Azure AI Search Indexer: Definition and Functionality

In the context of Azure AI Search, an **indexer** is a **crawler that extracts textual data from cloud data sources and populates a search index using field-to-field mappings**. It serves as an automated data ingestion mechanism that connects external Azure data sources to your search index without requiring custom code for data loading.[1]

Core Functionality

An indexer operates using what Microsoft refers to as a **"pull model"** where the search service pulls data from external sources rather than requiring you to push data to the index. This automation significantly reduces the complexity and amount of code needed for data ingestion when working with supported data sources.[2] [1]

Key Components and Architecture

An indexer requires three fundamental components:[2]

Component	Description
Data Source	Specifies the connection to external data (origin)
Target Index	Points to the destination search index
Name	Uniquely identifies the indexer in the indexer collection

Supported Workflows

Indexers support two primary workflows:[3] [2]

Raw Content Indexing

- Extracts strings and metadata from textual content for full-text search scenarios
- Extracts raw vector content for vector search (from sources like Azure SQL Database or Azure Cosmos DB)
- Processes only existing content from your data sources

Skills-based Indexing (AI Enrichment)

- Extends indexing through built-in or custom skills that create new searchable content
- Enables AI processing such as OCR on images, language detection, entity recognition, and key phrase extraction
- Generates new content that doesn't exist in your external data source through skillsets[1]

Supported Data Sources

Azure AI Search indexers work with various Azure data sources:[4]

Data Source	Capabilities
Azure Blob Storage	Extract blob metadata and content with automatic change detection
Azure SQL Database	Extract field values from tables/views with change tracking support
Azure Cosmos DB	Connect through SQL API to extract items from containers
Azure Table Storage	Extract rows from Azure Tables
Azure Data Lake Storage Gen2	Extract content from hierarchical directory structures

Scheduling and Execution

Indexers can run in multiple modes:[5]

- **On-demand execution**: Run immediately when created (default behavior)
- **Scheduled execution**: Run at regular intervals (minimum 5 minutes, maximum 24 hours)
- **Manual execution**: Triggered as needed

The search service runs **one indexer job per search unit**. If you need concurrent processing, ensure your service has sufficient replicas, as indexers don't run in the background and may impact query performance under heavy load.[1]

Change Detection and Incremental Updates

Most Azure data sources support built-in change detection, allowing indexers to automatically identify and process only new, updated, or deleted content during subsequent runs. This enables efficient incremental indexing rather than full reprocessing of all data.[5]

Indexers represent a powerful automation tool in Azure AI Search that streamlines the process of keeping your search indexes synchronized with external data sources while supporting both simple data extraction and complex AI-powered content enrichment scenarios.

⁂

1. https://learn.microsoft.com/en-us/azure/search/search-indexer-overview
2. https://learn.microsoft.com/en-us/azure/search/search-how-to-create-indexers
3. https://docs.azure.cn/en-us/search/search-how-to-create-indexers

4. https://learn.microsoft.com/en-us/azure/search/search-data-sources-gallery
5. https://learn.microsoft.com/en-us/azure/search/search-howto-schedule-indexers
6. https://blog.novanet.no/writing-an-azure-cognitive-search-indexer/
7. https://learn.microsoft.com/en-us/azure/search/search-what-is-an-index
8. https://learn.microsoft.com/es-es/azure/search/search-indexer-overview
9. https://www.thinktecture.com/azure/azure-search-index/
10. https://learn.microsoft.com/en-us/azure/search/search-what-is-azure-search
11. https://learn.microsoft.com/en-us/azure/search/search-how-to-index-sql-database
12. https://docs.azure.cn/en-us/search/search-file-storage-integration
13. https://learn.microsoft.com/en-us/azure/search/tutorial-skillset
14. https://stackoverflow.com/questions/74295203/custom-schedule-for-azure-search-indexer
15. https://learn.microsoft.com/en-us/azure/search/tutorial-multiple-data-sources
16. https://bradmcnaughton.com/custom-skill-for-azure-ai-search
17. https://docs.azure.cn/en-us/search/search-howto-run-reset-indexers
18. https://www.sharepointeurope.com/keeping-azure-search-index-date-azure-functions/
19. https://learn.microsoft.com/en-us/rest/api/searchservice/create-data-source
20. https://learn.microsoft.com/en-us/shows/learn-live/microsoft-learn-ai-skills-challenge-ep09-create-an-azure-ai-search-solution
21. https://docs.azure.cn/en-us/search/search-how-to-index-sql-database
22. https://learn.microsoft.com/en-us/rest/api/searchservice/create-indexer
23. https://learn.microsoft.com/en-us/azure/search/search-howto-indexing-azure-blob-storage