

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.*

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

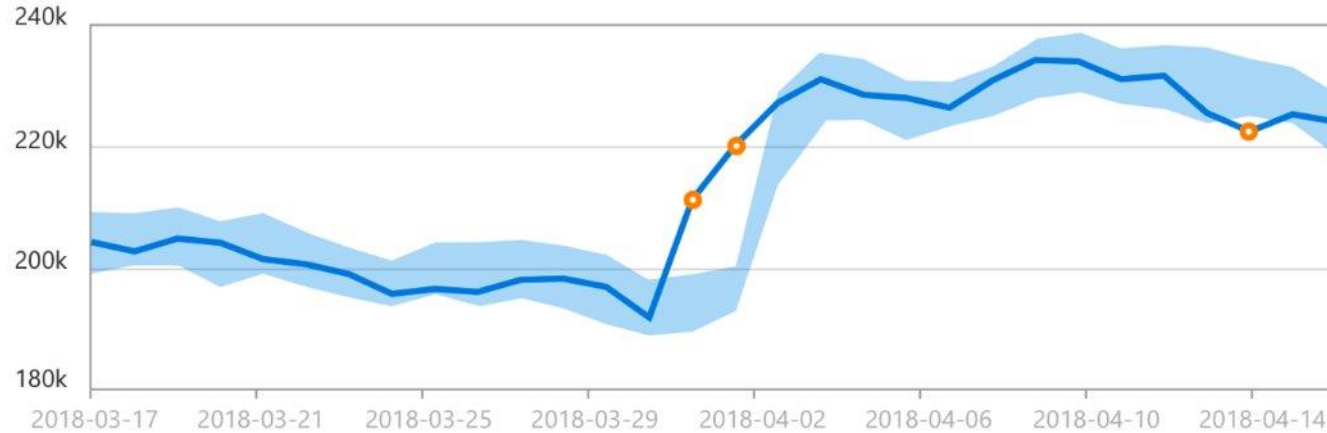
Vous allez apprendre les concepts suivants :

- 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

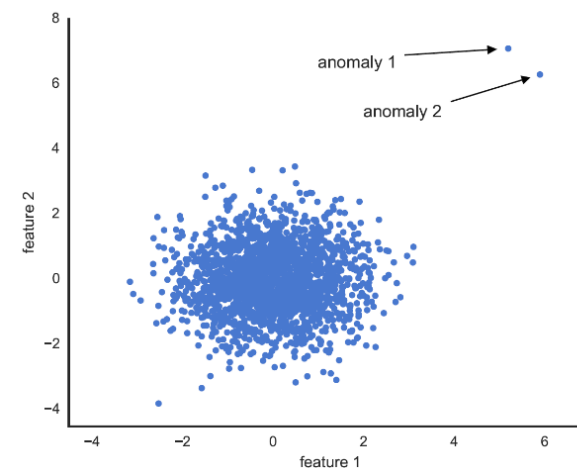


- comparaison d'un mois à l'autre des dépenses liées au coût des produits



Comment fonctionne t il?

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



$$upperBoundary = expectedValue + (100 - marginScale) \times upperMargin$$



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

A stylized logo for JSON, featuring the word "JSON" in white, bold, sans-serif font, enclosed within large, yellow, stylized curly braces. The background is a dark gray rectangle.

```
1  {
2    "granularity": "hourly",
3    "series": [
4      {
5        "timestamp": "2021-03-01T01:00:00Z",
6        "value": -10.56
7      },
8      {
9        "timestamp": "2021-03-02T02:00:00Z",
10       "value": -8.30
11     },
12     {
13       "timestamp": "2021-03-02T03:00:00Z",
14       "value": -10.30
15     },
16     {
17       "timestamp": "2021-03-02T04:00:00Z",
18       "value": 5.95
19     },
20   ]
21 }
```

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 réel

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?

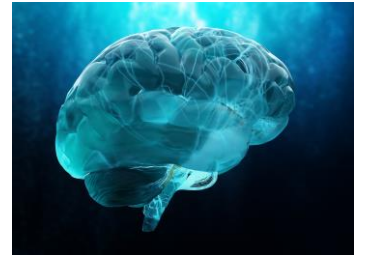
Document manuscript, les images....

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



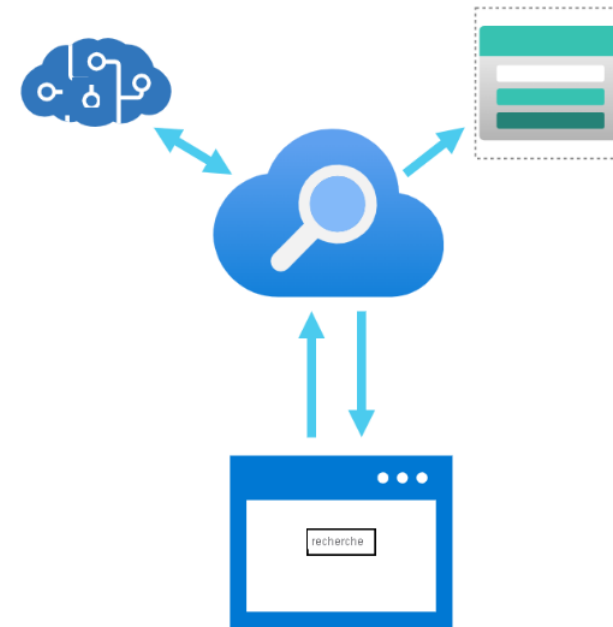
- Exploitez les modèles DNN de pointe de Bing
- Modèles formés sur des billions de requêtes de recherche utilisateur et de clics couvrant plusieurs domaines

Alors c'est quoi le "knowledge mining"?

Uncover latent insights

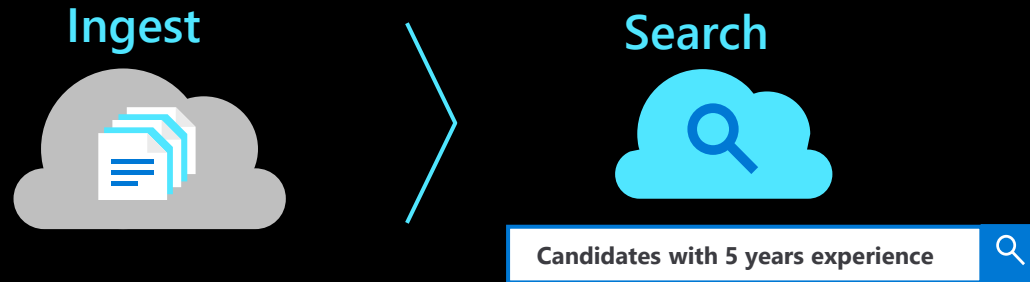
Recherche Cognitive Azure

- 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

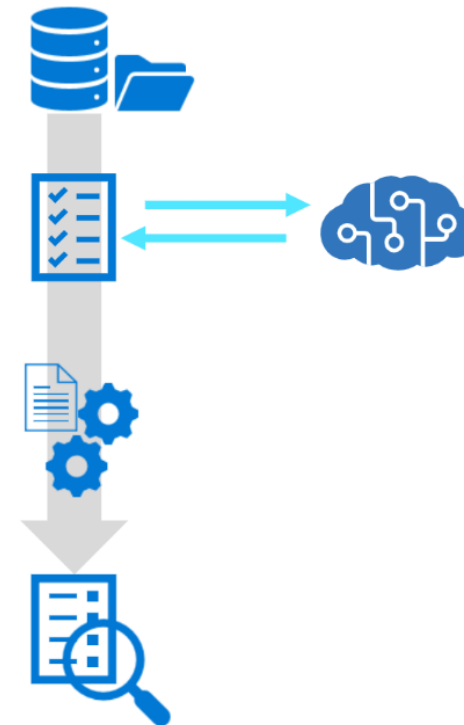
2. Apply AI to understand user intent
Use semantic search to improve the quality of
search results

Recherche
basée sur l'IA




Comment ça marche?

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










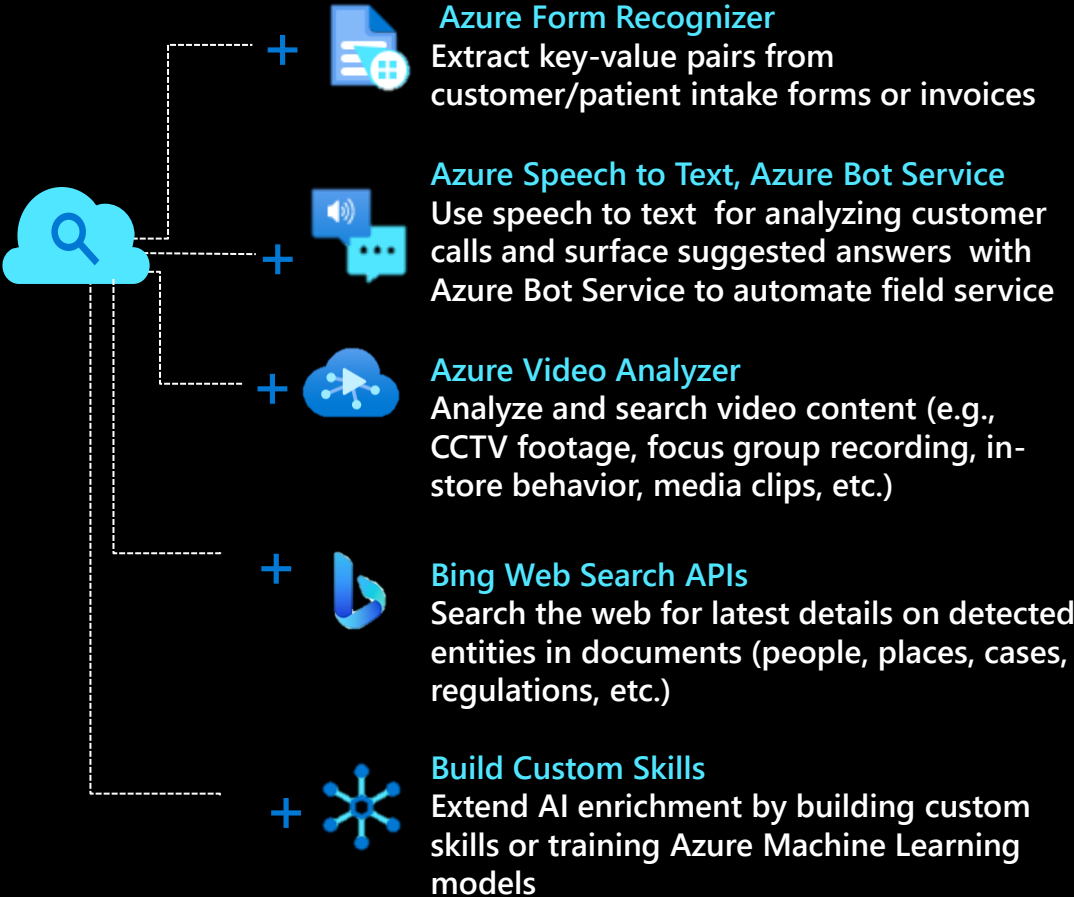


Image Processing
Based on Computer Vision API

BUILT-IN SKILLS		
 Key Phrase extraction	 Location entity extraction	 Sentiment analysis
 Organization entity extraction	 <u>Persons</u> entity extraction	 Language detection
 Face detection	 Celebrity recognition	 Image tag extraction
 Text Utilities	 Landmark detection	 Printed text recognition

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

(Examples of common service combinations)



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



What is the capital of France?



Paris is the capital city of France. It has a population of 2,138,551, and is located on a latitude of 48.85 and longitude of 2.35. **Paris** is also the political center of France, which is considered a Republic, and home to its Executive head of state.

[Where is France?](#)

<https://www.worldatlas.com/eu/fr/where-is-france.html>

[Where is France?](#)

<https://www.worldatlas.com/eu/fr/where-is-france.html>

Paris is the **capital** city of **France**. It has a population of 2,138,551, and is located on a latitude of 48.85 and longitude of 2.35. **Paris** is also the political center of **France**, which is considered a Republic, and home to its Executive head of state.

[Where is France?](#)

<http://www.worldatlas.com/eu/fr/where-is-france.html>

Paris is the **capital** city of **France**. It has a population of 2,138,551, and is located on a latitude of 48.85 and longitude of 2.35. **Paris** is also the political center of **France**, which is considered a Republic, and home to its Executive head of state.

[France](#)

<https://www.iaufrance.org/locations/france>

Paris, the ultra-cosmopolitan **capital** city, is at the heart of **French** culture and an ideal environment in which to learn and perfect **French** language skills. The city has a longstanding global reputation as a mecca of art, culture, fashion, gastronomy, history, and architecture.

Ranking Knobs

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

The screenshot displays the Microsoft Azure portal interface for a search service named 'clinical-trials'. The left sidebar contains navigation links for Overview, Activity log, Access control (IAM), Tags, Diagnose and solve problems, Settings (Knowledge Center, Keys, Scale, Search traffic analytics, Identity, Networking, Properties, Locks), Monitoring (Alerts, Metrics, Diagnostic settings, Logs), Automation (Tasks (preview), Export template), and Support + troubleshooting (Resource health). The main content area shows the service's status as 'Running' in the 'West Central US' location. It includes a table of properties: Resource group, Status, Location, Subscription, Subscription ID, Tags, Url, Pricing tier, Replicas, Partitions, and Search units. A 'Get Started' section at the bottom offers three options: 'Connect your data', 'Use AI to extract and enrich', and 'Explore your data', each with a 'Learn more' button.

Microsoft Azure [Report a bug](#) Search resources, services, and docs (G+)

Home > **clinical-trials** Search service

Search (Ctrl+/) << + Add index Import data Search explorer Refresh Delete Move >

Overview

Activity log

Access control (IAM)

Tags

Diagnose and solve problems

Settings

Knowledge Center

Keys

Scale

Search traffic analytics

Identity

Networking

Properties

Locks

Monitoring

Alerts

Metrics

Diagnostic settings

Logs

Automation

Tasks (preview)

Export template

Support + troubleshooting

Resource health

Get 99.9% availability guaranteed with 3 replicas or more. →

Essentials

Resource group (change) :

Status : Running

Location : West Central US

Subscription (change)

Subscription ID

Tags (change) : [Click here to add tags](#)

Url : <https://clinical-trials.search.windows.net>

Pricing tier : Basic

Replicas : 1 (No SLA)

Partitions : 1

Search units : 1

Would you like to try our new semantic search capabilities? [Sign up for the preview.](#)

Get Started Usage Monitoring Indexes Indexers Data sources Skillsets Debug sessions

Build a full-text search experience with AI and semantic search

Get started building a full-text search experience and learn how to integrate with your custom applications and other Azure services.

Connect your data

Start here to learn how to quickly connect to your data to build your first search index.

Import Data

Learn more

Use AI to extract and enrich

Add AI and custom skills to extract text from images, blobs and other unstructured data.

Learn more

Explore your data

Connect to apps, optimize search results. Leverage features like faceting, filtering, scoring profiles and more.

Launch Explorer

Learn more

Query

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

Microsoft Azure [Report a bug](#) Search resources, services, and docs (G+)

Home > **clinical-trials** Search service

Search (Ctrl+/) << + Add index Import data Search explorer Refresh Delete Move >

Overview

- Activity log
- Access control (IAM)
- Tags
- Diagnose and solve problems

Settings

- Knowledge Center
- Keys
- Scale
- Search traffic analytics
- Identity
- Networking
- Properties
- Locks

Monitoring

- Alerts
- Metrics
- D diagnostic settings
- Logs

Automation

- Tasks (preview)
- Export template

Support + troubleshooting

- Resource health

Get 99.9% availability guaranteed with 3 replicas or more. →

Essentials

Resource group (change) : Url : https://clinical-trials.search.wi

Status : Running Pricing tier : Basic

Location : West Central US Replicas : 1 (No SLA)

Subscription (change) Partitions : 1

Subscription ID Search units : 1

Tags (change) : [Click here to add tags](#)

Would you like to try our new semantic search capabilities? [Sign up for the preview.](#)

Get Started Usage Monitoring Indexes Indexers Data sources Skillsets Debug sessions

Build a full-text search experience with AI and semantic search

Get started building a full-text search experience and learn how to integrate with your custom applications and other Azure services.

Connect your data

Start here to learn how to quickly connect to your data to build your first search index.

Import Data

Learn more

Use AI to extract and enrich

Add AI and custom skills to extract text from images, blobs and other unstructured data.

Learn more

Explore your data

Connect to apps, optimize search results. Leverage features like faceting, filtering, scoring profiles and more.

Launch Explorer

Learn more

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



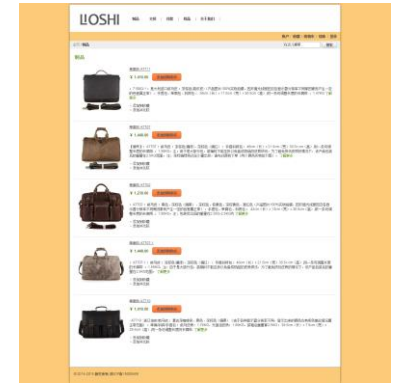
- Trouvez des réponses aux questions en langage naturel
- Trouvez des connaissances pertinentes dans un grand corpus de texte à l'aide de l'expansion sémantique
- Trouvez les bons paragraphes dans le corpus de texte pour répondre à des questions spécifiques

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

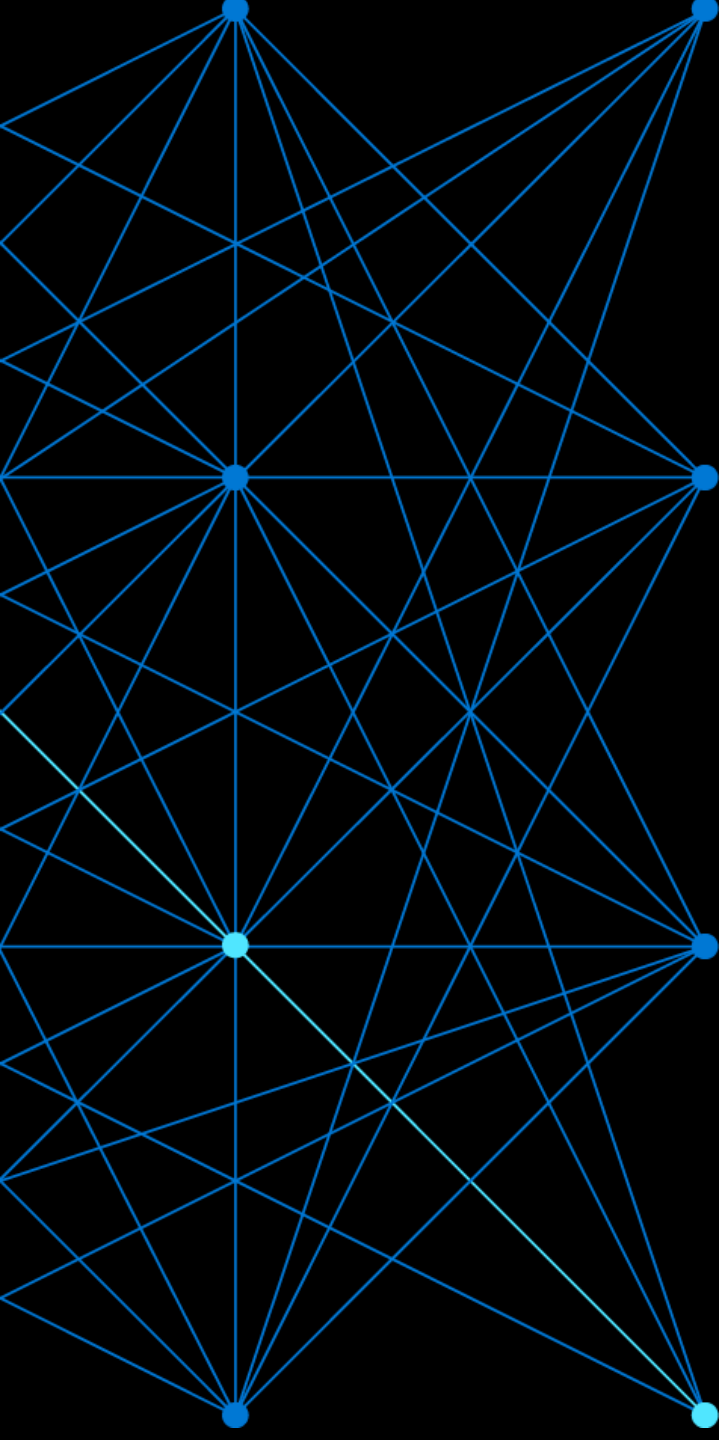
Recherche dans le catalogue (commerce électronique, applications Web et mobiles orientées client)



- 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

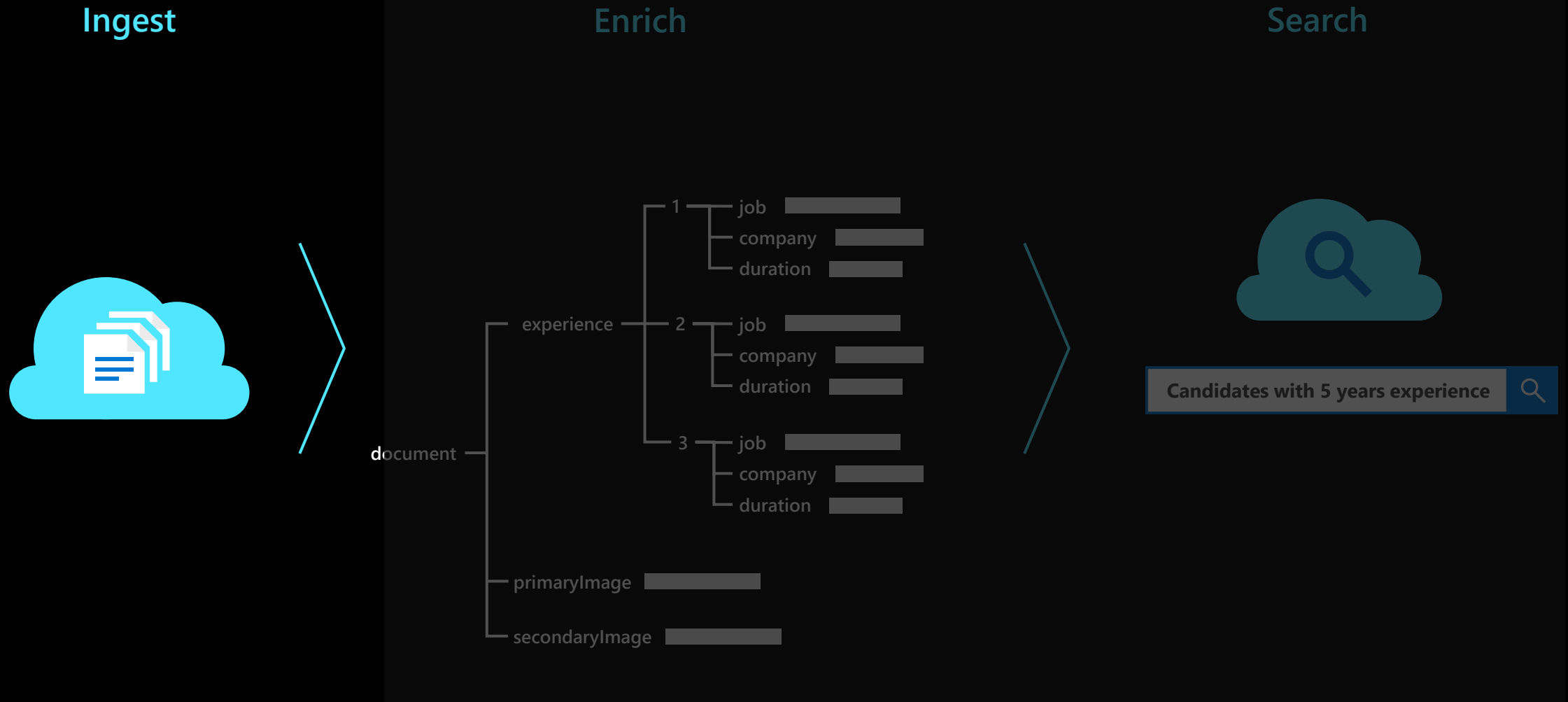
Demo

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:

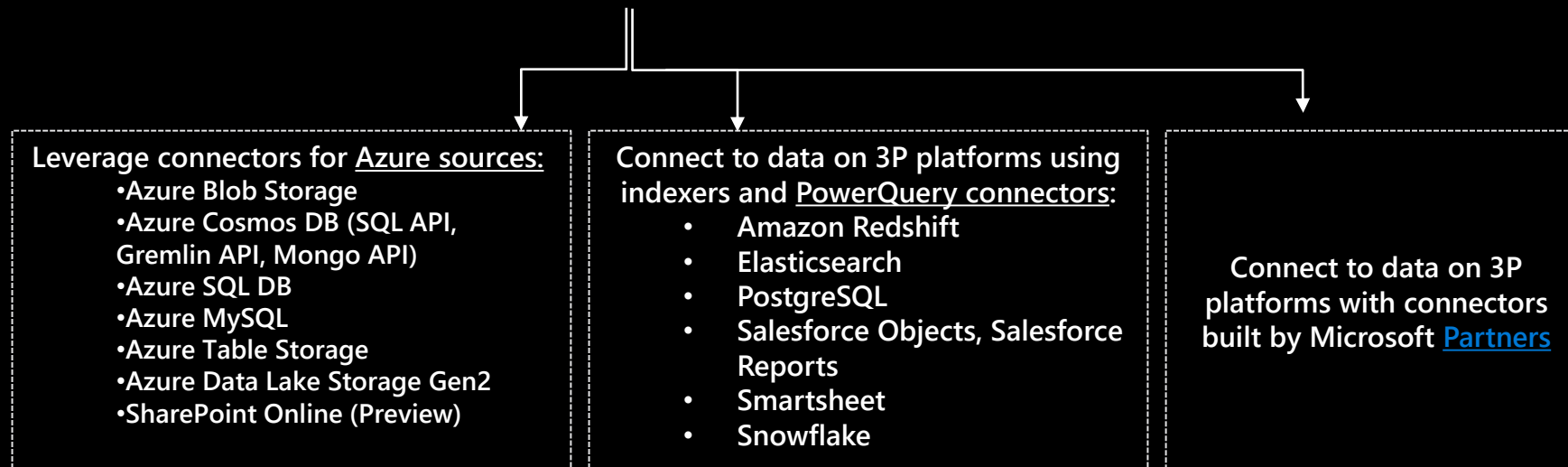
- ✓ If data is in custom apps and/or non-Azure stores

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)



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*, *facettable*, *sortable*)

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

Field name	Type	Retrievable	Filterable	Sortable	Facetable	Searchable	Analyzer	Suggester
HotelId	Edm.String	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	English - Microsoft	
HotelName	Edm.String	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	English - Microsoft	
Description	Edm.String	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	English - Microsoft	

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.

Lake Washington Boulevard South

1 - 50 of 249

type

Apartment

House (249)

price

0.0 < 69 < 1.0m <

sqft

9383442

This is a townhouse and is a short sale. This property has ocean views located close to a river and features a swimming pool, crown mouldings and a large walk in closet.

9382448

This is a duplex residence and is brand new. This property has ocean views located close to schools and features a swimming pool, beautiful bedroom floors and vaulted ceilings.

9383199

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:

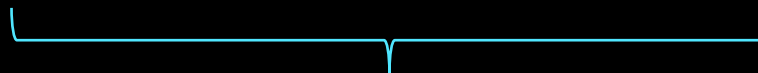
- **Language Analyzers:** support for 35 Lucene analyzers and 50 Microsoft natural language processing analyzers) ([link](#))
- **Specialized Analyzers** (language agnostic) use when text inputs require specialized processing ([link](#))
- **Custom Analyzers:** User-defined configuration of combinations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

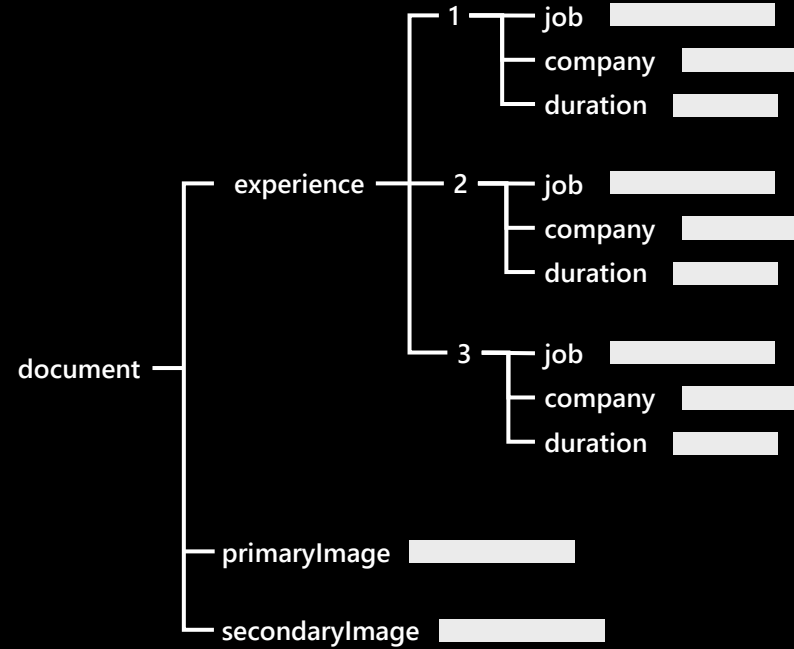
- ✓ If possible, use the same analyzer for indexing and querying
- ✓ If search fails or doesn't return expected results, the most likely scenario is token discrepancies between query term inputs and the tokenized items in the index
- ✓ Analyzers are best added when setting up fields (when building an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

Exploration des connaissances avec Azure Cognitive Search (2)

Ingest



Enrich



Search

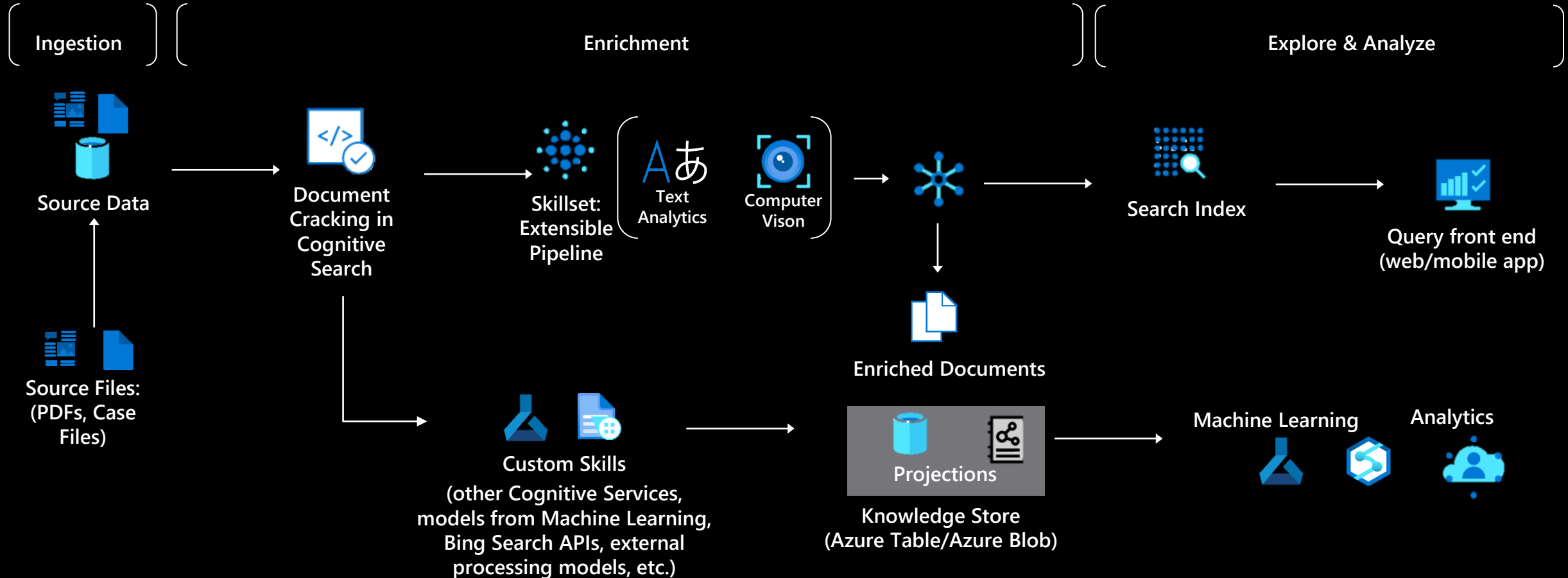


Candidates with 5 years experience



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

Cognitive Skill

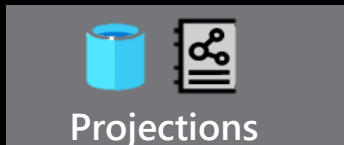


Skillset:
Extensible
Pipeline



Enriched Documents

Projections
determine whether
knowledge stores
contain tables,
objects or files



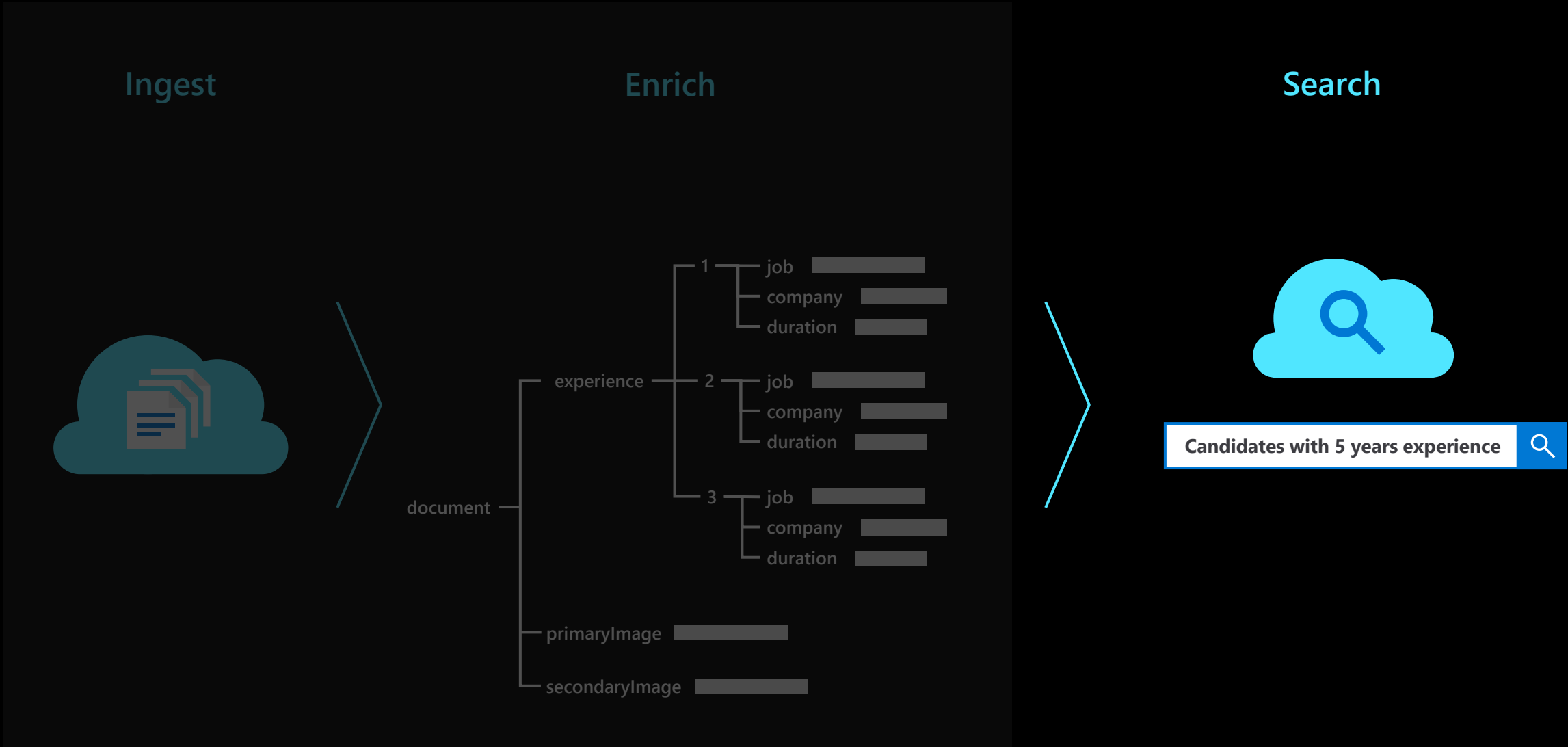
Projections
Knowledge Store

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

The screenshot shows the Azure Storage Explorer (preview) interface. The left sidebar displays the navigation pane with options: Overview, Activity log, Tags, Diagnose and solve problems, Access Control (IAM), Data migration, Events, and Storage Explorer (preview). The main pane shows the 'blobstorage' account. Under 'BLOB CONTAINERS', there are 'hotel-reviews-portal-ss-image-proj' and 'hotel-rooms'. Under 'FILE SHARES', there are 'hotel-reviews-portal-ss-image-proj' and 'hotel-rooms'. Under 'QUEUES', there are 'hotel-reviews-portal-ss-image-proj' and 'hotel-rooms'. Under 'TABLES', there are 'hotelReviewsPortalSsDocument', 'hotelReviewsPortalSsKeyPhrases', and 'hotelReviewsPortalSsPages'. The 'hotelReviewsPortalSsPages' table is selected, and its data is displayed in the right pane. The table has two columns: 'SENTIMENTSCORE' and 'TRANSLATEDTEXT'.

SENTIMENTSCORE	TRANSLATEDTEXT
0.9878631830215454	Really lovely hotel. Stayed on the very top floor and were surprise
0.9792299866676331	Pleasant 10 min walk along the sea front to the Water Bus. restaur
0.9925427436628613	Nice hotel , with very friendly staff and helpful - great choice for b
0.20990490913391113	Great stay... close to ferry, food not so good nearby
0.08420050144195557	It was a 10 min+ walk to water bus, would have liked it closer. Loc
0.9999998807907104	Stayed with parents, wife twin toddlers in two triple rooms. The ho
0	DON'T stay here unless you're less than 2 feet tall or like sleeping
0.2196826934814453	This hotel is in Lido which is a better choice than staying in crowd
0.9943383932113647	We had absolutely no problems whatsoever with this hotel and we
0.17032000422477722	Lovely hotel, 10 min walk to the water bus stop on lido. Away from
1	Located on the Lido I would recommend this hotel if you require s

Exploration des connaissances avec Azure Cognitive Search (3)



Visual Studio Code Extension



File Home View Modeling Help

Paste Cut Copy Format Painter Clipboard

Get Data Recent Sources Enter Data Edit Queries Refresh New Page New Visual Ask A Question Buttons Text box Image Shapes Custom visuals

Themes Relationships Calculations

Cognitive Search Content Analytics Sample

powered by Cognitive Search + Knowledge Store

Visual of Today by Entity Type

Entity: [Hotel] [Restaurant] [Bakery]

Visual of Content Segments by Size


Visual of Top Documents by Frequency

Document Content

Accueil chauxevue, en français Changement du linge de lit
Bardez sympatyczna obsługa, klimat hotelu, wietna azienka
Bra o lugntige Stor terrass. Nra till den hrigta Lidostrand.
DON'T stay here unless you're less than 2 feet tall or like s...
Ett mycket bra hotell. Det som drog ner betyget var att vi...

Visuals Filters Values Add data fields here Drillthrough Cross-report Off On

Instructions CognitiveSearch-KnowledgeStore-Analytics Keyphrase-Graph-Viewer Entity-Graph-Viewer



The Metropolitan Museum of Art

Art Explorer

Powered by cognitive search

[Explore](#)
[Collections](#)
[About](#)

×

Tags

Women

Men

Flowers

Cave

Forests

person

nature

drawing

illustration

Portraits

Female Nudes

Landscapes

Rivers

art

landscape

painting

animal

Drop


book

Medium


Department

Artist

Search Results (15 of 71 results)



Flowers in a Blue Vase
Adolphe Montbelli
Painting



The Court of the Princess
Adolphe Montbelli
Painting

Q&A



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/>

