

AWS re:Invent

NOV. 28 – DEC. 2, 2022 | LAS VEGAS, NV

ANT307

Improve search relevance with ML in Amazon OpenSearch Service

Jianwei Li

Senior Analytics Specialist TAM
AWS

Muthu Pitchaimani

Senior Specialist Solution Architect
AWS



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

Agenda

What is search relevance?

How to measure relevance?

Improving relevance

Text search versus semantic search

Semantic search architecture

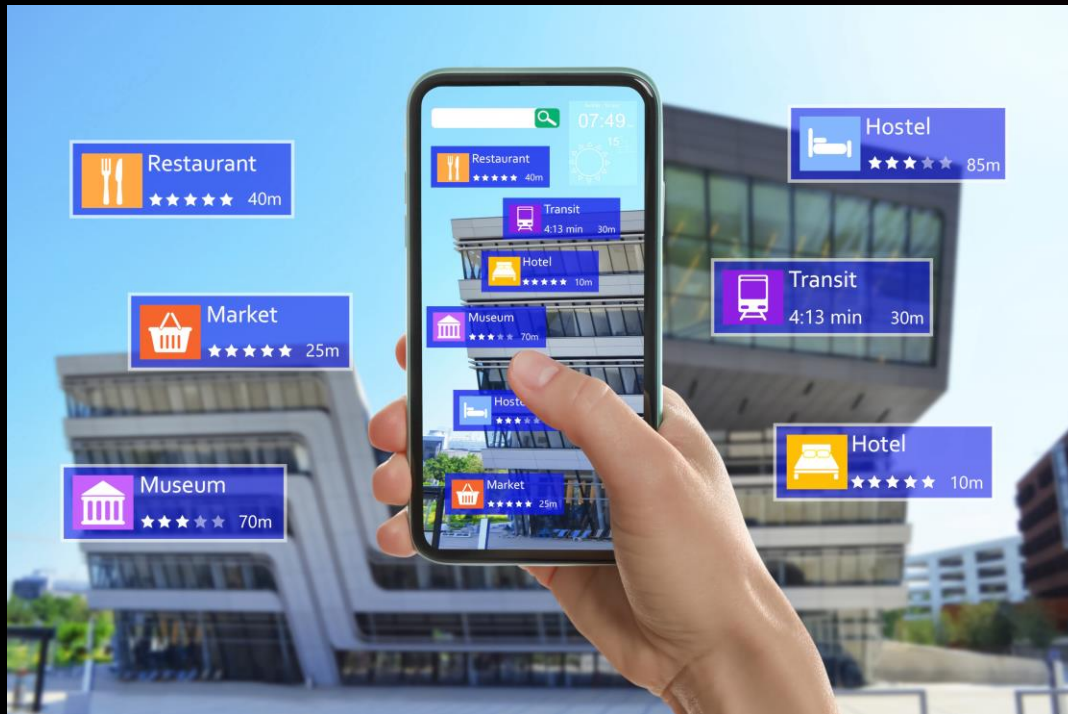
Workshop introduction

Get started

Q&A



What is search relevance?



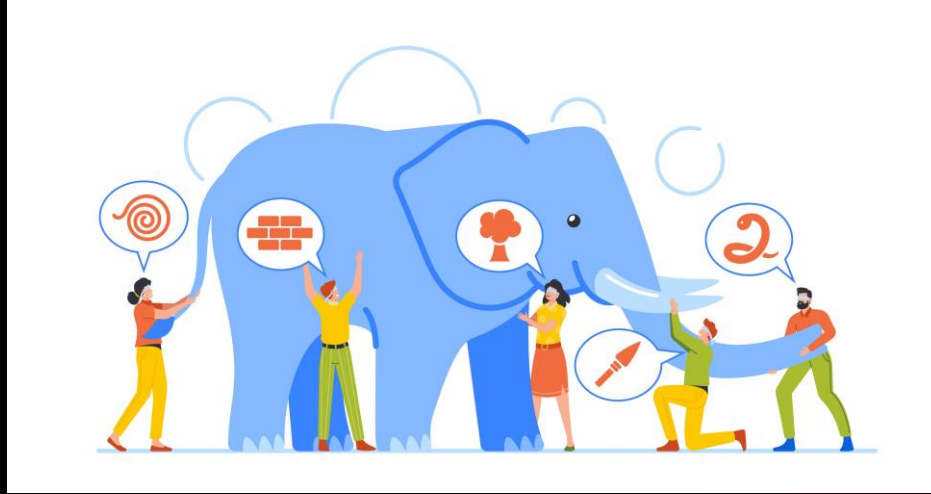
“Relevance is a measure of how good the results are in satisfying a particular information need”

Web search, shopping, food delivery, watch a movie...

How to measure relevance?



How do you tell whether something is relevant or not?



Relevance is inherently subjective
Search for "big bang theory"



Fraction of relevant results
vs.
Fraction of relevant results
retrieved

Improving relevance with ML

- Is ranking and scoring enough?
 - Matching, advanced statistics, synonyms (text search) can take you only so far
- Various techniques: learning to boost, learning to rank, vector similarity with ML
- Understanding meaning and context behind queries become important
- Vector similarity with machine learning (ML)
 - Not mutually exclusive with text search
 - Hybrid search expectedly higher relevance than simple matching

Text search

Query

does this work with my xbox?



Keyword/ Full text search



1. Inverted index
2. BM25 score

Result

does this work for xbox one S?



does this work for ps4?



does this work with pc?



does it work with Lucent 6416D?



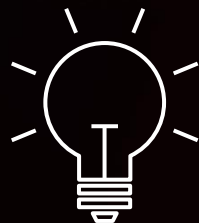
Semantic search

Query

does this work with my xbox?



Semantic search



1. Text to vector with semantic
2. Vector data search

Result

Do they **work with xbox** one?



Can I use this for the **xbox** one?



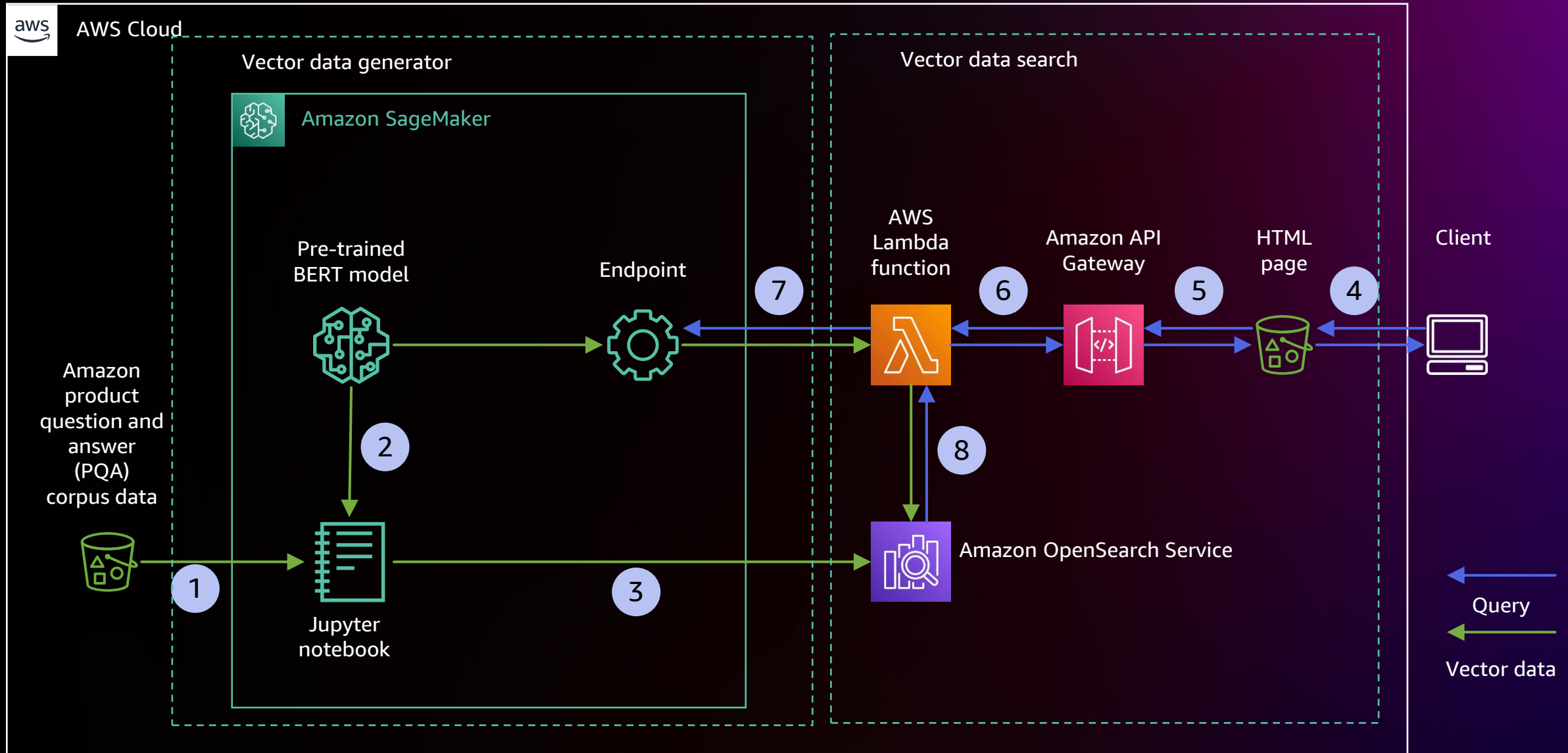
How do I use this headset **with my xbox** one?



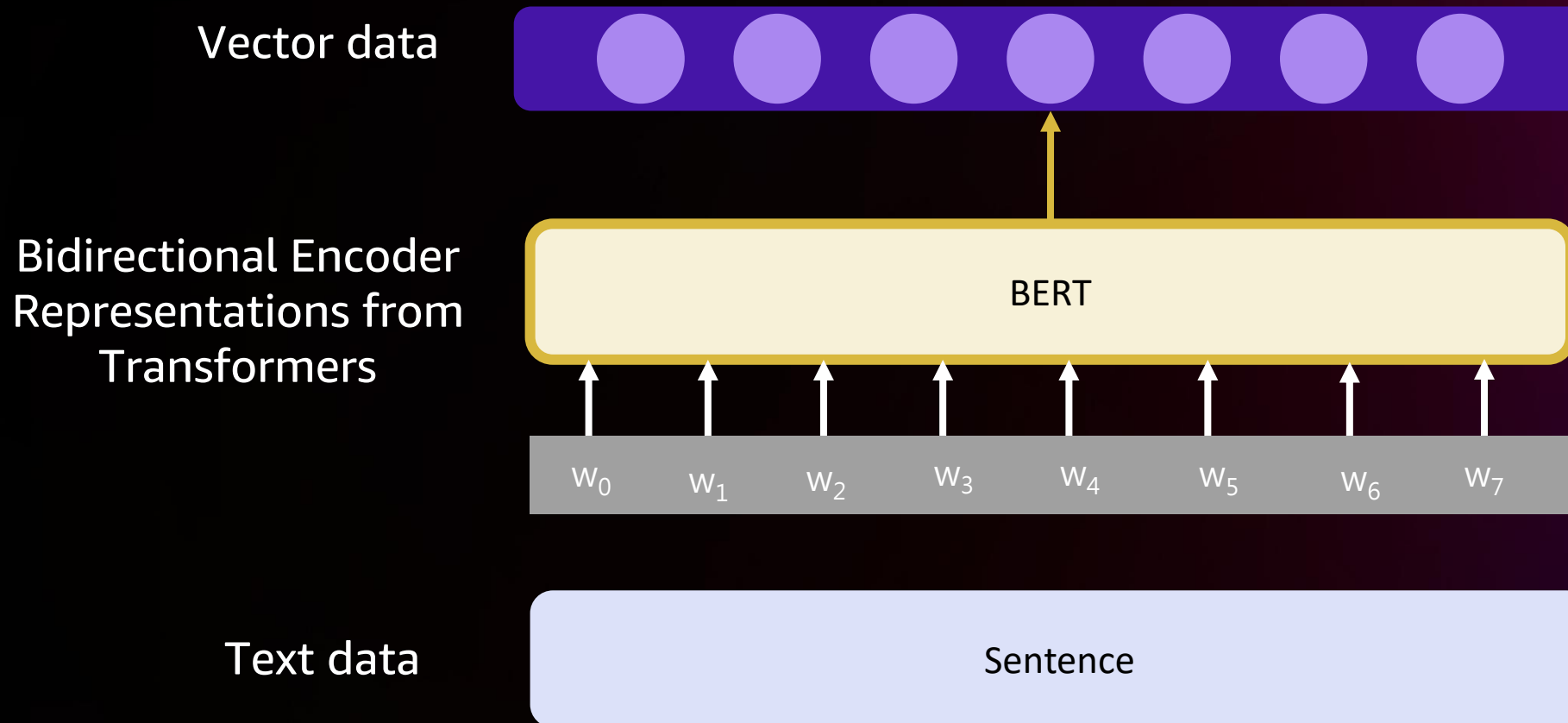
Is this headset compatible **with xbox 1S** or above?



Semantic search architecture

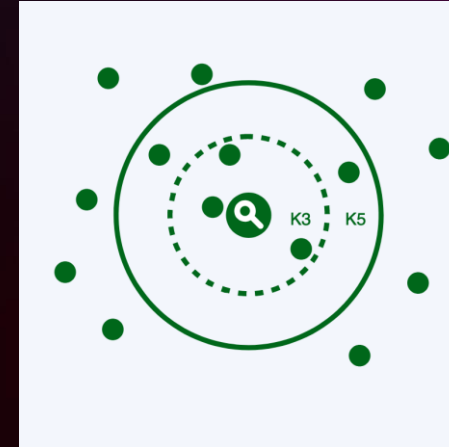


Vector data generator

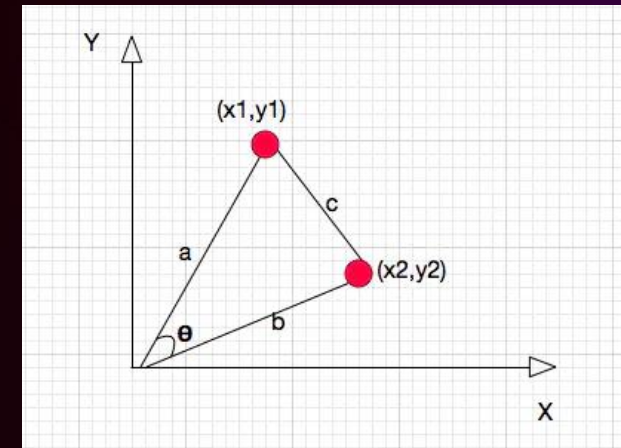


Vector data search: kNN

- Create vector data with feature data
- Calculate vector data similarity in the vector space
- Euclidean distance or cosine similarity

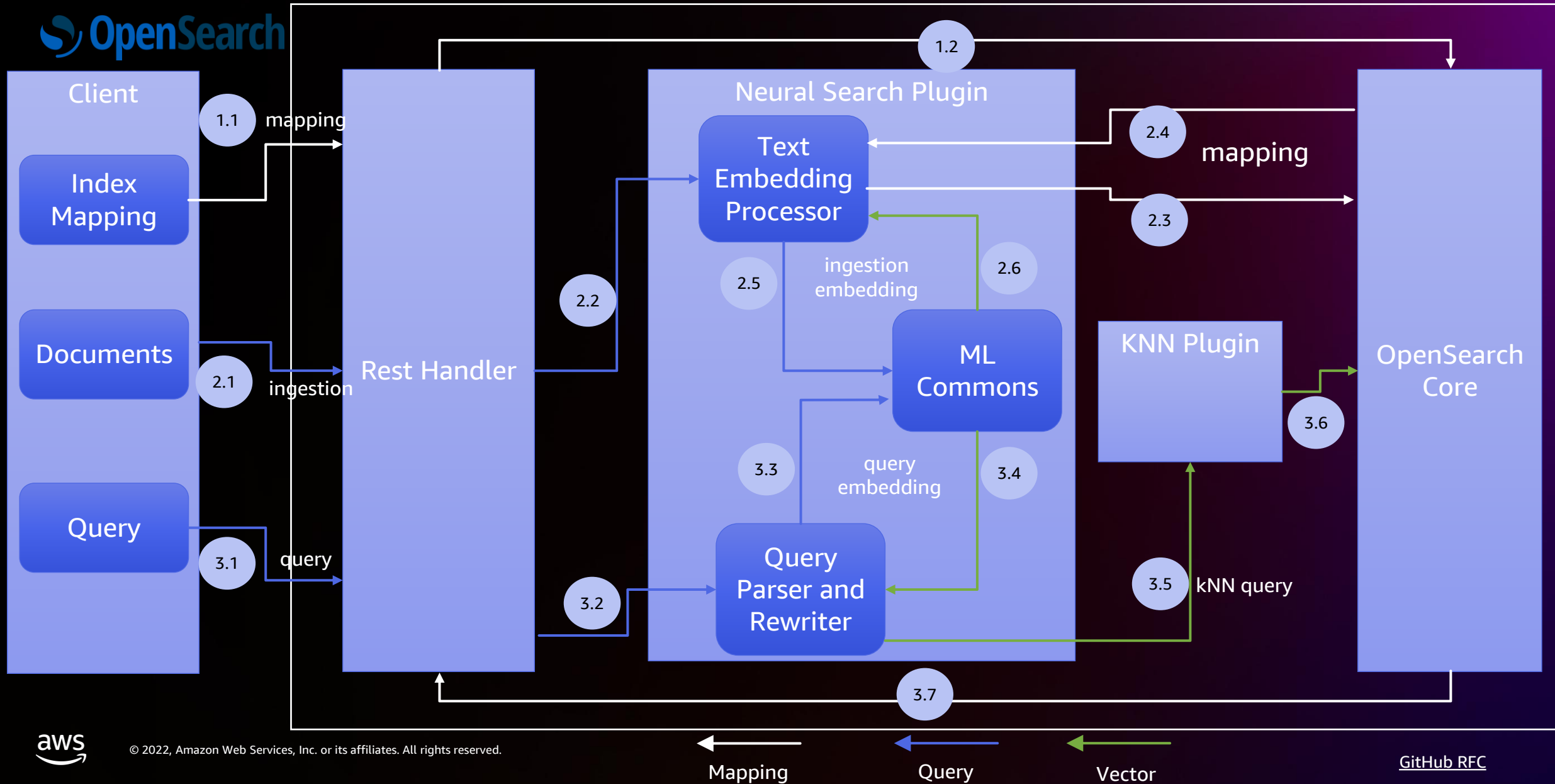


$$\text{similarity} = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}}.$$



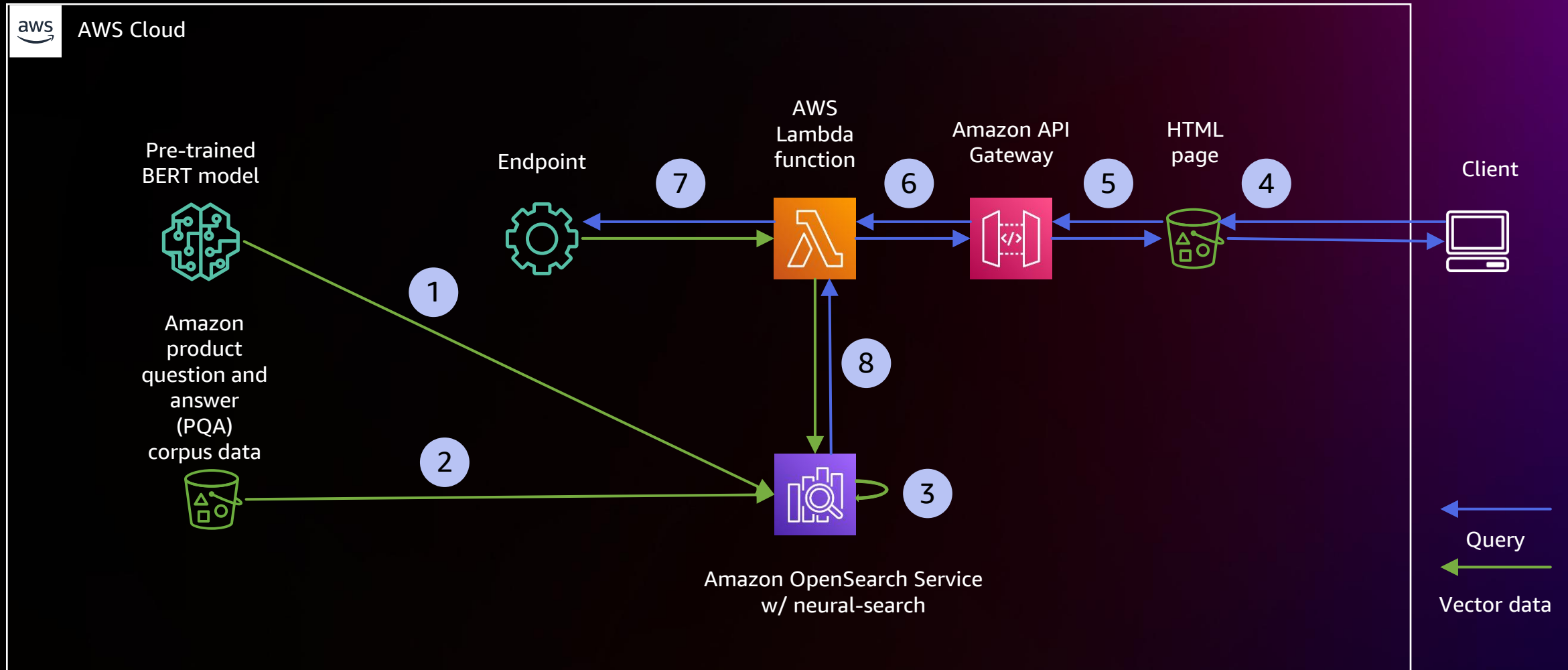
OpenSearch neural-search plugin

New in OpenSearch 2.4



Semantic search with Neural Search

New in OpenSearch 2.4



Workshop content

- Module 1: Fundamental search concepts
- Module 2: Text search
- Module 3: Semantic search
- Module 4: Full-stack semantic search
- Module 5: Semantic search with fine-tuned model (optional)

Get started with this workshop

- As a participant, you will have access to an AWS account with any optional pre-provisioned infrastructure and IAM policies needed to complete this workshop
- The AWS account will only be available for the duration of this workshop; you will lose access to the account thereafter
- The pre-provisioned infrastructure are deployed to a **us-east-1** Region
- Be sure to review the terms and conditions of the event; do not upload any personal or confidential information in the account

Step 1: Sign-in via OTP

<https://catalog.workshops.aws/join>



aws workshop studio

Workshop Studio > Sign in

Sign in

Choose a preferred sign-in method

Email one-time password (OTP)

Enter your personal or corporate email to receive a one-time password

Login with Amazon

Login with your Amazon.com retail account

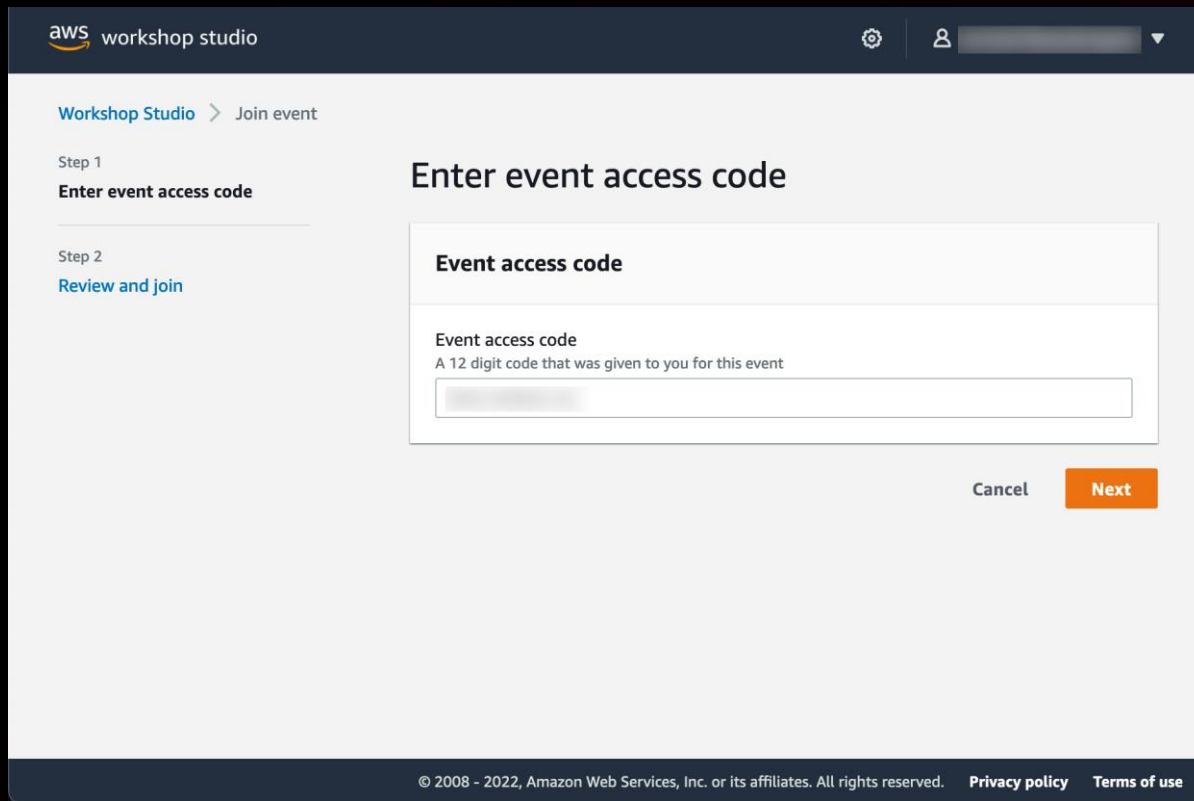
Amazon employee

Login with your Amazon Corporate account. Only for Amazon Employees.

© 2008 - 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved. [Privacy policy](#) [Terms of use](#)

Step 2: Enter event access code

Enter 12-digit event access code



The screenshot shows the AWS Workshop Studio interface. At the top, there's a header with the AWS logo and 'workshop studio'. Below the header, a breadcrumb trail shows 'Workshop Studio > Join event'. The main content area is titled 'Enter event access code'. On the left, there's a sidebar with 'Step 1: Enter event access code' (active) and 'Step 2: Review and join'. The main form area has a title 'Enter event access code' and a section 'Event access code' with a description: 'A 12 digit code that was given to you for this event'. Below this is a text input field. At the bottom right of the form are 'Cancel' and 'Next' buttons. The footer contains copyright information and links to 'Privacy policy' and 'Terms of use'.

7fcc-064e80-15

Step 3: Review terms and join event

[Workshop Studio](#) > [Join event](#)

Step 1

[Enter event access code](#)

Step 2

Review and join

Review and join

Event details

Name	Start time	Duration	Level
Improve search relevance with ML in Amazon OpenSearch Service event	10/13/2022 11:08 AM	2 hours	-

Description

Test event for content Improve search relevance with ML in Amazon OpenSearch Service

Terms and Conditions

Read and accept before joining the event

1. By using AWS Workshop Studio for the relevant event, you agree to the AWS Event Terms and Conditions and the AWS Acceptable Use Policy. You acknowledge and agree that are using an AWS-owned account that you can only access for the duration of the relevant event. If you find residual resources or materials in the AWS-owned account, you will make us aware and cease use of the account. AWS reserves the right to terminate the account and delete the contents at any time.
2. You will not: (a) process or run any operation on any data other than test data sets or lab-approved materials by AWS, and (b) copy, import, export or otherwise create derivative works of materials provided by AWS, including but not limited to, data sets.
3. AWS is under no obligation to enable the transmission of your materials through Event Engine and may, in its discretion, edit, block, refuse to post, or remove your materials at any time.
4. Your use of AWS Workshop Studio will comply with these terms and all applicable laws, and your access to AWS Workshop Studio will immediately and automatically terminate if you do not comply with any of these terms or conditions.

☒ I agree with the Terms and Conditions

[Cancel](#)

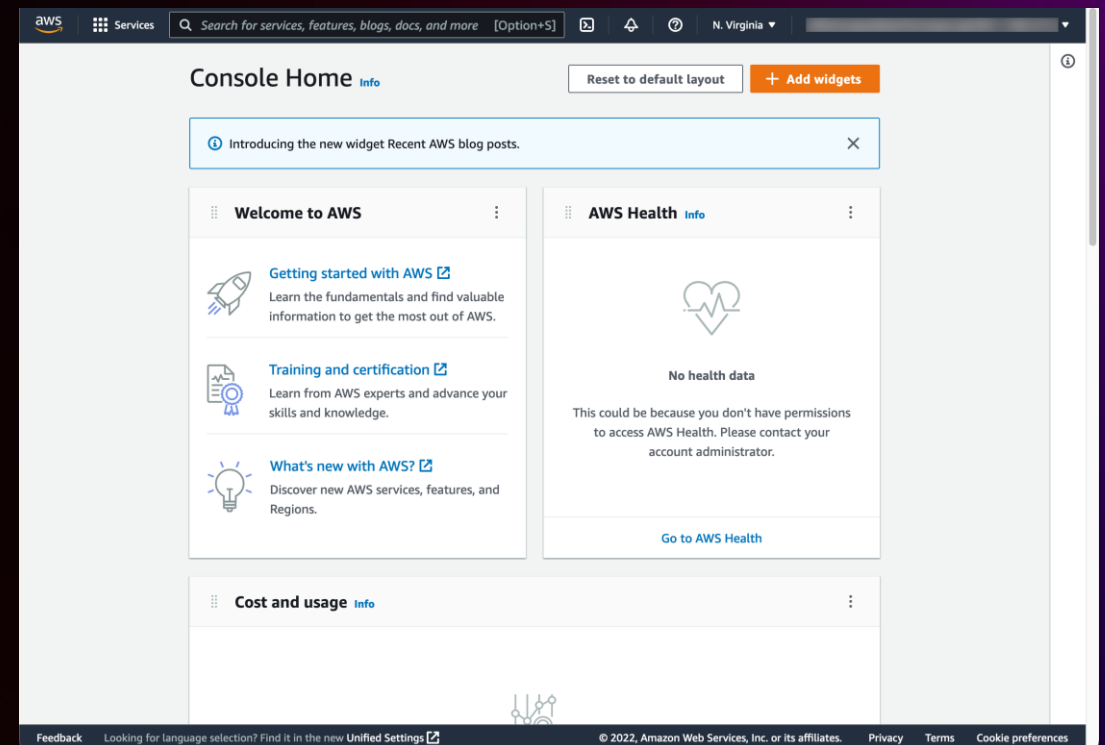
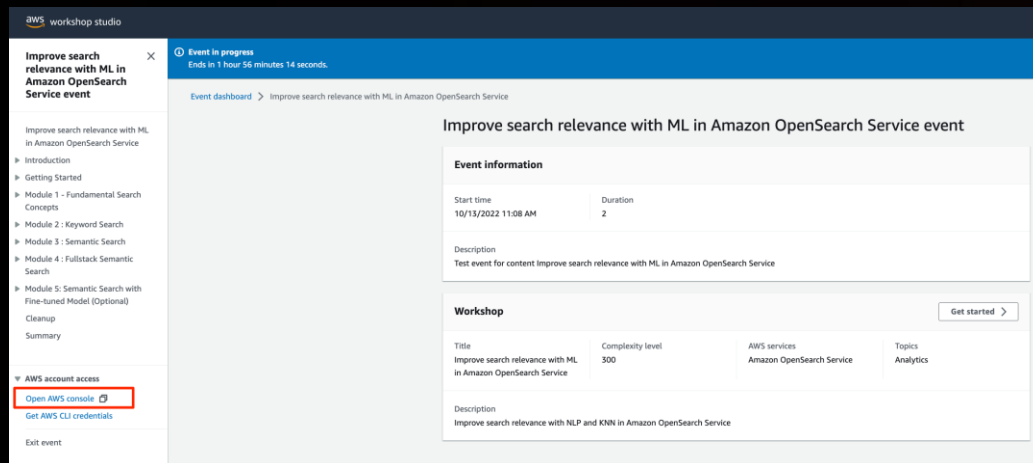
[Previous](#)

[Join event](#)



Step 4: Access AWS account

Access the AWS Management Console or generate AWS CLI credentials as needed



Step 5: Get started with the workshop

Event In progress
Ends in 1 hour 55 minutes.

Event dashboard > Improve search relevance with ML in Amazon OpenSearch Service

Improve search relevance with ML in Amazon OpenSearch Service event

Event information

Start time	Duration
10/13/2022 11:08 AM	2

Description
Test event for content Improve search relevance with ML in Amazon OpenSearch Service

Workshop Get started >

Title	Complexity level	AWS services	Topics
Improve search relevance with ML in Amazon OpenSearch Service	300	Amazon OpenSearch Service	Analytics

Description
Improve search relevance with NLP and KNN in Amazon OpenSearch Service



aws workshop studio

Event In progress
Ends in 1 hour 53 minutes 34 seconds.

Event dashboard > Improve search relevance with ML in Amazon OpenSearch Service

Improve search relevance with ML in Amazon OpenSearch Service

Keyword Search

Keyword/Full text search refers to searching some text inside extensive text data and returning results that contain some or all of the words from the query.

Semantic Search

Semantic search uses machine learning to understand the meaning of search queries, rather than simply literal word matches. Semantic search improves search accuracy by understanding the intent and contextual meaning of terms to generate more relevant results.

Workshop facilitators

- Jey Vell, Senior SA Manager
- Viral Shah, Principal GTM Specialist
- Josh Tow, Solution Architect
- Arun Lakshmanan, Analytics Specialist SA
- Mamta Vaidya, SA Manager
- Prashant Agrawal, Senior Analytics Specialist SA
- Dilip Rajan, Mgr, Partner Soln Architect
- Rafael Rodrigues, Enterprise Solutions Architect

Get started

Workshop for this event

<https://tinyurl.com/2navfkaf>

Event access code

7fcc-064e80-15

Workshop permanent link

<https://tinyurl.com/592ej6kw>



Thank you!

Jianwei Li

lijianwe@amazon.com

Muthu Pitchaimani

muthupmk@amazon.com



Please complete the session survey in the **mobile app**



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