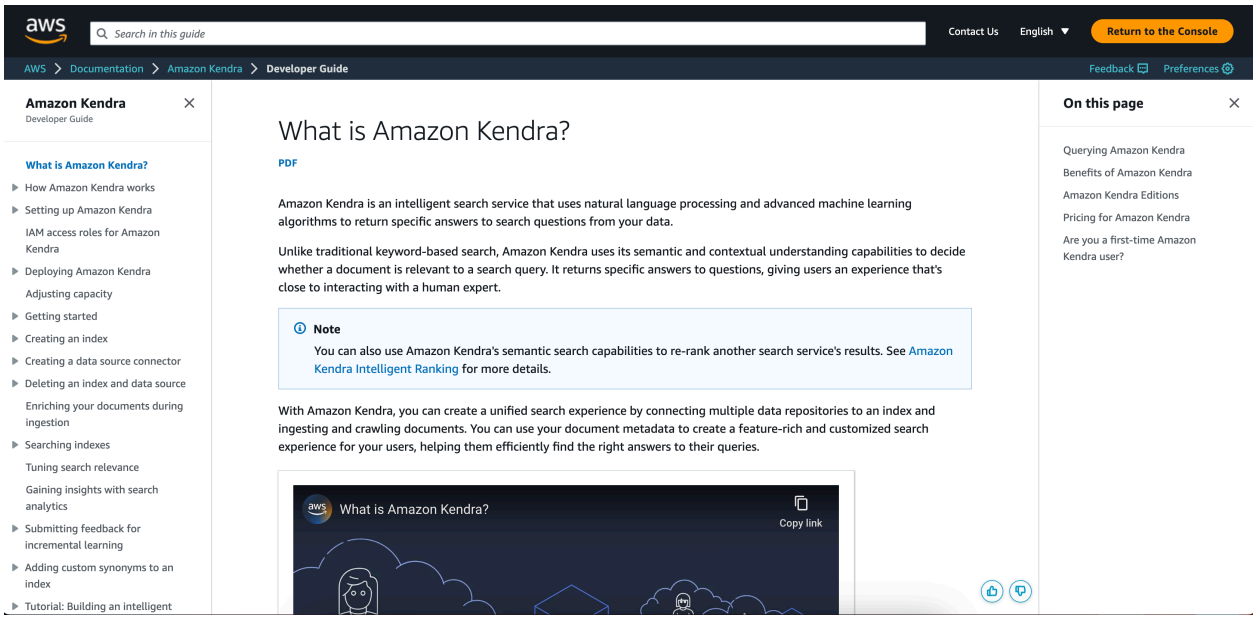
AWS Kendra learning research

I learnt this mostly from the official website of AWS.

AWS Kendra is to find answers faster with intelligent enterprise search powered by machine learning. I think it is like another ChatGPT within AWS.
I am currently using Atlassian Confluence as enterprise search tool in my current company, but probably Confluence is not so intelligent like AWS Kendra.

Here is the official Website of AWS Kendra tutorial:



I read through all of the pages of the tutorials and I want to list some important points here:

What is Amazon Kendra?
Amazon Kendra is an intelligent search service that uses natural language processing and advanced machine learning algorithms to return specific answers to search questions from your data.
Unlike traditional keyword-based search, Amazon Kendra uses its semantic and contextual understanding capabilities to decide whether a document is relevant to a search query. It returns specific answers to questions, giving users an experience that's close to interacting with a human expert.
With Amazon Kendra, you can create a unified search experience by connecting multiple data repositories to an index and ingesting and crawling documents. You can use your document metadata to create a feature-rich and customized search experience for your users, helping them efficiently find the right answers to their queries.

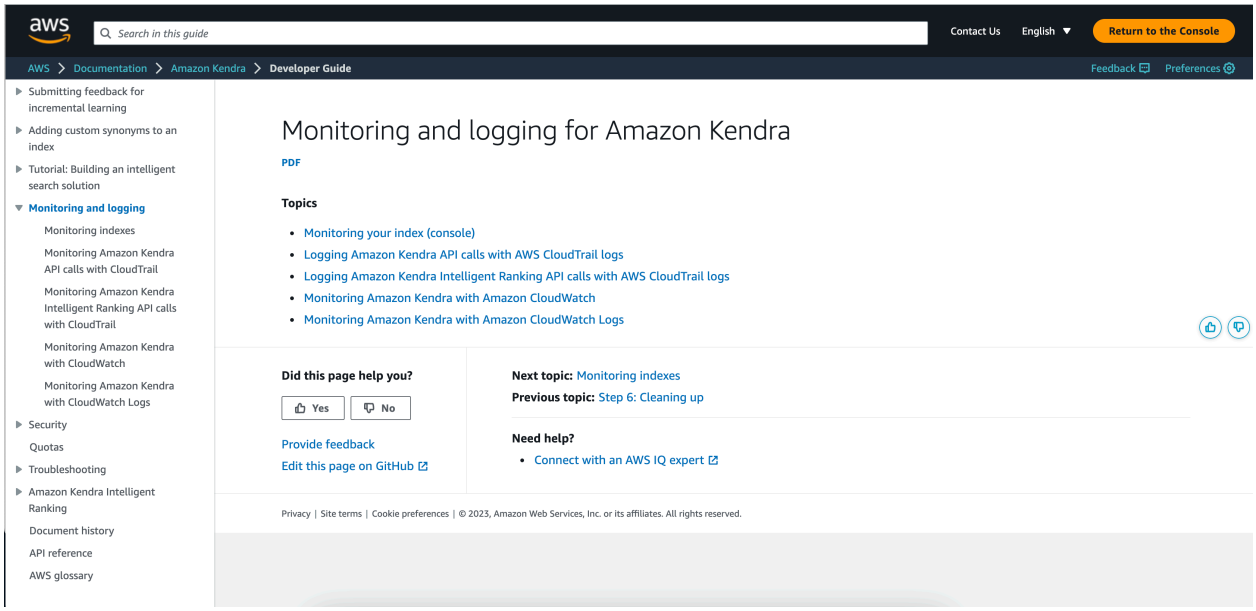
Querying Amazon Kendra
You can ask Amazon Kendra the following types of queries:
Factoid questions—Simple who, what, when, or where questions, such as Where is the nearest service center to Seattle? Factoid questions have fact-based answers that can be returned as a single word or phrase. The answer is retrieved from a FAQ or from your indexed documents.

Descriptive questions—Questions where the answer could be a sentence, passage, or an entire document. For example, How do I connect my Echo Plus to my network? Or, How do I get tax benefits for lower income families?

Keyword and natural language questions—Questions that include complex, conversational content where the intended meaning may not be clear. For example, keynote address. When Amazon Kendra encounters a word like “address”—which has multiple contextual meanings—it correctly infers the intended meaning behind the search query and returns relevant information.

Monitoring and logging for Amazon Kendra

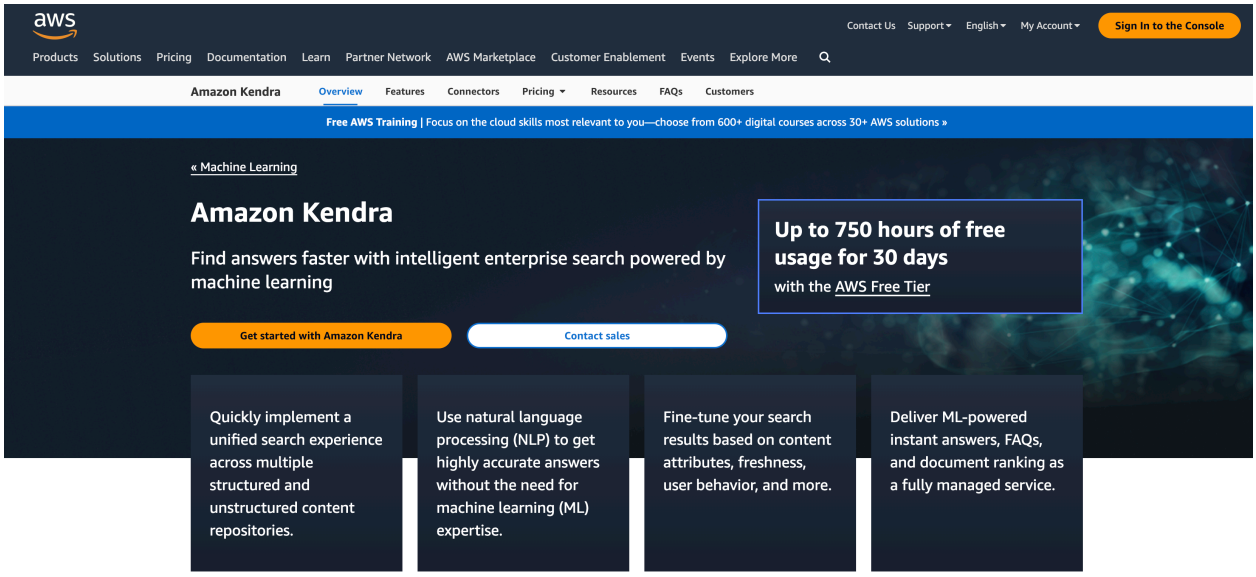
Here is the screenshot of where the tutorial is:



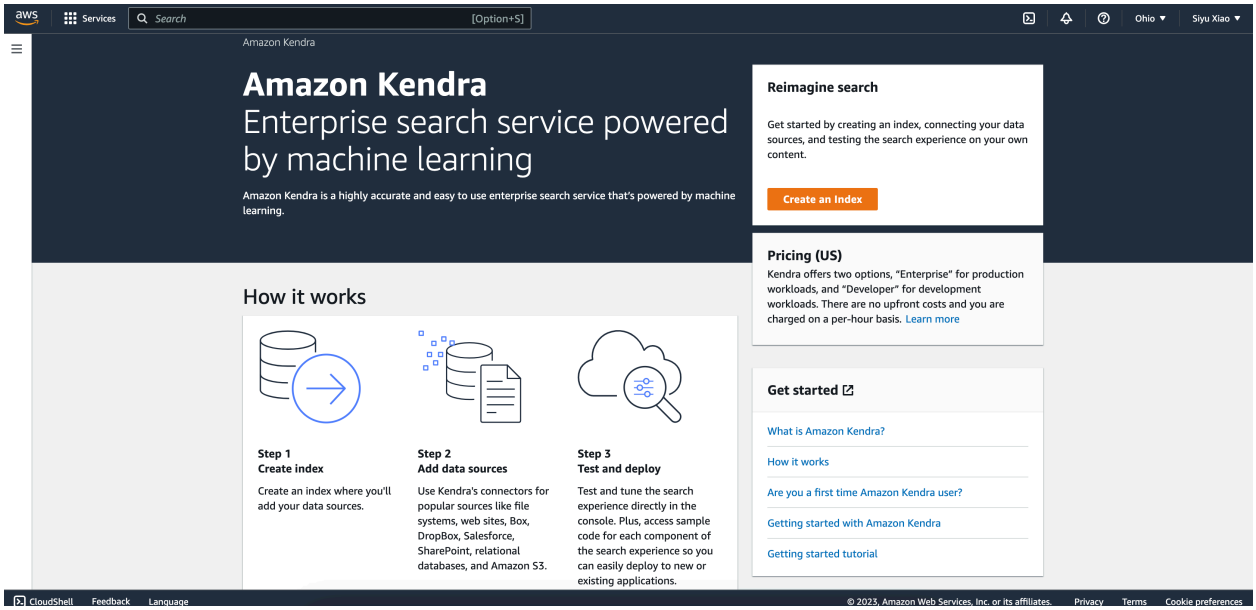
There are 3 topics of Kendra monitoring: Monitoring your index (console), Monitoring Amazon Kendra with Amazon CloudWatch, Monitoring Amazon Kendra with Amazon CloudWatch Logs. I have read these 3 topics and I will remember to find monitoring solutions when I need in the future.

Try to build a demo Kendra service

I actually learnt Kendra, not just copy and paste the tutorial here, I log in my AWS account and try to use Kendra, here is the proof:



Then I clicked on "Get started with Amazon Kendra" and get this:
Like you have seen here, firstly we need to create index:



aws | Services | Search [Option+S] | Ohio | Siyu Xiao

Amazon Kendra > Indexes > Create index

Step 1
Specify index details

Step 2
Configure user access control

Step 3
Provisioning details

Specify index details Info

Index details

Index name

siyu

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

Description - *optional*

Logging details Info

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 Info

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

IAM roles used for domains, data sources, or FAQs can't be used for indexes. If you are unsure if an existing role is used for a domain, data source, or FAQ, choose "Create a new role" to avoid an error.

Create a new role (Recommended)

Role name

Your role name will be prefixed with 'AmazonKendra-us-east-2-'.

AmazonKendra-us-east-2-siyu

aws | Services | Search [Option+S] | Ohio | Siyu Xiao

Amazon Kendra > Indexes > Create index

Step 1
Specify index details

Step 2
Configure user access control

Step 3
Provisioning details

Configure user access control Info

Configure document-level access control for indexed documents with automatic filtering based on user ID and user groups.

Access control settings Info

Use tokens for access control?

Choose whether to implement token-based access control for this index. You can change this later.

- (selected) **No** — Choose this option to make all indexed content searchable and displayable for all users. Any access control list is ignored, but you can filter on user and group attributes.
- **Yes** — Choose this option to enable token-based user access control. All documents with no access control and documents accessible to the user are searchable and displayable.

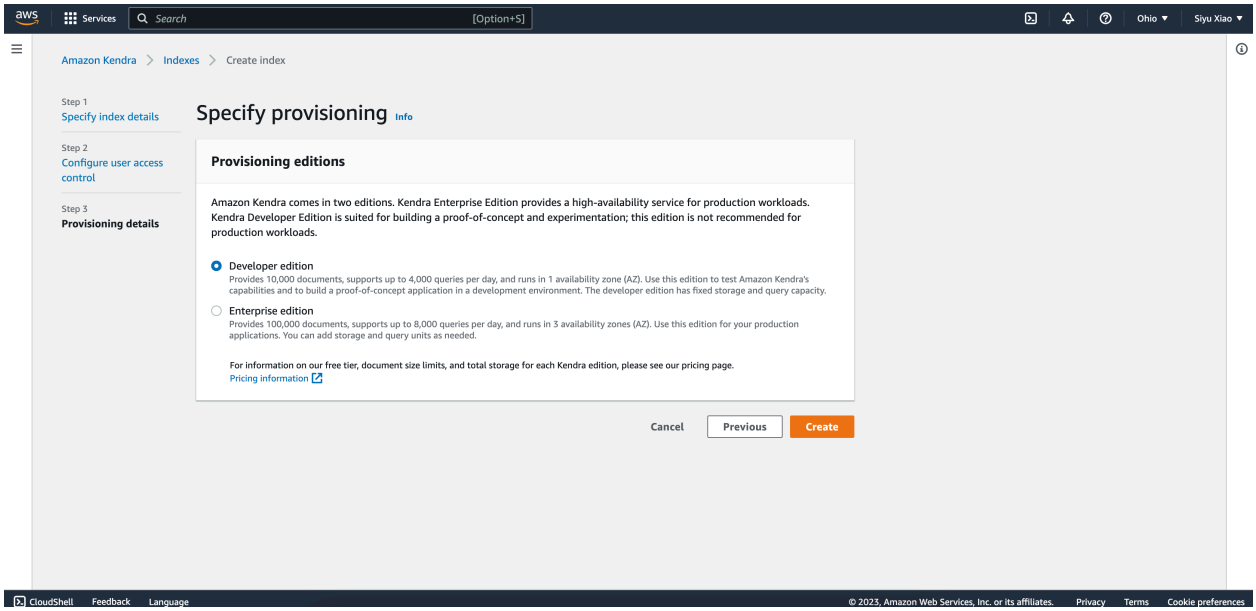
User-group expansion

Choose whether to look up user-groups with AWS IAM Identity Center integration.

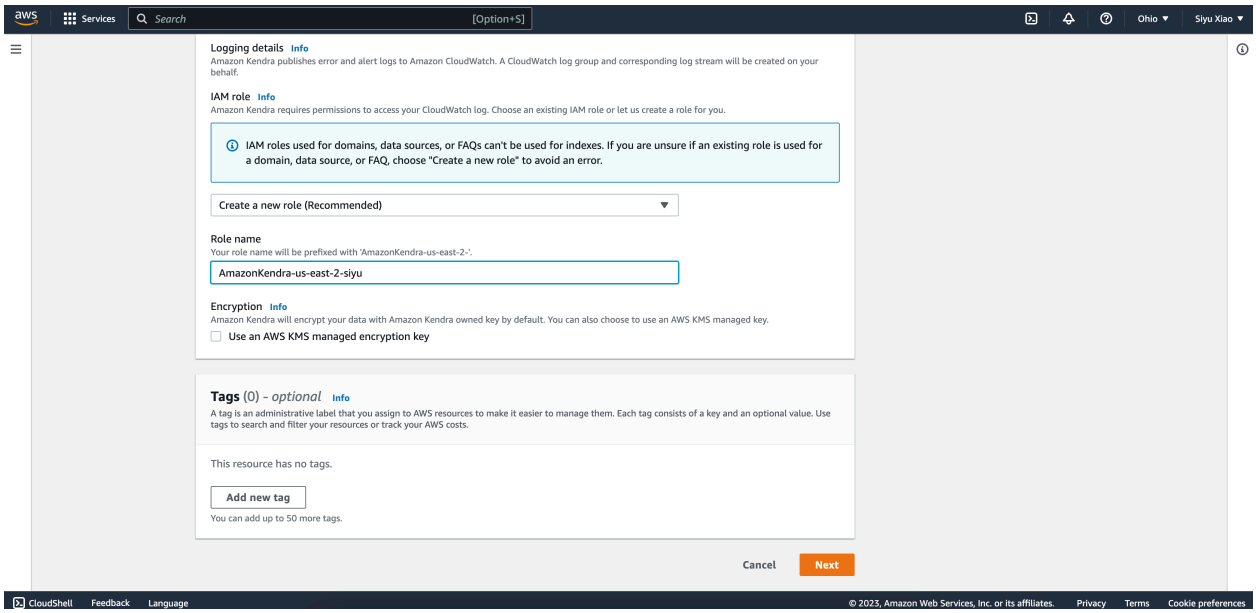
- (selected) **None** — Group information in the query request will be used to determine the group.
- **AWS IAM Identity Center** — Group information in Identity Providers (eg. Active Directory) will be used to determine the group. You first must enable AWS IAM Identity Center and create an organization to sync users and groups from your active directory.

Cancel | Previous | Next

CloudShell | Feedback | Language | © 2023, Amazon Web Services, Inc. or its affiliates. | Privacy | Terms | Cookie preferences

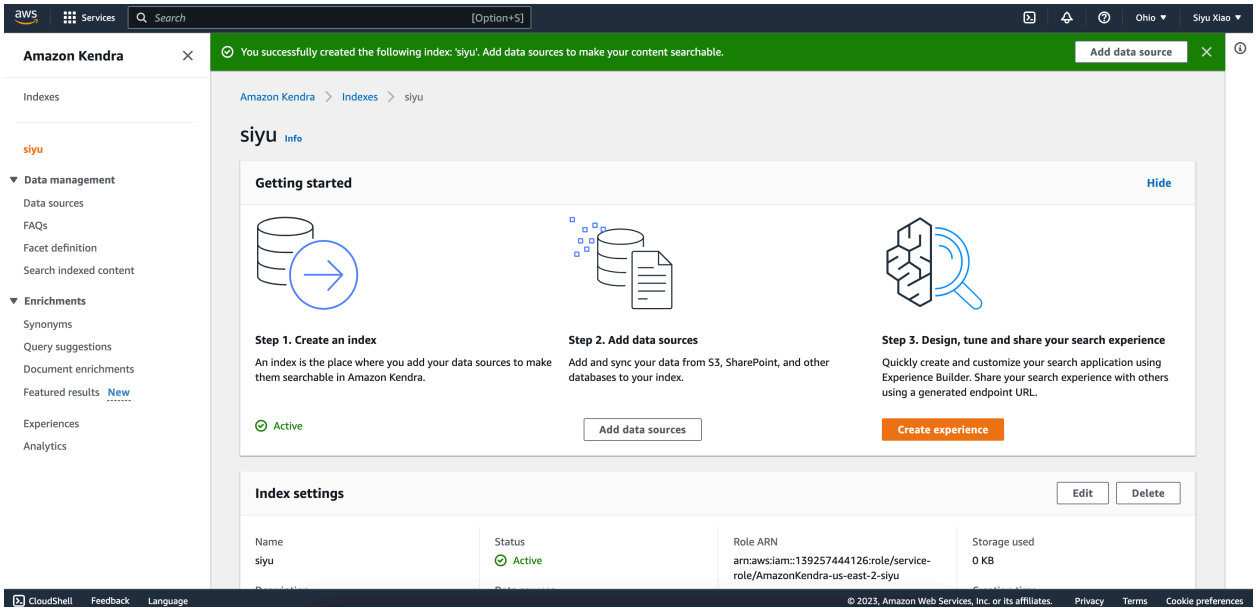


Here is an encryption option, which means Kendra will encrypt your data with Amazon Kendra owned key by default. Considering we are just learning to run a demo job and the data does not need to be encrypted, I did not select it.

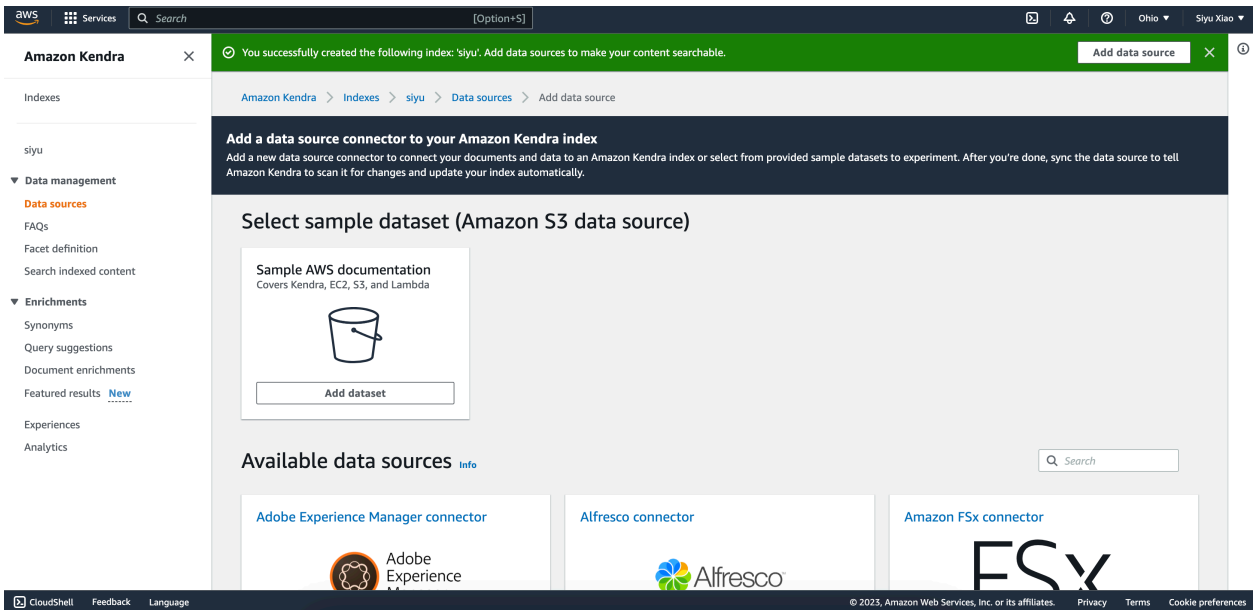


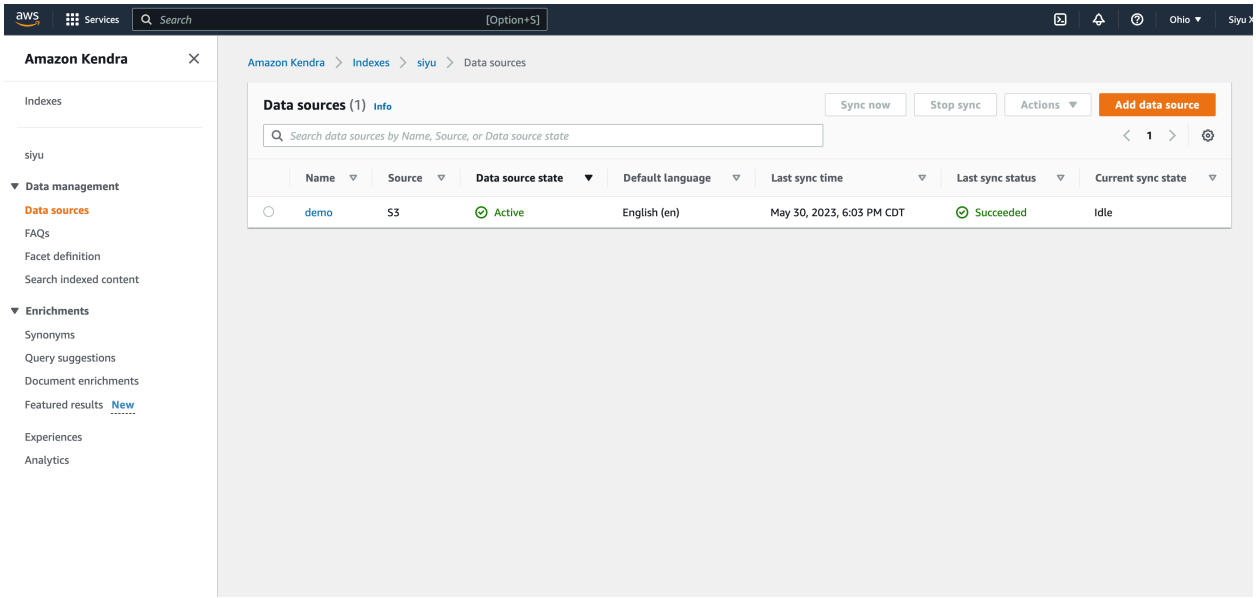
Then click Create. Here is when Kendra is creating the index "siyu":

Here you can see I successfully created it:

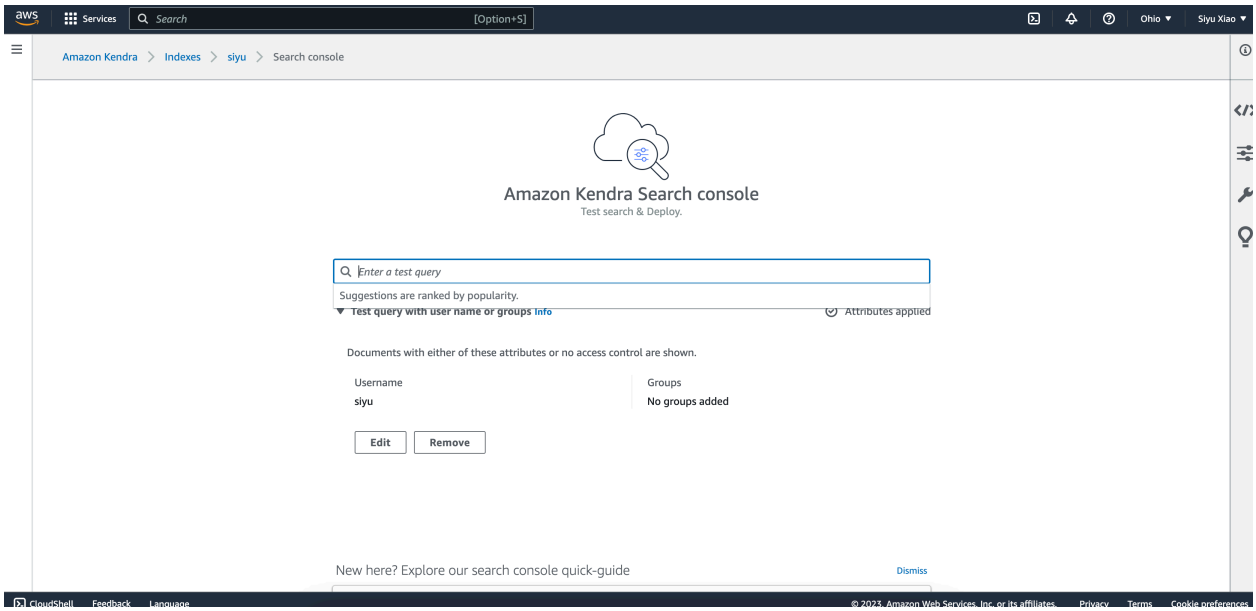


Then we can do the Step 2: Add data sources:
Here I click on Add dataset and follow the steps and created a data source successfully:

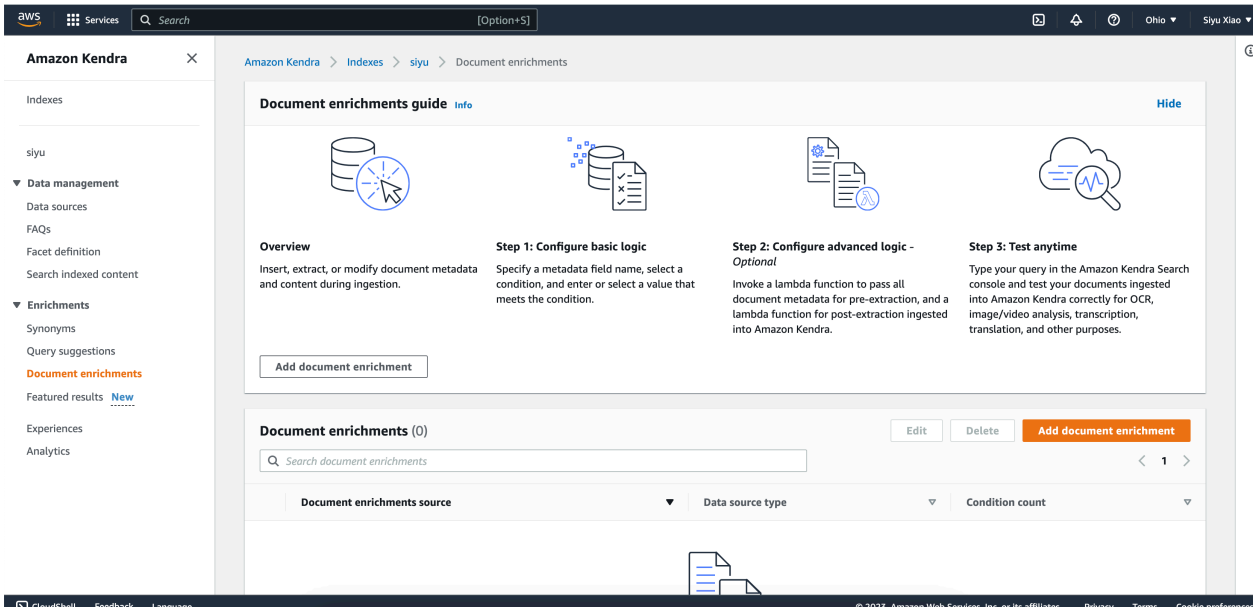




Then we can search in Kendra console:



We can use Document enrichments to preprocess the documents and add data that can be used for filtering or boosting.



I can also create experience using Experience Builder and share my search experience to others with the generated URL. There are too many screen shots for other functions in Kendra and I will not put all of them here.

So to sum up, we firstly create an index, then we add data sources to make content searchable, then we design, tune and share the search experience I want for end users. This is the most important and basic steps to use Amazon Kendra.