Security in Azure Cognitive Search - overview

08/01/2020 • 10 minutes to read • +3

In this article

This article describes the key security features in Azure Cognitive Search that can protect content and operations.

- At the storage layer, encryption-at-rest is built in for all service-managed content saved to disk, including indexes, synonym maps, and the definitions of indexers, data sources, and skillsets. Azure Cognitive Search also supports the addition of customer-managed keys (CMK) for supplemental encryption of indexed content. For services created after August 1 2020, CMK encryption extends to data on temporary disks, for full double encryption of indexed content.

- Inbound security protects the search service endpoint at increasing levels of security: from API keys on the request, to inbound rules in the firewall, to private endpoints that fully shield your service from the public internet.

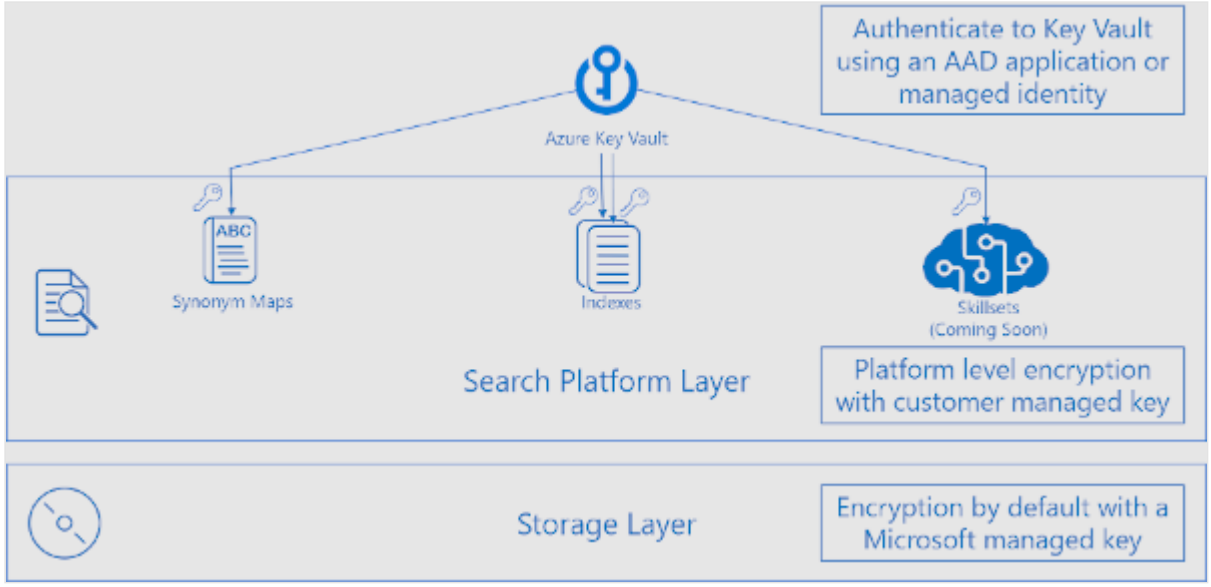
- Outbound security applies to indexers that pull content from external sources. For outbound requests, set up a managed identity to make search a trusted service when accessing data from Azure Storage, Azure SQL, Cosmos DB, or other Azure data sources. A managed identity is a substitute for credentials or access keys on the connection. Outbound security is not covered in this article. For more information about this capability, see Connect to a data source using a managed identity.

Watch this fast-paced video for an overview of the security architecture and each feature category.

28:09

Encrypted transmissions and storage

In Azure Cognitive Search, encryption starts with connections and transmissions, and extends to content stored on disk. For search services on the public internet, Azure Cognitive Search listens on HTTPS port 443. All client-to-service connections use TLS 1.2 encryption. Earlier versions (1.0 or 1.1) are not supported.



For data handled internally by the search service, the following table describes the data encryption models. Some features, such as knowledge store, incremental enrichment, and indexer-based indexing, read from or write to data structures in other Azure Services. Those services have their own levels of encryption support separate from Azure Cognitive Search.

Model	Keys	Requirements	Restrictions	Applies to
server-side encryption	Microsoft-managed keys	None (built-in)	None, available on all tiers, in all regions, for content created after January 24 2018.	Content (indexes and synonym maps) and definitions (indexers, data sources, skillsets)
server-side encryption	customer-managed keys	Azure Key Vault	Available on billable tiers, in all regions, for content created after January 2019.	Content (indexes and synonym maps) on data disks
server-side double encryption	customer-managed keys	Azure Key Vault	Available on billable tiers, in selected regions, on search services after August 1 2020.	Content (indexes and synonym maps) on data disks and temporary disks

Service-managed keys

Service-managed encryption is a Microsoft-internal operation, based on Azure Storage Service Encryption, using 256-bit AES encryption. It occurs automatically on all indexing, including on incremental updates to indexes that are not fully encrypted (created before January 2018).

Customer-managed keys (CMK)

Customer-managed keys require an additional billable service, Azure Key Vault, which can be in a different region, but under the same subscription, as Azure Cognitive Search. Enabling CMK encryption will increase index size and degrade query performance. Based on observations to date, you can expect to see an increase of 30%-60% in query times, although actual performance will vary depending on the index definition and types of queries. Because of this performance impact, we recommend that you only enable this feature on indexes that really require it. For more information, see Configure customer-managed encryption keys in Azure Cognitive Search.

Double encryption

In Azure Cognitive Search, double encryption is an extension of CMK. It is understood to be two-fold encryption (once by CMK, and again by service-managed keys), and comprehensive in scope, encompassing long term storage that is written to a data disk, and short term storage written to temporary disks. The difference between CMK before August 1 2020 and after, and what makes CMK a double encryption feature in Azure Cognitive Search, is the additional encryption of data-at-rest on temporary disks.

Double encryption is currently available on new services that are created in these regions after August 1:

- West US 2
- East US
- South Central US
- US Gov Virginia

- US Gov Arizona

Inbound security and endpoint protection

Inbound security features protect the search service endpoint through increasing levels of security and complexity. First, all requests require an API key for authenticated access. Second, you can optionally set firewall rules that limit access to specific IP addresses. For advanced protection, a third option is to enable Azure Private Link to shield your service endpoint from all internet traffic.

Public access using API keys

By default, a search service is accessed through the public cloud, using key-based authentication for admin or query access to the search service endpoint. An api-key is a string composed of randomly generated numbers and letters. The type of key (admin or query) determines the level of access. Submission of a valid key is considered proof the request originates from a trusted entity.

There are two levels of access to your search service, enabled by the following API keys:

- Admin key (allows read-write access for create-read-update-delete operations on the search service)
- Query key (allows read-only access to the documents collection of an index)

Admin keys are created when the service is provisioned. There are two admin keys, designated as *primary* and *secondary* to keep them straight, but in fact they are interchangeable. Each service has two admin keys so that you can roll one over without losing access to your service. You can regenerate admin key periodically per Azure security best practices, but you cannot add to the total admin key count. There are a maximum of two admin keys per search service.

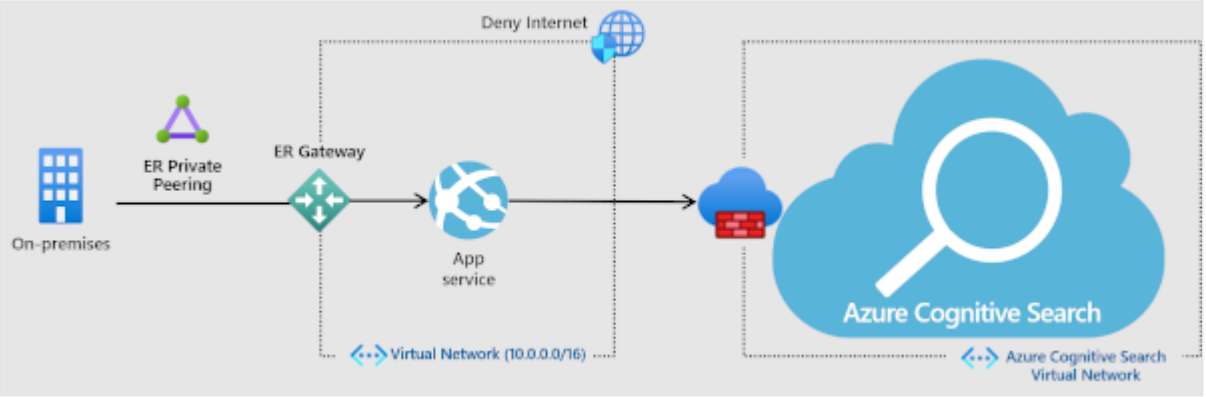
Query keys are created as-needed and are designed for client applications that issue queries. You can create up to 50 query keys. In application code, you specify the search URL and a query api-key to allow read-only access to the documents

collection of a specific index. Together, the endpoint, an api-key for read-only access, and a target index define the scope and access level of the connection from your client application.

Authentication is required on each request, where each request is composed of a mandatory key, an operation, and an object. When chained together, the two permission levels (full or read-only) plus the context (for example, a query operation on an index) are sufficient for providing full-spectrum security on service operations. For more information about keys, see Create and manage api-keys.

IP-restricted access

To further control access to your search service, you can create inbound firewall rules that allow access to specific IP address or a range of IP addresses. All client connections must be made through an allowed IP address, or the connection is denied.



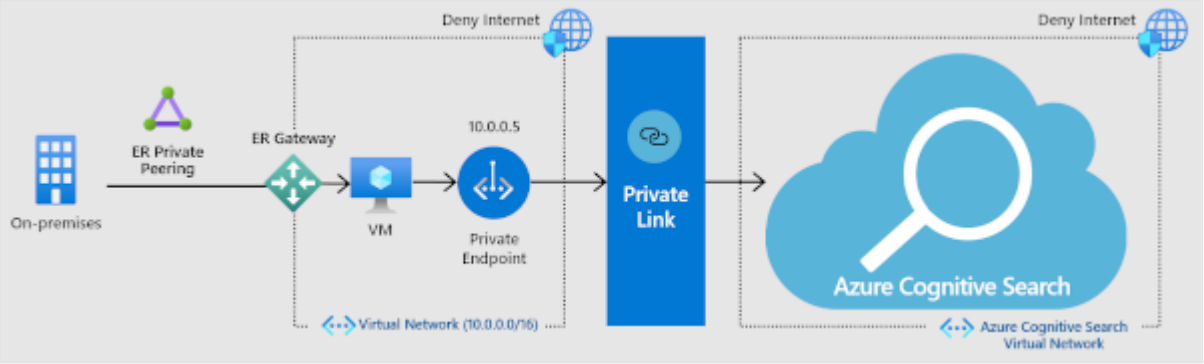
You can use the portal to configure inbound access.

Alternatively, you can use the management REST APIs. Starting with API version 2020-03-13, with the IpRule parameter, you can restrict access to your service by identifying IP addresses, individually or in a range, that you want to grant access to your search service.

Private endpoint (no Internet traffic)

A Private Endpoint for Azure Cognitive Search allows a client on a virtual network to securely access data in a search index over a Private Link.

The private endpoint uses an IP address from the virtual network address space for connections to your search service. Network traffic between the client and the search service traverses over the virtual network and a private link on the Microsoft backbone network, eliminating exposure from the public internet. A VNET allows for secure communication among resources, with your on-premises network as well as the Internet.



While this solution is the most secure, using additional services is an added cost so be sure you have a clear understanding of the benefits before diving in. or more information about costs, see the pricing page. For more information about how these components work together, watch the video at the top of this article. Coverage of the private endpoint option starts at 5:48 into the video. For instructions on how to set up the endpoint, see Create a Private Endpoint for Azure Cognitive Search.

Index access

In Azure Cognitive Search, an individual index is not a securable object. Instead, access to an index is determined at the service layer (read or write access to the service), along with the context of an operation.

For end-user access, you can structure query requests to connect using a query key, which makes any request read-only, and include the specific index used by your app. In a query request, there is no concept of joining indexes or accessing multiple indexes simultaneously so all requests target a single index by definition. As such, construction of the query request itself (a key plus a single target index) defines the security boundary.

Administrator and developer access to indexes is undifferentiated: both need write access to create, delete, and update objects managed by the service. Anyone with an admin key to your service can read, modify, or delete any index in the same service. For protection against accidental or malicious deletion of indexes, your in-house source control for code assets is the remedy for reversing an unwanted index deletion or modification. Azure Cognitive Search has failover within the cluster to ensure availability, but it does not store or execute your proprietary code used to create or load indexes.

For multitenancy solutions requiring security boundaries at the index level, such solutions typically include a middle tier, which customers use to handle index isolation. For more information about the multitenant use case, see Design patterns for multitenant SaaS applications and Azure Cognitive Search.

User access

How a user accesses an index and other objects is determined by the type of API key on the request. Most developers create and assign *query keys* for client-side search requests. A query key grants read-only access to searchable content within the index.

If you require granular, per-user control over search results, you can build security filters on your queries, returning documents associated with a given security identity. Instead of predefined roles and role assignments, identity-based access control is implemented as a *filter* that trims search results of documents and content based on identities. The following table describes two approaches for trimming search results of unauthorized content.

Approach	Description
Security trimming based on identity filters	Documents the basic workflow for implementing user identity access control. It covers adding security identifiers to an index, and then explains filtering against that field to trim results of prohibited content.
Security trimming based on Azure Active Directory identities	This article expands on the previous article, providing steps for retrieving identities from Azure Active Directory (Azure AD), one of the free services in the Azure cloud platform.

Administrative rights

Azure role-based access control (Azure RBAC) is an authorization system built on Azure Resource Manager for provisioning of Azure resources. In Azure Cognitive Search, Resource Manager is used to create or delete the service, manage API keys, and scale the service. As such, Azure role assignments will determine who can perform those tasks, regardless of whether they are using the portal, PowerShell, or the Management REST APIs.

In contrast, admin rights over content hosted on the service, such as the ability to create or delete an index, is conferred through API keys as described in the previous section.

> **Tip**
>
> Using Azure-wide mechanisms, you can lock a subscription or resource to prevent accidental or unauthorized deletion of your search service by users with admin rights. For more information, see **Lock resources to prevent unexpected deletion**.

Certifications and compliance

Azure Cognitive Search has been certified compliant for multiple global, regional, and industry-specific standards for both the public cloud and Azure Government. For the complete list, download the **Microsoft Azure Compliance Offerings** whitepaper from the official Audit reports page.

For compliance, you can use Azure Policy to implement the high-security best practices of Azure Security Benchmark. Azure Security Benchmark is a collection of security recommendations, codified into security controls that map to key actions you should take to mitigate threats to services and data. There are currently 11 security controls, including Network Security, Logging and Monitoring, and Data Protection to name a few.

Azure Policy is a capability built into Azure that helps you manage compliance for multiple standards, including those of Azure Security Benchmark. For well-known benchmarks, Azure Policy provides built-in definitions that provide both criteria as

well as an actionable response that addresses non-compliance.

For Azure Cognitive Search, there is currently one built-in definition. It is for diagnostic logging. With this built-in, you can assign a policy that identifies any search service that is missing diagnostic logging, and then turns it on. For more information, see Azure Policy Regulatory Compliance controls for Azure Cognitive Search.

See also

- Azure security fundamentals
- Azure Security
- Azure Security Center