Effective Utilization of Azure Cognitive Services in Search Operations

Suresh K C
Suresh.channegowda@wipro.com

Table of Contents

1. Problem Statement 2
2. Solution Approach 3
1.1. Azure Cognitive Search 3
1.2. Solution Architecture 4
1.3. End State Architecture 5
3. Technical Details and Implementation logics 6
3.1. Data/File source and Data inject 6
3.2. Data store 6
3.3. Data Preparation, Train & Indexing Azure Cognitive search 8
4. Sample Search 11
4.1. Sample Search on Chatbot 13
4.2. Sample Search for Computer Vision on Web 15
5. Business Benefits 17

1. Problem Statement

In today's world, have tons of data everywhere, around 80% of the information are in the form unstructured. Only limited organizations have capable to process these huge amounts of structured & unstructured data into Business values. These data processing also limited to data available in digital format.

This is mainly because of limitation of the today's search options. Search indexing are limited to keywords or combination of them with very limited use of AI, ML and NLP. These search index rankings are based on the frequencies of words than the user intent. Results user may not be well experienced with the existing search options.

Azure Cognitive search is revolutionizing the process of search both on structured & unstructured data which will helps the build the business value by empowering the data analyst & unlock the value of the knowledge. Azure Cognitive services has been boosting the search operations.

This blog will give insight on how to configure Azure Cognitive service along with Azure Cognitive Search to demonstrate how to gear-up the search operations

2. Solution Approach

Higher Accuracy – The search has evolved and since there is more unstructured data around, advanced technologies like Semantic Search, AI/ML based advanced learnings needed to provide results and insights with higher accuracy

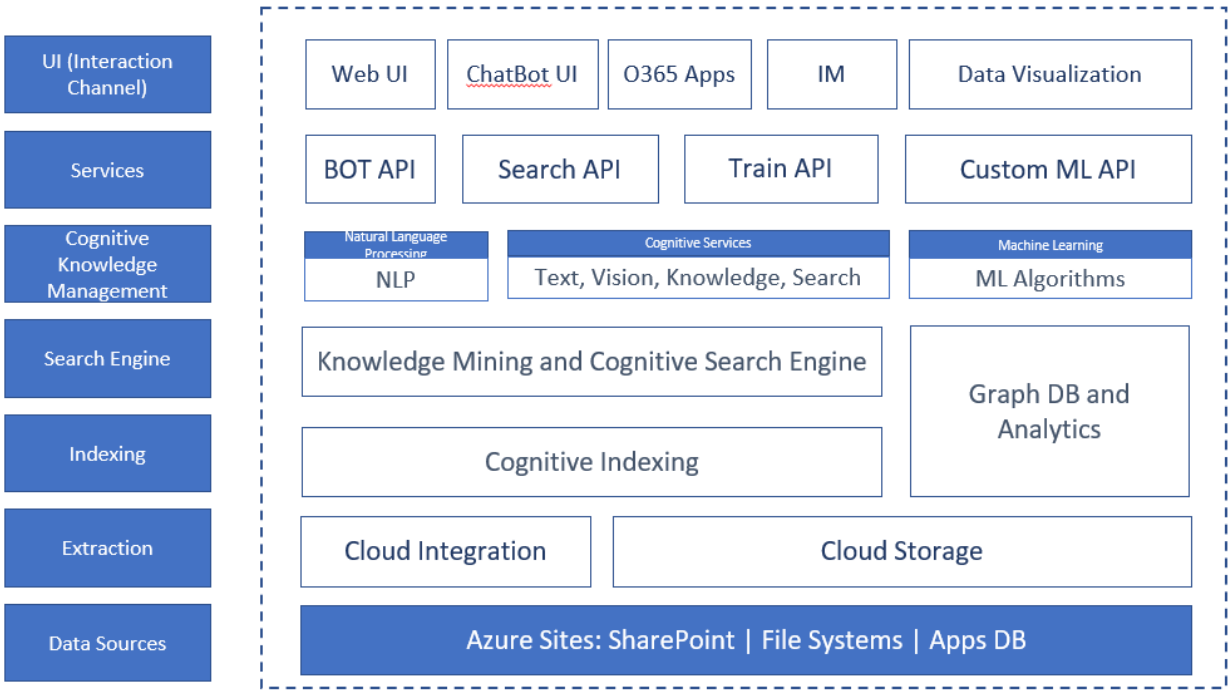
Contextual Search – Most of search engines works on assigning rankings and score to keywords based on frequency rather than a typical text-based search. Algorithmic search worked to an extent but failed as users are looking for learnings, insights from data

Adaptability – AI will help business to improve productivity

1.1. Azure Cognitive Search

- Consolidate heterogeneous content into user-defined search index. Offload indexing and query workloads onto a dedicated search service.
- Relevance tuning of search by **AI enrichment, knowledge mining, cognitive skillsets**
- Transform large undifferentiated documents, text or images, files stored in different data sources into searchable chunks.
- Add linguistic or custom text analysis.
- Configure analyzers to achieve specialized processing of raw content, such as filtering out diacritics, or recognizing and preserving patterns in strings.

1.2. Solution Architecture



Azure PaaS services are extensively useful to implement Azure Cognitive Search services. Service layer can be interacting with Azure search engine and cognitive service APIs to get the relevant search related info and supply to UI.

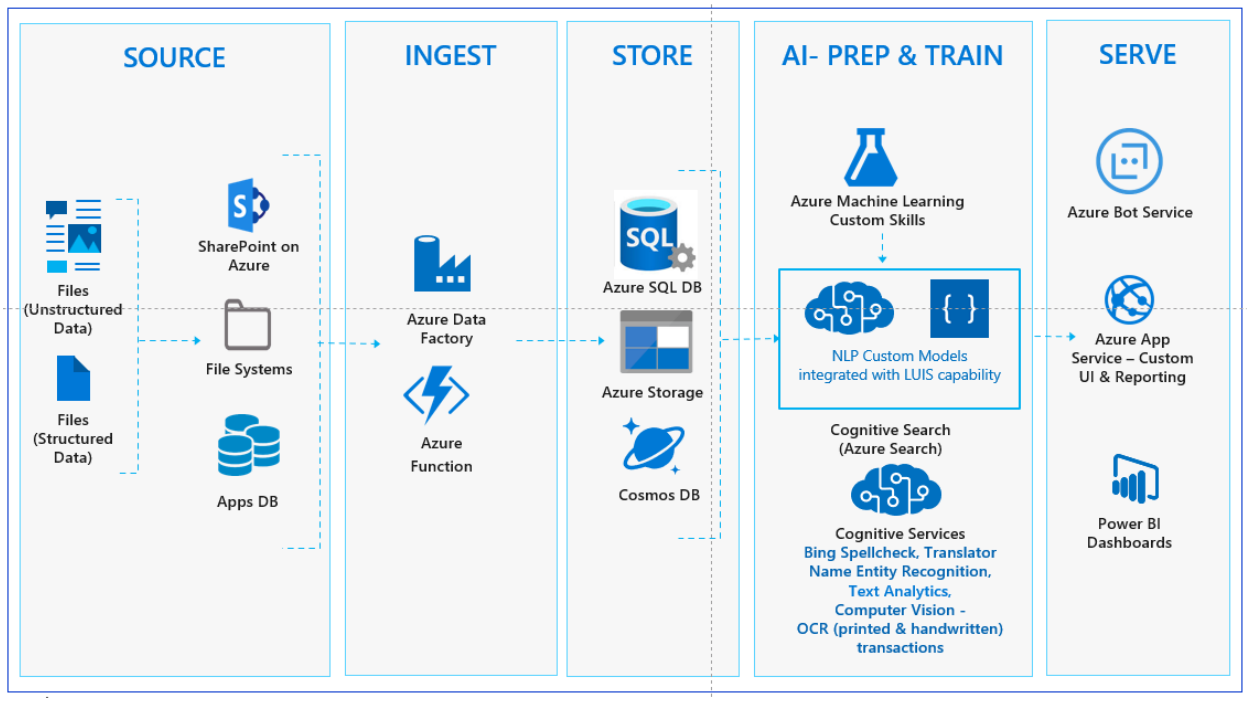
Data sources – Solution supports several kind of data sources like Blob, file system, SharePoint, application' s transactional database like Cosmos Db or SQL DB etc.

Extraction/Ingestion – Azure cognitive search service support only data sources which are available on Azure Cloud. So, searchable data should be available on Azure cloud for the purpose of Azure cognitive search indexing. Cloud Integration component will helps move searchable data/files to Azure cloud. Unstructured data will be moved to Azure Blob storage and structured data will be moved to Azure Table storage/Azure SQL or Cosmos DB.

Data preparation and Train – Indexing – Once the data available on Azure cloud, have to prepare data by applying cognitive skillset to enriched documents, which gains structure and substance as skill writes its output as nodes. Enriched documents are ready for indexing through output field mappings. Any raw content that to be transferred intact, from source to an index, is defined through field mappings. Data preparation can be automated and scheduled for new & delta updates of source content.

Services – Services are options to serve the search contents through the search API. These APIs can be integrated into end user applications like web, chatbot and any other custom applications. Search services accept the input queries for search & share the output search content along with source details as response.

1.3. End State Architecture



3. Technical Details and Implementation steps

3.1. Data/File source and Data inject

Azure cognitive search supports heterogenous data sources. Both structured & unstructured data like excel, CSV files, PDF, images etc can be considered for search operation.

3.2. Data store

Data files are moved from source to Azure storage account. This will help the Azure cognitive search for faster data preparation (enriching, indexing). Structured data stored in Azure table storage and Cosmos DB and unstructured data stored in Azure Blob storage.

The below screenshots depict the data placed on Azure storage account

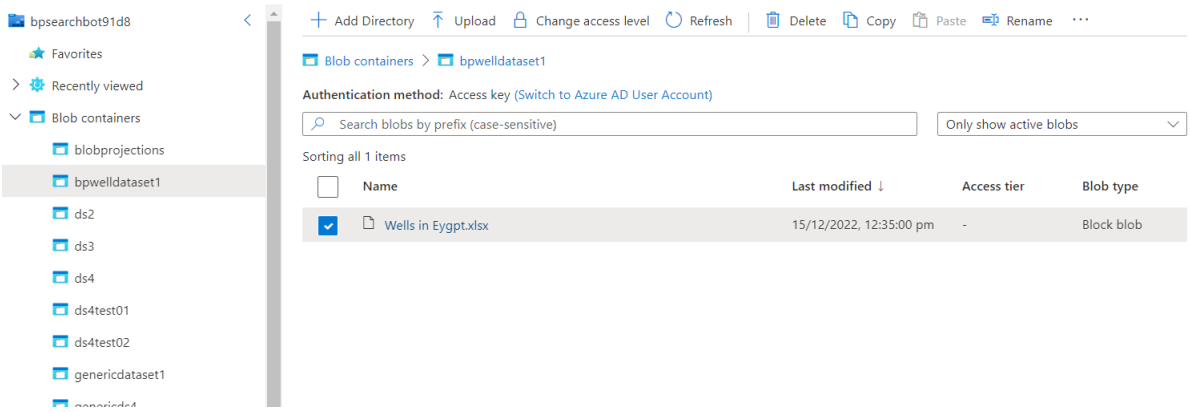


Fig 1. Sample data in the form of excel

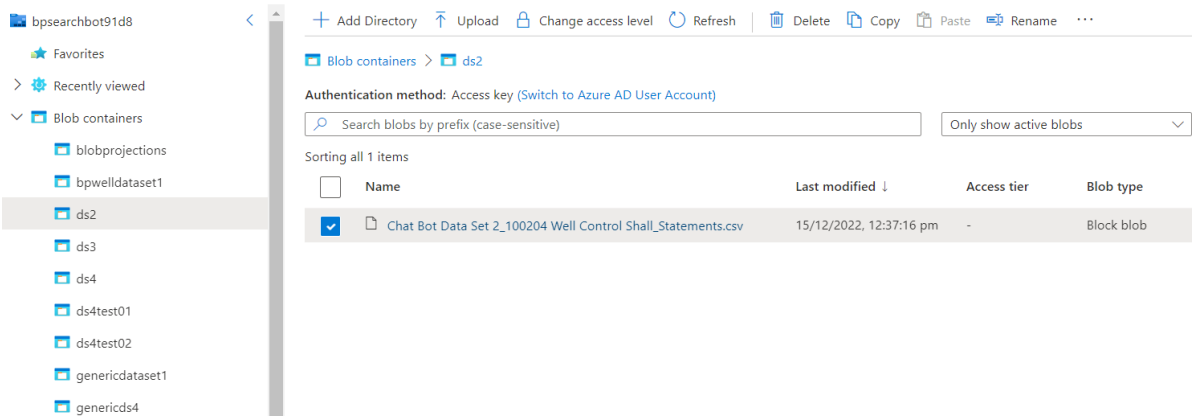


Fig 2. Sample data in the form of CSV

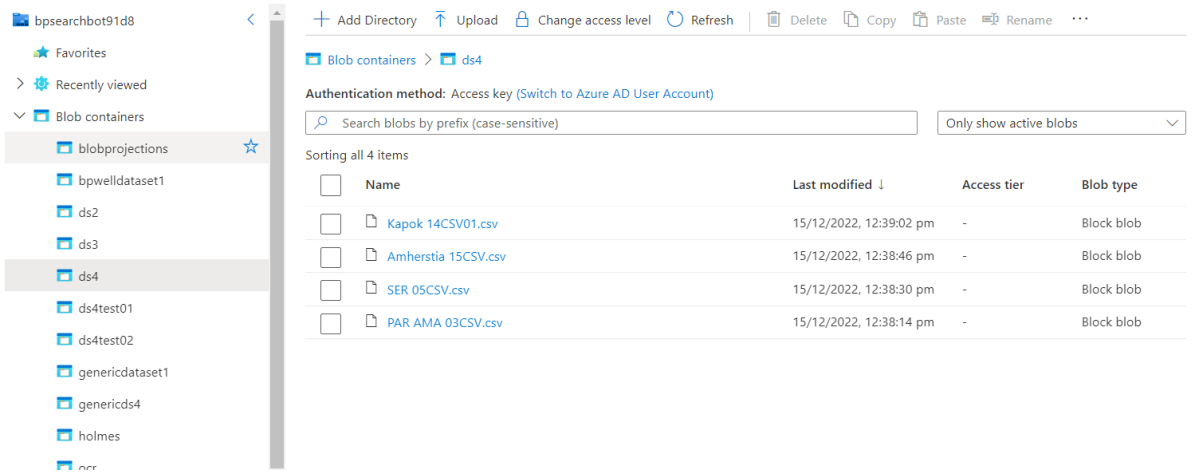


Fig 3. More sample data in CSV format

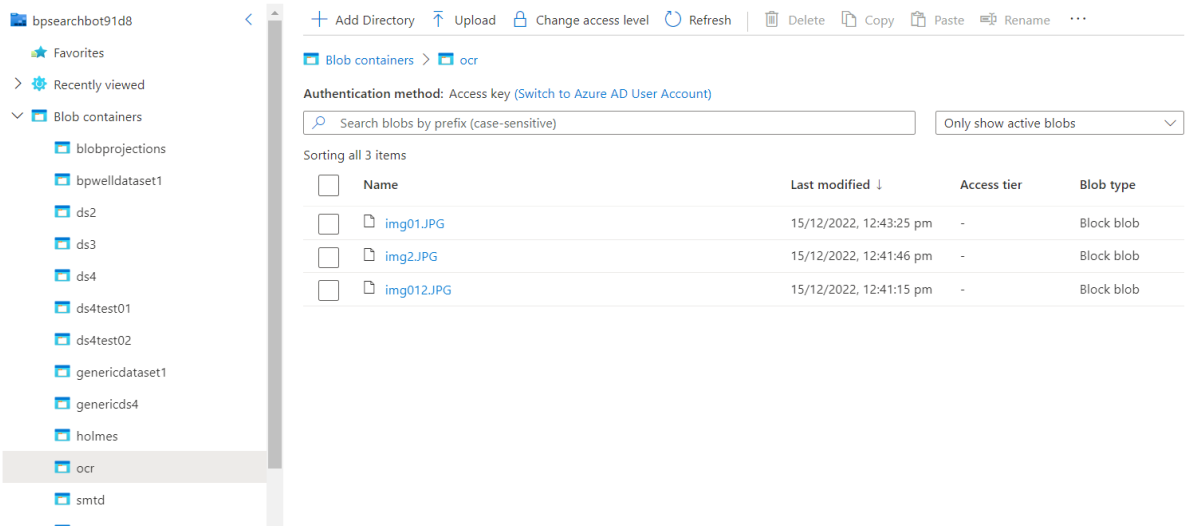


Fig 2. Sample images used for OCR operation with Computer Vision.

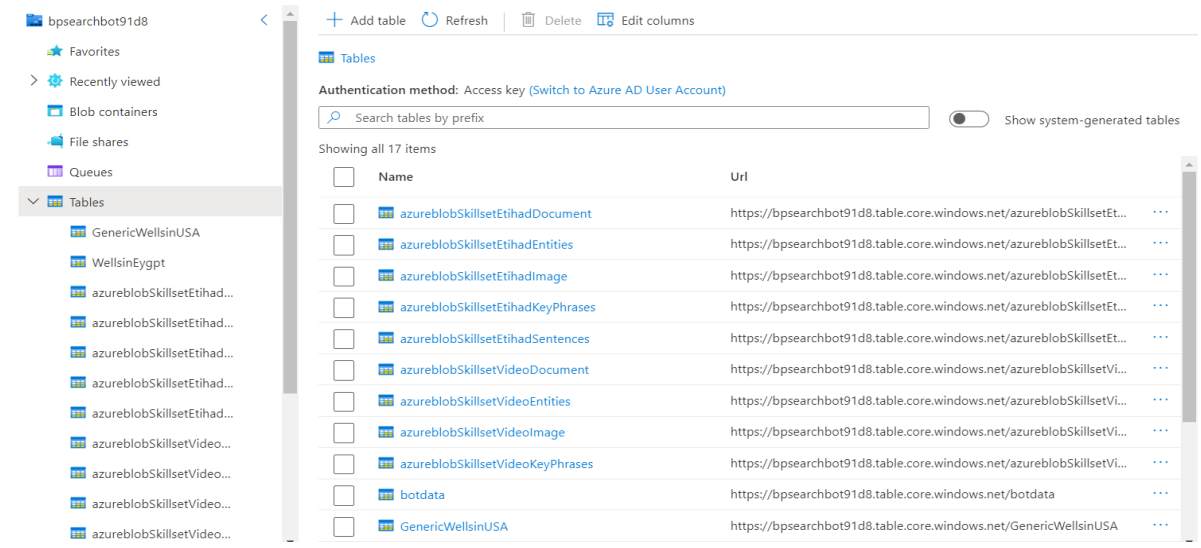
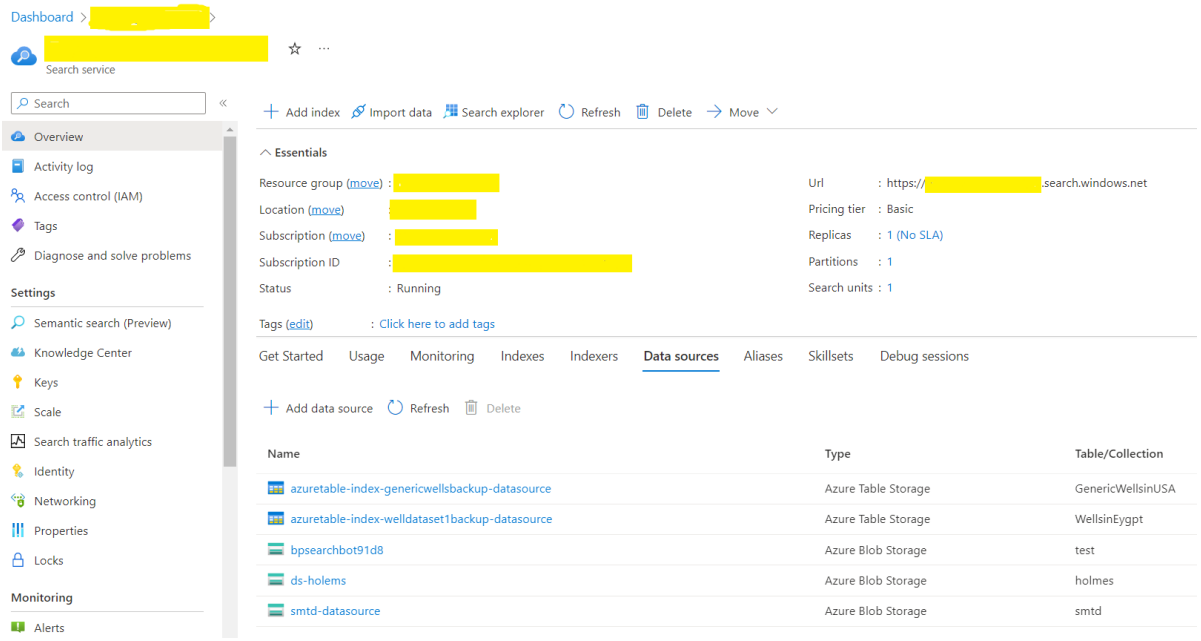


Fig 2. Sample structured data in Azure Table storage

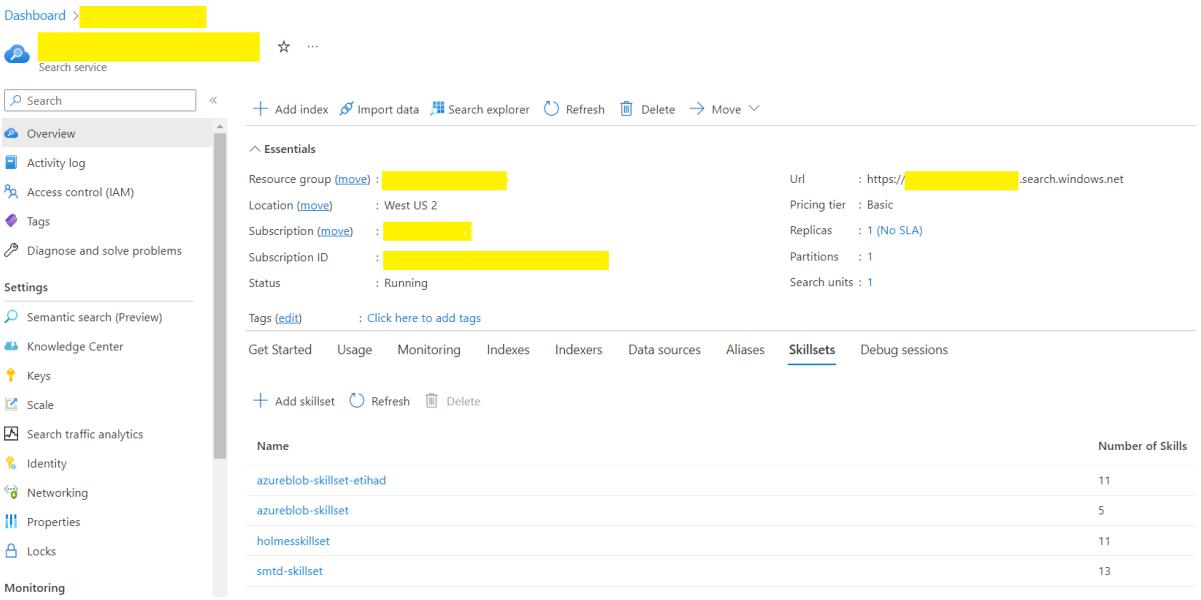
3.3. Data Preparation, Training & Indexing Azure Cognitive search

Once the search service is available. Below are some important steps to be configured as part of Azure Cognitive search Configuration

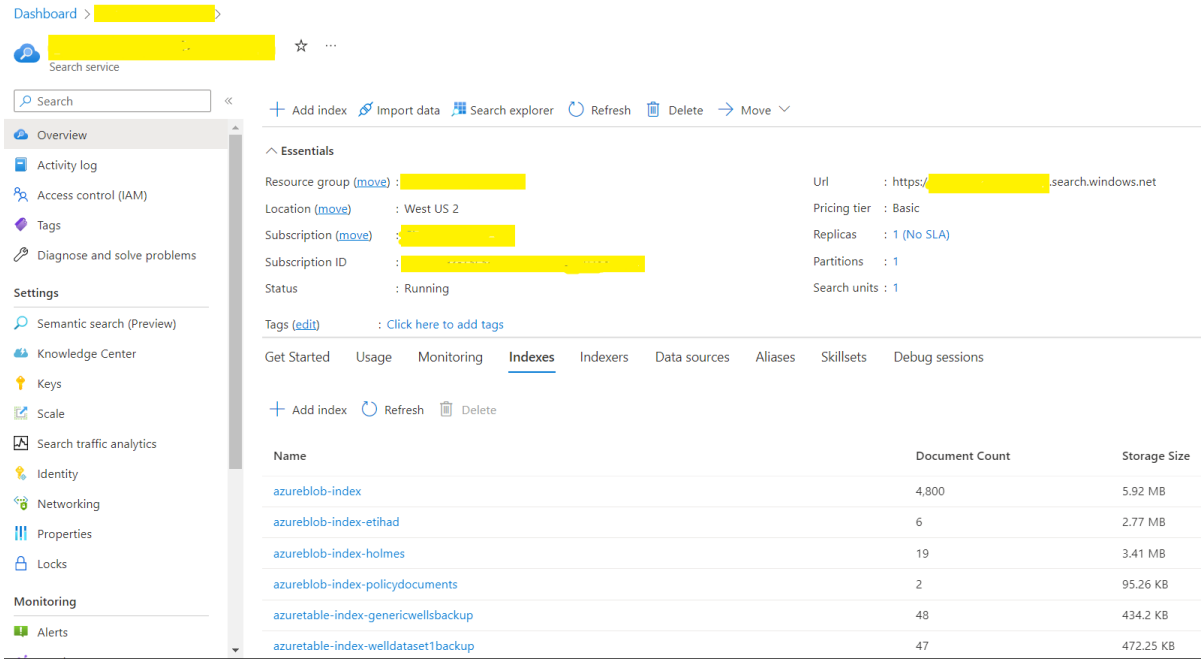
Data Sources – Configure data sources by choosing the “Add data source” option and providing connection string, storage/ container details. Should supply the identity details based on the authentication & authorization of access data sources



Skillset – skills helps to enrich the documents and gain the structure & substance as skill writes its output as nodes. Skillsets should be configured to improve search by using built-in or custom skills.



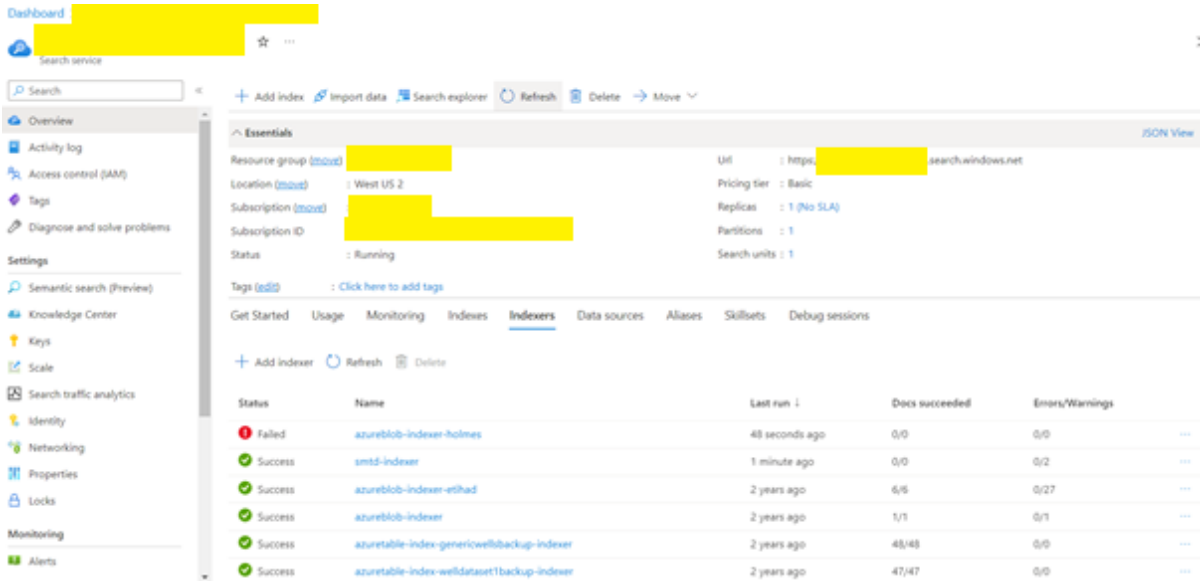
Indexes – Indexes are the base for any search. Create indexes by suppling the required field details & indexing properties like searchable, filterable, sortable, retrievable, etc.



Alias – an index alias is a secondary name that can be used to refer to an index for querying, indexing, and other operations. Optionally, an alias that maps to a search index and substitute the alias name in places where would otherwise reference an index name.

Indexer – indexer in Cognitive Search are crawler that extracts searchable content from data sources and populates search indexes using field-to-field mappings between source

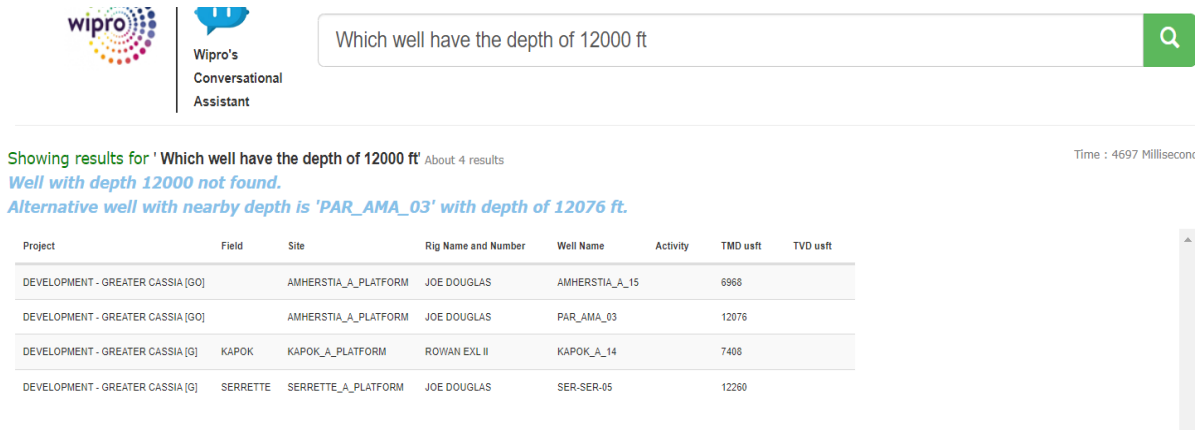
and search indexes. Indexer should be configured by providing the index, data source and optionally skillsets. Indexer can be run manually or scheduler which will help to update indexes on new data/files or on delta added.

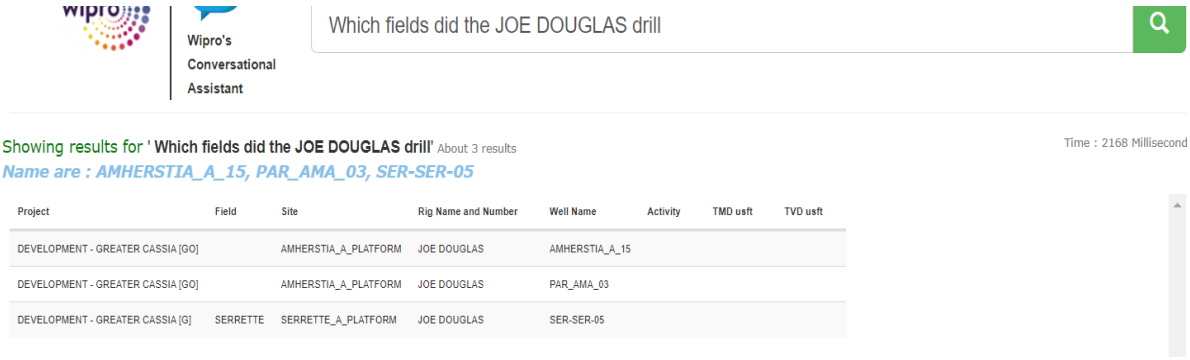


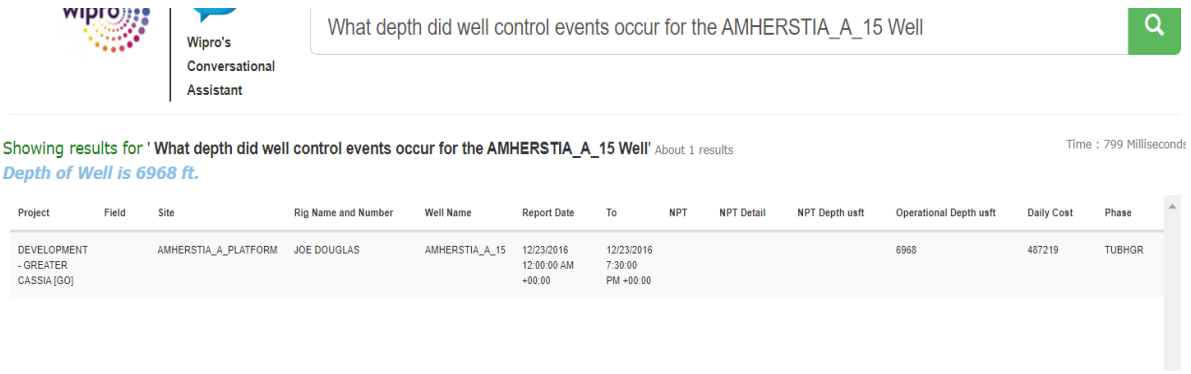
Services – once the Azure Cognitive search service configurated, it can be integrated with any custom application using search service API available with search service. Azure Search APIs can be integrated on heterogenous applications which supports API calls, ex. Web, chatbot or Power BI etc.

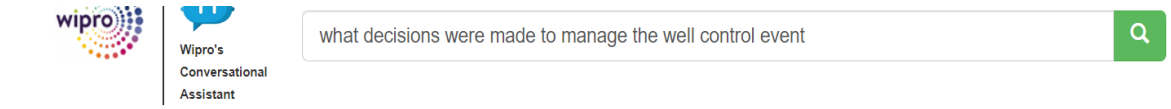
4. Sample Search

Azure search feature can be integrated on supportive applications. Sample screenshot below are captured from web and Chatbot application. This is to demonstrate that search feature can be integrated with variety of applications.









Showing results for ' **what decisions were made to manage the well control event**' About - results

Time : 1001 Milliseconds

Display Document

The forward plan from this point was to circulate the slug out of the hole in order to remove any additional complexity to the well control operation.
The team discussed 3 options that will allow the bit to go back to bottom: Option 1: Stripping through the annular Option 2: Stripping through the Weatherford RCD bearing Option 3: Placing a heavy mud cap in the riser and RIH conventionally After discussing the advantages and disadvantages of all options, it was decided to go with the mud cap solution, Option 3, as it is the simplest solution with minimal risk.
Option 2 was not considered as it relied on third party equipment and it was not appropriate to commence dynamically stripping operations for the first time in a well control situation.
Option 1 was not considered as it risks damaging the annular through excessive stripping.
After consultation with the GWO Well Control team, it was decided not to circulate the whole influx in one go.
In order to ensure having manageable gas quantities on surface, it was decided to do 2 circulations, one at the 13 3/8" shoe and one at TD.
Even though the influx was swabbed in hole, it was decided to raise the mud weight to 15.
Since the well was open now, and gas possibly still in the open hole, it was decided to put a hard line on the total volume pumped while washing and reaming to bottom.
It was decided to circulate out through the rig's chokes any gas brought up inside casing from 5050m.
It was decided to bullhead the pill through the BHA instead of pulling out to RIH with a DP stinger.

Conversational
Assistant

Showing results for ' **find well control incident for salamat 1 well**' About - results

Time : 2057 Millisecond

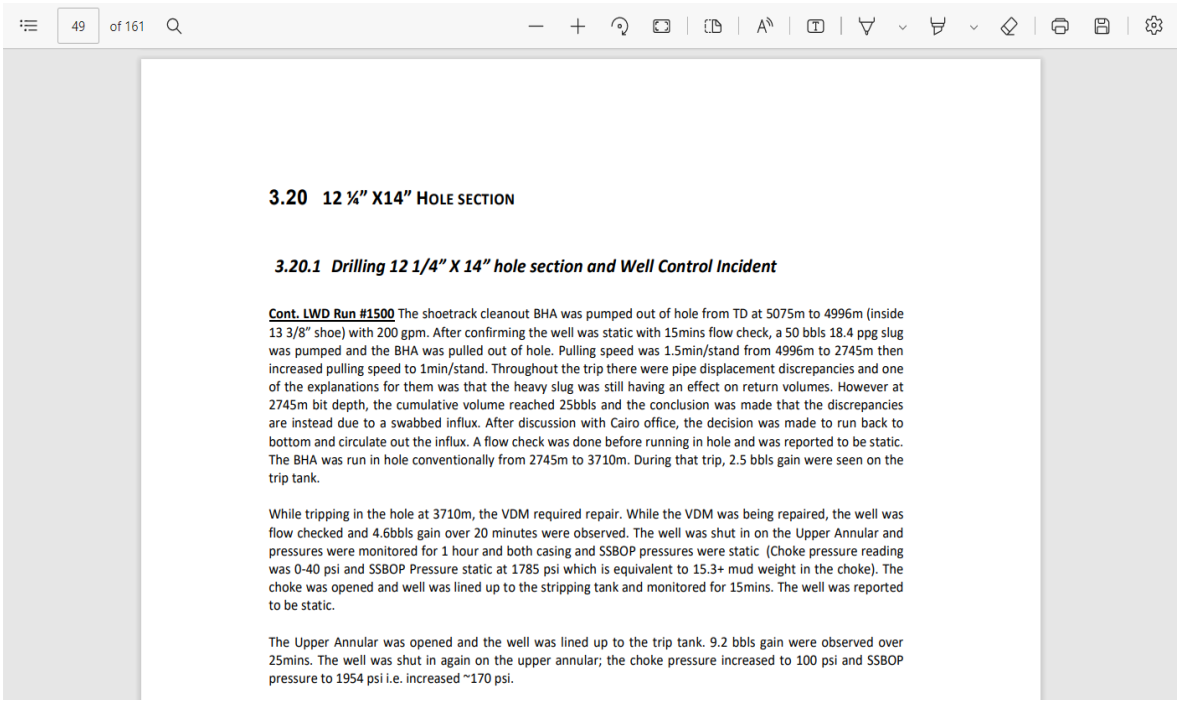
Display Document

3.20.1 Drilling 12 1/4" X 14" hole section and Well Control Incident

Cont. LWD Run #1500 The shoetrack cleanout BHA was pumped out of hole from TD at 5075m to 4996m (inside 13 3/8" shoe) with 200 gpm. After confirming the well was static with 15mins flow check, a 50 bbls 18.4 ppg slug was pumped and the BHA was pulled out of hole. Pulling speed was 1.5min/stand from 4996m to 2745m then increased pulling speed to 1min/stand. Throughout the trip there were pipe displacement discrepancies and one of the explanations for them was that the heavy slug was still having an effect on return volumes. However at 2745m bit depth, the cumulative volume reached 25bbls and the conclusion was made that the discrepancies are instead due to a swabbed influx. After discussion with Cairo office, the decision was made to run back to bottom and circulate out the influx. A flow check was done before running in hole and was reported to be static. The BHA was run in hole conventionally from 2745m to 3710m. During that trip, 2.5 bbls gain were seen on the trip tank.

While tripping in the hole at 3710m, the VDM required repair. While the VDM was being repaired, the well was flow checked and 4.6bbls gain over 20 minutes were observed. The well was shut in on the Upper Annular and pressures were monitored for 1 hour and both casing and SSBOP pressures were static (Choke pressure reading was 0-40 psi and SSBOP Pressure static at 1785 psi which is equivalent to 15.3+ mud weight in the choke). The choke was opened and well was lined up to the stripping tank and monitored for 15mins. The well was reported to be static.

The Upper Annular was opened and the well was lined up to the trip tank. 9.2 bbls gain were observed over



49 of 161

3.20 12 ¼" X14" HOLE SECTION

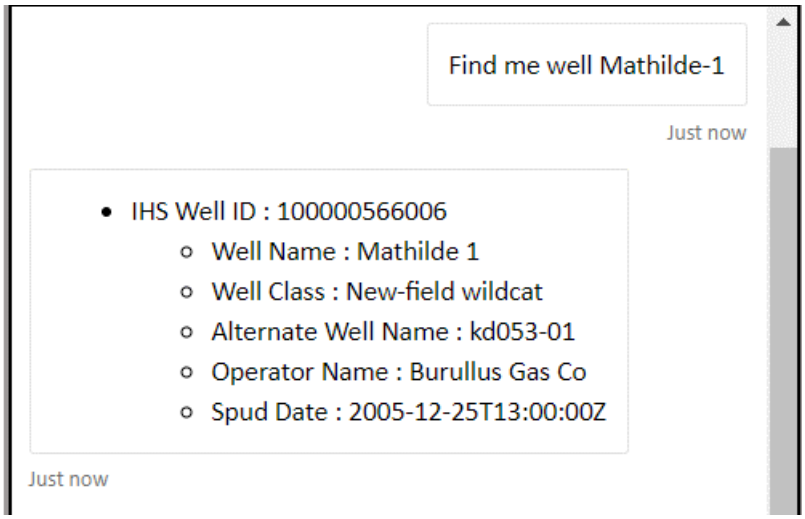
3.20.1 Drilling 12 1/4" X 14" hole section and Well Control Incident

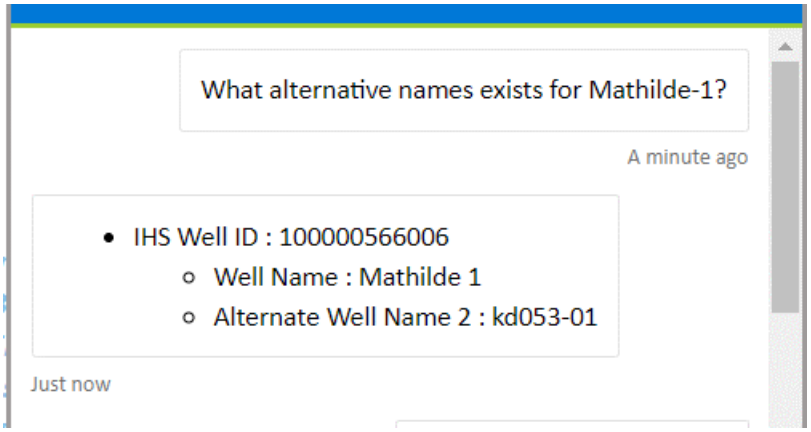
Cont. LWD Run #1500 The shoetrack cleanout BHA was pumped out of hole from TD at 5075m to 4996m (inside 13 3/8" shoe) with 200 gpm. After confirming the well was static with 15mins flow check, a 50 bbls 18.4 ppg slug was pumped and the BHA was pulled out of hole. Pulling speed was 1.5min/stand from 4996m to 2745m then increased pulling speed to 1min/stand. Throughout the trip there were pipe displacement discrepancies and one of the explanations for them was that the heavy slug was still having an effect on return volumes. However at 2745m bit depth, the cumulative volume reached 25bbls and the conclusion was made that the discrepancies are instead due to a swabbed influx. After discussion with Cairo office, the decision was made to run back to bottom and circulate out the influx. A flow check was done before running in hole and was reported to be static. The BHA was run in hole conventionally from 2745m to 3710m. During that trip, 2.5 bbls gain were seen on the trip tank.

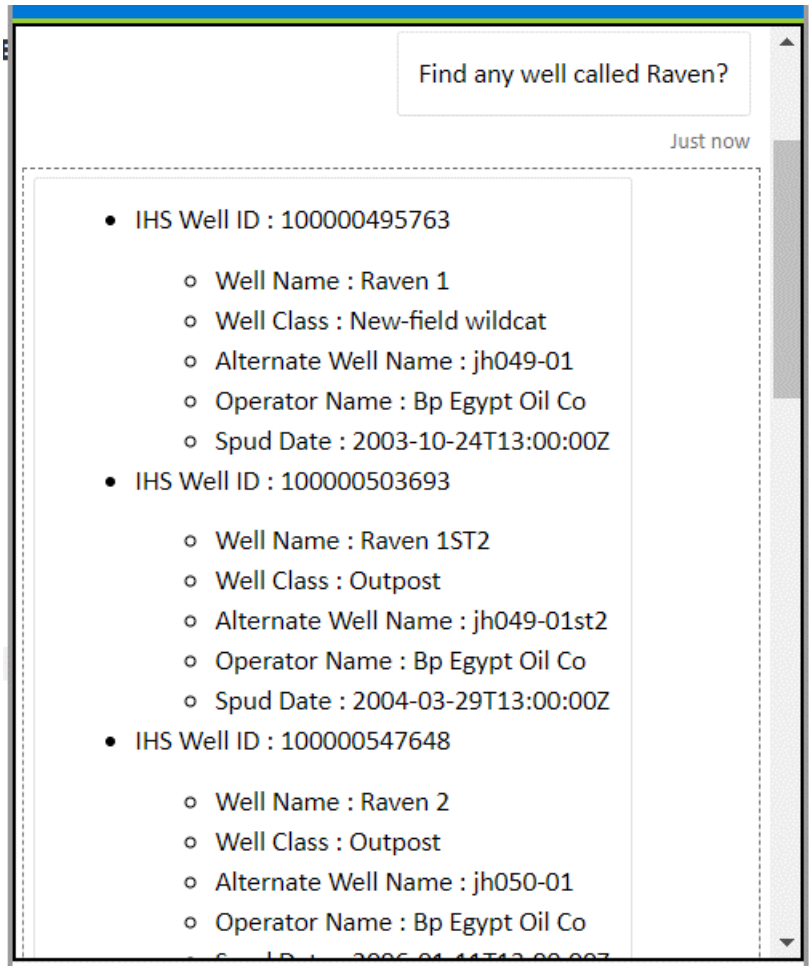
While tripping in the hole at 3710m, the VDM required repair. While the VDM was being repaired, the well was flow checked and 4.6bbls gain over 20 minutes were observed. The well was shut in on the Upper Annular and pressures were monitored for 1 hour and both casing and SSBOP pressures were static (Choke pressure reading was 0-40 psi and SSBOP Pressure static at 1785 psi which is equivalent to 15.3+ mud weight in the choke). The choke was opened and well was lined up to the stripping tank and monitored for 15mins. The well was reported to be static.

The Upper Annular was opened and the well was lined up to the trip tank. 9.2 bbls gain were observed over 25mins. The well was shut in again on the upper annular; the choke pressure increased to 100 psi and SSBOP pressure to 1954 psi i.e. increased ~170 psi.

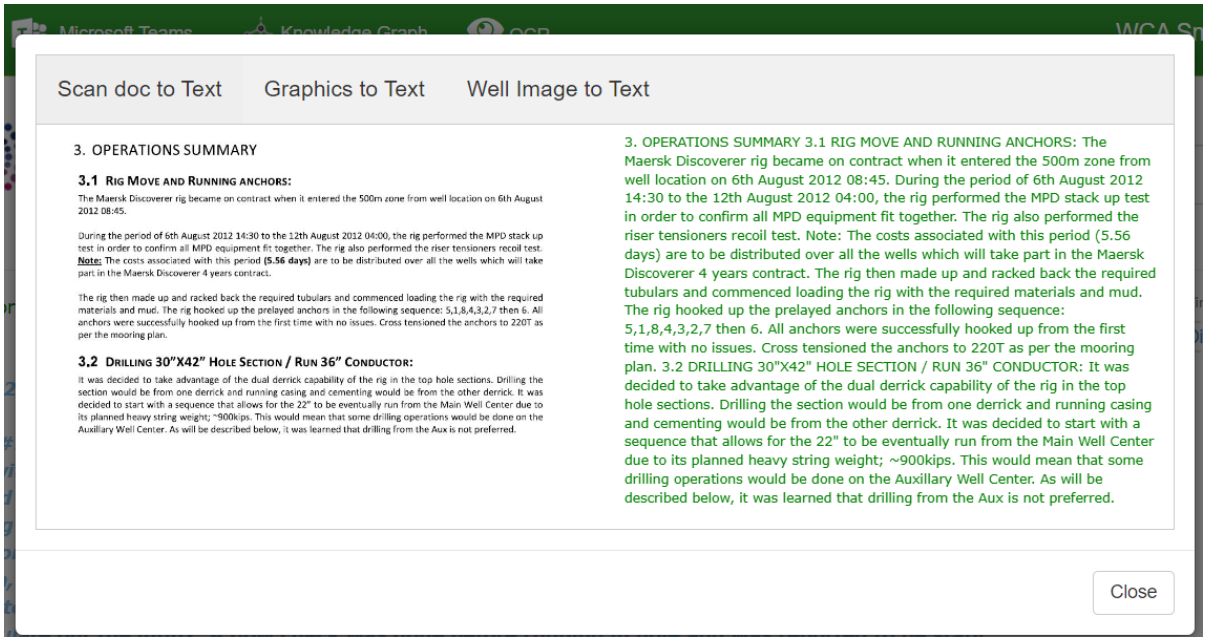
4.1. Sample Search on Chatbot

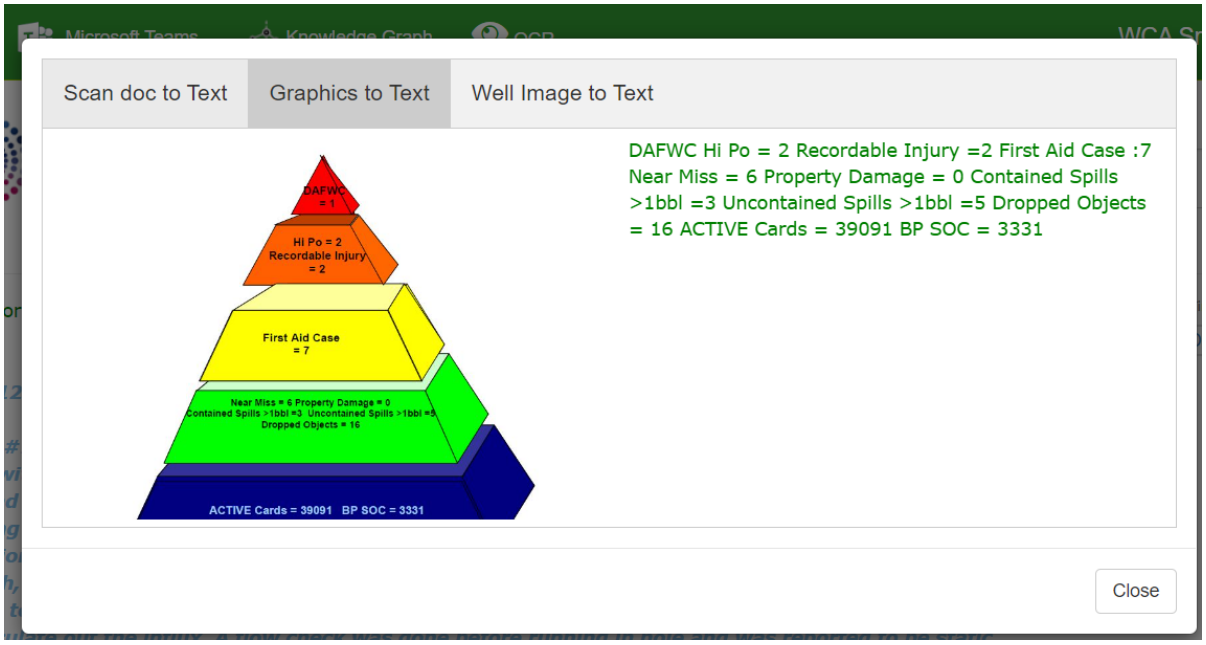


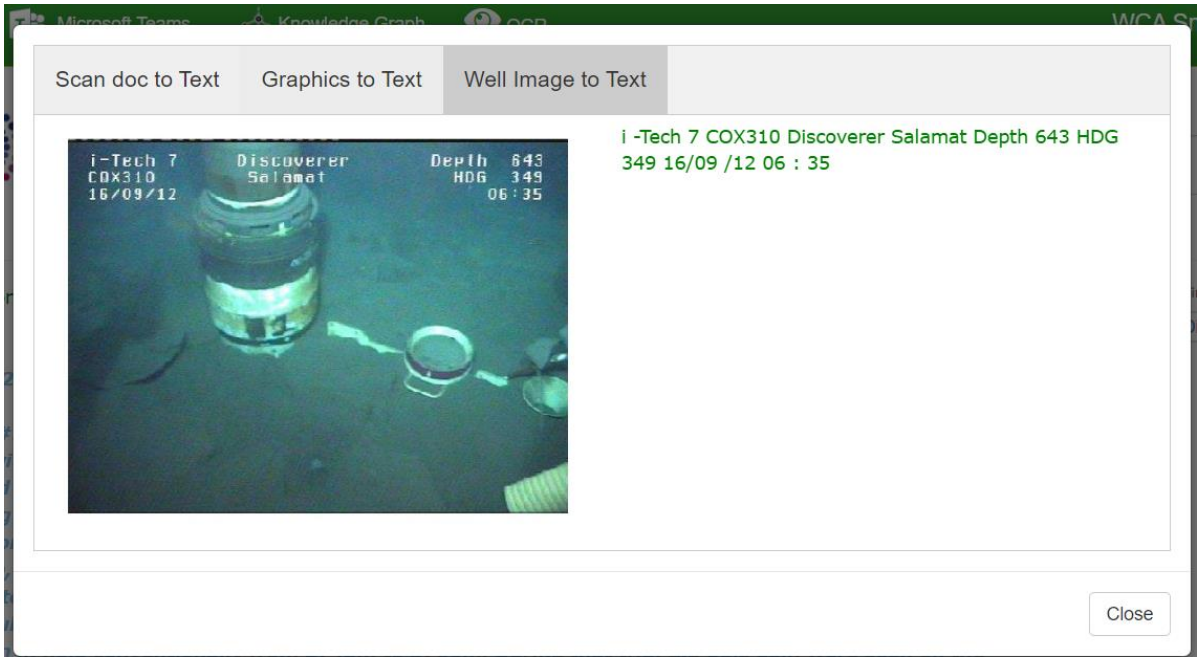

What alternative names exists for Mathilde-1?
A minute ago
IHS Well ID : 100000566006
Well Name : Mathilde 1
Alternate Well Name 2 : kd053-01
Just now


Find any well called Raven?
Just now
IHS Well ID : 100000495763
Well Name : Raven 1
Well Class : New-field wildcat
Alternate Well Name : jh049-01
Operator Name : Bp Egypt Oil Co
Spud Date : 2003-10-24T13:00:00Z
IHS Well ID : 100000503693
Well Name : Raven 1ST2
Well Class : Outpost
Alternate Well Name : jh049-01st2
Operator Name : Bp Egypt Oil Co
Spud Date : 2004-03-29T13:00:00Z
IHS Well ID : 100000547648
Well Name : Raven 2
Well Class : Outpost
Alternate Well Name : jh050-01
Operator Name : Bp Egypt Oil Co

4.2. Sample Search for Computer Vision on Web






Scan doc to Text
Graphics to Text
Well Image to Text
i-Tech 7
COX310
16/09/12
Discoverer
Salamat
Depth 643
HDG 349
06:35
i -Tech 7 COX310 Discoverer Salamat Depth 643 HDG 349 16/09 /12 06 : 35
Close

5. Business Benefits

Few highlighted business benefits –

Search feature can be adoptable for enterprise scale by building the search modules.

Amplified efficiency of organization – search functionality removes the necessity of using multiple application for different search purpose. Amalgamation of data tools allows organizations to streamline their business processes.

Enhanced employee experience and commitment – AI/ML algorithms that provide personalized suggestions can help users to find relevant data more quickly and the flexibility of cognitive search creates an improved user experience through personalization. Since an employee's search experience is improved, they're more likely to use the tools consistently.

Lower operational costs – Maximized productivity decreases an organization's operational costs since less time and resources are needed for gathering information and knowledge discovery. This is especially beneficial to industries such as healthcare and legal services that work with massive amounts of data.