#SPSBE

Using AI and Azure Cognitive Search to Enrich, Index and Search your Enterprise Data

Donald Hessing | @dhessing

BIWUG

SharePoint Saturday Belgium 2019

Thanks to our sponsors!

Platinum

Gold

Silver

Community

Who am I?

Donald Hessing

- Microsoft Lead Architect, Capgemini DCX Netherlands
- Microsoft Certified Master(MCM)
- Speaker Chair for ESPC17, ESPC18, ESPC19
- @dhessing

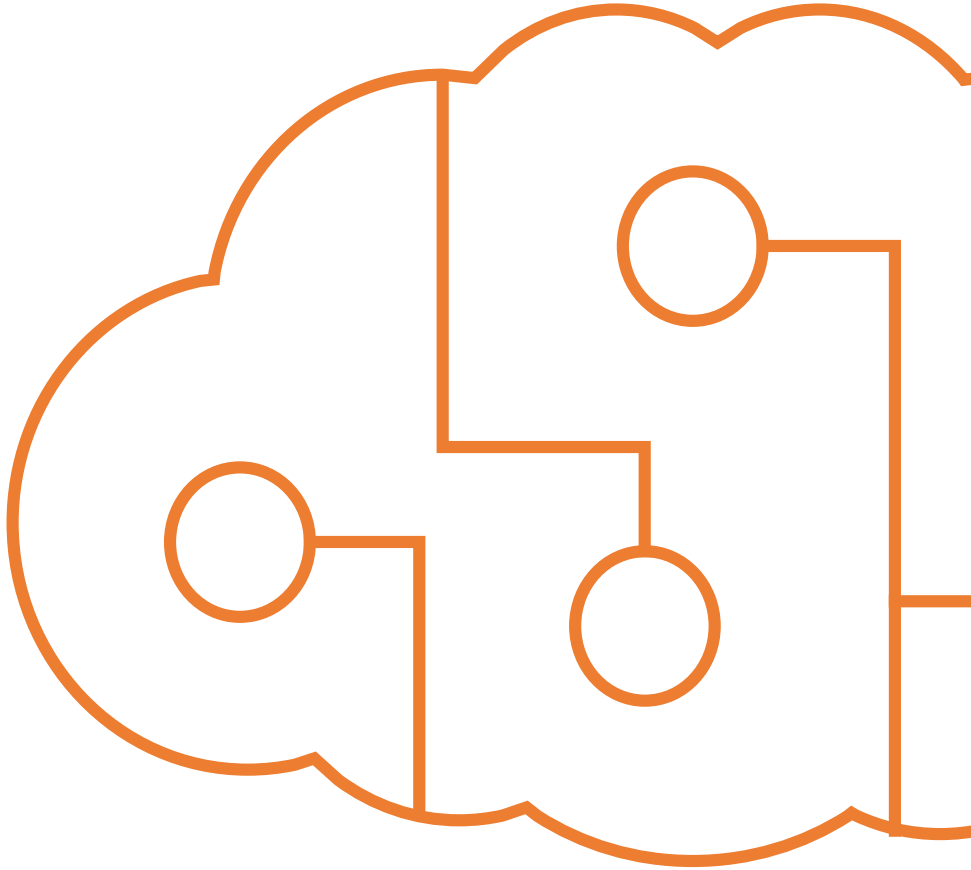
Agenda

Introduction to Azure Cognitive Search

Indexing Pipeline

Custom Skills with Azure Functions

Search UI



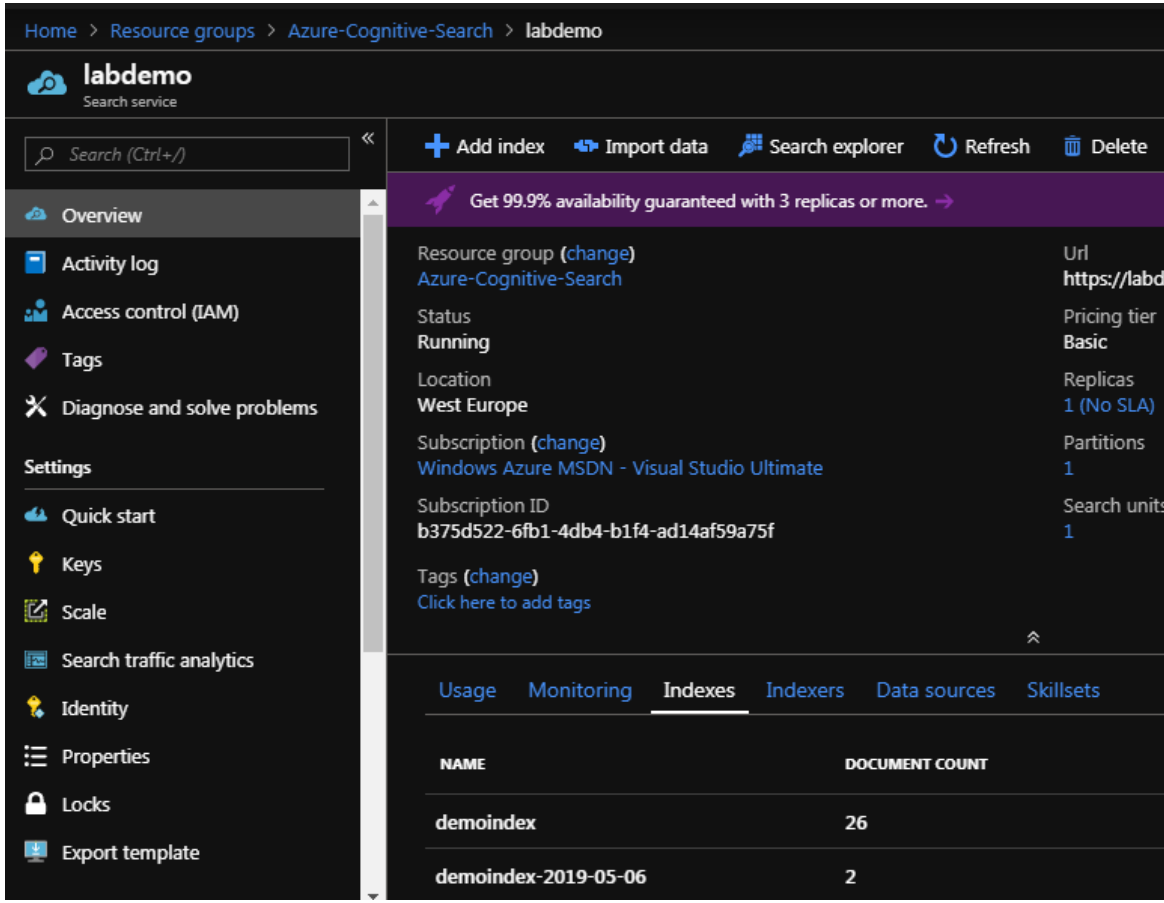
80%

of business relevant information is unstructured

Azure Search

Azure Search is a search-as-a-service cloud solution that gives developers APIs and tools for adding a rich search experience over private, heterogeneous content in web, mobile, and enterprise applications

- Management free
- Keyword search
- Faceting
- Language analyzers
- Geospatial support
- Suggestions/auto-complete
- Customizable scoring
- Proximity search
- Synonyms, etc.



Doc ID: 6606909

SECRET

M/R: On 26 July Mr. Yeates of the NSA Legislative Affairs Office, with concurrence from Mr. Rudolph, D/Chief V, asked Mr. Welday, Chief T12 to effect a search of certain 1962 SIGINT product records held by the SIGINT Repository; the results of which would be used to respond to a query from Mr. Blakey in connection with the House Select Committee on Assassinations investigation into the circumstances surrounding the death of President Kennedy.

Using the guidelines described in the memo the following microfilm copy of product was searched, by T1244 people.

1. QOC Series

Reports - 3/O/QOC/R1-130
Report #3-10, 86, 97, 110, 111 and 128 not held in Repository.

2/O/QOC/R1-13
Report #8 not held in Repository.

2X/O/QOC/R1-352
Report #48, 49, 228, 248, 266, 308 - 317 and 350 not held in Repository. Report #178 - 180 not readable on film (date of issue between 8 August and 20 September 1962 therefore did not recall H.C.)

Translations -
3/O/QOC/T1-150 Trnslation #99 not held in Repository.

2/O/QOC/T1-76

2X/O/QOC/T1-2957 Translation #1057 - 1065, 1182, 1313, 2102, 2128, and 2250 not held in Repository.

2. QOY Series

Reports - 3/O/QOY/R1-11 Report #3, 8, and 10 not held in Reporitory.

2/O/QOY/R1

2X/O/QOY/R1 Report #11 not held in Repository.

Translations -
3/O/QOY/T1-50

2/O/QOY/T2 Translation #1 not held in Repository.

2X/O/QOY/T1-519 Translation #42-45, 92, 177, 253, 260, 261, and 434 not held in Repository. Translations 500-507 not readable on film (date of issue Oct 1962, therefore did not recall H.C.)

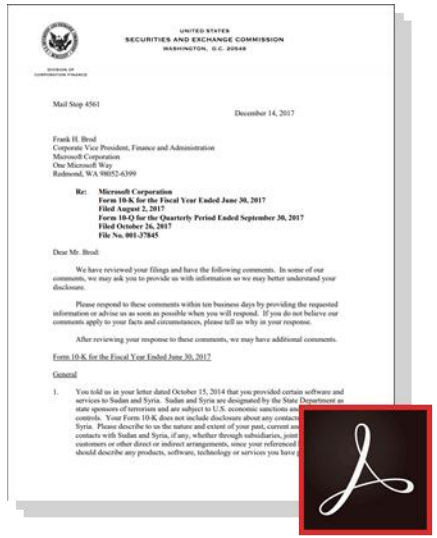
3. QOF Series

Reports - 3/O/QOF/R1-39

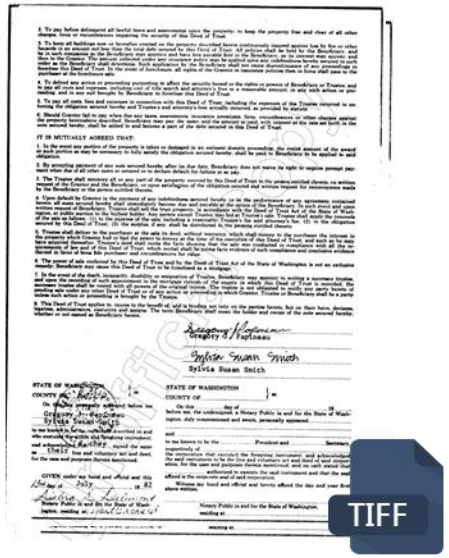
2/O/QOF/R1-225

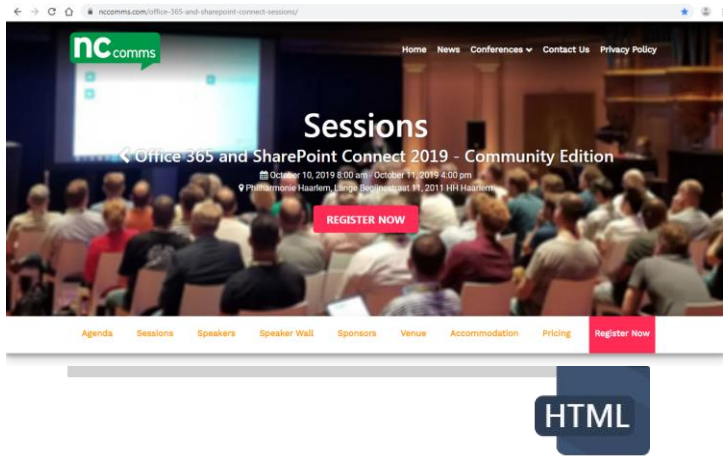
SECRET

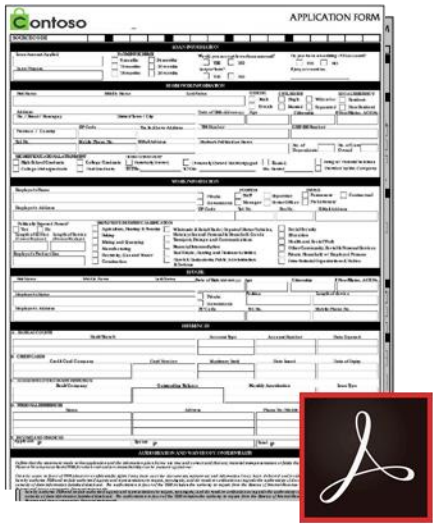
~~EXCLUDE VIA SIGINT CHANNELS ONLY~~

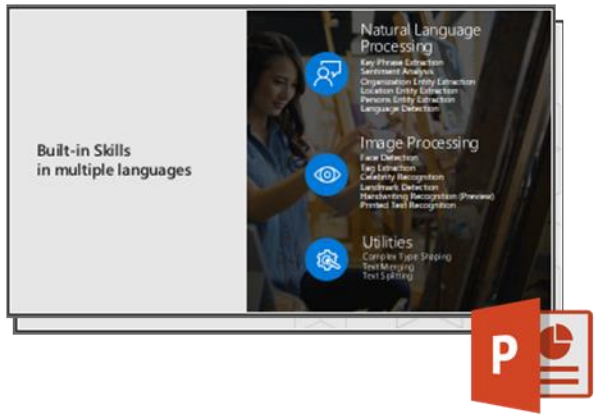

UNITED STATES
SECURITIES AND EXCHANGE COMMISSION
WASHINGTON, D.C. 20549
December 14, 2017
Frank H. Brod
Corporate Vice President, Finance and Administration
Microsoft Corporation
One Microsoft Way
Redmond, WA 98052-6399
Re: Microsoft Corporation
Form 10-K for the Fiscal Year Ended June 30, 2017
Filed August 2, 2017
Form 10-Q for the Quarterly Period Ended September 30, 2017
Filed October 26, 2017
File No. 001-37845
Dear Mr. Brod:
We have reviewed your filings and have the following comments. In some of our comments, we may ask you to provide us with information so we may better understand your disclosure.
Please respond to these comments within ten business days by providing the requested information or advise us as soon as possible when you will respond. If you do not believe our comments apply to your facts and circumstances, please tell us why in your response.
After reviewing your response to these comments, we may have additional comments.
Form 10-K for the Fiscal Year Ended June 30, 2017
General
1. You told us in your letter dated October 15, 2014 that you provided certain software and services to Sudan and Syria. Sudan and Syria are designated by the State Department as state sponsors of terrorism and are subject to U.S. economic sanctions and
Syria. Please describe to us the nature and extent of your past, current and
contacts with Sudan and Syria, if any, whether through subsidiaries, joint
customers or other direct or indirect arrangements, since your referenced
should describe any products, software, technology or services you have

team could also provide more quick resolution to cases with a more complete personalization platform.
High Level Components
Starting with the previously mentioned components, we can create a more complete personalization solution. It should be emphasized that a single platform that brings all these experiences together (omnichannel commerce platform) brings a much higher value to the end customer than each of the components/channels working separately.
Dynamics CRM
Dynamics AX
Self-checkout Application Framework
Personalization Platform
Recommendations
User Segmentation
Personalized Advertising
Multi World Testing Decision Service
Customer Churn
Buyer Propensity
Voice of the customer insights
Personalization can occur in multiple forums. It can be delivered digitally (mobile and web stores), on face-to-face interactions (such as those on a physical store), over phone-to-phone encounters (such as CRM experiences) or over other mediums of communication such as email/sms.
Mobile and Web experiences
CIRP integrates seamlessly with existing mobile and web stores. This is extremely important because the front-end stores are not only the output channel for personalization, but also the input channel for the signals required to personalize in the first place.
Microsoft doesn't own a mobile or online store front-end experience. This m
provide a complete turn key solution. The mitigation for this is to plug into e
platforms. Many of the top e-commerce platforms support plug-in market pl
certainly partner with other retail platforms.
certainly partner with other retail platforms.


Sylvia Susan Smith
Sylvia Susan Smith
STATE OF WASHINGTON
COUNTY OF


nccomms.com/office-365-and-sharepoint-connect-sessions/
nc comms
Home
News
Conferences
Contact Us
Privacy Policy
Sessions
Office 365 and SharePoint Connect 2019 - Community Edition
October 10, 2019 8:00 am / October 11, 2019 4:00 pm
Philharmonie Haarlem, Lange Begijnestraat 11, 2011 HH Haarlem
REGISTER NOW
Agenda
Sessions
Speakers
Speaker Wall
Sponsors
Venue
Accommodation
Pricing
Register Now


Contoso
APPLICATION FORM


Built-in Skills
in multiple languages
Natural Language Processing
Key Phrase Extraction
Sentiment Analysis
Organization Entity Extraction
Location Entity Extraction
Persons Entity Extraction
Language Detection
Image Processing
Face Detection
Tag Extraction
Celebrity Recognition
Landmark Detection
Handwriting Recognition (Preview)
Printed Text Recognition
Utilities
Text Merging
Text Splitting

...

Scott Guthrie

Title:
Executive Vice President, C+E

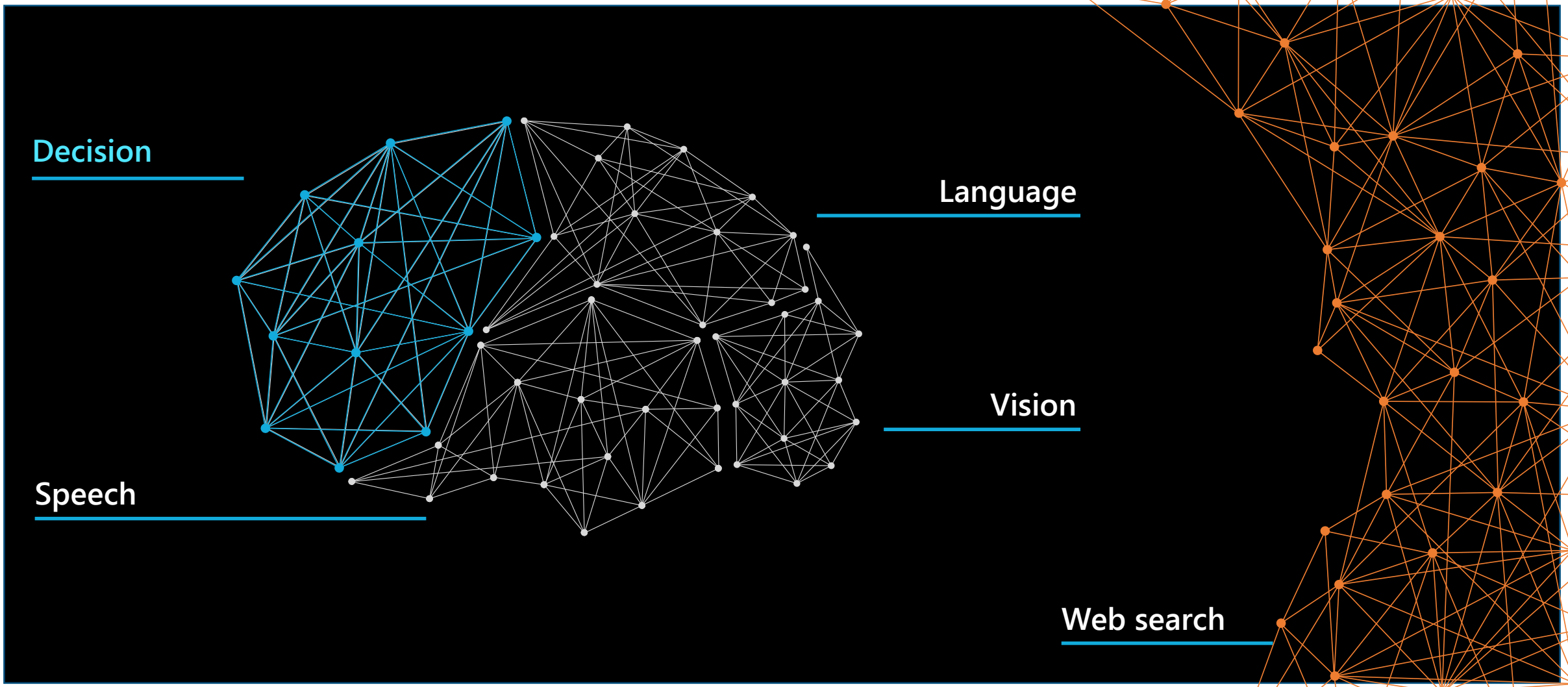
Company:
Microsoft

accent color: blue?

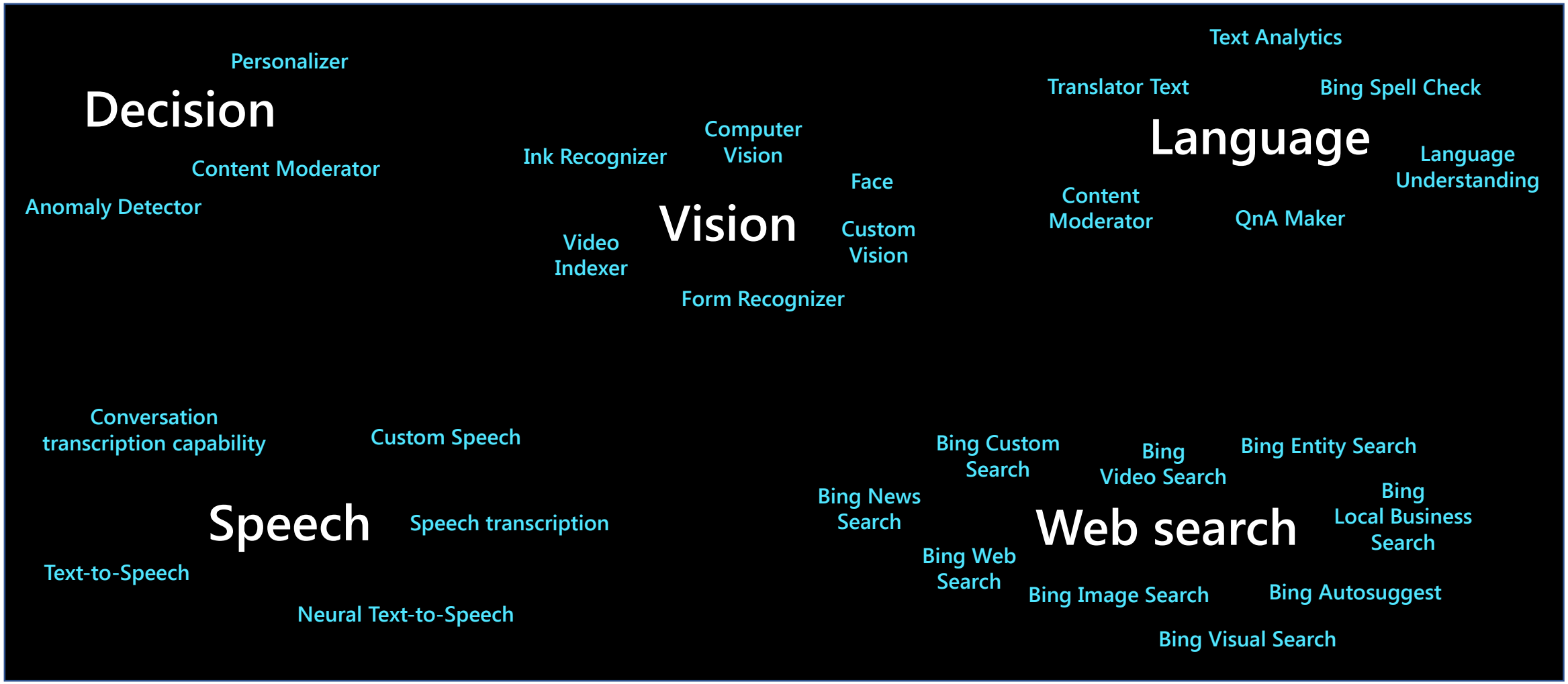
Azure Cognitive Search

DEMO

Azure Cognitive Search



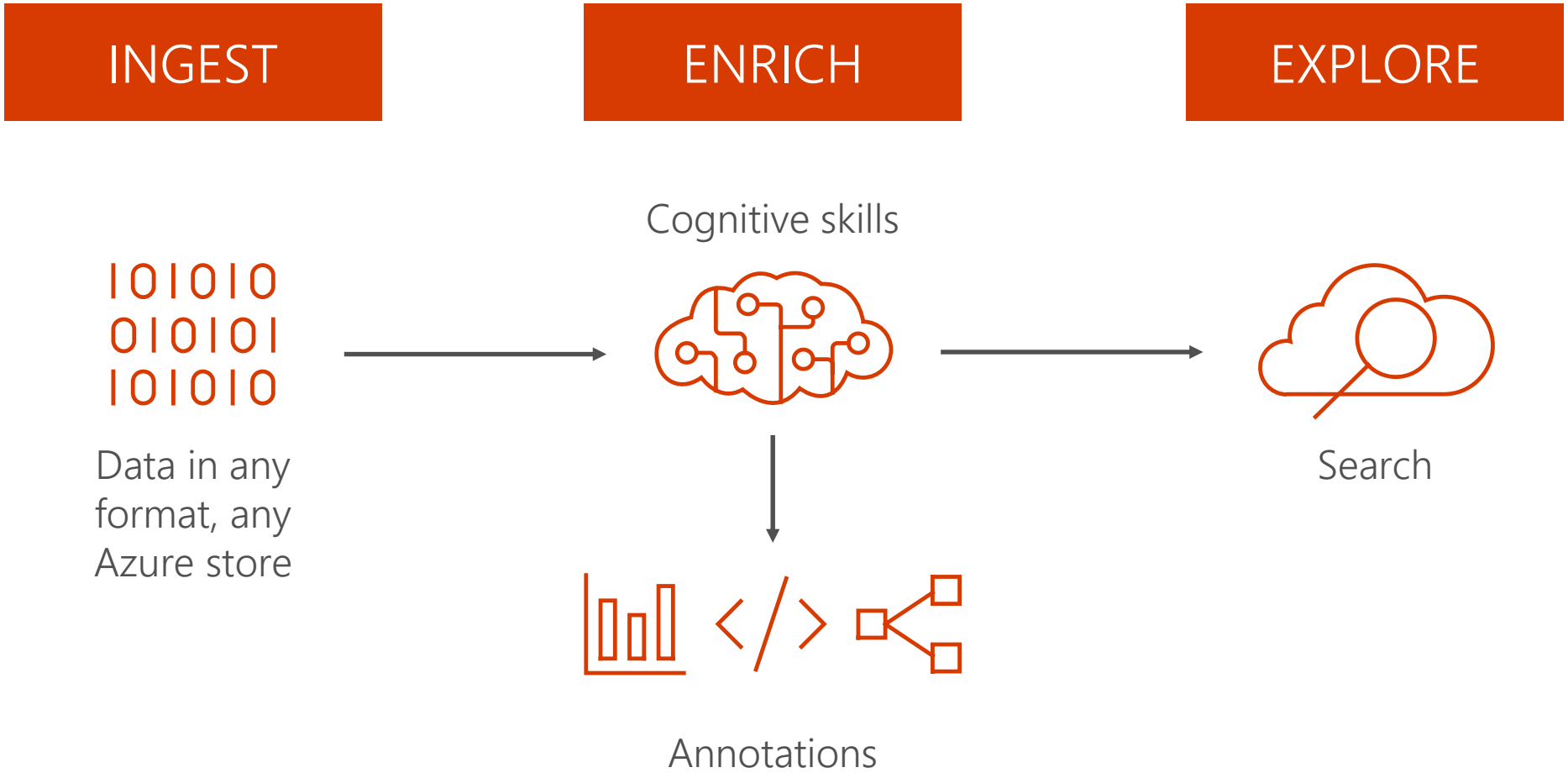
Azure Cognitive Services



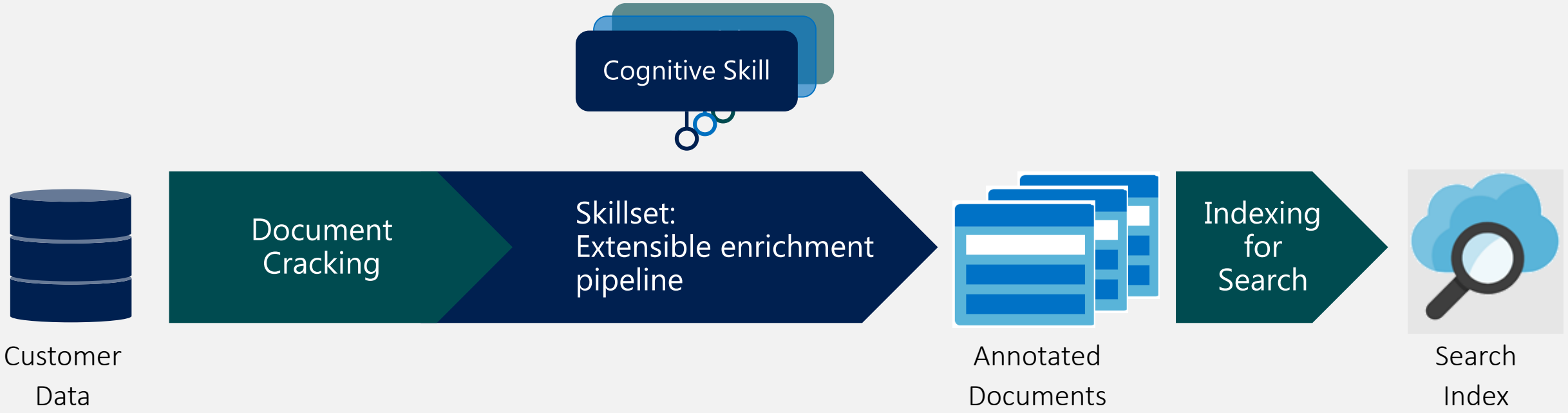
Cognitive Services

DEMO

At high level

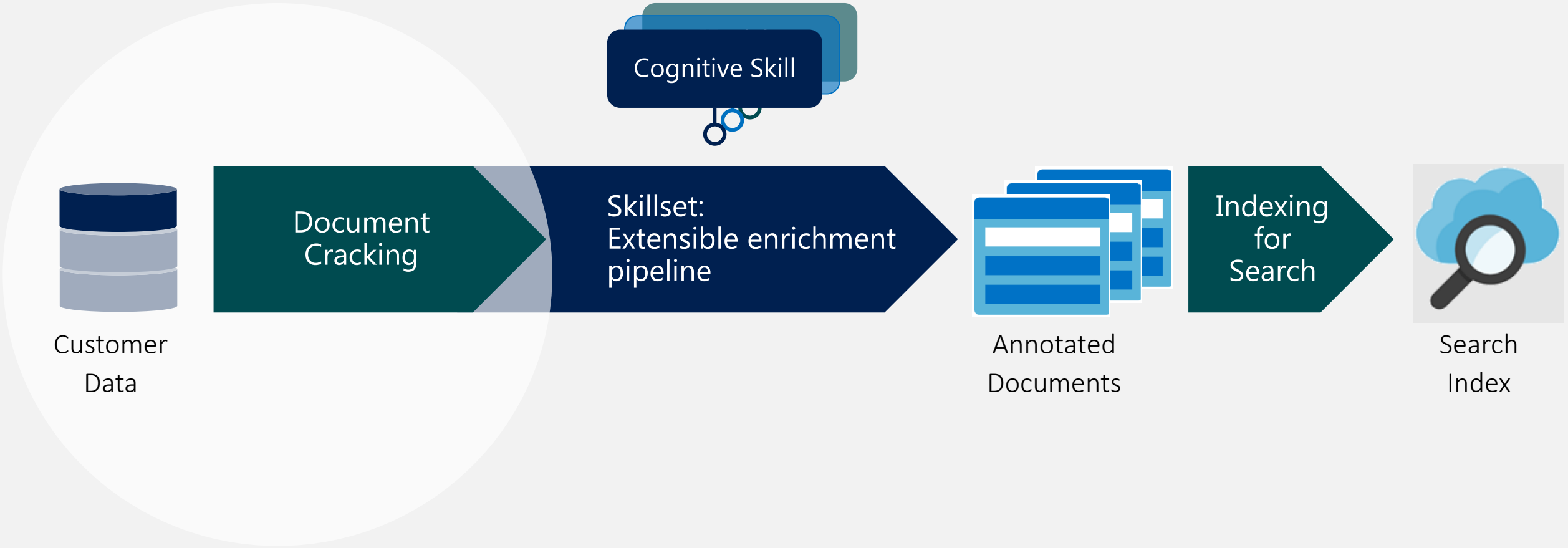


Cognitive Search Architecture



INGEST | ENRICH | EXPLORE

Document Cracking



INGEST | ENRICH | EXPLORE

Document Cracking

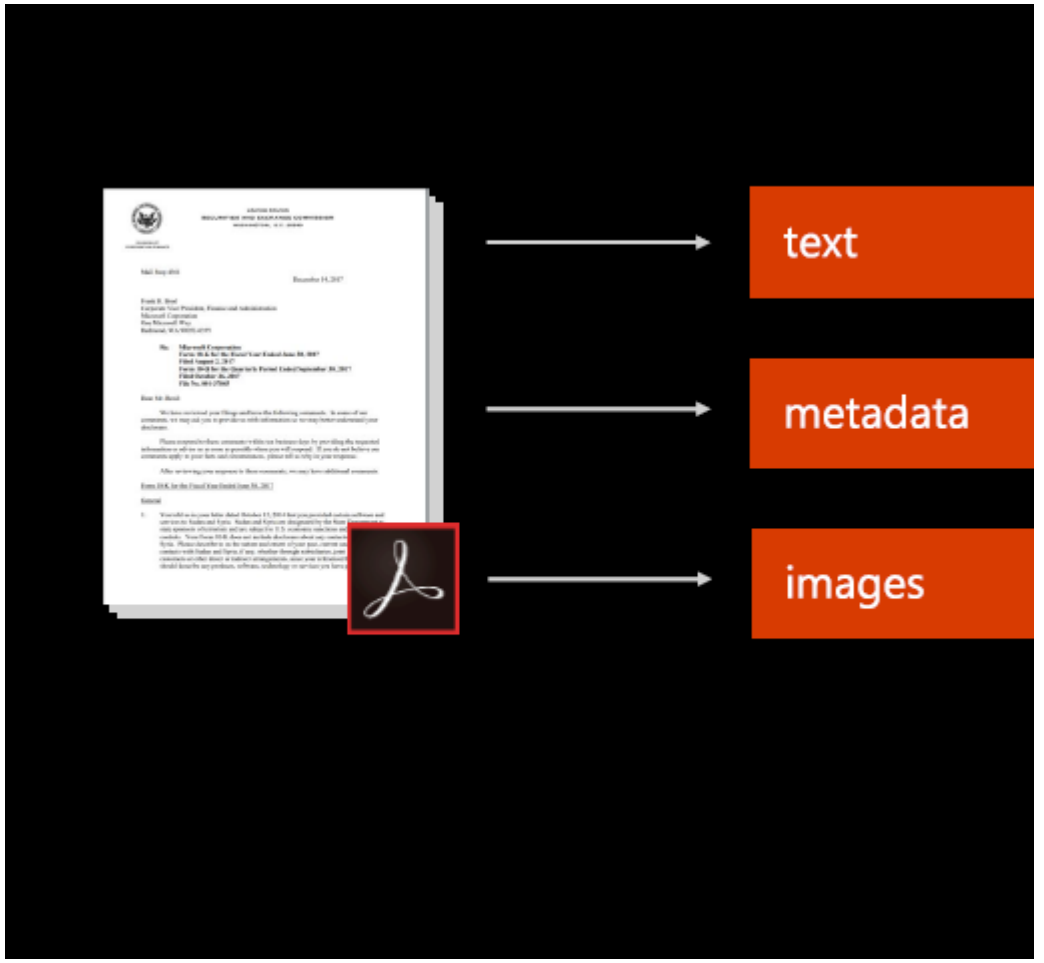
Different types of data sources:

- Azure Blob Storage
- Azure SQL Databases
- SQL Server on Azure VM
- Azure Cosmos DB
- Azure Tables

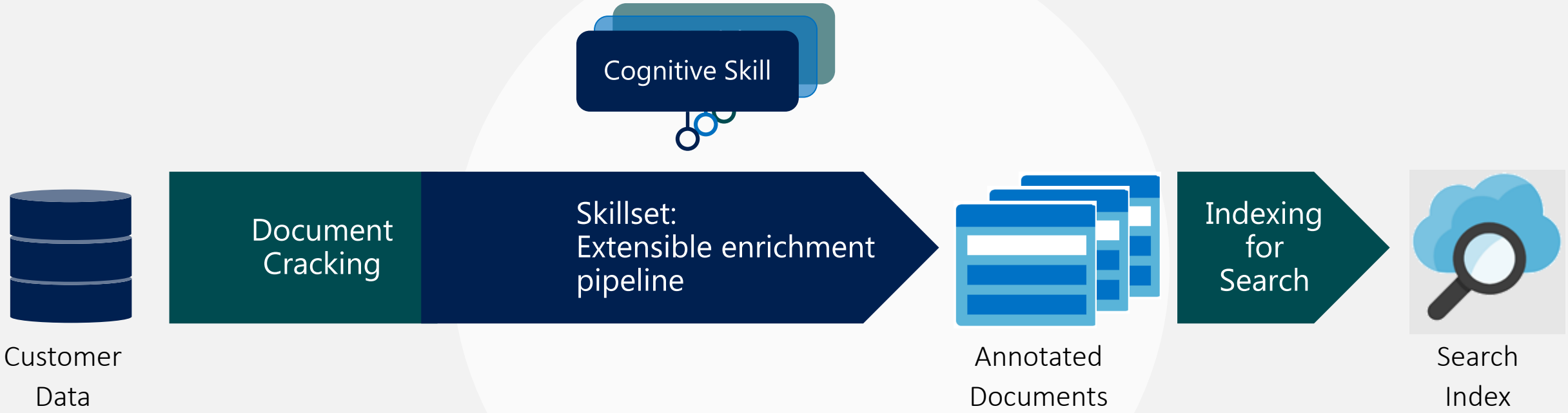
File formats supported in Blob storage:

- pdf, xml,
- png , rtf, json, html, doc, ppt, xls

Data push via .NET SDK for non Azure sources



Indexing Pipeline



INGEST | ENRICH | EXPLORE

Azure Cognitive Skills

Built-in skills

Key Phrase extraction

Location entity extraction

Sentiment analysis

Organization entity extraction

Persons entity extraction

Language detection

Face detection

Celebrity recognition

Image tag extraction

Text Utilities

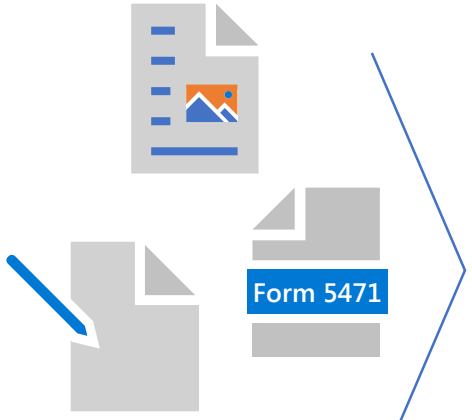
Landmark detection

Printed text recognition

Custom skills

Azure Search + Cognitive Services

Information



Pre-trained ML Models

- Key Phrase extraction
- Location entity extraction
- Sentiment analysis
- Organization entity extraction
- Persons entity extraction
- Language detection
- Face detection
- Celebrity recognition
- Tag extraction
- Custom skills
- Landmark detection
- Printed text recognition

Fully text-searchable rich index

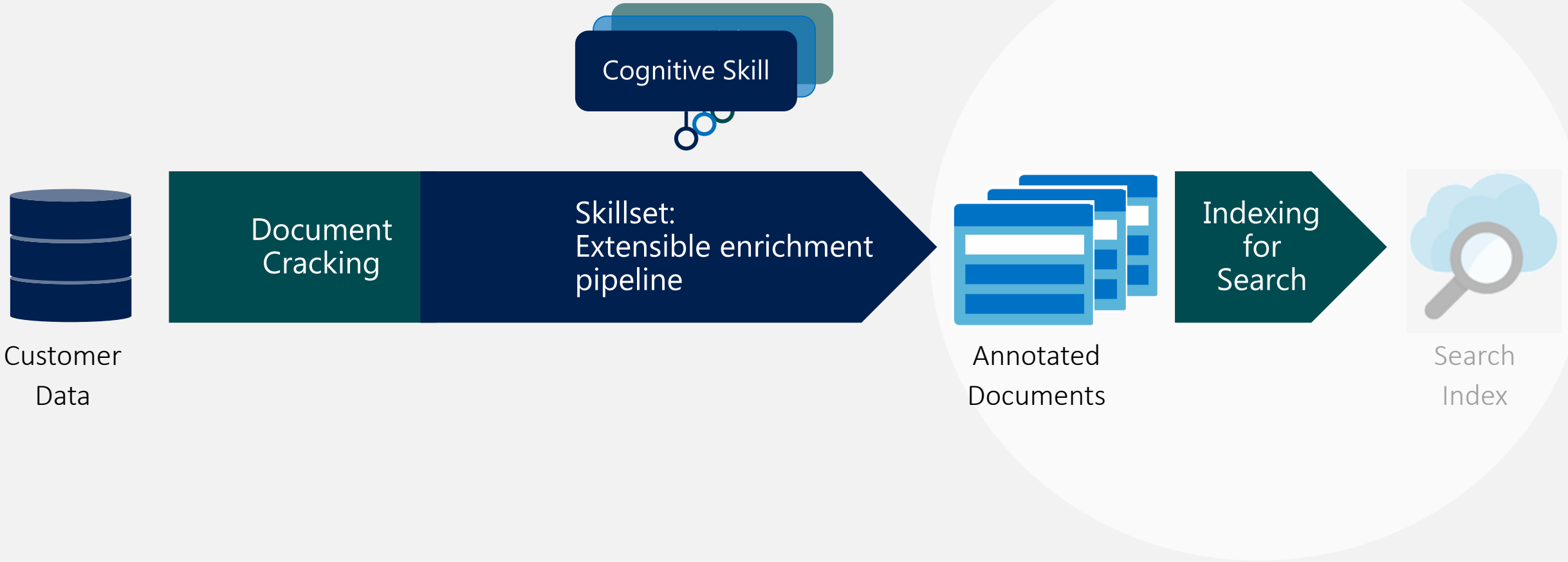
Azure Search

INGEST | ENRICH | EXPLORE

Azure Search Portal

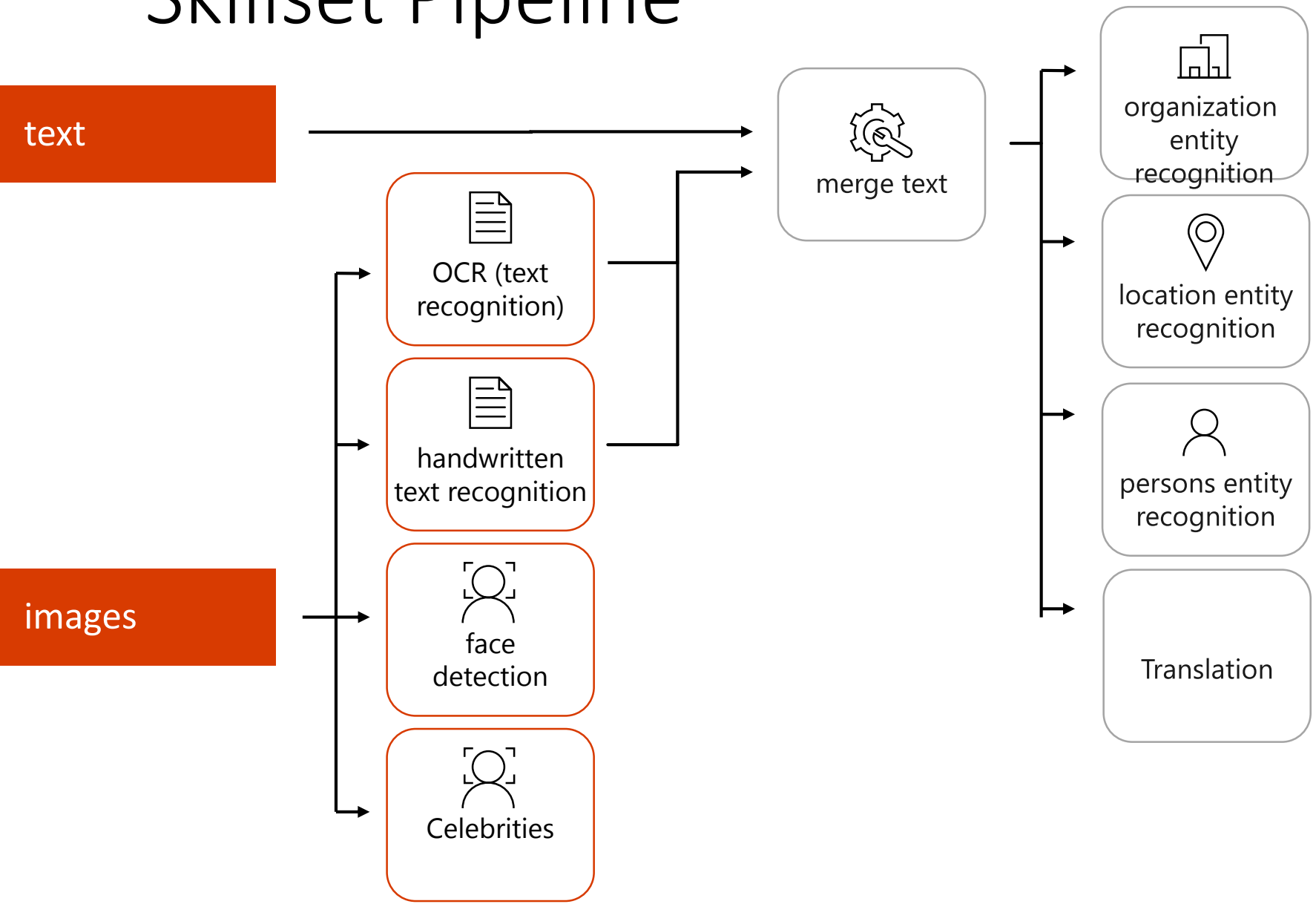
DEMO

Index and document annotation



INGEST | ENRICH | EXPLORE

Skillset Pipeline



Define the Skillset

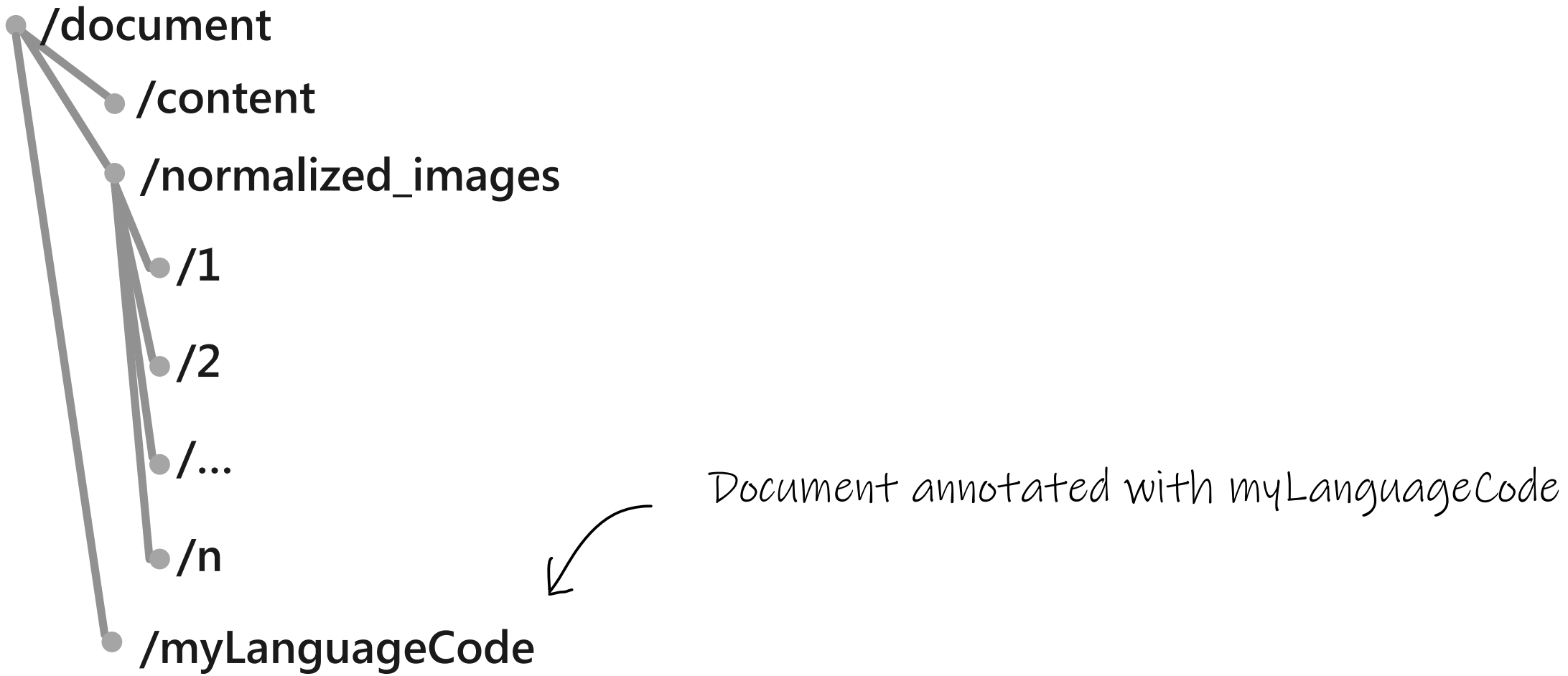
```
PUT https://demo.search.windows.net/skillsets/demoskillset?api-version=2019-05-06

"skills": [

  {
    "@odata.type": "#Microsoft.Skills.Text.LanguageDetectionSkill",
    "inputs":
    [
      {  "name": "text", "source": "/document/content" }
    ],
    "outputs":
    [
      {  "name": "languageCode", "targetName": "myLanguageCode" },
      {  "name": "languageName", "targetName": "myLanguageName" }
    ]
  },
```

Annotate the output in the document

Document Annotations



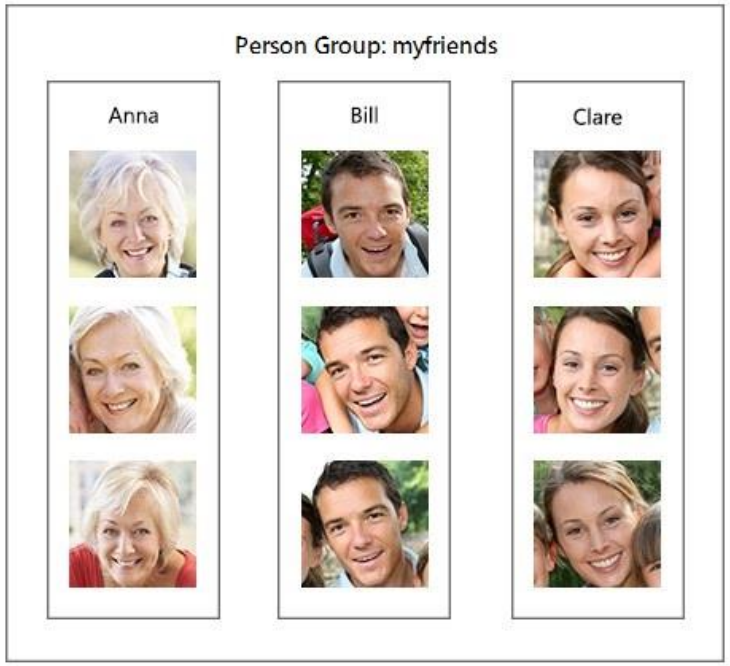
Skillset

```
…,
{
  "@odata.type": "#Microsoft.Skills.Text.NamedEntityRecognitionSkill",
  "categories": [ "Organization" ],
  "defaultLanguageCode": "en",
  "inputs":
  [
    {  "name": "text", "source": "/document/content" },
    {  "name": "languageCode", "source": "/document/myLanguageCode" }
  ],
  "outputs":
  [
    {  "name": "organizations", "targetName": "organizations" }
  ]
},
```

Now you can use it as an input in another skill!

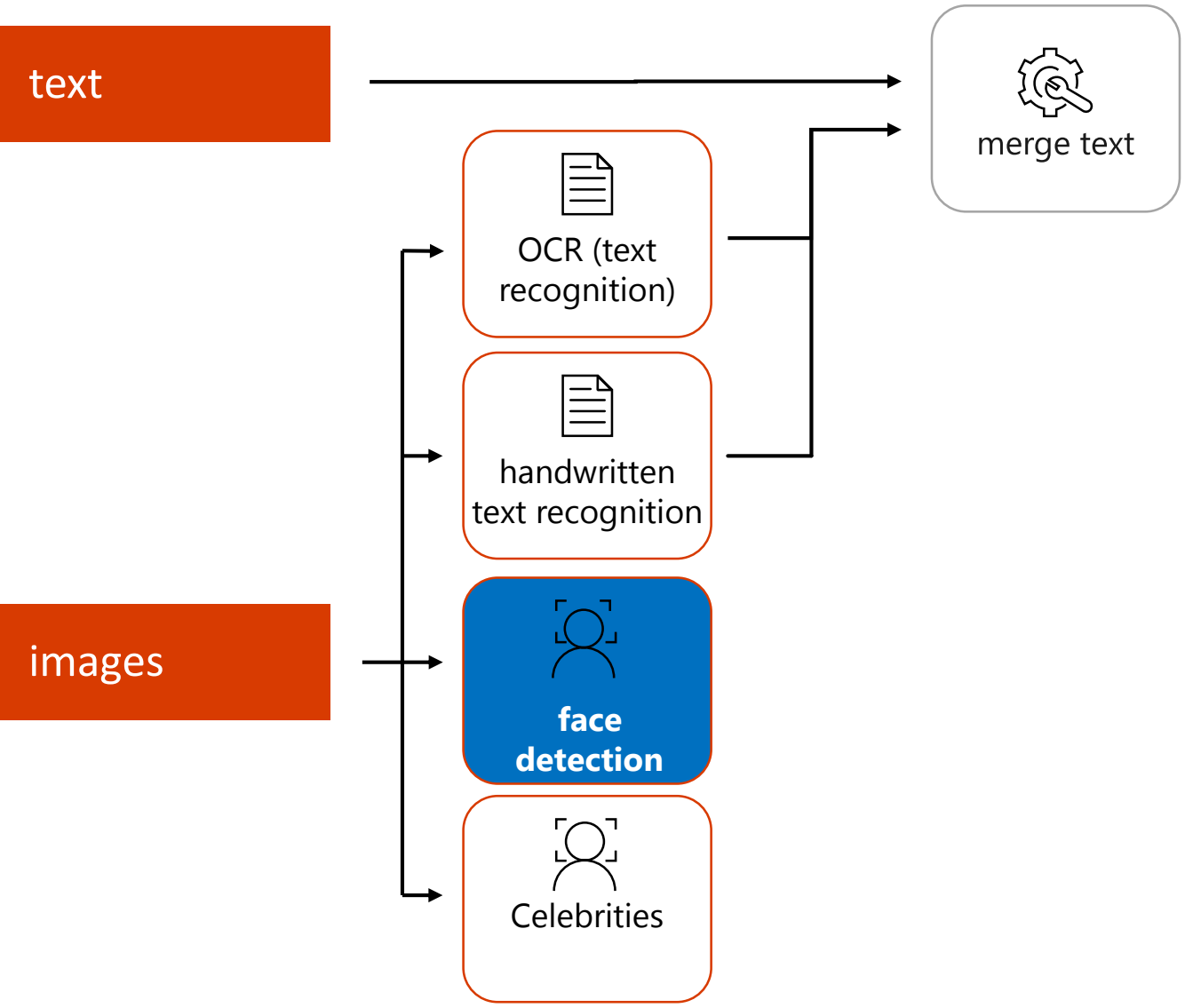
Building the Skillset, Index, Indexer

DEMO



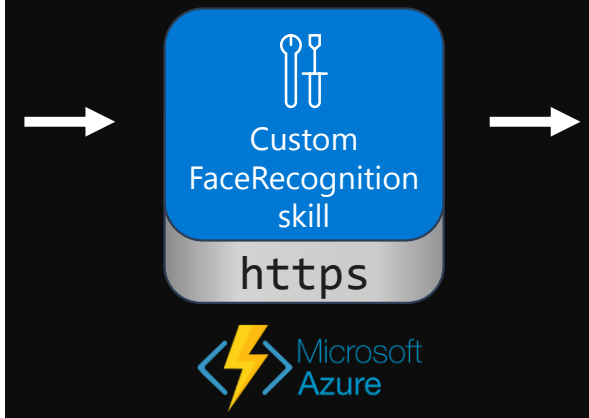
The Azure Face API is a cognitive service that provides algorithms for detecting, recognizing, and analyzing human faces in images

Skillset Pipeline



Custom Skill with Azure Functions

```
{
  "values": [
    {
      "recordId": "7cad2",
      "data":
         {
           “text":
               ImageBase64Encoded"
         }
    },
    {
      "recordId": "7cad3",
      "data":
         {
           “text":
             “ImageBase64Encoded",
         }
    },
    …
  ]
}
```




```
{
  "values": [
    {
      "recordId": "7cad2",
      "data":
         {
         “text":
           “Donald Hessing"
         }
    },
    {
      "recordId": "7cad3",
      "data":
         {
           “text":
             “Jussi Roine"
         }
    },

    …
  ]
}
```

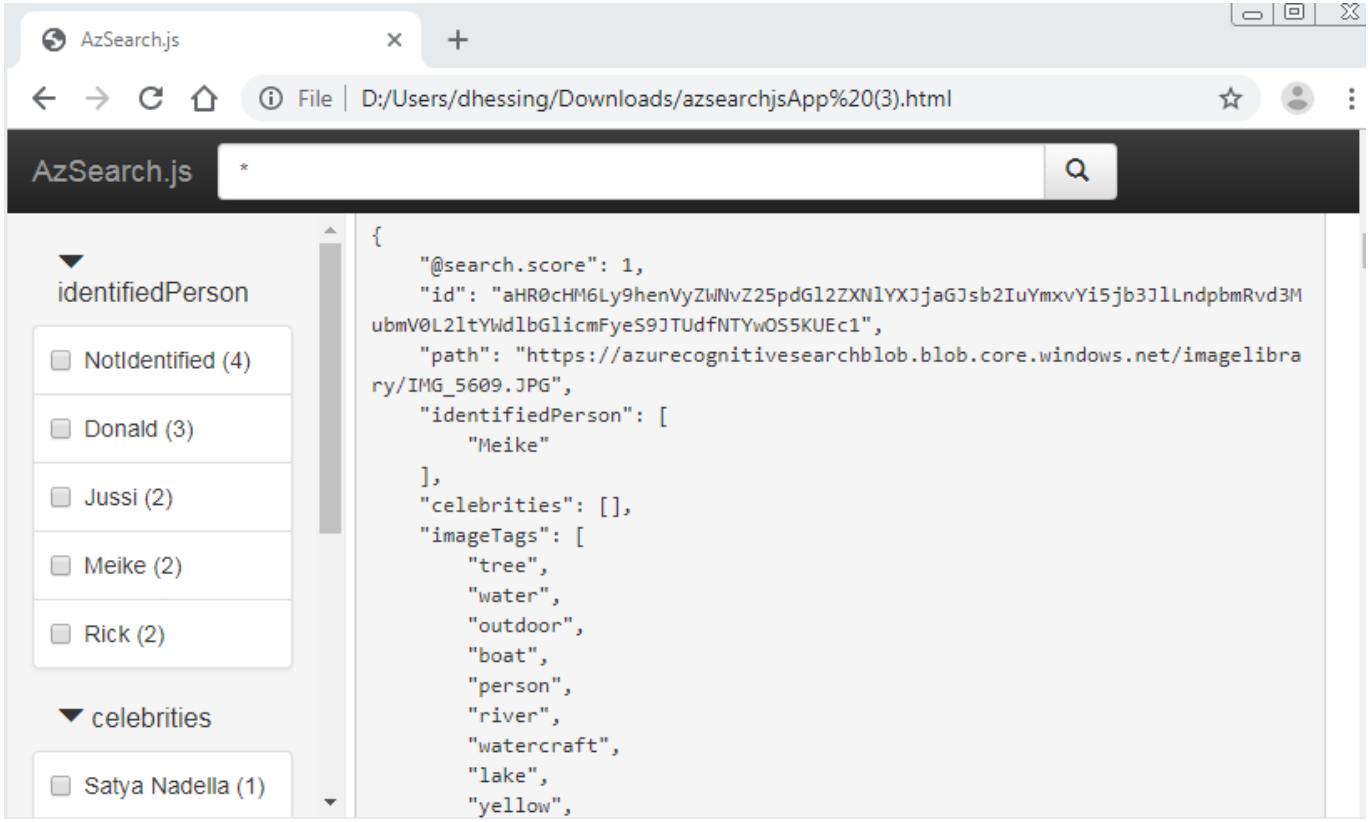
Custom Skill – Using the FACE API with pre-trained images

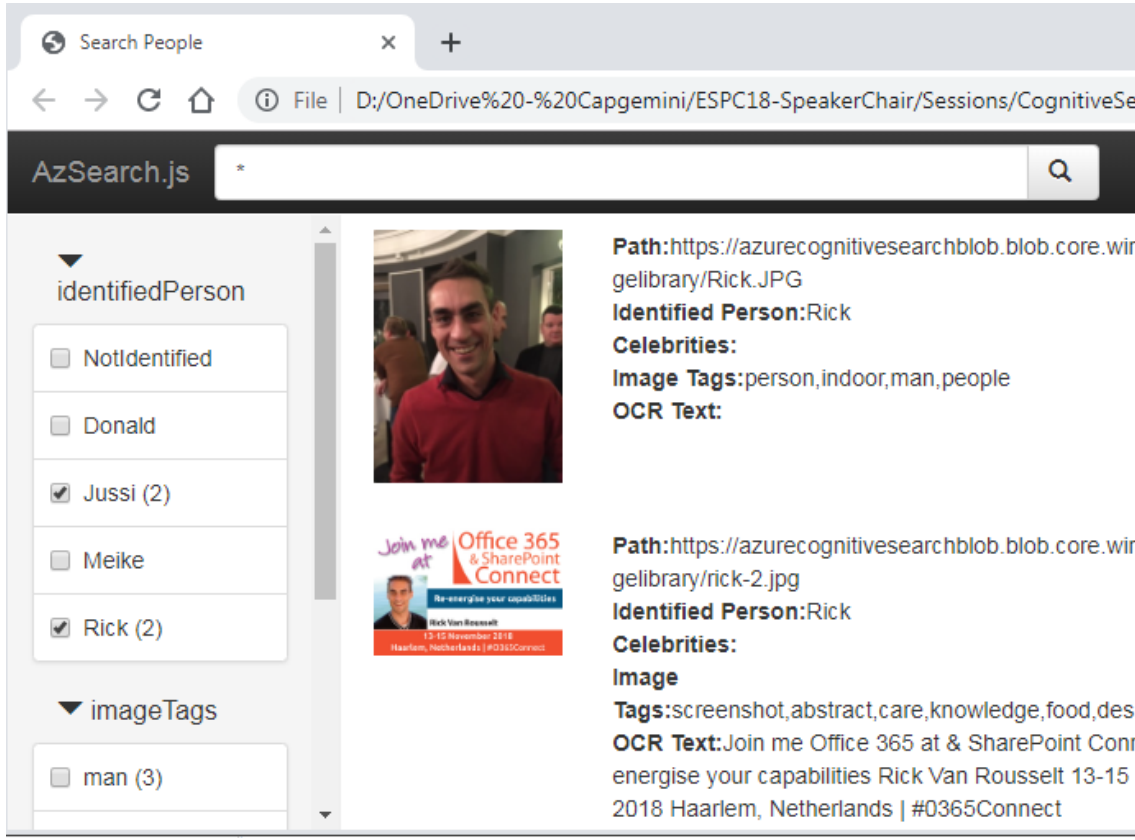
```
{
  "@odata.type": "#Microsoft.Skills.Custom.WebApiSkill",
  "description": "Custom face recognition skill",
  "uri":"https://customskillfaceapi.azurewebsites.net/api/RecognizeFace?code=[Key]"
  ,
  "batchSize": 1,
  “context": “/document",
  "inputs":
  [
    {  "name": "text", "source": "/document/normalized_images/*" },
    {  "name": “path", , "source": "/document/path" }
  ],
  "outputs":
  [
    {  "name": “text", "targetName": "identifiedPerson" }
  ]
}
```

Custom Skill People recognition

DEMO

Building the Azure Search UI - JavaScript





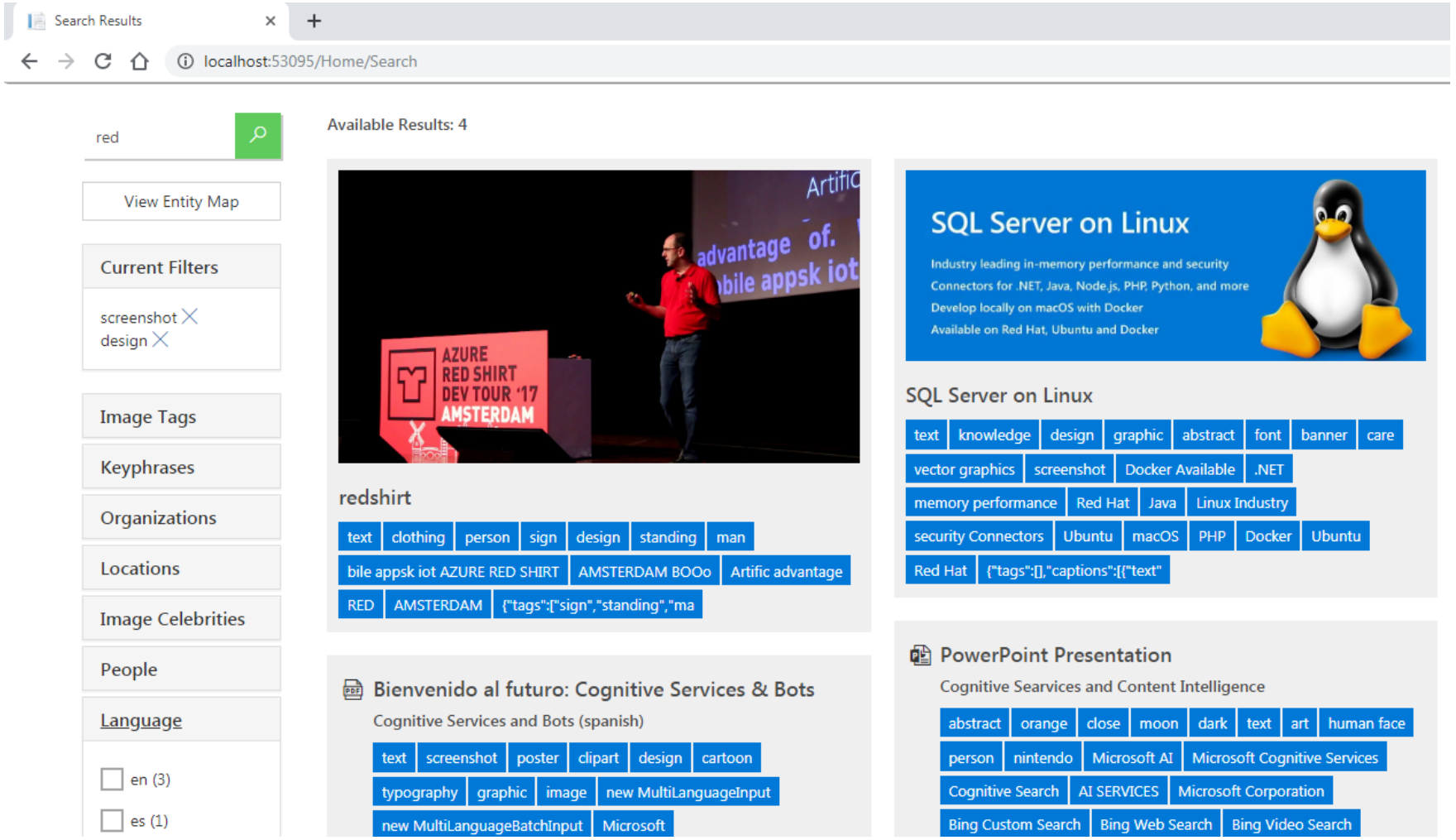
Generate your UI:

http://azsearchstore.azurewebsites.net/azsearchgenerator/index.html

Update the generated your UI:

https://github.com/Yahnoosh/AzSearch.js

Building the Search UI – ASP.NET



https://github.com/Azure-Samples/azure-search-knowledge-mining/tree/master/02%20-%20Web%20UI%20Template

#SPSBE

BIWUG

SharePoint Saturday Belgium 2019

#SPSBE

http://spsbe.be

Please rate this session!

BIWUG