Microsoft

AI-900 Azure AI Fundamentals

N'oubliez pas de vous connecter avec vos noms et prénoms

Pensez bien à couper vos micros !!

Session 4

Détecteur d'anomalies et Recherche cognitive Azure

La session commence à 12h00

Notification GDPR pour ce webinar Teams

*En vous connectant à cette session par le biais de **Microsoft Teams**, votre nom, e-mail, numéro de téléphone et/ou titre peuvent être vus par les autres participants.*

Microsoft

DISCLAIMER

This presentation features Microsoft pre-release product or features which may be substantially modified without notice before commercial release. It represents a product view and is not indicative of final licensing of individual features. This presentation does not provide you with any legal rights to any intellectual property in any Microsoft product.

MICROSOFT MAKES NO WARRANTIES, EXPRESS OR IMPLIED, WITH RESPECT TO THE INFORMATION PROVIDED.

Sessions de l’académie AI-900

Descriptif des sessions
Session 1 : Lancement et Introduction au machine learning
Session 2 : Vision par ordinateur
Session 3 : Traitement automatique du langage naturel (NLP)
Session 4 : Détecteur d’anomalies et Recherche cognitive Azure
Session 5 : Klaxoon

Pour obtenir le support de cette présentation

Présentations AI-900 Academy

(Le lien est mis dans le chat de Teams)

Objectifs pédagogiques

Vous allez apprendre les concepts suivants :

- Detecteur d'anomalie
  - Ce qu'est le detecteur d'anomaly
  - Comment fonctionne t il?
  - Son cas d'usage
- Recherche cognitive Azure
  - Recherche cognitive , c'est quoi?
  - Comment ça marche?
  - Comment creer un index et l'enrichir?

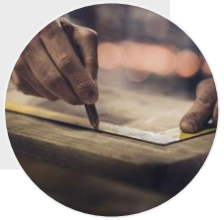
Détecteur d’anomalie

Détecteur d'anomalie ?

Applications pour la detection d’anomalie

- Surveillance de la pression sanguine

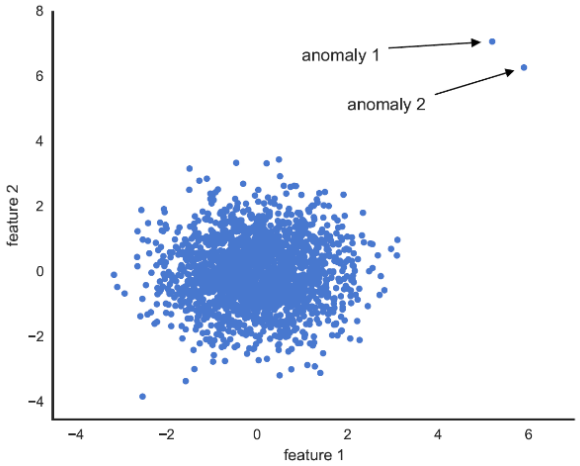
- évaluation du temps moyen entre défaillances des produits matériels



- comparaison d’un mois à l’autre des dépenses liées au coût des produits

Comment fonctionne t il?

Le service Détecteur d'anomalies identifie les anomalies qui existent en dehors de l'étendue d'une limite.



upperBoundary = expectedValue + (100 – marginScale) × upperMargin

Format de données

- Le service Détecteur d'anomalies accepte les données au format JSON
- Les principaux aspects des données envoyées sont
  - la granularité,
  - un horodateur
  - la valeur enregistrée pour cet horodateur.

Le service prend en charge au maximum 8 640 points de données


```
{
    "granularity": "hourly",
    "series": [
      {
        "timestamp": "2021-03-01T01:00:00Z",
        "value": -10.56
      },
      {
        "timestamp": "2021-03-02T02:00:00Z",
        "value": -8.30
      },
      {
        "timestamp": "2021-03-02T03:00:00Z",
        "value": -10.30
      },
      {
        "timestamp": "2021-03-02T04:00:00Z",
        "value": 5.95
      },
    ]
}
```

Type de detection

Par lots (Batch)

- algorithme à toute une série de données en même temps
- Utilisez votre série chronologique pour détecter d'éventuelles anomalies dans l'ensemble de vos données

Bonne utilisation

- Données de série chronologique à tendance stable affichant des pics ou des creux occasionnels
- Données de série chronologique saisonnière affichant des anomalies occasionnelles

Temps reel

utilise des données de streaming en comparant des points de données vus précédemment au dernier point de données pour déterminer si votre dernier est une anomalie

utile pour analyser les exigences de stockage critiques qui doivent être traitées immédiatement

Demo

Creation du service

Exemple

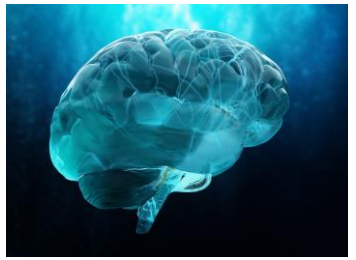
Recherche Cognitive Azure

Recherche Cognitive… c'est quoi ???

La recherche… c'est simple on dispose d'un index.

Quid des informations qui ne sont pas dans l'index?
Document manuscript, les images….

C'est le terrain de jeu des outils de "knowledge mining



Ou les extracteurs d'information

C'est là que se trouve
Azure Cognitive Search

Azure Cognitive Search: En résumé

Platform as a Service (PaaS) recherche de niveau entreprise

Mise en œuvre sans effort

- Ingérer des données structurées et non structurées + extraire facilement des magasins de données Azure
- Configuration rapide, aucun codage requis
- Évolutivité et maintenance faciles
- Fonctionnalités de recherche riches, prêtes à l'emploi

Intégration facile

- Services de vision intégrés : OCR, analyse d'images, extraction de balises
- Services linguistiques intégrés : extraction de phrases clés, détection de langue, analyse des sentiments, reconnaissance d'entités
- Orchestrez la manipulation de texte, personnalisez les compétences et créez un magasin de connaissances pour des données enrichies
- Offres de conformité de premier plan et intégration avec Azure Active Directory pour un accès géré par l'utilisateur

Facilement personnalisable

- Intégrez vos propres données pour créer des modèles et des classificateurs personnalisés
- Étendez-vous à votre secteur d'activité et à votre entreprise spécifiques
- Étendez-vous à votre secteur d'activité et à votre entreprise spécifiques

Recherche sémantique

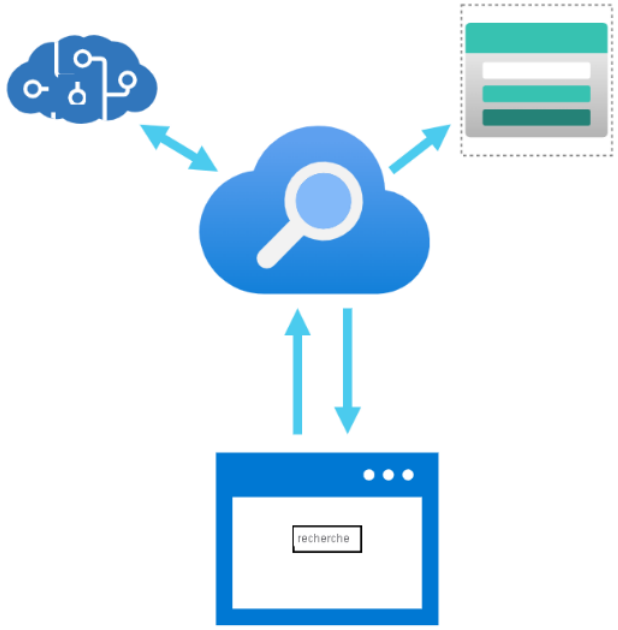
- Exploitez les modèles DNN de pointe de Bing
- Modèles formés sur des billons de requêtes de recherche utilisateur et de clics couvrant plusieurs domaines

Alors c’est quoi le “knowledge mining”?

Recherche Cognitive Azure

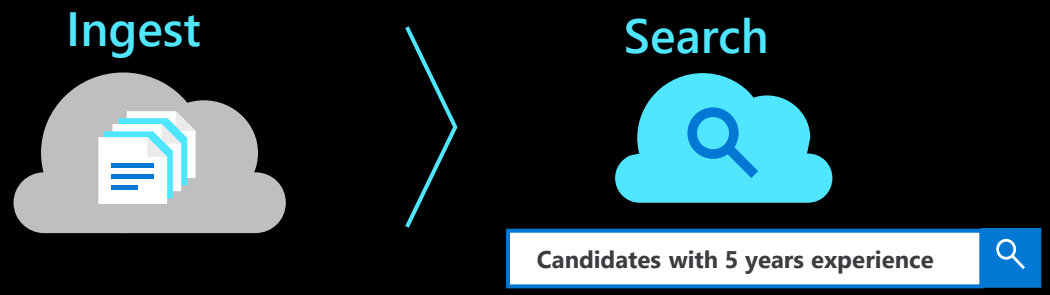
- Fournit l'infrastructure et les outils nécessaires pour créer des solutions de recherche
- PaaS (Platform as a Service)

- Fonctionnalités:
    - Données issues de n'importe quelle source
    - Recherche en texte intégral et analyse
    - Recherche basée sur l'intelligence artificielle
    - Multilingue
    - Prise en charge de la géolocalisation
    - Expérience utilisateur configurable



L’IA peut aider à rendre la recherche intelligente

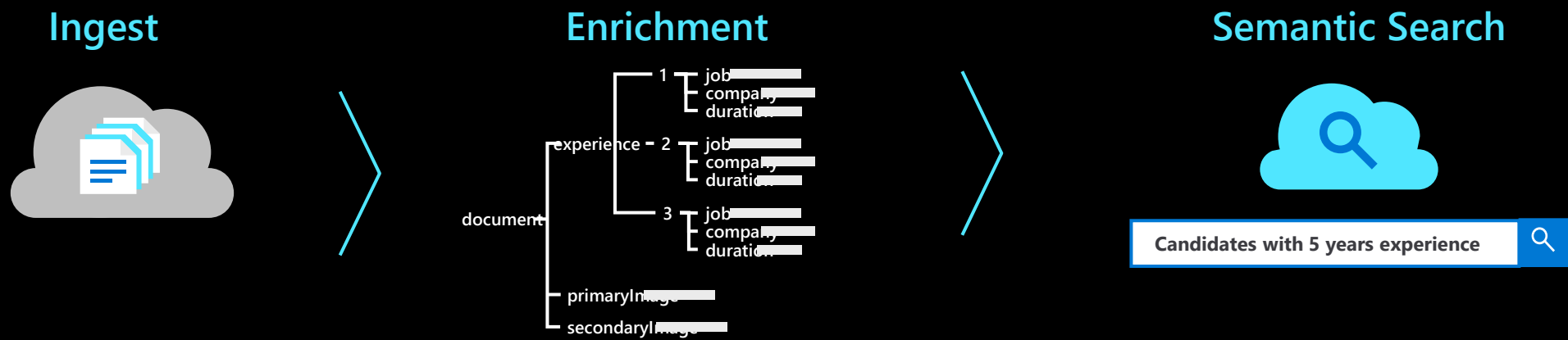
Méthodes de recherche traditionnelles



1. Use AI for Knowledge Mining
Apply NLP, OCR, language detection, etc. to “mine” documents and make them structured

2. Apply AI to understand user intent
Use semantic search to improve the quality of search results

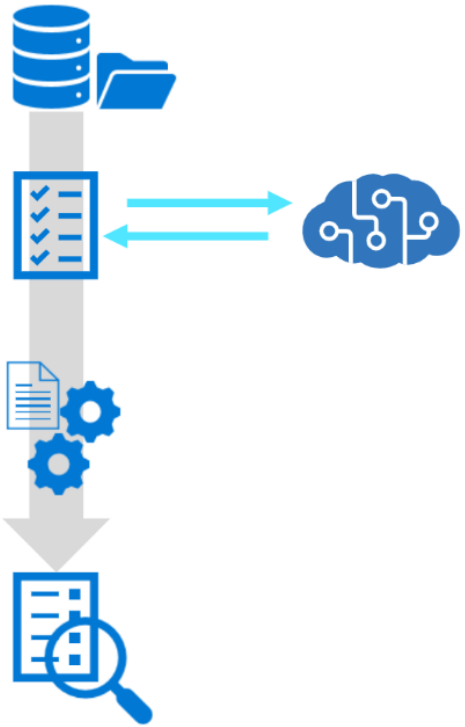
Recherche basée sur l’IA



Comment ça marche?

- source de données
  - stockage Azure
  - Azure SQL DB
  - CosmosDB
  - Format: JSON

  - Indexeur ingere les données automatiquement, et enrichissent
  - les données à l'aide de l'IA (aka les services cognitifs)
- Assure la connection et la serialization



Enrichissement de l'IA avec la recherche cognitive

Built-in skills in Azure Cognitive Search are based on pre-trained machine learning models in Cognitive Services APIs and fall into two categories:

Natural Language Processing
Based on Text Analytics API

Image Processing
Based on Computer Vision API

Customers can also integrate with other Azure Cognitive Services or build custom skills and models:

(Examples of common service combinations)

Azure Form Recognizer
Extract key-value pairs from customer/patient intake forms or invoices

Azure Speech to Text, Azure Bot Service
Use speech to text for analyzing customer calls and surface suggested answers with Azure Bot Service to automate field service

Azure Video Analyzer
Analyze and search video content (e.g., CCTV footage, focus group recording, in-store behavior, media clips, etc.)

Bing Web Search APIs
Search the web for latest details on detected entities in documents (people, places, cases, regulations, etc.)

Build Custom Skills
Extend AI enrichment by building custom skills or training Azure Machine Learning models

Compétences intégrées (liste complète)


Custom entity lookup


PII masking

Brand detection


Document extraction


Azure Machine Learning


Key Phrase extraction


Location entity extraction


Sentiment analysis


Text translation


Organization entity extraction


Person entity extraction


Language detection


PII extraction


Face detection


Celebrity recognition


Tag extraction


Object detection


Custom skills


Landmark detection


Printed text recognition

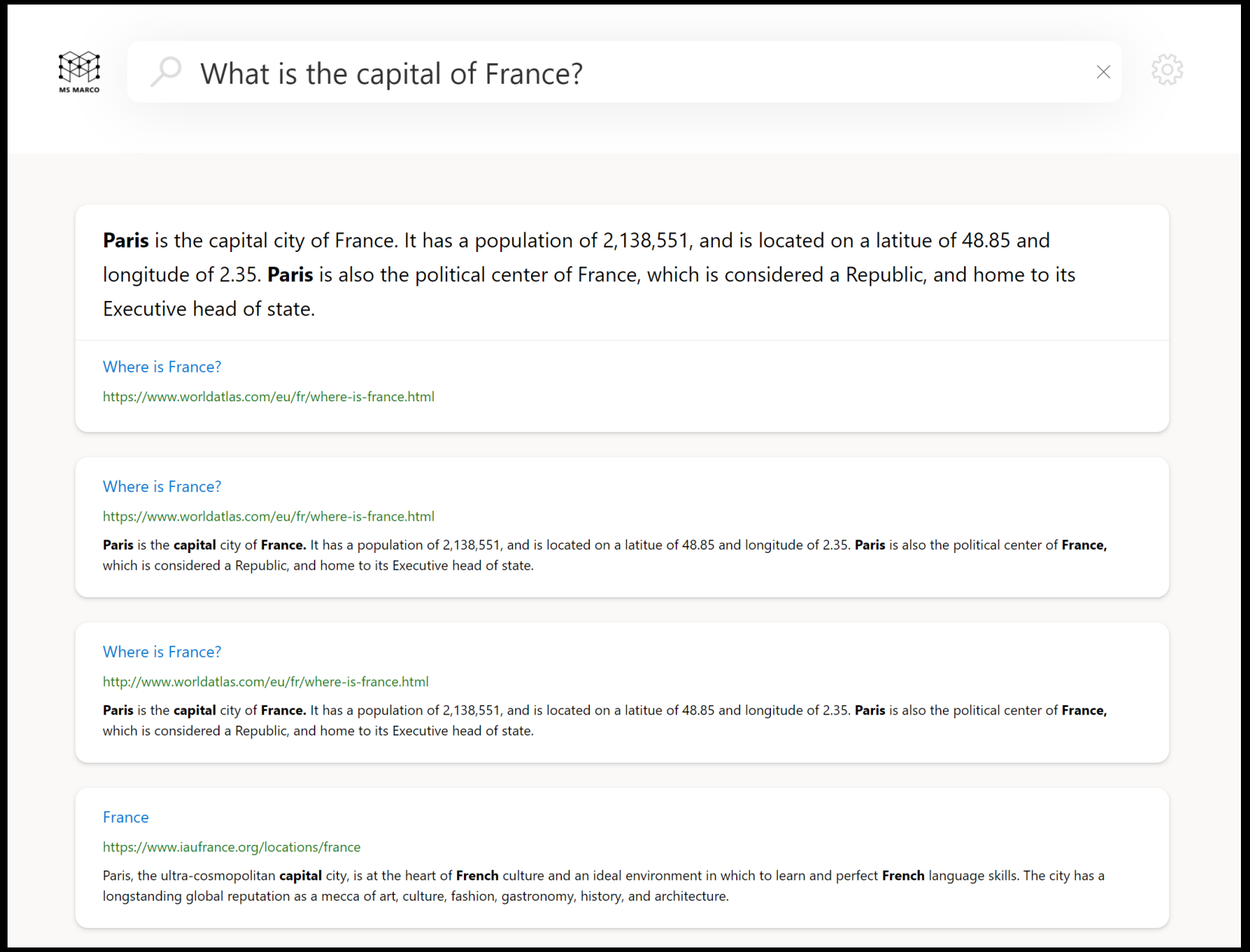
Semantic Search

Semantic Ranking

Semantic Answers

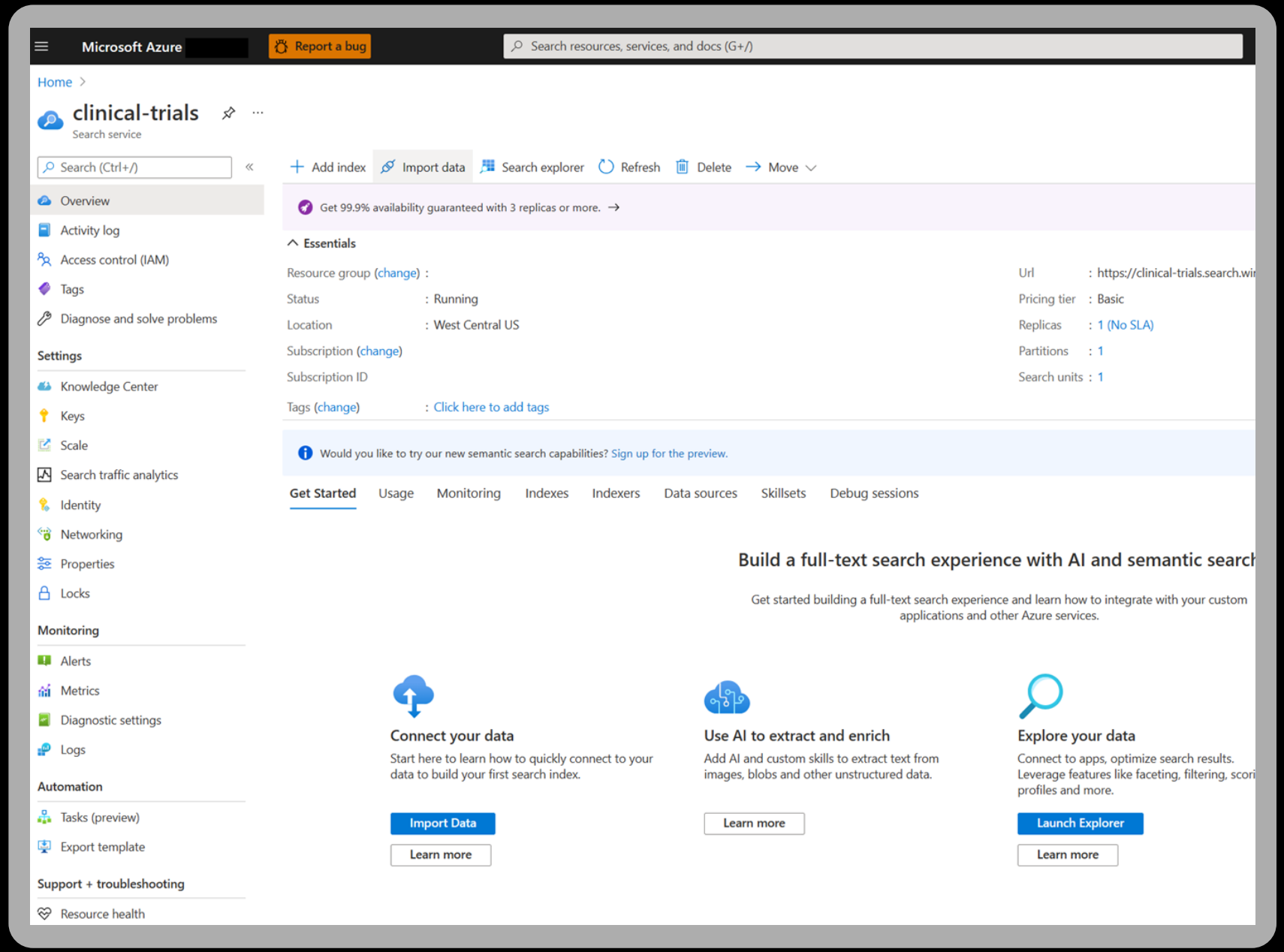
Semantic Captions

Spell Check



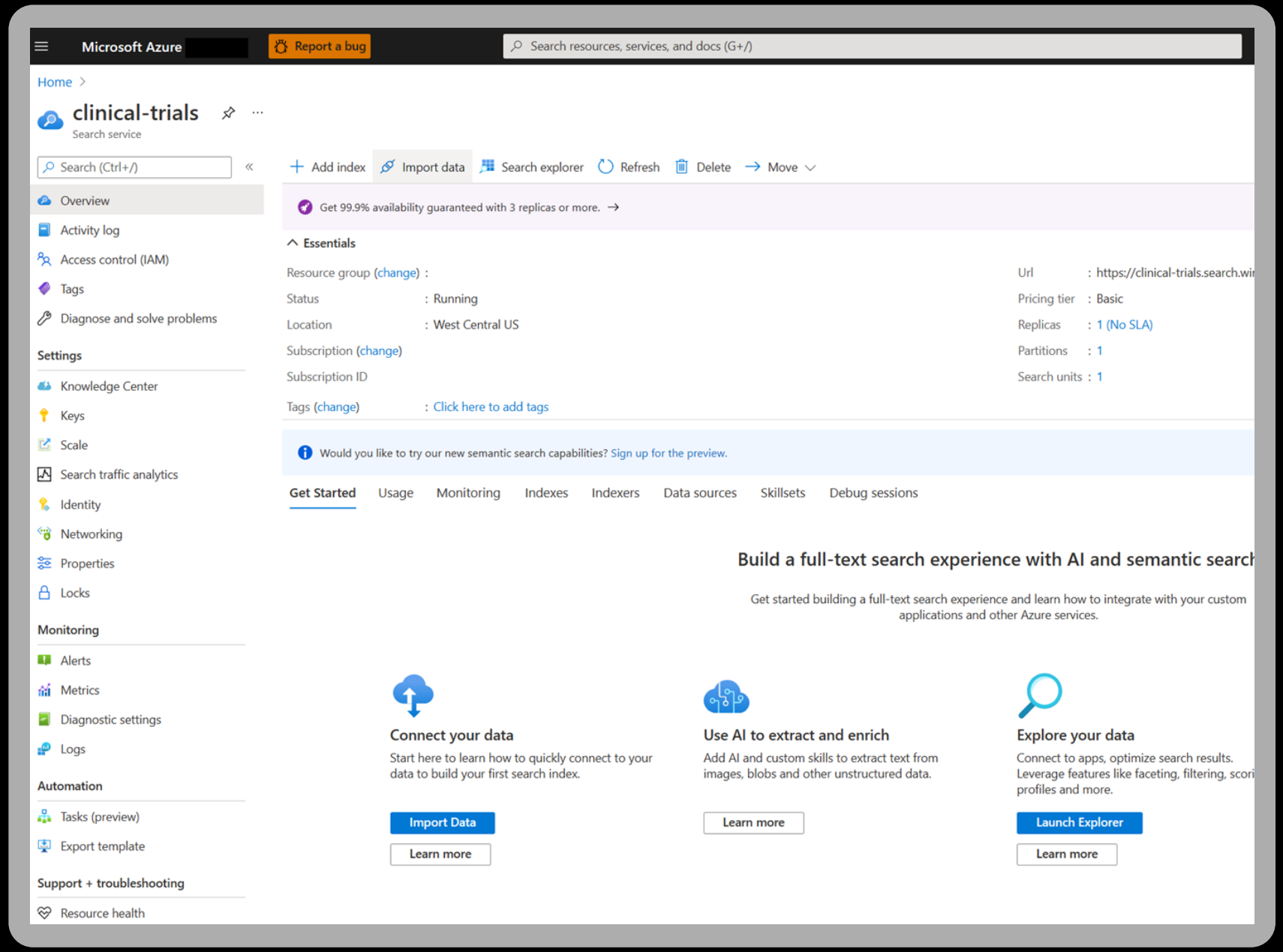
Ranking Knobs

language analyzers
custom analyzer
tokenizers
scoring profiles
customizable scoring
field selection and boosting
synonyms



Query

faceting
filtering
full text
autocomplete
suggestions
partial term matching
fuzzy
proximity
wildcard
lucene syntax



Applications courantes de la recherche cognitive

Recherche d'entreprise (Trouvez le bon document)

- Rechercher le bon document à partir d'un référentiel volumineux
- Augmentez la productivité de vos équipes
- Enrichir les documents avec l'IA : classification, extraction d'entités, OCR, etc.

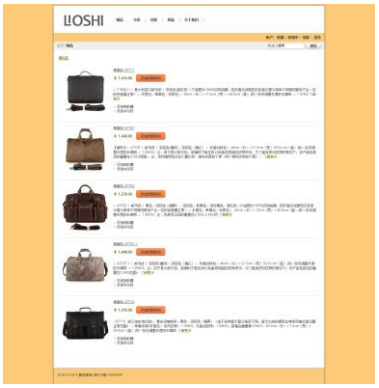
Exploration des connaissances (Trouvez le bon

- Trouvez des réponses aux questions en langage naturel
- Trouvez des connaissances pertinentes dans un grand corpus de texte à l'aide de l'expansion sémantique
- Trouvez les bons paragraphes dans le corpus de texte pour répondre à des questions

Renseignement sur les documents (Numériser les actifs)

- Indexer les documents structurés tels que les contrats, les factures, les commandes client, etc.
- Extraire les entités importantes
- Trouver le document pertinent
- Trouvez les bonnes informations dans les documents

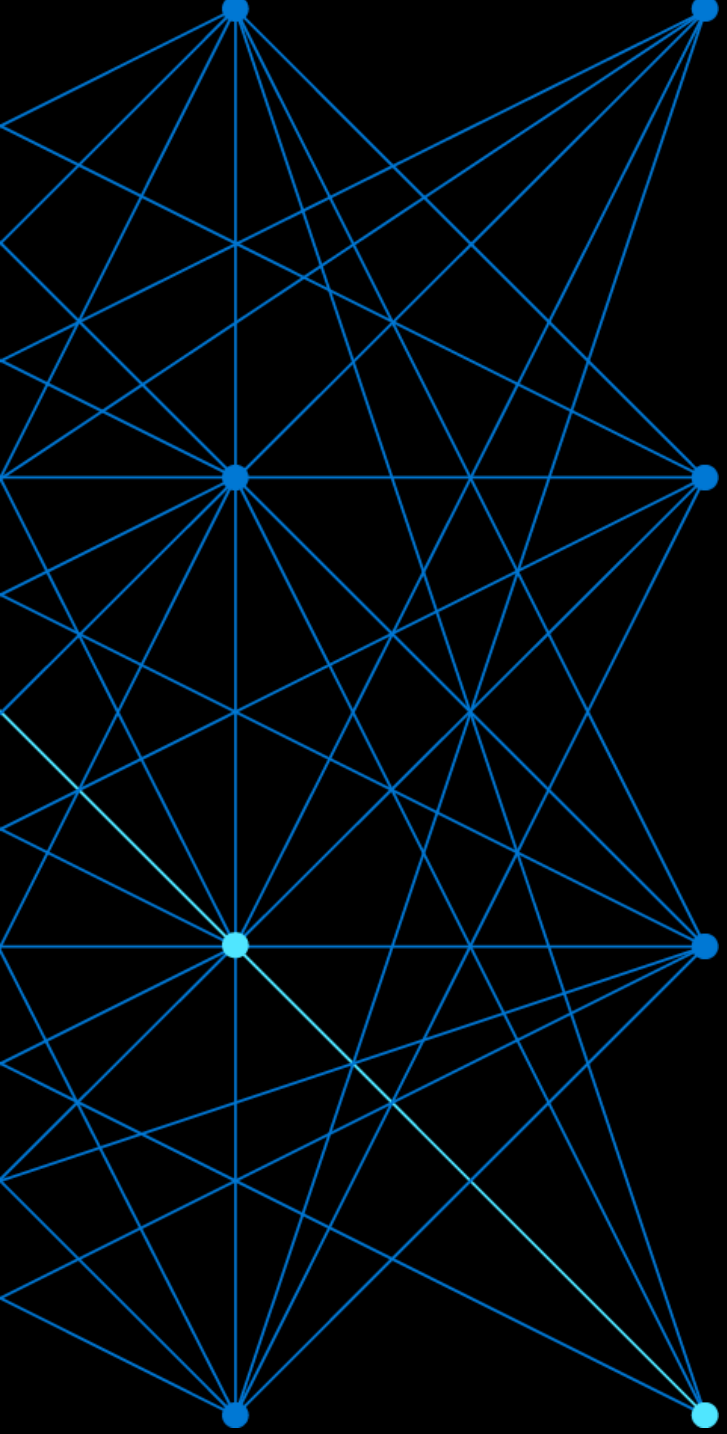
Recherche dans le catalogue (commerce électronique, applications Web et mobiles orientées client)



- Augmenter la pertinence des résultats des produits
- Empêcher « zéro résultat de recherche »
- Fournir des recommandations de produits intelligentes en fonction de l'intention de l'utilisateur
- Augmenter les taux de clics et de conversion

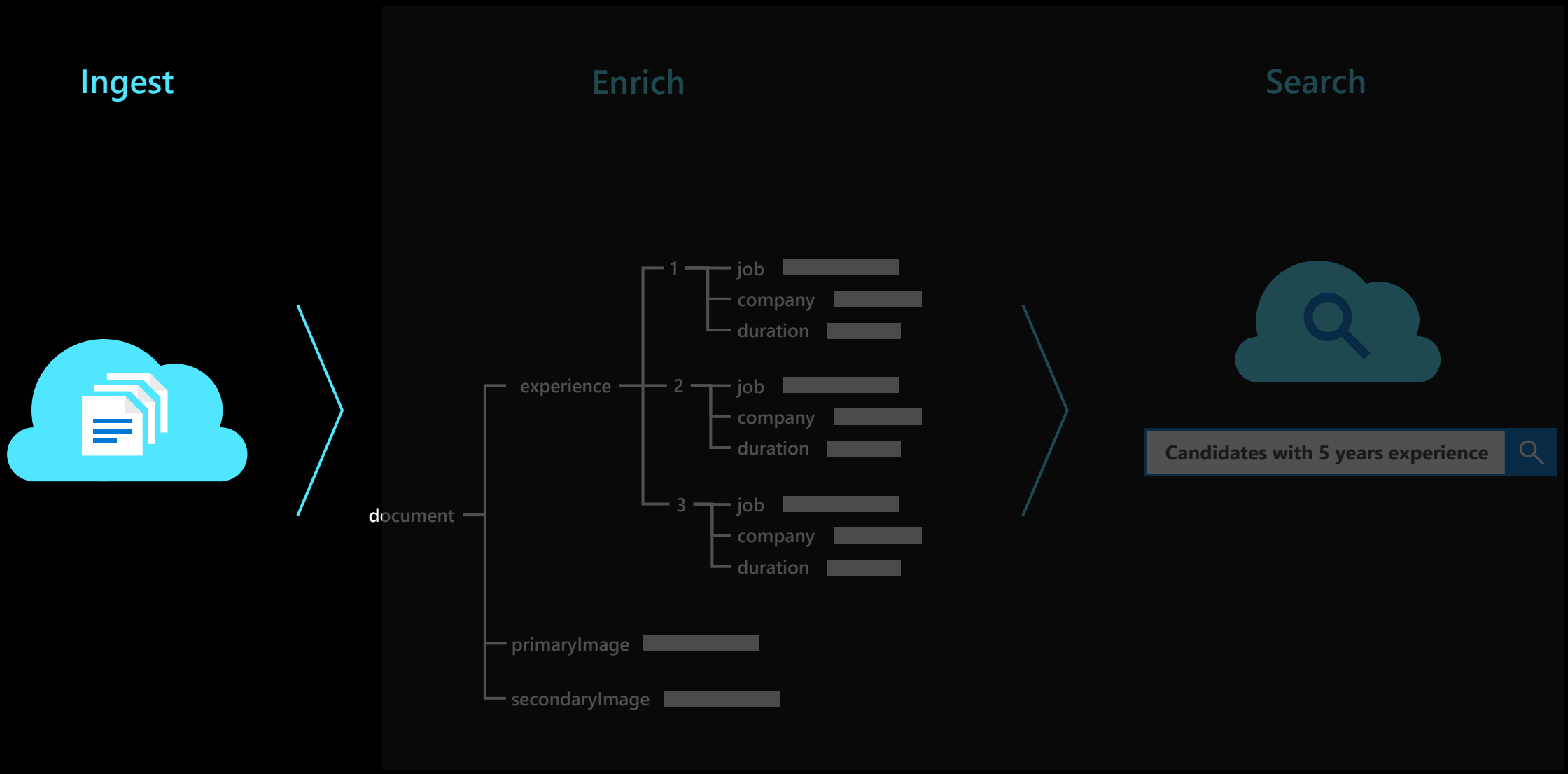
Demo

Exemple



Let’s dive !

Exploration des connaissances avec Azure Cognitive Search(1)



Ingérer des données à partir de sources de données externes

Data Types:

MSFT Office formats,
PDF, PNG
RTF, JSON
HTML, XML, KML, ZIP, GZ,
EML, RTF

Push data to an index

- Make use of custom apps to push JSON datasets into search indexes (any source)
- Submit individual docs or batches (1000 per batch or 16MB)
- Use Rest APIs or SDKs (.NET, Java, Python, JavaScript/Typescript)
- Push also available with data copied into Azure Data Factory or Synapse Analytics

Use if:
✓ If data is in custom apps and/or non-Azure stores

Pull data into an indexer

- Indexers crawl only supported Azure data sources (below)
- "Pulls data in" (no need to add code to add data into an indexer)
- Use Portal, .NET SDK or Rest APIs

Use if:
✓ If your data is in an Azure store
✓ You wish to leverage AI enrichment capabilities (external processing en route to an index)

Leverage connectors for Azure sources:
- Azure Blob Storage
- Azure Cosmos DB (SQL API, Gremlin API, Mongo API)
- Azure SQL DB
- Azure MySQL
- Azure Table Storage
- Azure Data Lake Storage Gen2
- SharePoint Online (Preview)

Connect to data on 3P platforms using indexers and PowerQuery connectors:
- Amazon Redshift
- Elasticsearch
- PostgreSQL
- Salesforce Objects, Salesforce Reports
- Smartsheet
- Snowflake

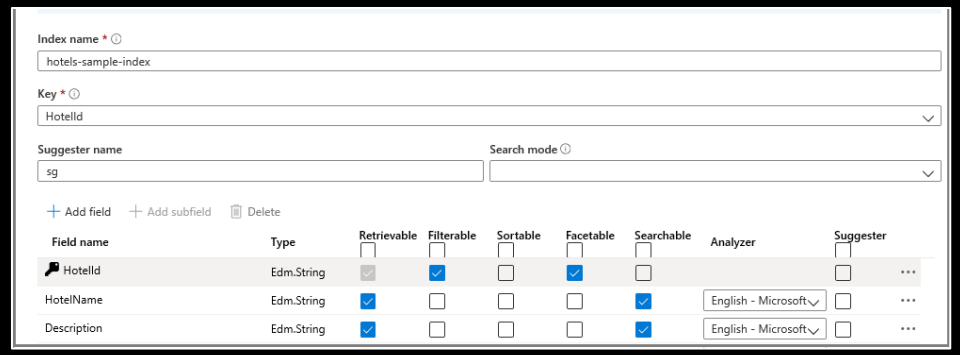
Connect to data on 3P platforms with connectors built by Microsoft Partners

Examen plus approfondi : indexation de vos données

Indexing is an intake process that loads content into to your search service and makes it searchable

An index is a store of searchable text (created & stored by the customer). An index is defined by its schema, consisting of "fields" with specified attributes (*retrievable, searchable, filterable, facetable, sortable*)

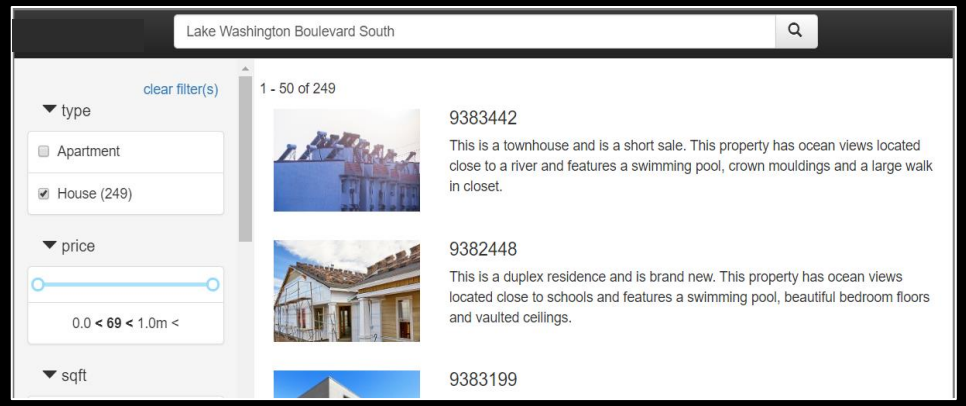
For POC/Testing, customers are advised to create indexes in the Azure Portal or using SDKs.



Search index
(view from Import Data Wizard)

Query Requests ↑ ↓ Search Responses

Client App



"Retrievable"
= selected fields show up in search results

"Filterable" / "Sortable" / "Facetable"
= selected fields are used in a filter, sort, or faceted navigation structure

"Searchable"
= selected fields are included in full text search (multilanguage support via Lucene & Microsoft language analyzers)

Querying execution is over the data in the search index.

In your client app, the search experience is defined using APIs from Azure Cognitive Search, and can include relevance tuning, autocomplete, synonym matching, fuzzy matching, pattern matching, filter, and sort.

Analyseurs

Analyzers perform lexical analysis using the linguistic or pre-defined rules. Every searchable string field has an analyzer property. The default analyzer is the Apache Lucene Standard Analyzer (no configuration required).

Built-in Analyzers:

- Language Analyzers: support for 35 Lucene analyzers and 50 Microsoft natural language processing analyzers) (link)
- Specialized Analyzers (language agnostic) use when text inputs require specialized processing (link)
- Custom Analyzers: User-defined configuration of combinations of existing elements

..“COVID-19”….

..“Chemise bleue”….

..“jsmith@gmail.com”

..“(321) 555-0199”

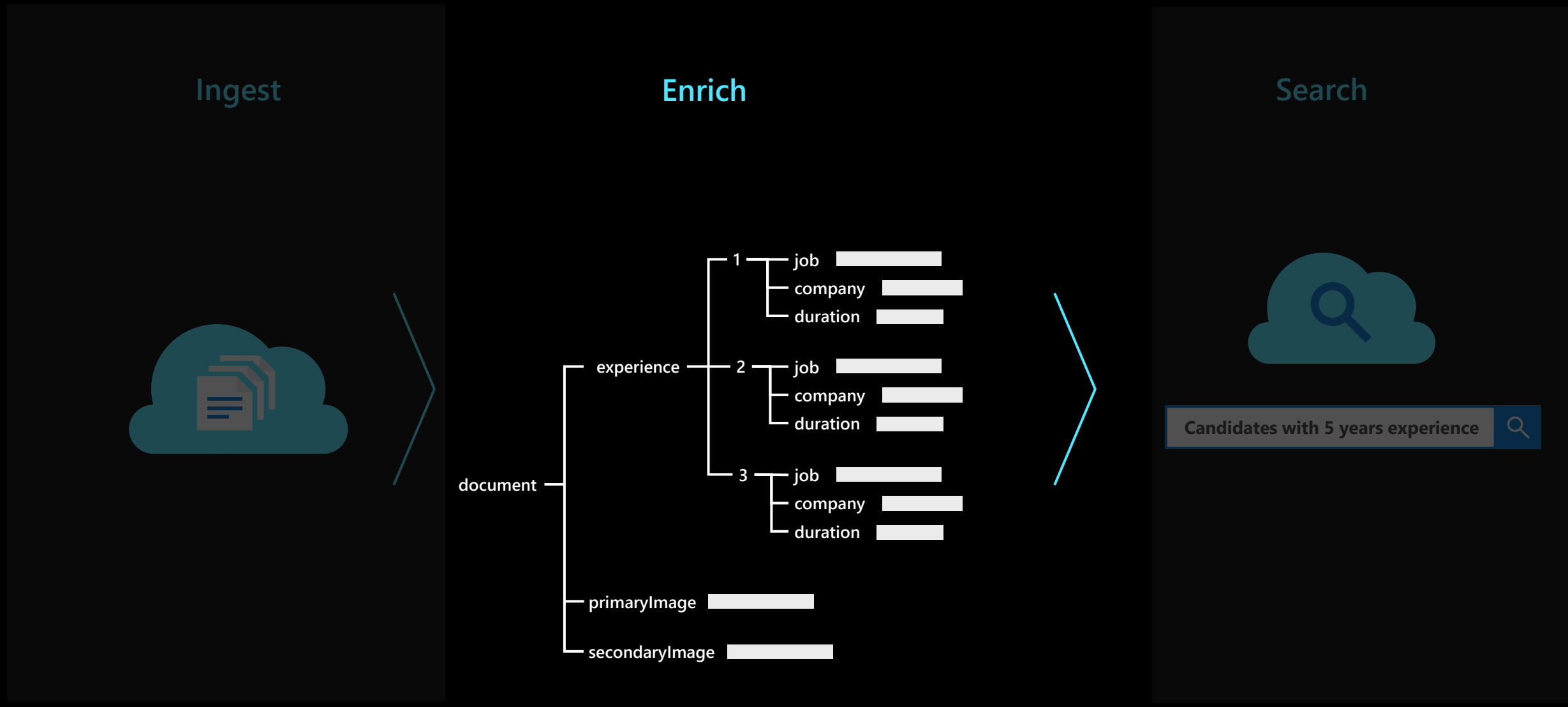
Consider other language analyzers if…

- Awareness of word or sentence *structure* adds value to text parsing
- If needing rich language support
- If content consists of non-Western language strings

Best practices:

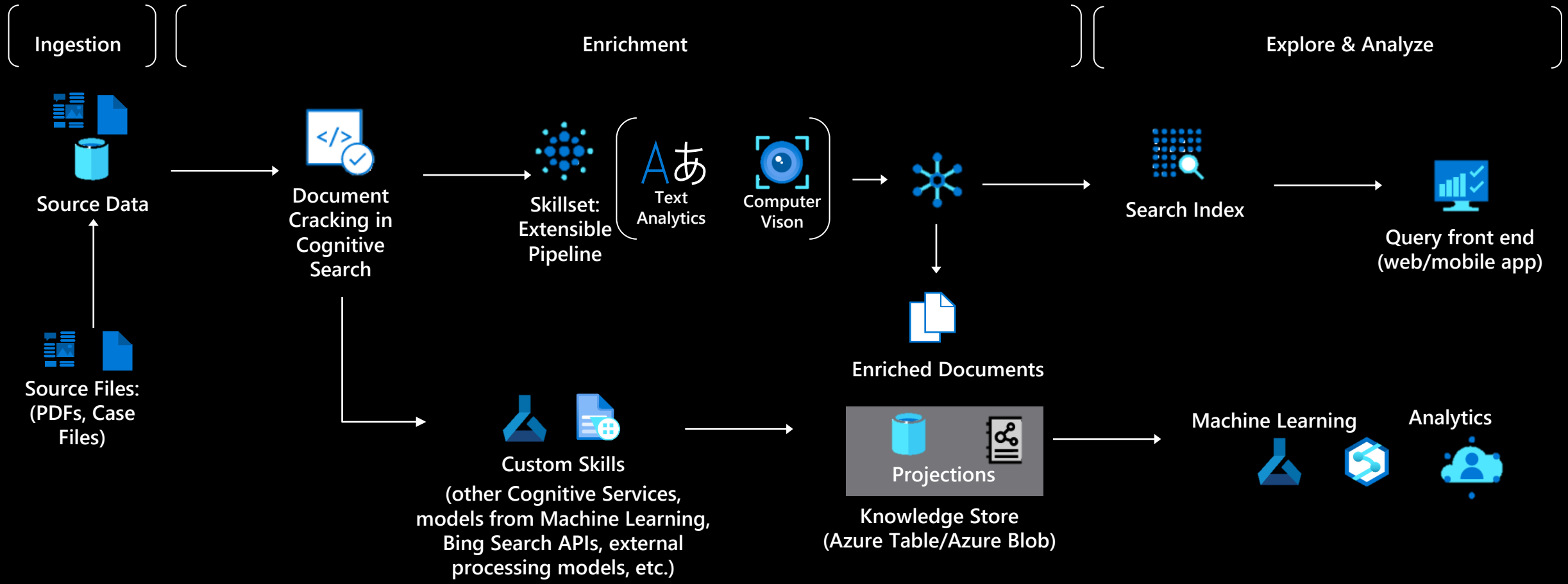
- ✓ If possible, use the same analyzer for indexing and querying
- ✓ If search fails or doesn’t return expected results, the most likely scenario is token discrepancies between query term inputs and the tokenized items in the index
- ✓ Analyzers are best added when setting up fields (when building an index)

Exploration des connaissances avec Azure Cognitive Search (2)



Enrichissement de l'IA avec des compétences cognitives

In Azure Cognitive Search, AI enrichment refers to built-in cognitive skills and custom skills that add analysis, transformations, and content generation during indexing. Enrichment is defined by a skillset that is attached to an indexer (requires pull method)

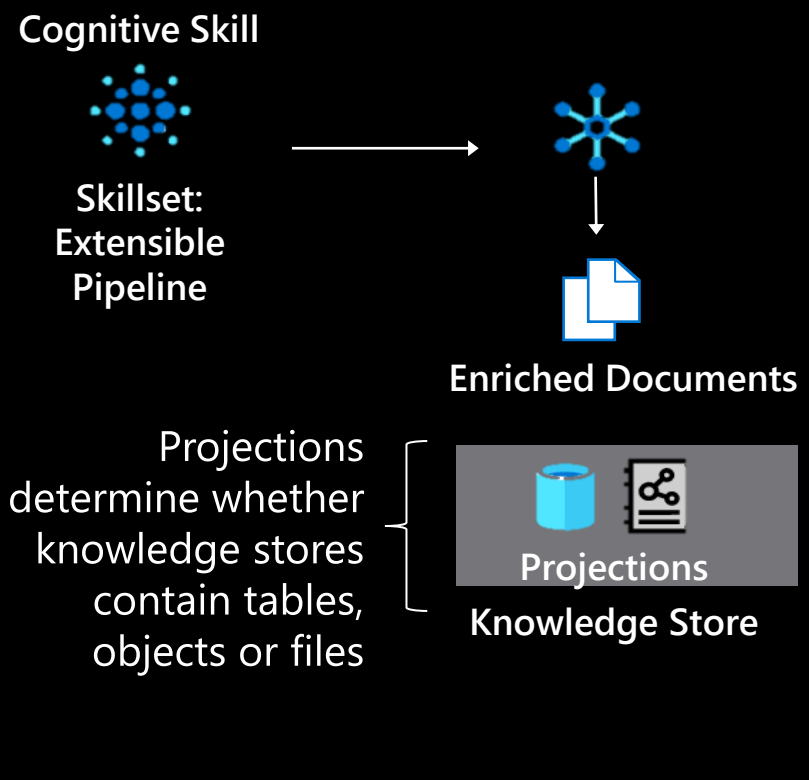


Enrichments create new information where none previously existed: extracting information from images, detecting sentiment, key phrases, and entities from text, etc. Enrichments also add structure to undifferentiated text.

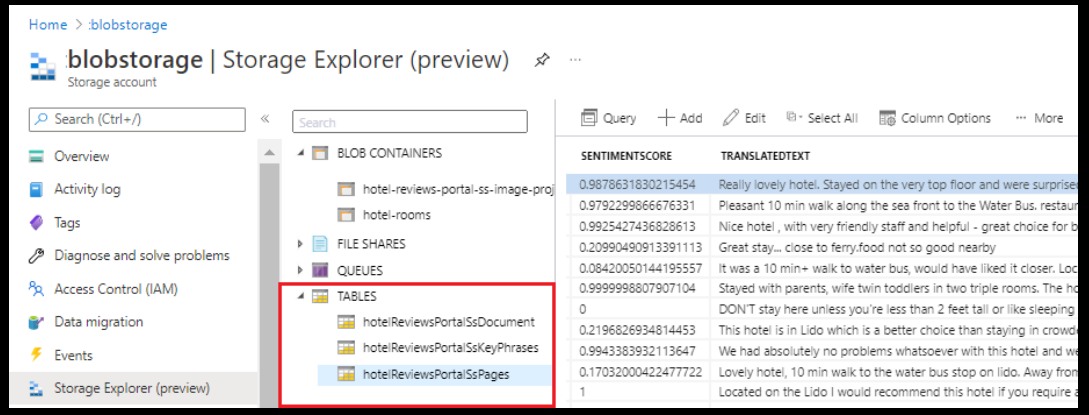
All of these processes result in making previously unsearchable content available to full text search scenarios. The output of an enrichment pipeline is either a search index or a knowledge store.

Magasins de connaissances

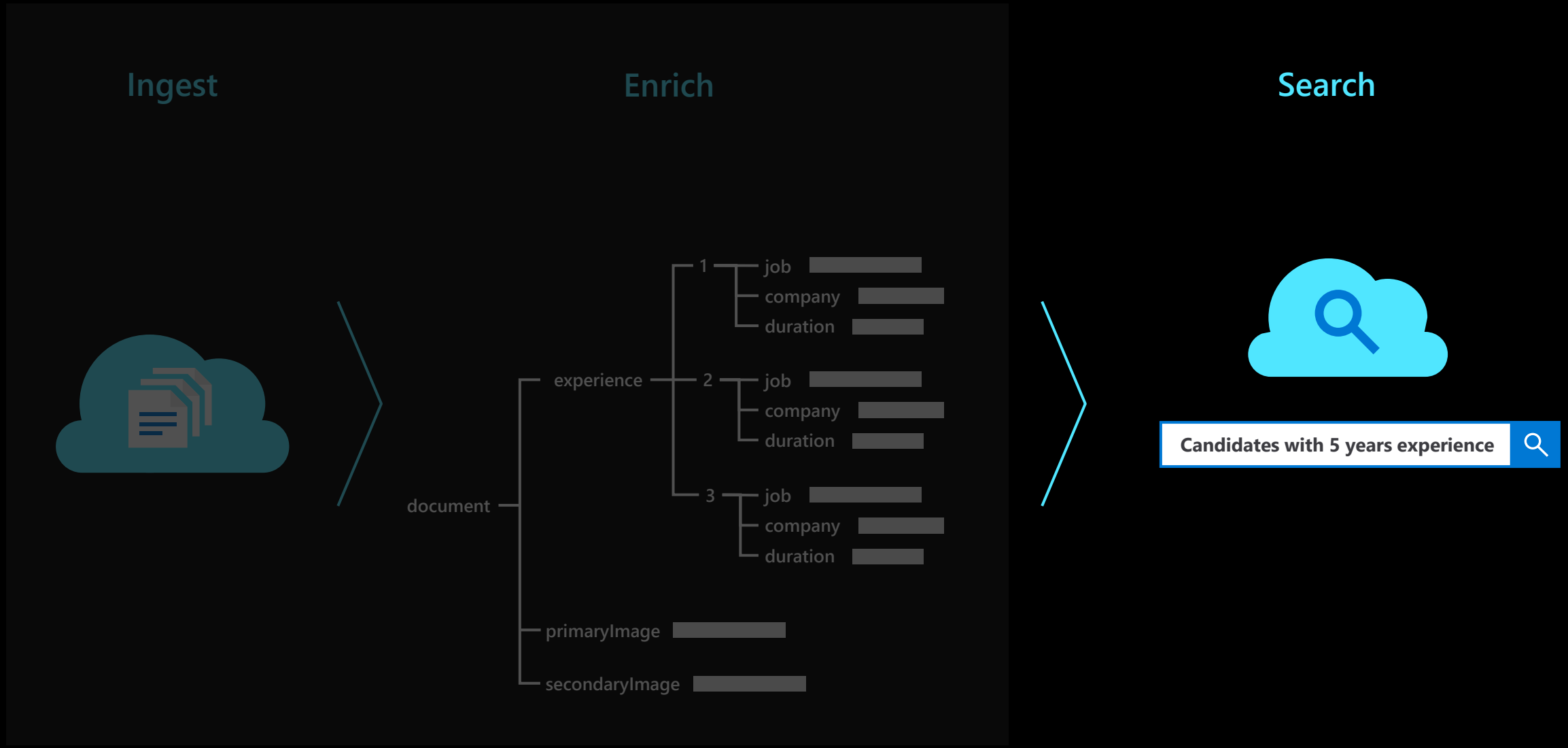
A knowledge store is a data sink created by the Cognitive Search AI Enrichment pipeline that stores enriched content for analysis or downstream processing in non-search scenarios



Enriched content is stored in Azure Table Storage or Azure Blob Storage

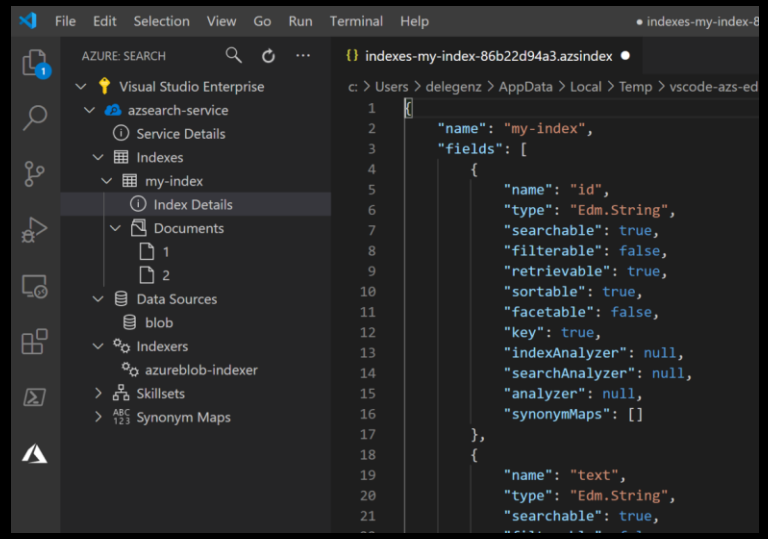


Exploration des connaissances avec Azure Cognitive Search (3)



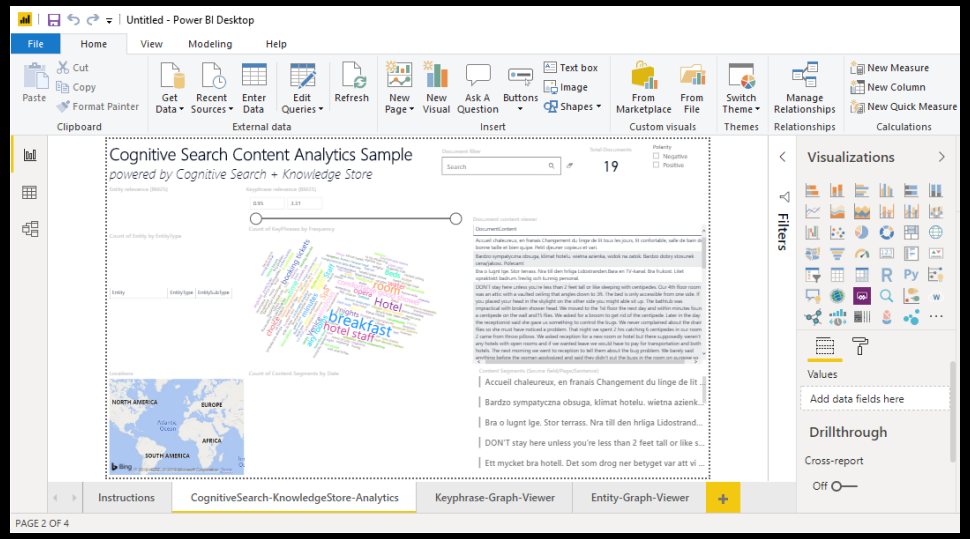
Explorer et analyser les requêtes dans des environnements familiers

Visual Studio Code Extension

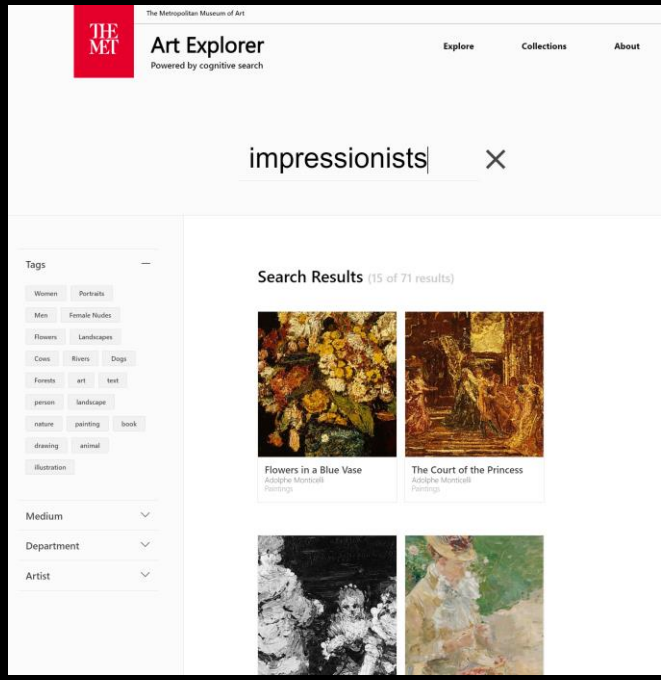


https://aka.ms/vscode-search

PowerBI & Power Apps



Web Apps




Microsoft
Q&A
SM

Allez plus loin sur Microsoft Learn

Explorez la détection d’anomalies dans Microsoft Azure
https://learn.microsoft.com/en-us/azure/cognitive-services/anomaly-detector/

Explorez la recherche cognitive dans Microsoft Azure
https://learn.microsoft.com/en-us/azure/search/