



Amazon Comprehend

Amazon Comprehend is a fully managed natural language processing (NLP) service by AWS that helps you uncover insights and relationships in text using machine learning. It's designed to process large amounts of unstructured data and provides capabilities to extract key information from the text, such as the overall sentiment, entities, key phrases, and more. You don't need prior expertise in machine learning to use it, and it automatically scales based on the volume of text.

Key Features:

1. Sentiment Analysis:

- Identifies the sentiment (positive, negative, neutral, or mixed) within the text. It is commonly used in applications like customer feedback analysis to understand user opinions.

2. Entity Recognition:

- Detects and categorizes entities such as people, locations, organizations, dates, and quantities within the text.

3. Key Phrase Extraction:

- Extracts important phrases from a document, helping to summarize the core information.

4. Language Detection:

- Automatically detects the language of the input text. This is useful in multi-language environments to route text to the appropriate processing service.

5. Text Classification:

- Classifies documents into predefined or custom categories. For example, it can categorize support tickets by type, product, or urgency.

6. Custom Entity Recognition:

- You can train the service to recognize custom entities specific to your use case (e.g., product codes, industry terms).

7. Topic Modeling:

- Helps you analyze a collection of documents by grouping them into different topics or themes based on the contents of the text.

How It Works:

Amazon Comprehend uses machine learning models trained on a large corpus of text data. You can either use the pre-trained models or train your own custom models using your data. Once the text is submitted to the service, Comprehend processes the information and returns insights such as detected entities, sentiments, and key phrases.

Common Use Cases:

- **Customer Support:** Analyze support tickets to detect common complaints and themes.
- **Sentiment Analysis:** Gauge customer sentiment from social media, reviews, or feedback forms.
- **Text Classification:** Automatically sort emails or documents into specific categories.
- **Document Search and Organization:** Enhance search functionality by tagging documents with topics, keywords, and entities.
- **Media Monitoring:** Analyze large-scale text sources such as news articles or social media posts for relevant mentions or sentiment trends.

Integration with Other AWS Services:

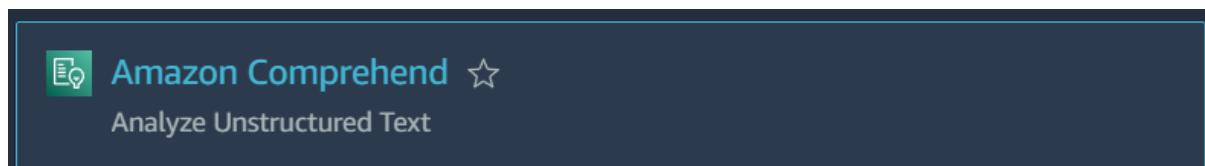
Amazon Comprehend can be integrated with other AWS services, such as:

- **Amazon S3:** Store text data in S3 and use Comprehend to process the files.
- **AWS Lambda:** Trigger text analysis based on events.
- **Amazon Kinesis:** Stream text data for real-time analysis.

Overall, Amazon Comprehend is a powerful tool for businesses looking to derive actionable insights from large volumes of text, allowing them to improve customer service, monitor brand reputation, and organize information more effectively.

To begin with this Lab:

1. In your AWS Console, search for Amazon Comprehend and choose the service accordingly.



2. Below you can see the dashboard for Amazon Comprehend. Click on launch.

Machine learning

Amazon Comprehend

Natural Language Processing and Text Analytics

Amazon Comprehend is a natural language processing (NLP) service that uses machine learning to find insights and relationships in text.

Start analyzing text

Jump in and try our APIs for Amazon Comprehend.

[Launch Amazon Comprehend](#)

How it works



1. Need to extract insights

Social media posts, emails, web pages, documents, phone transcripts and medical records.

Pricing (US)

With Amazon Comprehend, you pay only for what you use. You are charged based on the amount of text processed on a monthly basis, and there are no minimum fees and no upfront commitments.

[Learn more](#)

3. This tool does the real time analysis of our text. Basically, what you can do is generate a random input text and then paste it here and click on Analyze.

Input data

[Supported languages](#)

Analysis type [Info](#)

Built-in

View real-time insights based on AWS built-in models.

Custom

View real-time insights based on custom models from an endpoint you've created.

Input text

456 Elm Street
Anytown, ST 12345

Dear Alex,

I am writing to enthusiastically recommend the Sustainability Summit 2024, which will take place on November 15, 2024, at the Anytown Convention Center. As someone who has been involved in the planning and execution of numerous successful events, I can confidently say that this summit promises to be an outstanding

1811 of 5000 characters used.

[Clear text](#)

[Analyze](#)

4. Once it has analyzed the text it will give you the entities, key phrases, language, PII, Sentiment, Targeted sentiment and Syntax for your given text.

Insights Info

Entities Key phrases Language PII Sentiment Targeted sentiment Syntax

▼ Analyzed text

Jordan Smith
Event Coordinator
Innovative Solutions Inc.
123 Main Street Anytown, ST 12345
jordan.smith@email.com
(555) 123-4567
October 21, 2024

5. Below is the result for entities. Similarly, you can find different results for different options it is providing to you.

▼ Results

Search

Entity	Type	Confidence
Jordan Smith	Person	0.98
Innovative Solutions Inc.	Organization	0.98
123 Main Street Anytown, ST 12345	Location	0.97
jordan.smith@email.com	Other	0.99+
(555) 123-4567	Other	0.98
October 21, 2024	Date	0.99+
Alex Johnson	Person	0.99+
Director	Person	0.66
Community Engagement Green Future Initiative	Organization	0.96
456 Elm Street Anytown, ST 12345	Location	0.97

6. So, this is the result for Key phrases.

▼ Results

Key phrases	Confidence
Innovative	0.54
October	0.49
2024	0.92
Street	0.65
Summit 2024	0.96
place	0.99+
November 15, 2024	0.99+
the Anytown Convention Center	0.99+
the planning and execution	0.99+
numerous successful events	0.99+

7. For language it is confident that this language is 99% English.

▼ Results

Language

English, en

0.99 confidence

8. In PII you have two options offsets and Labels.

Insights Info

Entities | Key phrases | Language | **PII** | Sentiment | Targeted sentiment | Syntax

▼ Personally identifiable information (PII) analysis mode

Offsets

Identify the location of PII in your text documents.

Labels

Label text documents with PII.

9. In the sentiments, you can see that it 86% neutral and 13% positive. Similarly, you can view all the results.

▼ Results

Sentiment

Neutral

0.86 confidence

Positive

0.13 confidence

Negative

0.00 confidence

Mixed

0.00 confidence

10. Then we used another text and analyzed it, here we can see that it has 98% negative sentiments.

▼ Results

Sentiment

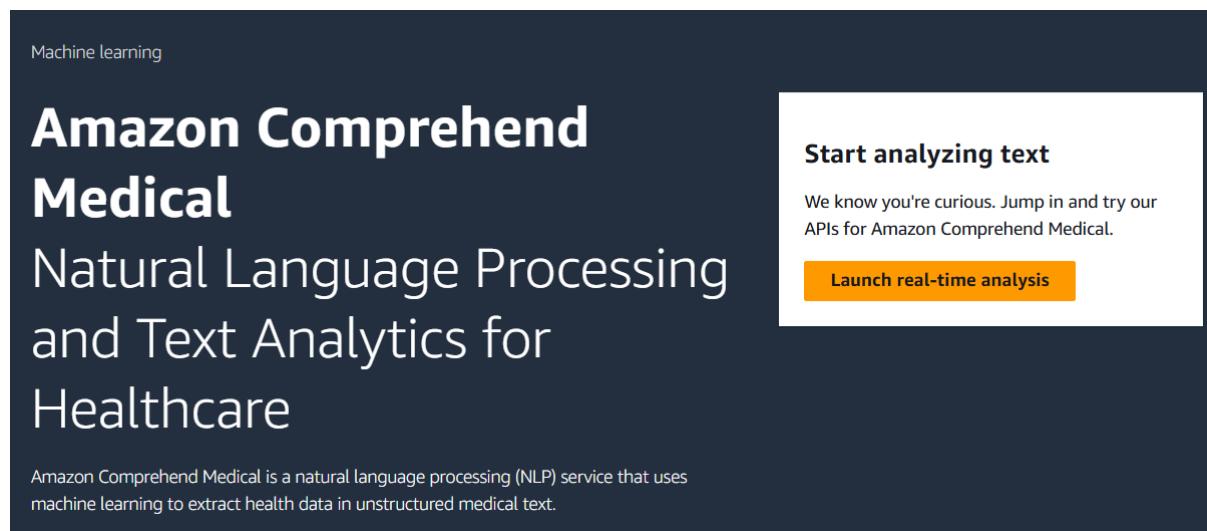
Neutral
0.00 confidence

Positive
0.00 confidence

Negative
0.98 confidence

Mixed
0.01 confidence

-
11. Now there is another service Amazon Comprehend Medical. Here you can analyze medical text. Click on launch.



The screenshot shows the Amazon Comprehend Medical landing page. At the top left, it says "Machine learning". The main title is "Amazon Comprehend Medical" in large white font, followed by a subtitle "Natural Language Processing and Text Analytics for Healthcare". Below the title, a description states: "Amazon Comprehend Medical is a natural language processing (NLP) service that uses machine learning to extract health data in unstructured medical text." To the right, there is a white callout box with a dark border containing the text "Start analyzing text" and "We know you're curious. Jump in and try our APIs for Amazon Comprehend Medical." Below this is a yellow button labeled "Launch real-time analysis".

12. In this we have a predefined text which is already analyzed.

Real-time analysis Info

See how Comprehend Medical recognizes entities related to the healthcare domain. To analyze your text, type or paste it in the text box.

Input text

Supported languages

Pt is 87 yo woman, highschool teacher with past medical history that includes
 - status post cardiac catheterization in April 2019.
 She presents today with palpitations and chest pressure.
 HPI : Sleeping trouble on present dosage of Clonidine. Severe Rash on face and leg, slightly itchy.
 Meds : Vyvanse 50 mgs po at breakfast daily,
 Clonidine 0.2 mgs -- 1 and 1 / 2 tabs po qhs
 HEENT : Boggy inferior turbinates, No oropharyngeal lesion.

566 of 10000 characters used.

[Clear text](#)

[Analyze](#)

13. And in the entities, you have a visual type of representation of your text. You can view all the things you want.

[Entities](#)

[RxNorm concepts](#)

[ICD-10-CM concepts](#)

[SNOMED CT concepts](#)

Analyzed text

