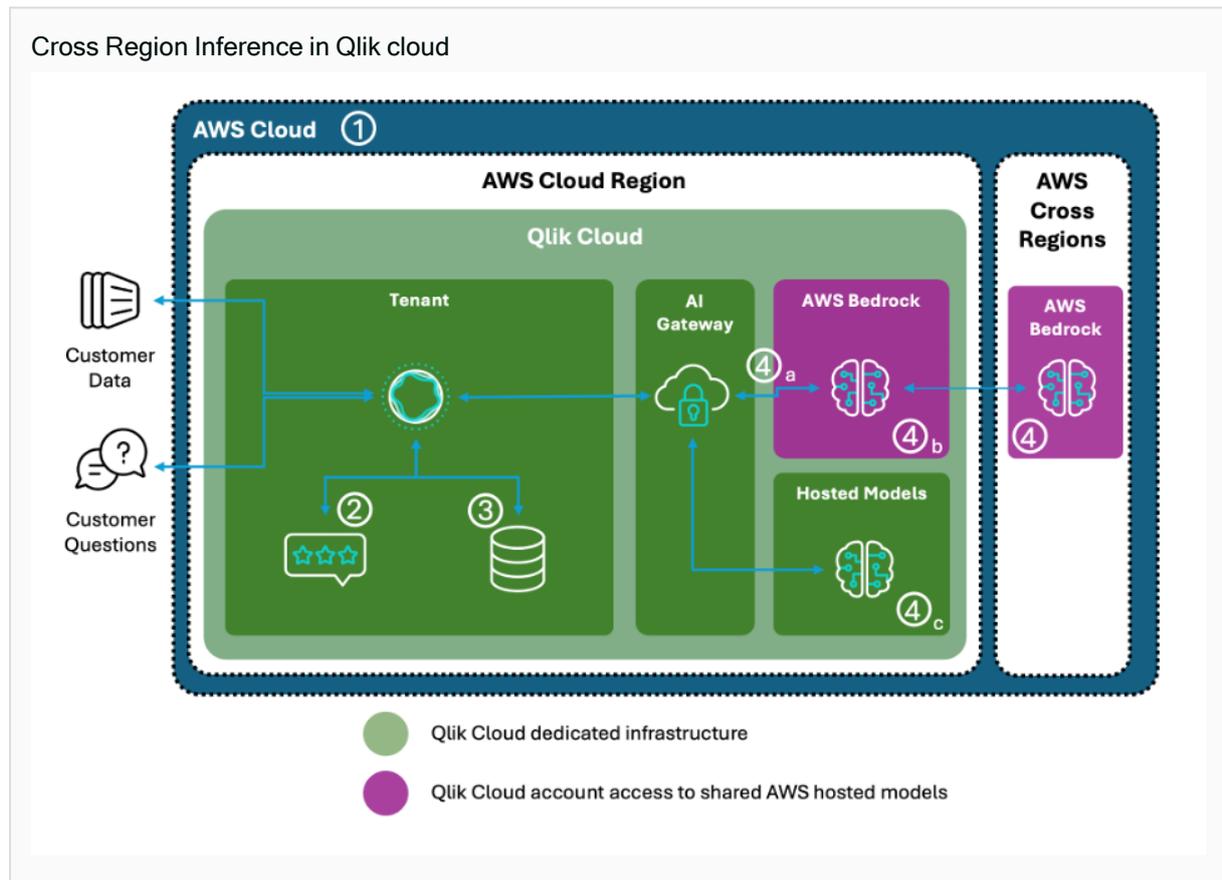
Qlik Cloud is a native AWS SaaS platform. Its AI capabilities are powered by Amazon Bedrock, AWS's fully managed service for accessing foundation models, including Claude from Anthropic. Because both Qlik Cloud and Bedrock operate on the same AWS infrastructure, this is an AWS-to-AWS architecture rather than a multi-cloud integration. The remote region used in Cross Region inference is referred to as the **inference region** throughout this document.



Your Qlik Cloud tenant resides within the AWS network at all times. When AI inference occurs in a different region, data moves between two components of the same AWS infrastructure – connected by AWS's private fiber backbone – rather than being sent to an external vendor over the internet.

This architecture means you have a single trust boundary, a single Data Processing Agreement (DPA), and a unified compliance framework, rather than the multiple agreements and reconciled policies required in traditional multi-cloud AI deployments.

In more detail:



1. All traffic stays within the same AWS Virtual Private Network.
2. Questions and responses are logged within the customer tenant using tenant encryption.
3. Indexed customer data for applications and knowledge bases are encrypted at rest in the index with tenant encryption.
4. Requests to generate AI representations are made to the LLM model through cross-region inference:
 - Data is encrypted in transit via the secure AWS network.
 - No data is retained by AWS or AI model providers; no customer data is used to train LLMs.
 - Some AI models are hosted within the Qlik Cloud VPC.

How Cross-Region Inference Works

The following describes what happens when a user action initiates an AI request and the inference request is routed from a home region (for example, Australia) to an inference region (for example, Singapore).

While the example below relates to Qlik Answers, the process is the same for all AI services using cross-region inference.

Step	Description	Data State	Example
1 – User query (home region)	User asks a question in Qlik Answers. Qlik retrieves relevant data from the application and holds it in memory.	In memory - home region	User asks: <i>"What were our top sales drivers last quarter?"</i> Data in memory: [CustomerID, Revenue, Product, Region...]
2 – Construct prompt (home region)	Qlik assembles the inference payload: the user's question, relevant contextual data from the app, and system instructions. Payload exists as plaintext in memory at this stage.	Plaintext - in memory only	<pre>{ "question": "What were our top sales drivers...", "context": { "Q3_sales": [{"Product": "Widget A", "Revenue": 500000}, {"Product": "Widget B", "Revenue": 350000}] } }</pre>
3 – Encrypt & transmit	TLS 1.3 encryption applied to entire payload. IAM role authentication established. Transmitted over AWS private fiber backbone – never the public internet.	TLS 1.3 encrypted - AWS private network	Home Region —[AWS Private Network]—> Inference Region – ciphertext only, no plaintext in transit
4 – Arrive at Amazon Bedrock (inference region)	TLS session established and verified. IAM credentials validated. Request payload decrypted by Bedrock so the model can process it.	△ Plaintext at inference endpoint - required for model processing	Decrypted request now in Bedrock memory: { "question": "What were our top sales drivers...", "context": { ...original data... } }
5 – AI model inference (in memory, ephemeral)	Bedrock passes plaintext to AI model. Model processes request in memory only. No disk writes. Inference duration: 1-5 seconds. Natural language response generated.	In memory only - 1 to 5 seconds	Model response: <i>"Based on the Q3 data provided, Widget A was the top sales driver with \$500K revenue, primarily from the APAC region..."</i>
6 – Purge memory & return response	Original request data immediately purged from memory. Response encrypted with TLS 1.3 and sent back to	TLS 1.3 encrypted - no data retained in	Inference Region —[AWS Private Network]—> Home Region – encrypted response only, original request data gone

Step	Description	Data State	Example
	home region via AWS private network. No data retained in the inference region.	inference region	
7 – Display results (home region)	Response decrypted by Qlik Cloud and rendered to user in natural language. User sees insights without raw data exposure.	Home region - response displayed to user	User sees: <i>"Widget A was your top sales driver last quarter with \$500K in revenue."</i>

Common Misconceptions

The following table addresses questions that frequently arise during security assessments of Qlik Cloud's AI features.

Misconception	Reality
Data leaves AWS and goes to Anthropic's servers.	Incorrect. Claude is hosted on Amazon Bedrock, which runs entirely on AWS infrastructure. Anthropic does not operate the servers used for Bedrock customers.
Cross-region means data goes over the public internet.	Incorrect. AWS owns and operates a private fiber optic network connecting its regions globally. Traffic between regions never traverses public internet routes.
AI services can access our data at any time.	Incorrect. Data is transmitted only during an active inference request (typically 1-5 seconds), processed in memory, then immediately purged. There is no standing access to your data.
Our data is logged and stored in the inference region.	Incorrect. Amazon Bedrock logs metadata only – timestamps, model version, token counts, and error codes. Data content is never written to logs.
Encryption means our data is always protected, including during inference.	Requires context. Encryption protects data in transit and at rest, but AI models must operate on plaintext – decryption occurs at the inference endpoint. This is a fundamental requirement of any AI service, not specific to Qlik Cloud.

Security Controls

Network Controls

All cross-region traffic between Qlik Cloud and Amazon Bedrock travels over AWS's private fiber optic backbone. This network is owned and operated by AWS and is separate from the public internet – there is no BGP routing through third-party internet service providers and no mixing with public internet traffic.

Connectivity between services uses AWS PrivateLink or VPC peering. VPC security groups restrict traffic to authorized endpoints, and network ACLs provide subnet-level filtering.

Encryption

All data in transit between Qlik Cloud and Bedrock is encrypted using TLS 1.3. Each session uses unique encryption keys, and certificate-based authentication prevents man-in-the-middle attacks.

Data at rest within your Qlik Cloud tenant is encrypted using AES-256, with keys managed through AWS Key Management Service (KMS). Customer-managed keys (BYOK) are used if configured by the customer.

Important: While data is encrypted in transit, it must be decrypted at the inference endpoint for the AI model to process it. This is an unavoidable requirement of all AI inference services. Encryption cannot be maintained end-to-end during active model processing.

Access Controls

Qlik Cloud authenticates to Amazon Bedrock using AWS IAM roles with temporary, automatically rotated credentials – not long-lived API keys. The principle of least privilege is applied throughout. End users never interact directly with Bedrock; all requests are mediated by Qlik Cloud, which enforces your existing access controls including role-based access, row-level security, and section access.

Data Handling and Retention

Data processed in an inference region exists in memory only for the duration of the inference request – typically 1 to 5 seconds. After the response is generated, memory is immediately cleared. No disk writes of your data occur in inference regions. Data passed to the inference region does not identify the user, customer, or the license related to the request.

Amazon Bedrock does not log data content. Only operational metadata is captured: timestamps, model version used, token counts, and error codes. Your data is never used to train or fine-tune AI models. This is contractually guaranteed in the AWS Data Processing Agreement.

Compliance and Audit

Because Qlik Cloud and Amazon Bedrock operate within the same AWS infrastructure, they share a unified compliance posture. See [Qlik Trust and Security](#) for currently held compliance certifications.

Data Storage by Location

Your data is only ever stored in your home region. The inference region performs ephemeral, in-memory processing only.

Data type	Home region	Inference region	Retention
Raw application data	✓ Stored	✗ Never stored	Permanent in home region
User questions	✓ Stored	✗ Not stored	Session-based only
Inference request payload	Transient during API call	In memory only (1-5 sec)	Purged immediately after inference
AI model response	✓ Stored	✗ Not stored	Displayed to user; session-based

Data type	Home region	Inference region	Retention
Embedding vectors (semantic search)	✓ Stored for search	In memory during generation only	Permanent in home region; purged in inference region
Data content in logs	✗ Not logged	✗ Never logged	N/A

Storage vs. in-memory processing: *Storage* means data is written to disk and persists across server restarts. *In-memory processing* means data is held in RAM temporarily and is gone when the process completes or the server restarts. Your data is only ever *stored* in your home region.

Amazon Bedrock as a Security Layer

Amazon Bedrock is AWS's fully managed service for accessing foundation models from multiple AI providers. When Qlik Cloud uses Large Language Models via Bedrock, the model is hosted on AWS infrastructure – not on the model vendor's own servers. The model vendor does not operate the infrastructure for Bedrock customers and does not have access to your inference requests or data. AWS handles all provisioning, scaling, security, and networking.

This is architecturally significant. Accessing Claude directly via Anthropic's API would involve data leaving the AWS security boundary and traversing the public internet to Anthropic's servers, under a separate compliance framework and DPA. Using Bedrock keeps the entire flow within AWS, under a single governance framework.

Data isolation guarantees: Bedrock enforces a no-training policy – your data is never used to improve or fine-tune the underlying models. This is contractually enforced through agreements between AWS and its model providers, including Anthropic.

Enterprise governance features: Bedrock provides model version pinning, metadata logging to CloudWatch, VPC integration for network isolation, and configurable guardrails for content safety.

Compliance inheritance: Bedrock inherits all AWS compliance certifications. For security assessments, this means evaluating AWS's security posture – which organizations already trust for Qlik Cloud – rather than assessing a separate external AI vendor.

Why Cross-Region Inference Is Necessary

Cross-region AI inference is a practical requirement arising from the hardware constraints of modern large language models, not a product design preference.

Large language models require specialized GPU hardware – particularly NVIDIA H100 and A100 GPUs – that is in limited global supply. AWS cannot provision Bedrock inference capacity in every region simultaneously. New data centers take 2-3 years to build, and GPU availability lags further behind. As a result, Bedrock capacity is concentrated in a small number of high-capacity regions.

Even in regions where capacity exists, cross-region routing is necessary for high availability. During peak business hours, a regional Bedrock deployment may be highly utilized. Routing to a different region ensures consistent response times and prevents timeouts.

Example: At 10 AM AEST (peak business hours in Sydney), Sydney Bedrock capacity may be 85% utilized while Singapore capacity is 40% utilized due to the different time zone. Routing to Singapore maintains a 2-3 second response time rather than a 30+ second wait or timeout.

As AWS expands Bedrock availability to more regions over time, cross-region routing may become less frequent however this is a quickly evolving area and we are highly dependent on our vendors. However, load balancing across regions will remain a feature of high-availability service design.

Security Assessment Guidance

The following guidance is intended to help security and compliance teams determine how to treat cross-region AI inference risk within their organizations.

When This Risk Is Generally Acceptable

Cross-region inference within AWS is consistent with the risk posture of organizations that already use Qlik Cloud on AWS, accept momentary in-memory data exposure during inference (1-5 seconds), and rely on AWS's DPA for contractual data protection. It is appropriate where service availability and consistent performance are priorities.

Risk Mitigation Options

Organizations can reduce exposure by excluding highly sensitive fields from AI-enabled features using section access or field exclusions, or by pre-processing data before querying – for example, tokenizing PII or aggregating values.

When This Risk May Not Be Acceptable

Cross-region inference may be incompatible with data sovereignty laws that prohibit any cross-border processing (common in certain government and healthcare contexts), regulatory requirements that mandate processing within specific geographic boundaries, or threat models that treat any in-memory plaintext exposure during inference as unacceptable.

In these cases, use the tenant-level setting to disable cross-region data processing, which is the default setting for Qlik Cloud tenants.

2.5 Qlik MCP Server™

What is the Model Context Protocol?

The Model Context Protocol (MCP) is an open standard developed by Anthropic that defines how AI assistants – such as Claude, GitHub Copilot, and others – connect to external data sources, tools, and applications. Qlik's MCP server implements this standard, exposing Qlik Cloud's analytical capabilities – including apps, data models, insights, and automation – as a set of tools that any MCP-compatible AI assistant can discover and use.

Rather than requiring each AI tool to build its own proprietary integrations, MCP provides a universal interface: a structured way for AI models to discover what a system can do, request data from it, and take actions within it. Think of MCP as a common language between AI assistants and enterprise platforms. Just as REST APIs standardized how applications communicate with each other over the web, MCP standardizes how AI models communicate with the tools and data sources they need to be useful in a business context.

Business value

The Qlik MCP server fundamentally changes how people interact with analytics. Instead of requiring users to navigate to a Qlik Application, locate the right chart, and interpret the output themselves, AI assistants can now retrieve, interpret, and surface Qlik insights directly within the tools people already work in – whether that is a chat interface, a coding environment, or an enterprise AI assistant.

This delivers value in several ways:

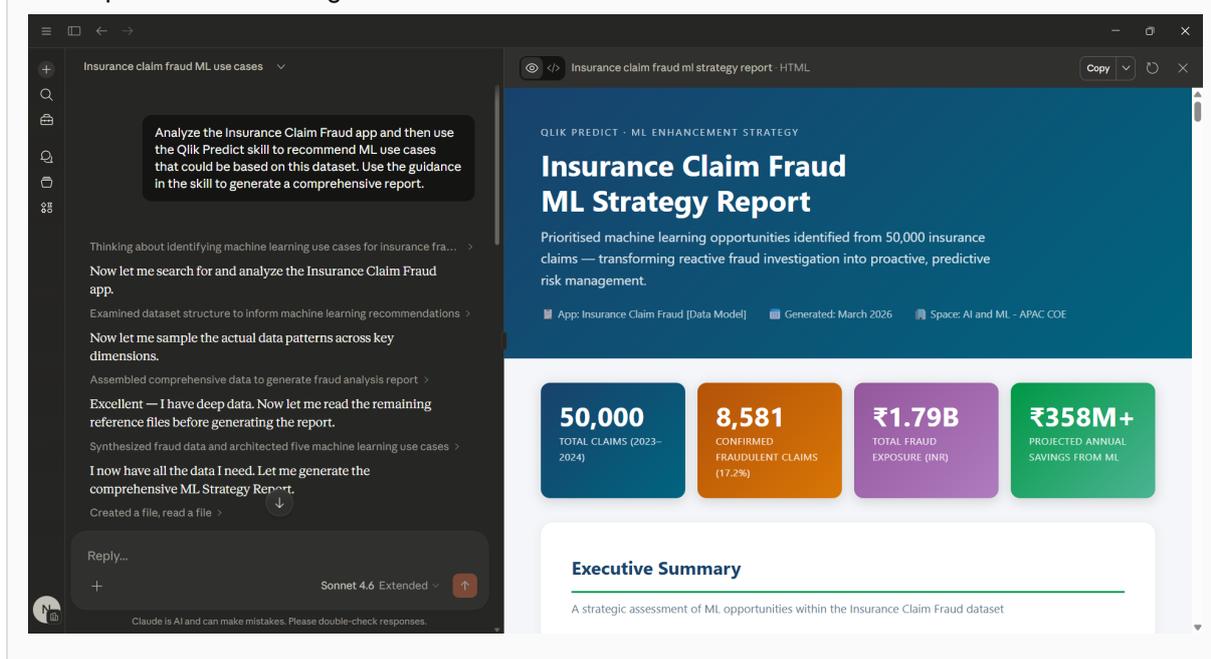
Analytics where work happens. Business users no longer need to context-switch between their workflow tools and Qlik. An AI assistant embedded in a chat platform or productivity tool can answer data questions by querying Qlik in the background and returning a plain-language answer – without the user ever opening a Qlik app.

Faster time to insight. Analysts and data teams can use AI assistants to interrogate Qlik data models, explore app content, and generate insights conversationally, significantly reducing the time from question to answer.

Broader reach for existing analytics investments. Organizations that have built robust Qlik apps and data models can extend the value of those assets to new audiences and new interfaces, without rebuilding or replicating data pipelines.

Reduced dependency on self-service navigation. Not every employee who needs data insights is comfortable building or navigating a BI dashboard. MCP-connected AI assistants lower the barrier, making Qlik's analytical depth accessible to a much wider range of users through natural language.

Anthropic's Claude working with Qlik's MCP Server



Use cases and opportunities

Conversational analytics in enterprise chat platforms

AI assistants connected to Qlik via MCP can answer business questions in platforms like Microsoft Teams or Slack. A sales manager asking What were our top five accounts by revenue last quarter? receives a direct, data-grounded answer – pulled in real time from Qlik – without leaving the chat tool.

AI-assisted data exploration for analysts

Data analysts using MCP-compatible development environments or AI coding assistants can query Qlik app structures, explore data models, and retrieve field-level data to accelerate analysis and dashboard development. Tasks that previously required manual navigation through the Qlik interface can be performed conversationally.

Automated reporting and narrative generation

MCP enables AI assistants to retrieve current data from Qlik and incorporate it into generated content – such as executive briefings, performance summaries, or operational reports – automatically updated with the latest figures from your Qlik apps.

Intelligent workflow automation

By connecting Qlik's analytical capabilities to agentic AI workflows, organizations can build automated processes that monitor data conditions, surface alerts, and trigger actions – all grounded in the governed, trusted data that lives in Qlik Cloud.

Developer and IT tooling

Development teams building internal AI tools or custom agents can use the Qlik MCP server to give their applications access to Qlik's data and insights layer, without building bespoke API integrations for each use case.

User personas

Business analyst

Business analysts are the primary builders of Qlik apps and data models. With MCP, they can use AI assistants to explore Qlik app structures more quickly, prototype new analyses conversationally, and validate data logic – reducing the manual effort involved in iterative development. MCP effectively gives analysts a faster, more intuitive interface to the platform they already work in.

Business user / information consumer

This persona consumes analytics outputs but may not be comfortable with traditional BI interfaces. The Qlik MCP server opens up Qlik's analytical depth to this audience by making it accessible through natural language, in the tools they use every day. A finance manager, operations lead, or HR business partner can ask data questions and receive grounded answers without needing to understand how a Qlik app is structured.

Data engineer / data architect

Data engineers responsible for Qlik's underlying data models can use MCP-connected AI tools to inspect app metadata, understand field lineage, and accelerate troubleshooting. Rather than manually navigating the Qlik interface to audit a data model, they can query it conversationally through a developer AI assistant.

AI / application developer

Developers building internal AI assistants, agents, or productivity tools can integrate Qlik's analytical capabilities into their applications using the MCP server, without building custom API integrations. MCP provides a standardized interface that significantly reduces development effort when Qlik data or insights need to be part of an AI-powered workflow.

IT administrator / platform owner

Platform owners responsible for Qlik Cloud governance will interact with the MCP server in terms of access control, monitoring, and configuration. They determine which AI assistants and tools are permitted to connect to Qlik via MCP, and what data those connections can access – ensuring that the same governance controls that apply to direct Qlik usage extend to AI-mediated access.

2.6 Qlik MCP Server™ - Security Architecture and Controls

Document Purpose: This document describes the security architecture and controls governing the Qlik Model Context Protocol (MCP) server. It clarifies how authentication, authorization, and data access work when third-party MCP clients interact with Qlik Cloud data – and where Qlik's responsibility ends.

Executive Summary

Qlik's MCP server enables third-party MCP clients – such as Claude Desktop/Cowork/Code, VS Code + GitHub Copilot, Cursor, or ChatGPT – to query Qlik Cloud data using the user's own credentials and permissions.

This document clarifies:

- **Authentication is OAuth 2.0** – the MCP server never accepts unauthenticated connections; all requests are bound to a verified Qlik user identity
- **Data access is governed by Qlik's Section Access controls** – applied at the account level, users can only retrieve data they are already authorized to see
- **MCP is a tool-calling interface, not an AI service** – Qlik exposes governed data access via the MCP protocol; it does not perform AI inference, generate completions, or process prompts on behalf of the client
- **Data returned to the user's LLM client is governed by that provider's policies** – once data leaves the Qlik MCP server boundary, Qlik's controls no longer apply

Key Insight: Qlik's MCP server functions as a governed data access layer. Qlik controls identity, entitlements, and what data is made accessible – equivalent to API-level access to Qlik Cloud. The user's choice of LLM client determines what happens to that data thereafter.

What MCP Is / What MCP Is Not

MCP Is

A protocol-based interface for external tools to call Qlik Cloud APIs

Governed by the same OAuth 2.0 and Section Access controls as any Qlik Cloud interaction

An open interface where the user chooses which LLM client to connect

Architecturally equivalent to a governed API or data export

MCP Is Not

An AI inference service – Qlik does not run models or generate completions

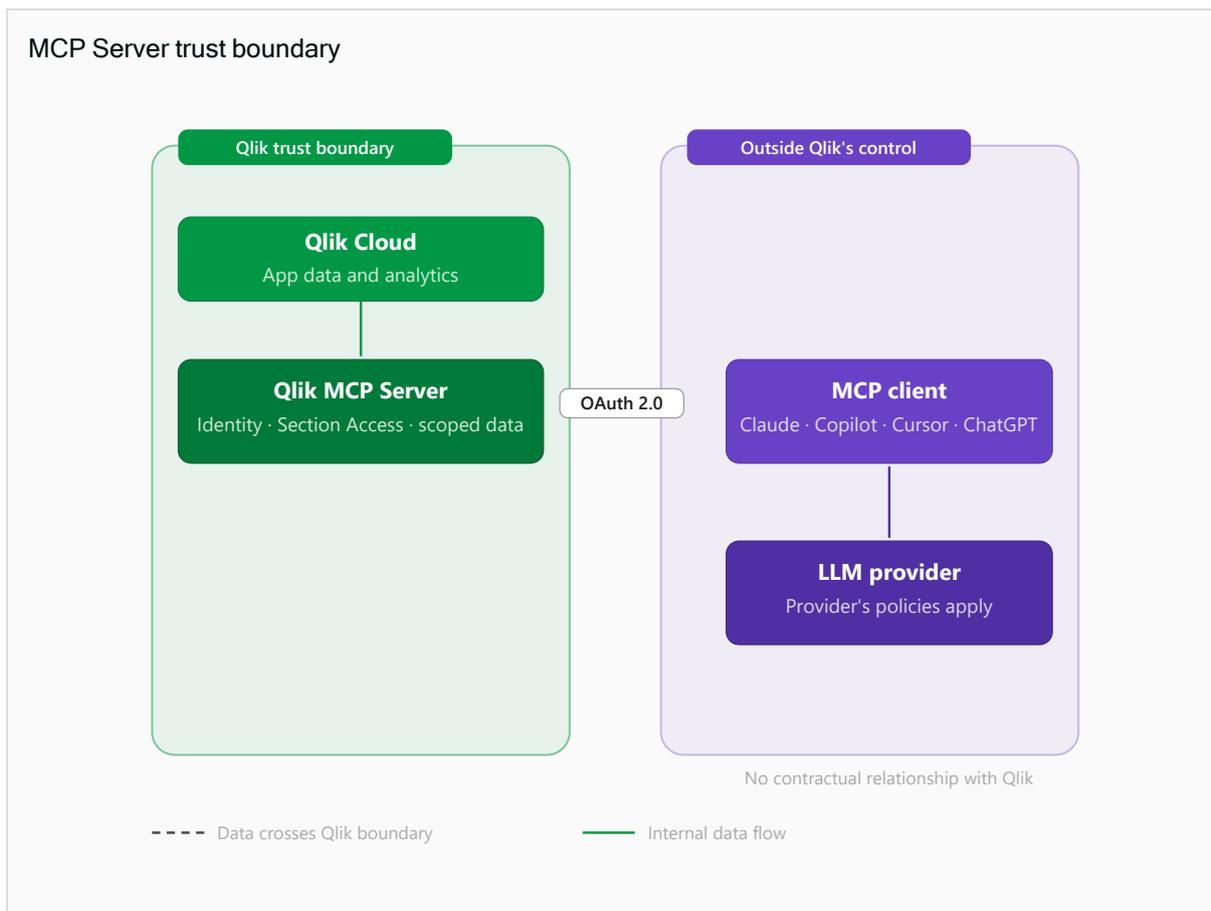
A bypass of existing Qlik security – no privilege escalation path exists

A closed-loop pipeline – data exits Qlik's boundary when returned to the client

A data warehouse or bulk extraction tool – results are scoped, per-request responses

For a detailed technical overview of the MCP standard, see the [Architecture Overview – Model Context Protocol](#).

Qlik's Trust Boundary



The external LLM provider – whether ChatGPT (OpenAI), Claude (Anthropic), Gemini (Google), Cursor, or others – is outside Qlik's trust boundary. Data handling beyond the MCP response is governed by that provider's policies. No contractual relationship exists between Qlik and the LLM provider.

Common Misconceptions vs. Reality

Misconception	Reality
Anyone can query Qlik data via MCP	✗ False. All MCP connections require OAuth 2.0 authentication. Unauthenticated requests are rejected.
MCP bypasses Qlik's data access controls	✗ False. Section Access is enforced at the account level. Users retrieve only data they are authorized to see – identical to their Qlik Cloud access.
Qlik controls what the external LLM does with the data	✗ False. Once data is sent to the user's MCP client, it is governed by the user's chosen LLM provider's policies – it is equivalent to a data export.

Security Controls Deep Dive

1. Authentication – OAuth 2.0

All MCP connections require OAuth 2.0 authorization:

- **User-bound tokens** – access tokens are tied to a specific Qlik user identity; no service account or shared credentials
- **Scoped permissions** – OAuth scopes restrict what operations the MCP client can perform
- **Token expiry and rotation** – short-lived tokens reduce the window for credential misuse
- **No static API keys** – eliminates long-lived credential risks associated with API key models

Implication: A compromised or malicious MCP client can only access data that the authenticated user is authorized to see. There is no privilege escalation path via MCP.

2. Data Entitlements – Section Access

Section Access is Qlik's row- and field-level security framework, enforced at the user account level:

- **Applied before data is returned** – MCP queries are subject to the same Section Access rules as any other Qlik interaction; no bypass mechanism exists
- **Account-level enforcement** – not a client-side filter; enforcement occurs server-side within Qlik Cloud before data leaves the platform
- **Consistent with existing Qlik access** – a user accessing Qlik via MCP sees exactly the same data they would see in the Qlik Cloud browser interface; nothing more

3. The External LLM Boundary

This is the key architectural consideration for any security assessment of MCP.

What Qlik Controls

Control	Coverage
User authentication	✓ OAuth 2.0 – all requests
Data access entitlements	✓ Section Access – enforced server-side
Data in transit (Qlik → MCP client)	✓ TLS encryption
Content logging within Qlik	✓ No content logged

What Qlik Does Not Control

Factor	Responsibility
LLM provider's data handling policies	User's chosen LLM provider
Whether the LLM provider logs request content	User's chosen LLM provider
Whether the LLM provider uses data for model training	User's chosen LLM provider
Data sovereignty once data reaches the LLM provider	User's chosen LLM provider

Implication: The MCP server cannot expose data beyond what the user is already entitled to access within Qlik Cloud.

Analogy: This is equivalent to a user exporting a data extract from Qlik Cloud and uploading it to an external AI tool. Qlik governs what data they are permitted to export (Section Access); the user and their chosen tool's provider govern what happens next.

Enterprises wishing to control which LLM clients their users may connect to Qlik via MCP should address this through internal AI usage policies, LLM provider vetting, and endpoint controls – in the same way they would govern other data export or third-party AI tool usage.

What Gets Stored Where

Data Type	Qlik Cloud (Home Region)	MCP Client / External LLM	Persistence
Raw application data	✓ Stored in Qlik Cloud	⚠ Governed by LLM provider policy	Permanent in home region
MCP query results (data values, measures)	✓ Transient during API response	⚠ Governed by LLM provider policy	Qlik: session-based only
OAuth tokens	✓ Managed per session	⚠ Stored in MCP client	Token lifetime only
Content logs	✗ Not logged by Qlik	⚠ LLM provider policy	N/A for Qlik

MCP Risk Profile

Risk Factor	MCP Coverage
User authentication	✓ OAuth 2.0
Data access entitlements	✓ Section Access
Data stays within Qlik boundary	⚠ Qlik-side only; external LLM is outside Qlik's control
Single compliance framework	⚠ Qlik's framework applies to data at rest and in transit; LLM provider's framework applies thereafter
No model training on data	⚠ LLM provider governs what happens with query-time data
Content logging	✓ Qlik: no content logged; ⚠ LLM provider policy applies for query content
Audit trail	✓ Qlik logs; ⚠ LLM provider audit trail is separate

Summary and Recommendations

Key Takeaways

- **Strong access controls at the boundary:** OAuth 2.0 authentication and Section Access enforcement ensure that the Qlik MCP server cannot expose data beyond a user's existing entitlements.

- **MCP is data access, not AI inference:** The MCP server exposes governed, scoped data to external tools via a standard protocol. It does not run models, generate completions, or process prompts.
- **Honest boundary acknowledgment:** Once data is returned to the user's LLM client, Qlik's controls end. This is architecturally equivalent to data export and should be treated as such in risk assessments.
- **Organizational controls are the appropriate mitigation:** Qlik cannot – and should not – restrict which LLM providers users may connect to via MCP. Enterprises should address external LLM risk through internal AI usage policies and approved provider lists.

For Security Assessments

Disposition	Conditions
✓ Accept this risk if...	Your org permits users to access Qlik data via API or export mechanisms. You trust users to operate within acceptable LLM provider boundaries. Internal AI usage policies govern which LLM providers are approved for use with business data.
⚠ Mitigate this risk by...	Establishing and communicating an approved LLM provider list. Ensuring Section Access rules are correctly configured. Reviewing the data handling and training policies of likely LLM providers. Enabling audit logging within Qlik Cloud to monitor MCP tool usage patterns.
✗ Reject this risk if...	Data sovereignty laws prohibit any transfer of data to non-approved infrastructure. Regulatory requirements mandate all AI processing occur within a specific compliance boundary. Your threat model cannot accept data leaving Qlik's environment under any circumstances.

Configuration option: Organizations may restrict or disable MCP server access at the tenant level. This prevents any external LLM connectivity while retaining full Qlik Cloud native functionality.

Document Version: 2.0 | Last Updated: March 2026

For questions or additional security documentation, please contact your Qlik account team or consult the [Qlik Help Center](#).

For background on the Model Context Protocol standard referenced in this document, see [What is the Model Context Protocol \(MCP\)?](#) and the [MCP Architecture Overview](#).

2.7 Qlik Discovery Agent™

Overview

Discovery Agent is Qlik's AI-powered anomaly detection and monitoring capability built into Qlik Cloud. It continuously scans your Qlik Cloud applications for meaningful changes in your data and delivers prioritized, AI-generated insights directly to users – without requiring manual monitoring, predefined rules, or static thresholds.

Discovery Agent is part of Qlik's broader Agentic AI framework, where assistants answer questions and agents monitor and act. While Qlik Answers responds to questions users ask, Discovery Agent proactively surfaces what users need to know, even when they haven't thought to ask.

Value and Business Impact

Discovery Agent enables organizations to move from reactive reporting to proactive, data-driven action. It enables:

Early detection – Emerging issues are caught before they escalate into costly problems. Teams are informed at the moment something meaningful changes, not after the fact.

Reduced manual work – Repetitive monitoring tasks are eliminated, freeing analyst time for higher-value work. Discovery Agent handles continuous surveillance across all monitored applications automatically.

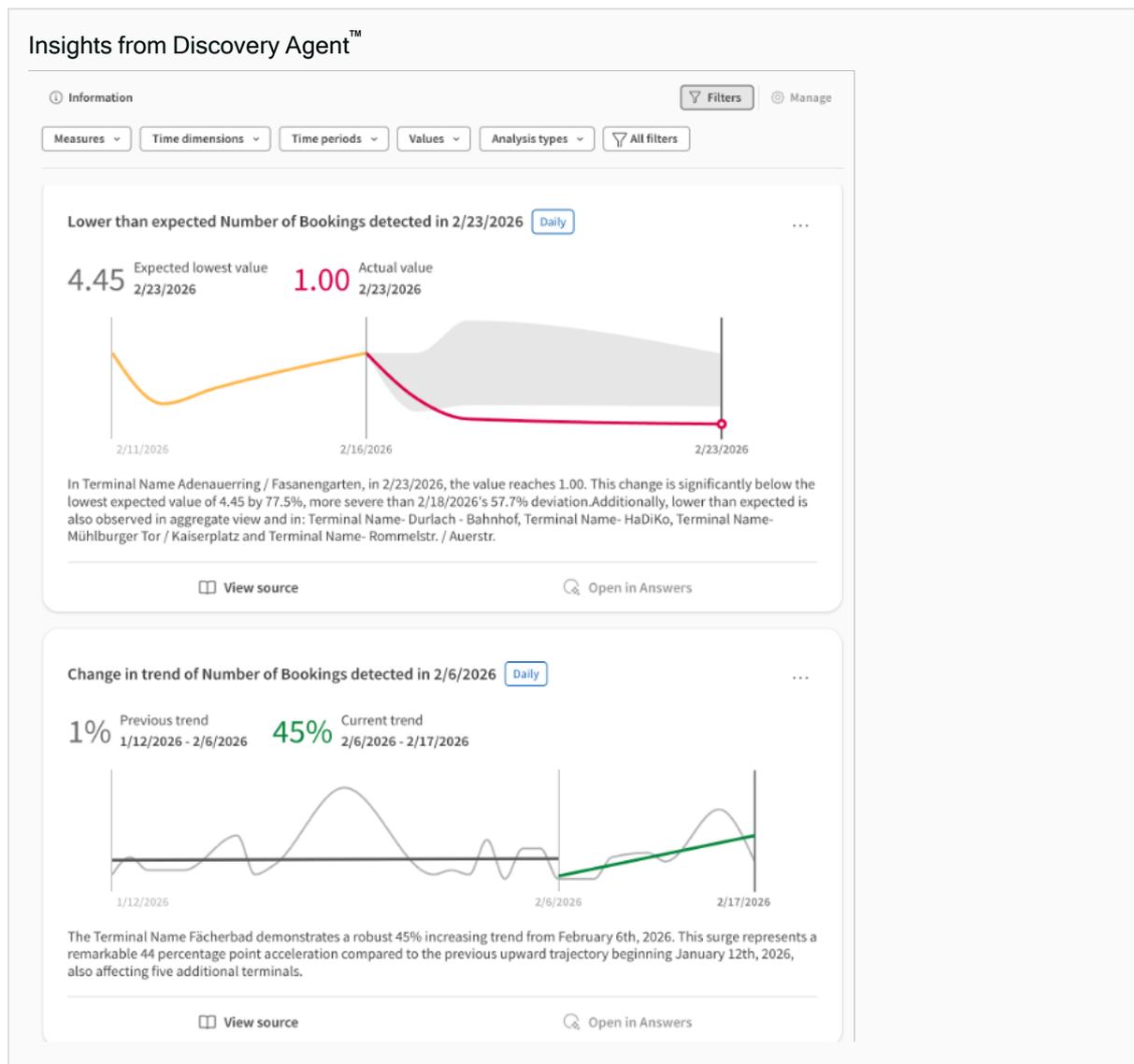
Business clarity – Rather than presenting every data point equally, Discovery Agent surfaces the changes that truly matter, ranked by significance and presented in plain language.

Faster decision-making – Teams across the organization can act on insights immediately, regardless of their familiarity with the underlying Qlik applications or data models.

Operational responsiveness – Automated, proactive alerting means organizations can react faster to changing conditions, protecting performance and uncovering new opportunities sooner.

What Discovery Agent Does

Discovery Agent continuously monitors your Qlik Cloud applications in the background. It automatically detects meaningful changes in your data, evaluates their significance, and delivers clear, prioritized insights to users in plain language – with no configuration of thresholds or alert conditions required. These insights are provided to users through their own personal Feed in the analytics hub.



How Discovery Agent Prioritizes What Matters

A core challenge in any monitoring system is distinguishing signal from noise. Discovery Agent addresses this through three interconnected mechanisms:

Dynamic baselines - Rather than relying on static thresholds, Discovery Agent uses forecasting to establish an expected range for each metric. Deviations from this forecast trigger evaluation, and there are no rules to manually maintain as the business changes.

Intelligent thresholding - Discovery Agent evaluates trends, spikes, and shifts against the dynamic baseline and filters out changes that fall within normal variation. Only anomalies that are genuinely significant are surfaced.

Adaptive sensitivity - Discovery Agent automatically adjusts its sensitivity based on the volume of insights it is generating. If too many insights are being delivered, sensitivity is reduced to focus on the most significant changes. If too few are being generated, sensitivity increases to ensure meaningful signals are not missed. This continuous self-tuning keeps the insight feed balanced and actionable without manual intervention.

Where Discovery Agent Fits in Qlik Cloud

Discovery Agent is part of Qlik's Agentic AI framework within Qlik Cloud Analytics. Within this framework, assistants – such as Qlik Answers – respond to questions that users ask. Agents, by contrast, monitor conditions and act autonomously on behalf of users.

Discovery Agent occupies the monitoring role within this framework: it watches your data continuously, determines what is significant, and brings that information to users proactively. This complements Qlik Answers, which remains the appropriate tool when users want to ask specific questions of their data. Together, these capabilities represent a shift from a model where analytics requires active user engagement to one where the platform brings relevant insight to users as conditions change.

Security architecture

Access Control

Discovery Agent fully respects application level access controls. The applications a user with see insights from are restricted to those applications they have permissions to see through the hub. Support for row-level section access is planned for a future release.

Cross-Region Inference

Discovery Agent requires cross-region inference to function. This means that when Discovery Agent processes AI requests, data may be temporarily sent to a different AWS region for inference processing. This is consistent with how other AI services in Qlik Cloud operate and is subject to the same security controls described in the Qlik Answers security architecture documentation. All processing occurs within the AWS private network, no data is persisted in the inference region, and customer data is never used to train AI models. Customers must opt in to cross-region data processing to enable Discovery Agent. This setting is disabled by default for Qlik Cloud tenants.

Discovery agent follows the same process as is used for access to structured data with Qlik Answers. This process is described in detail in the section [Qlik Answers - Security Architecture and Controls](#) . For the latest information on inference processing regions by product feature and home region, see [Enabling cross-region data processing](#) on Qlik Help.

3 About Qlik Evaluation Guides

The content provided herein is provided for informational purposes. Due to Qlik Cloud's continuous release process, at times the content herein may differ from actual platform functionality. Please refer to [Qlik Cloud Help](#) for the product documentation for Qlik Cloud.

Any statement about future plans or intentions for the Qlik Cloud platform contained herein is not a commitment to deliver those features or functionalities, as the development, release, and timing of any features or functionality described for Qlik's products remain at our sole discretion.

For additional information regarding Qlik Cloud, please see [Qlik Cloud](#) or contact your Qlik representative.

3.1 Document history

This content has been developed to assist customers and prospective customers to understand and evaluate the Qlik Cloud platform and its related services. Traditionally this content has been published in document format only as a PDF; however, it is now primarily published as web content with PDF files available if required.

Over its history, this content has been known by the following names:

- Qlik Technical papers
- Qlik White papers
- Qlik Technical overview

This documentation supersedes the above documents.

3.2 Changelog

The PDF documents are generated from the evaluation guides at [Qlik Help](#) . The changelog for this evaluation guide is shown below.

Changelog – AI in Qlik Cloud

March 2026

First update

- Add discovery Agent
- remove CRIS requirement for MCP Server

Initial release

- Add MCP
- Add Platform
- Add Answers
- Add Security content



About Qlik

Qlik's vision is a data-literate world, where everyone can use data and analytics to improve decision-making and solve their most challenging problems. Our cloud-based Qlik Active Intelligence Platform® delivers end-to-end, real-time data integration and analytics cloud solutions to close the gaps between data, insights and action. By transforming data into Active Intelligence, businesses can drive better decisions, improve revenue and profitability, and optimize customer relationships. Qlik does business in more than 100 countries and serves over 38,000 active customers around the world.

[qlik.com](https://www.qlik.com)