



aws INNOVATE

AI/ML EDITION

24 February 2022

Amazon Kendra 통해 자연어 검색으로 원하는 답변을 빠르고 정확하게 얻어보기

김영진

Solutions Architect
Amazon Web Services





우리의 미션은 모든 개발자, 데이터 과학자 및
사용자에게 **사용 편리한** 기계 학습을 제공

Critical information is getting harder to find



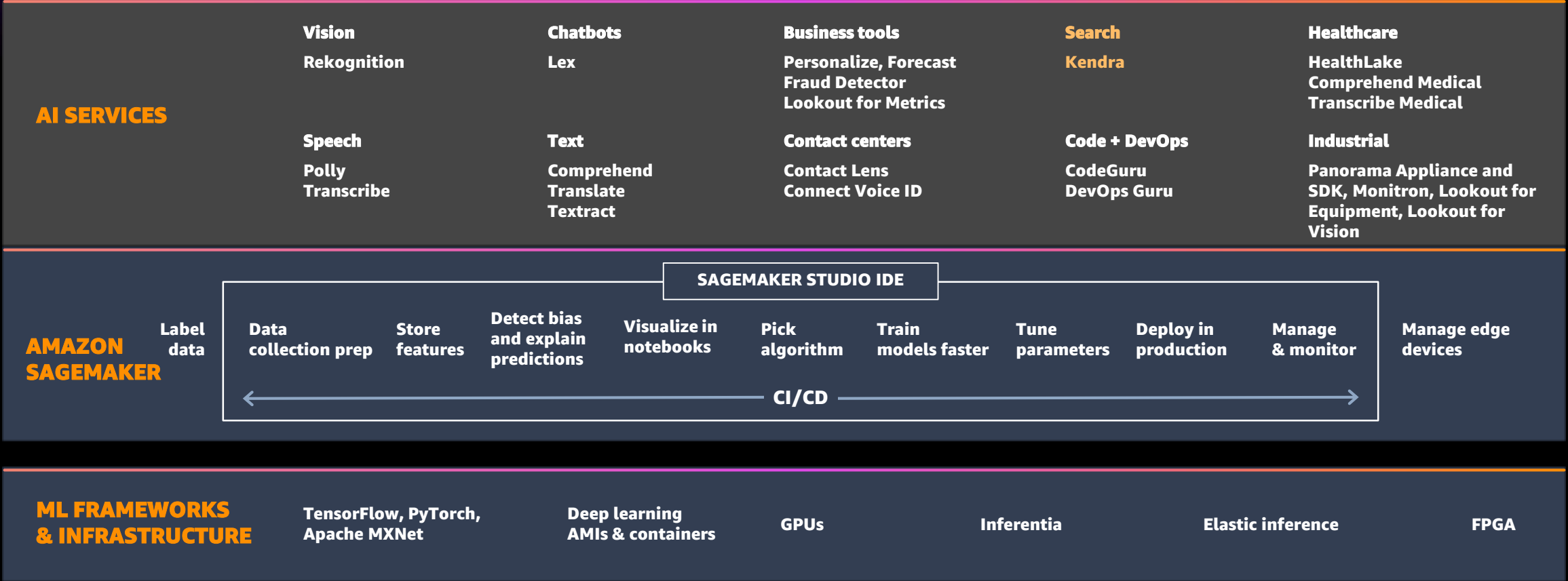
기업의 데이터는 급증하고
있으며 종종 흩어져 있고
액세스할 수 없습니다.

80% of all data will be
unstructured by 2025

Source: <https://www.capgemini.com/2018/08/reorganizing-unstructured-data/>

The AWS ML stack

가장 광범위하고 완전한 기계 학습 기능 세트



Amazon Kendra



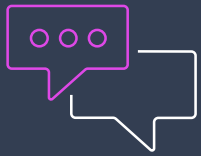
Amazon Kendra

Amazon Kendra는 기계 학습을 기반으로 하는 사용자가 자연어를 사용하여 구조화/구조화되지 않은 데이터를 검색할 수 있도록 하는 매우 정확한 **지능형 검색 서비스**입니다.

Amazon Kendra

정확한 정보를 쉽고 빠르게 찾아보세요

EASY TO FIND WHAT YOU ARE LOOKING FOR



Natural language queries

구조화되지 않은
콘텐츠의 문맥 검색



NLU & machine learning core

Reading comprehension
FAQ matching
Document ranking



Broad domain expertise

사전훈련된 14개
주요 도메인에
최적화



Continuous improvement

사용자 피드백을
통해 점진적으로
학습

QUICK SETUP



Native & partner connectors

Automate ingestion
어디에서나 최고의 답변을
찾기

SECURE



Secure search

전송 및 저장 시 암호화된
토큰 기반 액세스 제어

Amazon Kendra

Natural language queries



Natural language queries

Natural language queries

"Where is the IT support desk in AWS Korea?" → "1st floor"

Keyword queries

"Health benefits" or "parking allowance"

what are less common COVID-19 symptoms ?

1-10 of about 94,600

Sort by Best match

Top suggested answer

Optimal Control Measures to Combat COVID19 Spread in Sri Lanka: A Mathematical Model Considering the Heterogeneity of Cases

Common symptoms of COVID 19 disease have been fever, cough and fatigue. **There are some less common symptoms including sputum production, headache, hemophiliacs, and diarrhea [3]**. According to WHO, COVID 19 has spread for more than 210 countries and independent territories while Italy, Spain, United States and Iran are the hardest hit apart from China where the disease is known to be emerged but now significantly controlled and stable. In numbers, currently more than 30 million people have been infected while there are 200 000 reported deaths worldwide [4] .

▼ More suggested answers (2)

Suggested answer (1 of 2)

Long-Term Neurological Threats of COVID-19: A Call to Update the Thinking About the Outcomes of the Coronavirus Pandemic

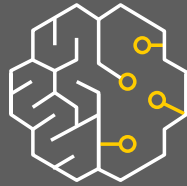
While the most common symptoms of COVID-19 at the onset of illness include fever, fatigue, dry cough, myalgia, and dyspnea, other **less common symptoms** are **headache, abdominal pain, diarrhea, nausea, and vomiting (10)**. Furthermore, it's been recently reported that most patients also complain of impairment of both olfactory and gustatory perception (11) and those are being considered early markers of COVID-19 infection. Though there is longstanding evidence that human coronaviruses, such as SARS-CoV-2, can spread to the brain from the respiratory tract (5, 12, 13), the occurrence of gastrointestinal **symptoms**

Amazon Kendra

NLU and ML Core



Natural
language
queries



NLU and
ML core

Better answers

- Reading comprehension
- FAQ matching
- Document ranking(문서순위)

Where is the it support desk in kumo?

Kendra's suggested answer

1st floor

... our IT help desk, deskside, which are located all around our buildings and open for support at most hours. The one in Kumo is on the **1st floor** and is open from 12:30 to 5 p.m. daily.

WEB : <https://w.amazon.com/bin/view/ILDRTEST/>

What is Kendra's suggested answer? [info](#)

Frequently asked questions

Where do I get IT help in Kumo?

What are the IT support hours in Dopler?

Where can I get IT help Seattle campus?

What are frequently asked questions? [info](#)

1-10 of 486 Recommended documents

AWS Kumo

...this is how we're doing it. Access controls Access to this host is governed by the permissions associated with the AWS-SUPPORT-KUMO hostclass. Permissions to the aws-support-kumo group... [\(More\)](#)

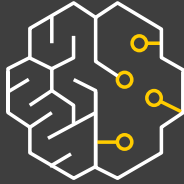
WEB : <https://w.amazon.com/bin/view/kumo-support>

Amazon Kendra

Domain expertise



Natural
language
queries



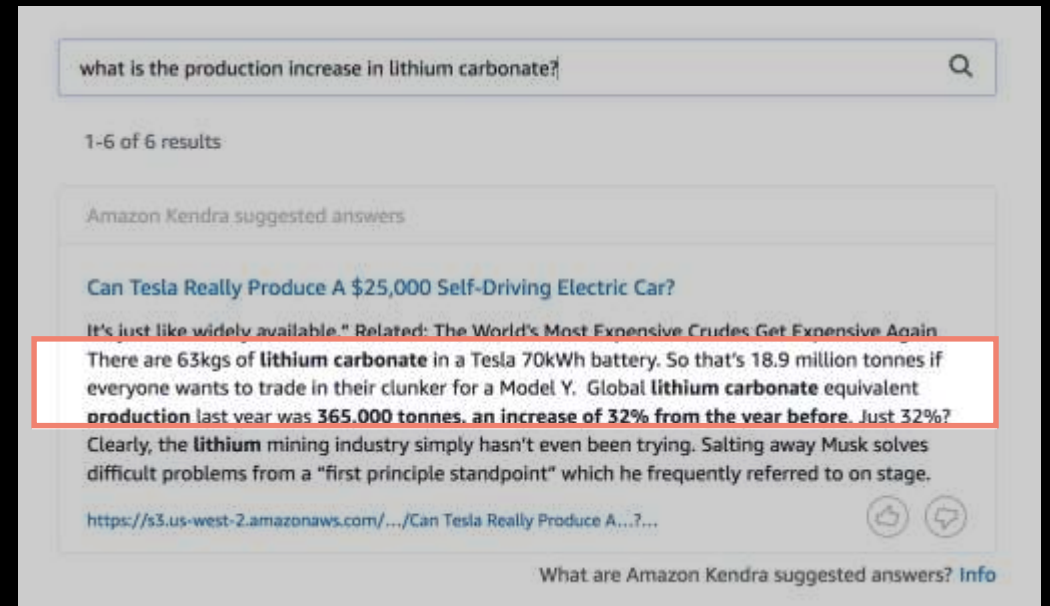
NLU and
ML core



Domain
expertise

14개 주요 도메인에
최적화

IT, Financial Services,
Insurance, Pharma,
Industrial, Energy,
Legal, Media and
Entertainment, Travel
and Hospitality,
Health, HR, News,
Telecommunications,
and Automotive

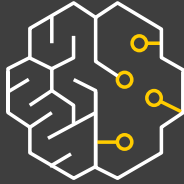


Amazon Kendra

Incremental learning



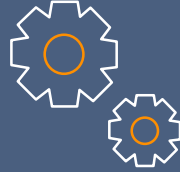
Natural
language
queries



NLU and
ML core



Domain
expertise

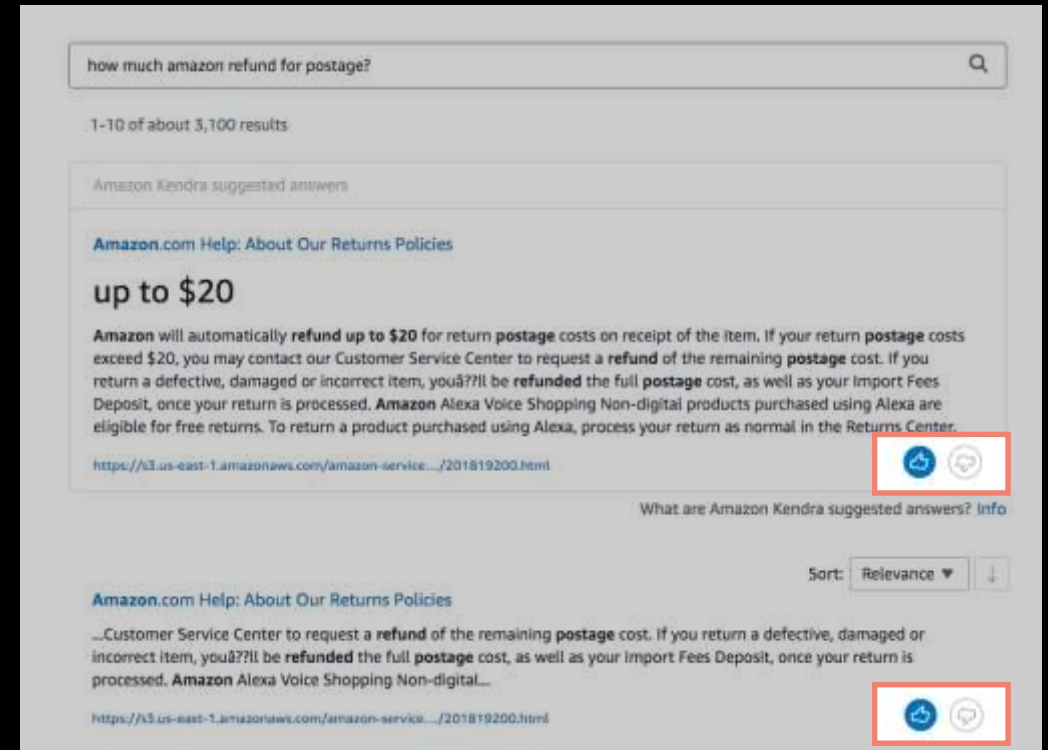


Continuous
improvement

Kendra는 시간이
지남에 따라 자동으로
향상됩니다.

Captures click-
through and
user feedback

- Retrains models
periodically



Amazon Kendra

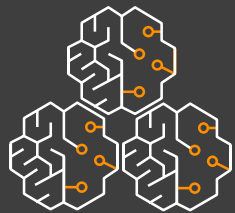
Native connectors



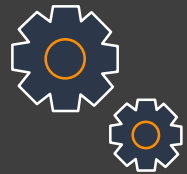
Natural
language
queries



NLU and
ML core



Domain
expertise



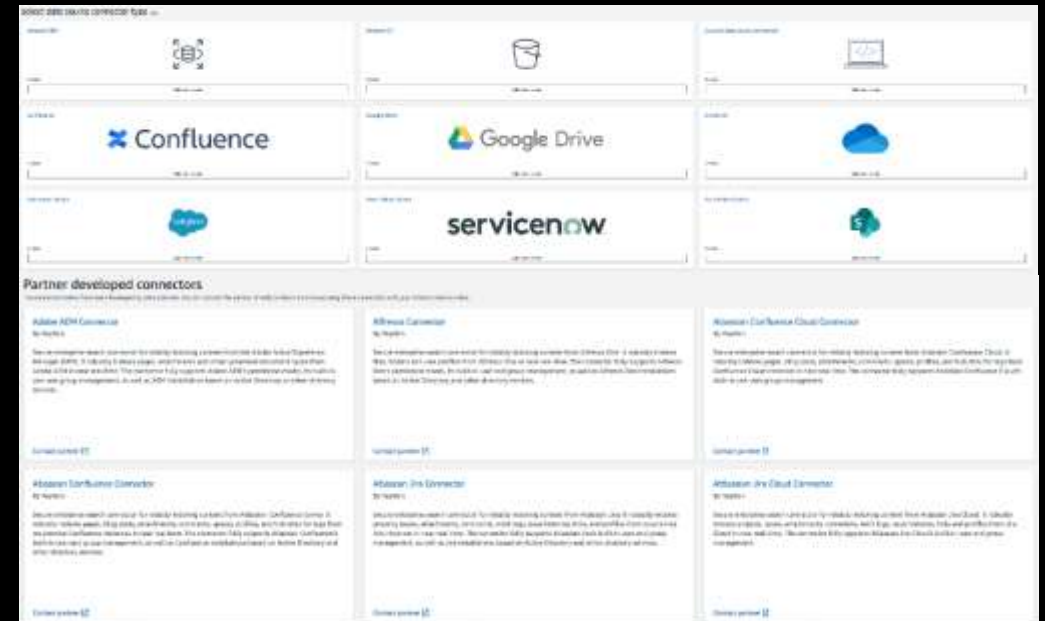
Continuous
improvement



Native & partner
connectors

➤ 다음을 포함한 50개의 인기 있는 데이터 소스
Microsoft SharePoint, Atlassian Confluence and Jira, Adobe Experience Manager, ServiceNow, and Salesforce

- Add credentials, point and shoot

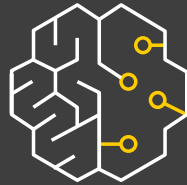


Amazon Kendra

Highly secure



Natural
language
queries



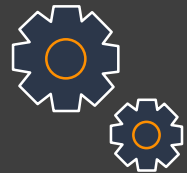
NLU and
ML core



Native & partner
connectors



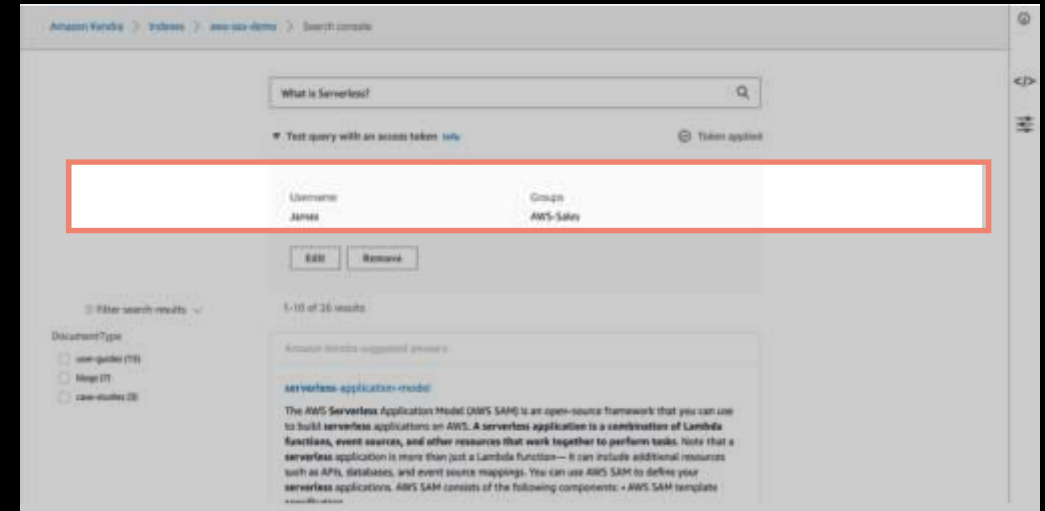
Domain
expertise



Continuous
improvement



Secure
search



Data encrypted in transit and rest



Token-based user access control

How it works

1

Create index

Index details

Index name

my-index-name

Maximum of 1000 alphanumeric characters. Can include hyphens (-), but not spaces.

Description - *optional*

Logging details

Amazon Kendra publishes error and alert logs to Amazon CloudWatch. A CloudWatch log group and corresponding log stream will be created on your behalf.

IAM role

Amazon Kendra requires permissions to access your CloudWatch log. Choose an existing IAM role or let us create a role for you.

Choose an option



Encryption

Amazon Kendra will encrypt your data with Amazon Kendra owned key by default. You can also choose to use an AWS KMS managed key.

☐ Use an AWS KMS managed encryption key

2

Select data source connector type [정보](#)

Amazon RDS



4 steps

Add connector

Amazon S3



3 steps

Add connector

Confluence



5 steps

Add connector

Custom data source connector



3 steps

Add connector

Google Drive



5 steps

Add connector

OneDrive



5 steps

Add connector

Salesforce Online



5 steps

Add connector

ServiceNow Online



5 steps

Add connector

SharePoint



5 steps

Add connector

WebCrawler



4 steps

Add connector

WorkDocs



5 steps

Add connector

2


FAQs

FAQs (0)

DeleteAdd FAQ

Q Search FAQs

< 1 > ⚙

Name ▾	Status ▾	Creation time ▾	Last Modified ▾
<div><p>You have no FAQs</p><p>Upload commonly asked questions and answers from your data set to improve search results for your end users.</p><div>Add FAQ</div></div>			

3



Kendra Search console

Test search & Deploy

Where is the i



Where is the it **help desk day1**

Where is the it **help desk nearest me**

Where is the it **support desk dopler**

Where is the it **support desk kumo**

Where is the it **support desk blackfoot**

Reading Comprehend

FAQ matching

Document Ranking

Kendra's suggested answer

1st floor

... our IT help **desk**, deskside, which are located all around our buildings and open for **support** at most hours. The one in **Kumo** is on the **1st floor** and is open from 12:30 to 5 p.m. daily.

WEB | <https://w.amazon.com/bin/view/ILDRTEST/>

What is Kendra's suggested answer? [Info](#)

Frequently asked questions

Where do I get IT help in Kumo? ▶

What are the IT support hours in Dopler? ▶

Where can I get IT help Seattle campus? ▶

What are frequently asked questions? [Info](#)

1-10 of 486 Recommended documents

Sort: Relevancy ▼

AWS Kumo

...this is how we're doing it. Access controls Access to this host is governed by the permissions associated with the AWS-SUPPORT-KUMO hostclass. Permissions to the aws-support-kumo group... [\(More\)](#)

WEB | <https://w.amazon.com/bin/view/kumo-support>

Kumo Support Center

helpful to developers working on the platform. Follow this guide setup SC_SAPI_CMS to setup the whole kumo-owned casemangement workflow on your desktop to expedite... [\(More\)](#)

PDF | <https://w.amazon.com/bin/view/TWHDKKTTS.pdf>

Filter search results ▼

Search in

Everywhere (486) Wiki (123)

Broadcast (265) Sage (98)

Date range:

Categories

☒ HR (23)

☐ IT (5)

☐ Marketing (5)

Authors

☐ miyingah (7)

☐ michealnb (4)

☐ angelak (4)

☐ masono (1)

☐ kimfonta (1)

☐ rwitaman (1)

Show more...

amazon.com employee discount



Kendra's suggested answer

New Hire Information

For more information, visit the Amazon US Benefits Inside Amazon page. **Employee Discount** Employees receive a **10% discount** (up to the value of \$100 annually) for selected products available on **Amazon.com**. **Employee discount** code numbers will be available once new hire documentation is complete... [\(More\)](#)

WEB | https://w.amazon.com/bin/view/AmazonStudios/New_Hire_Information/



What is Kendra's suggested answer? [Info](#)

Frequently asked questions

[How do I apply the amazon.com discount code?](#)



[How often can I use the amazon.com discount?](#)



[What items can I apply my employee discount?](#)



What are frequently asked questions? [Info](#)

1-10 of 98 Recommended documents

Sort: Relevancy ▼

Employee Offer

With AT&T This wiki explains how Amazon employees can redeem their employee benefit...Note: Please refer to the **Employee Extras** page for current **discount** details. Some of the information below... [\(More\)](#)

WEB | <https://w.amazon.com/bin/view/key-benefits-offers>



Relevance tuning (optional)

- 특정 비즈니스 컨텍스트가 있는 경우
관련성(relevance)을 광범위하게 조정

- 신뢰할 수 있는 데이터 소스
- 저자(Authors) 또는 부서(departments)
- 최신문서(Document freshness)
- 찬성(Upvotes) and 인기도(popularity)

Boosting results

Boost text, date and numeric field types on a 1-10 scale to increase their searchability.

Search fields for boosting

Text Date Numeric

► title Boost: 1

► author Boost: 1

► description Boost: 1

► category Boost: 1

► tags Boost: 1

► body_text Boost: 1

► data_source Multi: 5, 1, 1, ...

Data source boost
Boost relevance of individual data sources....

marketing-documentation 5

Source: sharepoint ID: 38937487343

wiki-hr

Data sources

Relevance tuning

Boost text, date and numeric field types on a 1-10 scale to increase their searchability.

Search fields for boosting

Text Date Numeric

► like_count Boost: 10

Numeric boost
Boost searchability based on ascending or descending order of numbers in this field.

☒ Ascending
☐ Descending

10

View/like counts

Relevance tuning

Boost text, date and numeric field types on a 1-10 scale to increase their searchability.

Search fields for boosting

Text Date Numeric

▼ creation_date Boost: 5

Freshness boost
Boost freshness of creation_date field.

5

▼ **Additional boosting**
Fine tune time sensitivity

Sensitivity range
Adjust slider to capture range of sensitivity. Smaller the range, the more sensitive results are to daily changes.

Document
freshness

Traditional vs Intelligent Search

Employees and customers want answers, not choices


Traditional search

The image shows a side-by-side comparison of two search interfaces. On the left is a 'Traditional search' interface titled 'Intranet Search'. It features a search bar with the text 'it support desk', a 'Your recent searches' dropdown, and a 'Not finding relevant results' message. Below the search bar, it says 'SEARCH IN:' and 'Displaying results 1 - 10 of 21'. The search results are listed under 'Everything (21)', 'Wiki (17)', 'Email List Archive (3)', and 'Show more...'. The results include 'IT_Support_Training_Program.Wiki', 'Com_Support_Wiki.Web', 'OperationalBestPractices.EventM', and 'Corp_Wiki_Pending.Web'. On the right is an 'Intelligent search' interface titled 'RESULTS PAGE'. It features a search bar with the text 'Where is the it support desk?'. Below the search bar, it says 'Kendra's suggested answer' and '1st floor'. The suggested answer text reads: '... our IT help desk, deskside, which are located all around our buildings and open for support at most hours. The one in Seattle is on the 1st floor and is open from 12:30 to 5 p.m. daily.' Below the suggested answer, it says 'WEB | https://w.organization.com/bin/view/ILDRTEST/'. At the bottom, it says 'Frequently asked questions' and lists three questions: 'Where do I get IT help?', 'What are the IT support hours?', and 'Where can I get IT help corporate campus?'. A large blue arrow points from the traditional search interface to the intelligent search interface.

Intelligent search

What's the difference for OpenSearch & Amazon Kendra?

DUS DEMO
DOCUMENT UNDERSTANDING SOLUTION

powered by 

[Document list](#) | [Upload your own documents](#)

[Log Out](#)

Search

Elasticsearch
Keyword Search Results

Amazon Kendra
Semantic Search Results

Elasticsearch and Amazon Kendra
Compare Search Technologies

Elasticsearch

Keyword search results

FOUND ABOUT 50 RESULTS ACROSS 18 DOCUMENTS

[How_Coronavirus_Spreads___CDC.pdf](#)

... 19 How COVID-1 Spreads COVID-19 is thought to spread mainly through close contact from perso...
...In respiratory droplets from someone who is infected...
...These droplets can land in the mouths or noses of people who are nearby or possibly be inhaled int...
...Some recent studies have suggested that COVID-19 may be spread by people who are not showing ...

[How_to_Protect_Yourself_And_Others___CDC.pdf](#)

... The virus is thought to spread mainly from person-to-person...
...These droplets can land in the mouths or noses of people who are nearby or possibly be inhaled int...
...Some recent studies have suggested that COVID-19 may be spread by people who are not showing ...
...Cover your mouth and nose with a cloth face cover when around others You could spread COVID-19 ..

What's the difference?

Amazon Kendra

Semantic search results

1-10 OF 164 RESULTS

Frequently asked questions [More info](#)

Can someone who has had COVID 19 spread the illness to others?

▶

Can someone who has been quarantined for COVID 19 spread th...


▶

Can the virus that causes COVID 19 be spread through food, incl...

▶

Can mosquitoes or ticks spread the virus that causes COVID 19?

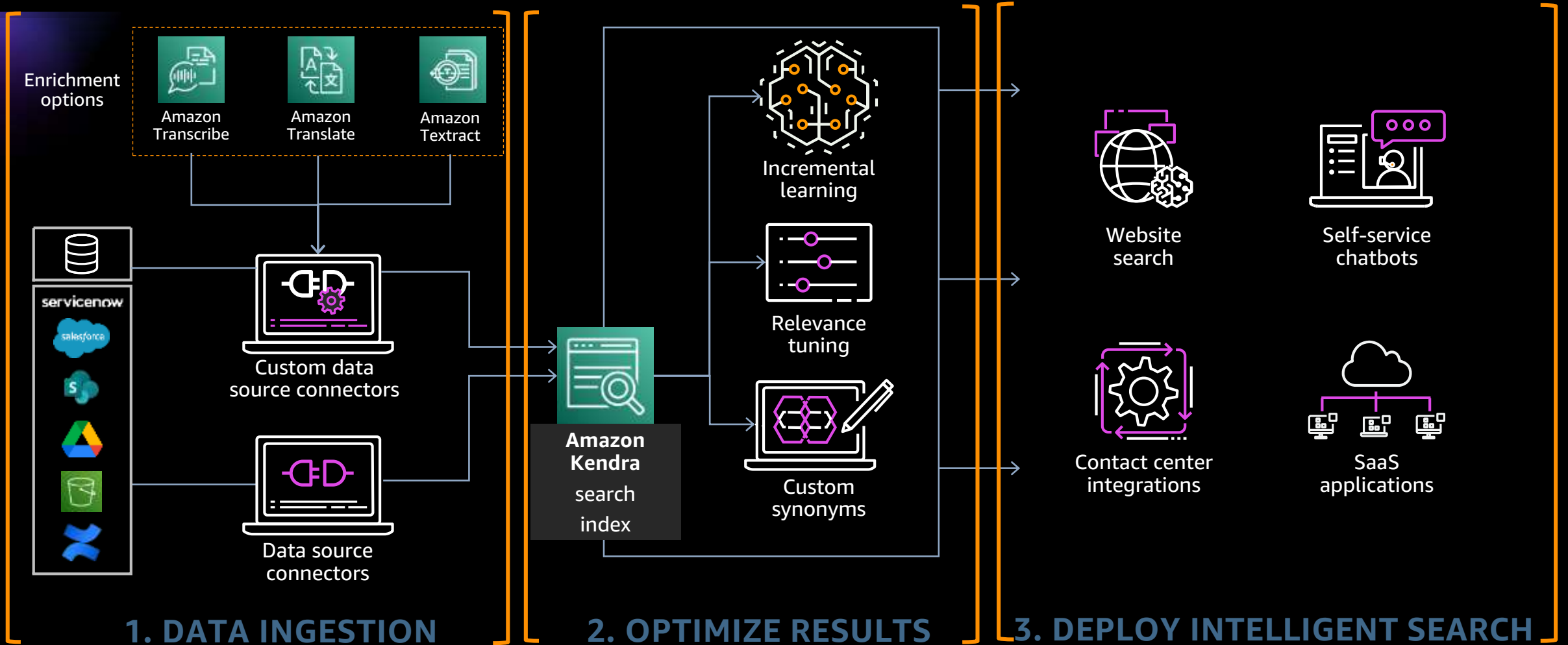
▶



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon Kendra

지능형 검색 애플리케이션 구축



What Connectors Does Kendra Support?



Amazon S3



File systems (SMB)



Web crawler



Databases



Sharepoint Online



Sharepoint 2013,
2016, 2019



Box



Dropbox



Exchange



OneDrive



Google Drive



Salesforce



Confluence



Jira



ServiceNow



Zendesk

Jive

...and more coming in 2021, 2022

DEMO



Demo 1 – Create & query an index w/ Amazon Kendra

Developer Center / Getting Started / Hands-on / ...

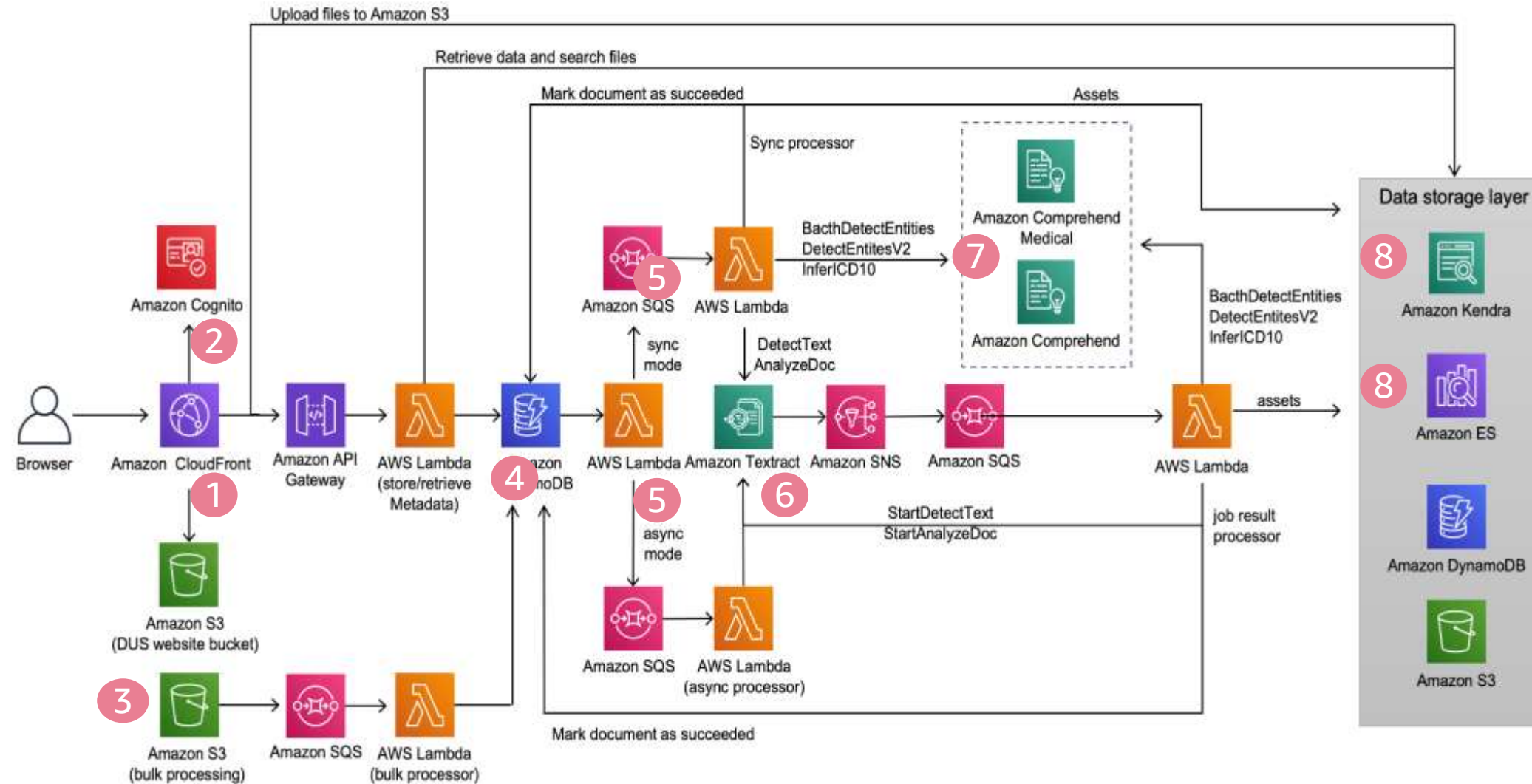
Create and query an index with Amazon Kendra

About this Tutorial

Time	1 hour
Cost	Less than \$1
Use Case	Machine Learning
Products	Amazon Kendra
Audience	Developer
Level	Beginner
Last Updated	September 11, 2020

https://aws.amazon.com/ko/getting-started/hands-on/create-query-index-with-amazon-kendra/?nc1=h_ls

Demo 2 – Document Understanding Solution



Document Understanding Solution

버전 1.0.3

릴리스 날짜: 2021년 11월

작성자: AWS

예상 배포 시간: 30~60분

[예상 비용](#)

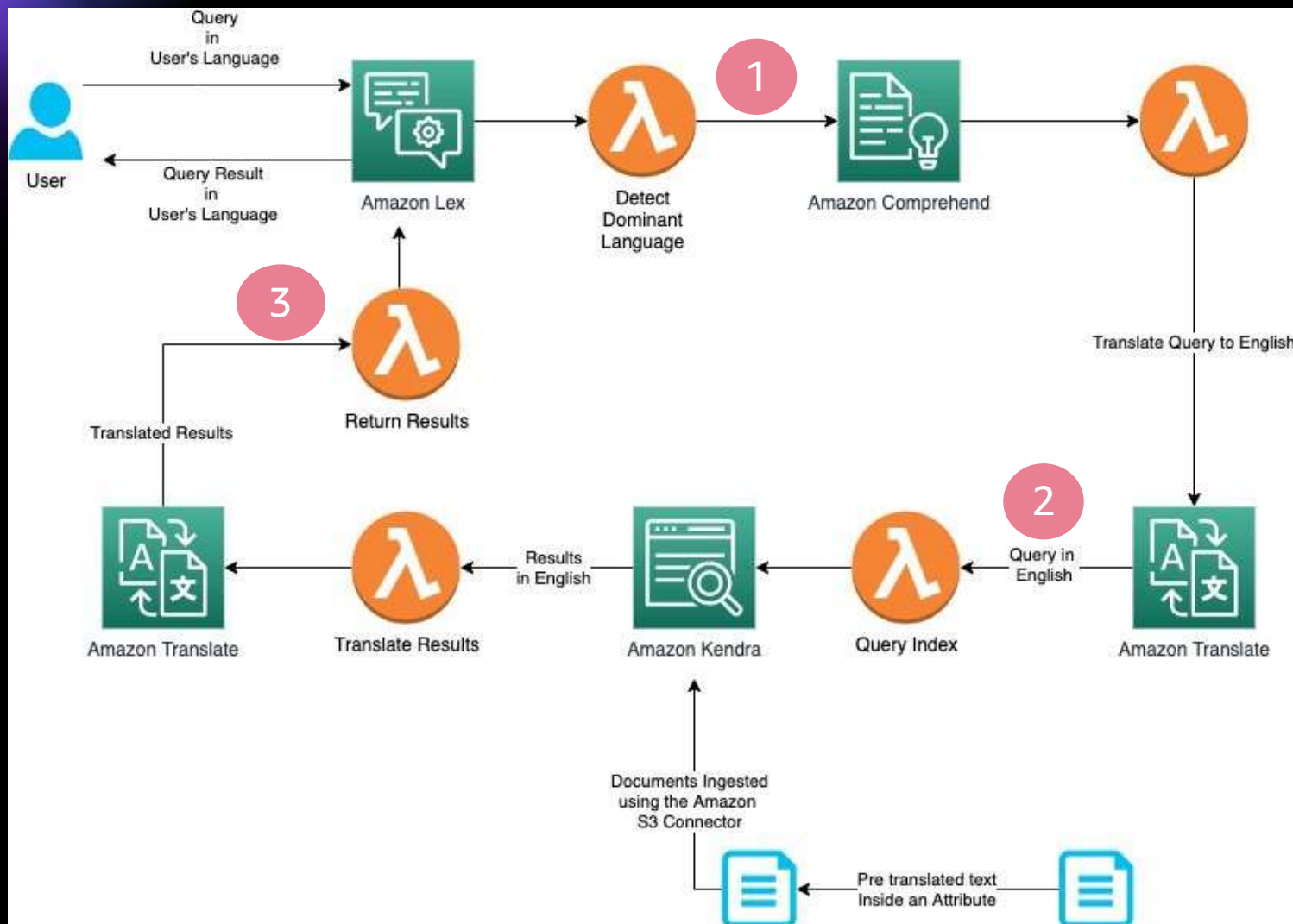
[소스 코드](#)

[CloudFormation 템플릿](#)

[구현 가이드 보기](#)

[AWS 콘솔 시작](#)

Amazon Translate 를 사용하여 Amazon Kendra에 다국어 언어지원 Architecture



(1) 사용자 언어감지

Amazon Lex을 사용하여 빌드된 챗봇은 Amazon Comprehend로 사용자 언어를 식별

(2) 쿼리 제출

Amazon Translate로 사용자언어를 영어로 번역하여 Kendra로 인덱스 쿼리 요청

(3) 번역된 결과 반환

Kendra의 영어로 된 검색결과를 Translate를 이용하여 사용자 언어로 번역한후 Amazon Lex 통해 최종리턴

<https://aws.amazon.com/ko/blogs/machine-learning/using-amazon-translate-to-provide-language-support-to-amazon-kendra/>

FAQ

1.내 데이터에 Amazon Kendra가 찾고 있는 정확한 답변이 포함되어 있지 않으면 어떻게 합니까?

A: 데이터에 질문에 대한 정확한 답변이 포함되어 있지 않으면 Kendra는 딥 러닝 모델에 따라 순위가 매겨진 가장 관련성이 높은 문서 목록을 반환합니다.

2.Amazon Kendra는 어떤 파일 형식을 지원합니까?

A: Kendra는 **.html, MS Office(.doc, .ppt), PDF 및 텍스트 형식의 비정형 및 반정형 데이터를** 지원합니다.
MediaSearch 솔루션을 Amazon Kendra와 함께 사용하여 오디오 및 비디오 파일을 검색할 수도 있습니다.

3.Amazon Kendra는 어떤 유형의 질문에 답변할 수 없습니까?

A: Kendra는 문서 간 통과 집계 또는 계산이 필요한 답변을 아직 지원하지 않습니다.

4.내 콘텐츠가 다른 고객을 위한 품질개선에 사용되나요?

Amazon Kendra가 처리하는 콘텐츠는 다른 고객을 위한 Amazon Kendra 또는 기타 Amazon 기계 학습/인공지능 기술을 개발하거나 품질을 개선하는 데 사용되지 않습니다.

Reference

1. Create and query an index with Amazon Kendra(Beginner)

<https://aws.amazon.com/ko/getting-started/hands-on/create-query-index-with-amazon-kendra/>

2. Amazon Kendra Immersion Day

<https://immersionday.com/kendra-immersion-day/>

3. Welcome to Kendra Essentials+

<https://catalog.us-east-1.prod.workshops.aws/v2/workshops/df64824d-abbe-4b0d-8b31-8752bceabade/en-US/100-getting-started/111-creating-an-index>

4. Amazon DUS 사이트 실행(300)

<https://aws.amazon.com/ko/solutions/implementations/document-understanding-solution/>

5. (Blog) AWS launches machine learning enabled search capabilities for COVID-9 Dataset

<https://aws.amazon.com/ko/blogs/publicsector/aws-launches-machine-learning-enabled-search-capabilities-covid-19-dataset/>

6. (Blog)Using Amazon Translate to provide language support to Amazon Kendra

<https://aws.amazon.com/ko/blogs/machine-learning/using-amazon-translate-to-provide-language-support-to-amazon-kendra/>

7. Amazon Kendra Developer Guide

<https://docs.aws.amazon.com/kendra/latest/dg/kendra-dg.pdf>



AI & ML 리소스 허브

AWS가 제공하는 AI 및 ML에 관한 다양한 자료들을 통해 더욱 심층적으로 학습해보세요!

- 기계 학습 여정 가이드
- 기계 학습의 7가지 주요 사용 사례
- 데이터, 분석 및 기계 학습을 위한 전략 플레이북
- 올바른 클라우드 서비스 및 인프라를 통한 기계 학습 혁신 가속화 전략 가이드
- 기계 학습에 적합한 컴퓨팅 인프라 선택 가이드
- 컨택트 센터의 서비스 개선 및 비용 절감 방법
- + 외의 다양한 동영상 학습 자료 및 기술 학습 자료!



<https://bit.ly/3yUk0Kx>

리소스 허브 방문하기



AWS Innovate - AI/ML 특집에 참석해주셔서 대단히 감사합니다.

저희가 준비한 강연, 어떻게 보셨나요?
더 나은 세미나를 위하여 **설문을 꼭 작성해 주세요!**



aws-korea-marketing@amazon.com



twitter.com/AWSKorea



facebook.com/amazonwebservices.ko



youtube.com/user/AWSKorea



linkedin.com/company/amazon-web-services



twitch.tv/aws

Thank you!