# NLP on public Clouds: Compare NLP-related services provided by AWS, Google Cloud, and Azure

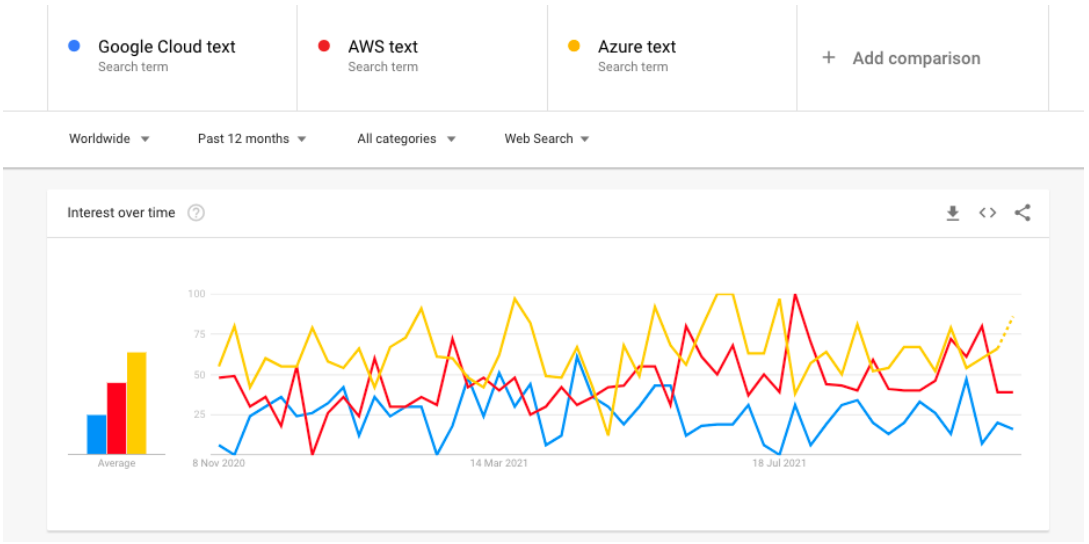
## Introduction

Natural language processing (NLP) is a broad category of artificial intelligence technologies that enable machines to understand human languages. Common NLP tasks include article classification, sentimental analysis, translation, etc.

With the rapid development of Cloud computing services, an increasing number of NLP tasks are moved to Clouds. That is, the application developers (cloud service consumers) focus on their own business logic, and leverage the computing power and intelligence of the Cloud platforms to perform NLP tasks. The three top public Cloud providers, Amazon Web Service (AWS), Google Cloud, and Microsoft Azure all offer NLP related services. This article gives an overview of the NLP services on public Cloud platform, compare different services, and make some recommendations to Cloud customers.

## Service Overview

The Cloud NLP services feature stong machine learning power, easy usage (“with no machine learning experience required”[1]) and privacy (“Your data is yours”[2]). Here is a simplified use case: A group of mobile app developers want to add entity recognition power to the app. The expected behavior is when certain entities (e.g. Mountain Elbrus) appear in an article, the word or phrase is highlighted and clicking them will trigger a jump to the search page. However, the developers do not have time, resources, expertise, or willingness to develop a solution by themselves. Instead, they buy a named entity recognition service, sending article text to Cloud servers via REST requests, and get back a list of entities to highlight.



The search trends above provide a sense of the popularity of Cloud NLP services. The following overview will follow the order, starting from the most popular one.

## Microsoft Azure Text Analytics

*“A collection of features from Cognitive Service for Language that extract, classify, and understand text within documents.” [2]*

Azure text analytics provides the following services [3]:

- Language detection
- Key phrase extraction
- Named entity recognition
- Sentiment analysis and opinion mining
- Question answering
- Text analytics for health

The pricing is per 1,000-character unit in the input (e.g. $0.001 for every 1000-character unit).

Azure also provides customizable services like custom text classification, but only for preview.

## AWS Comprehend

*“Derive and understand valuable insights from text within documents.” [1]*

AWS provides the following services [4]:

- Syntax detection
- Dominant language detection
- Key phrase extraction
- Named entities recognition
- Detecting PII (Personal identifiable information)
- Sentiment analysis
- Custom classification
- Event detection

The pricing is per 100-character unit in the input. Common tasks are charged $0.0001 per 100-character unit [5].

AWS also provides a batch processing interface [6].

## Google Cloud Natural Language AI

*“Derive insights from unstructured text using Google machine learning.” [7]*

Google Cloud provides the following services [7]:

- Syntax analysis
- Named entities recognition
- Sentiment analysis

- Entity Sentiment analysis
- Text classification

The pricing is per 1,000-character unit in the input (e.g. $0.001 for every 1000-character unit) [8].

A distinguishing feature is AutoML, which gives users more control and allows easy training of user's own model [9].

## Comparisons

The following table compares the NLP services provided by the three platforms:

| | Azure | AWS | Google Cloud |
|---|---|---|---|
| Features | Least coverage | Most coverage | Medium coverage, supports custom model training |
| Pricing* | Discount for high usage | Cheap if inputs are small (~300 characters) | Discount for high usage |
| Prerequisite | Minimum | Medium | Medium |
| API | Restful, C# SDK | Restful, CLI, various per-language SDK (C++, Python, Java, .Net, etc.) | Restful, gRPC, SDK for Go, Java, Node.js, Python |
| Documentation | Basic starter guide | Rich docs and examples | Clear docs and examples |

* Pricing is as for West US.

Based on the comparison, here are some tips for deciding which service to use:

- If you want to get a taste of NLP on Clouds, choose Google Cloud. The instructions and starter guide have the best readability.
- If you have heavy need on NLP services, choose AWS. It has the best feature richness and language support. Besides, it integrates well with other AWS services.

## Conclusion

With the rapid evolvement of NLP technologies and the growing Cloud market, it is foreseeable that the NLP services on public Clouds will be more and more powerful. I believe the easily accessible Cloud NLP services will enhance AI's usability and helps to empower various areas in our lives.

# References

[1] AWS Comprehend, https://aws.amazon.com/comprehend/
[2] Text Analysis | Amazon Azure,
https://azure.microsoft.com/en-us/services/cognitive-services/text-analytics/#overview
[3] Language Service pricing,
https://azure.microsoft.com/en-us/pricing/details/cognitive-services/language-service/
[4] Getting started with APIs,
https://docs.aws.amazon.com/comprehend/latest/dg/get-started-api.html
[5] Amazon Comprehend pricing, https://aws.amazon.com/comprehend/pricing
[6] Using the Batch APIs,
https://docs.aws.amazon.com/comprehend/latest/dg/get-started-batch.html
[7] Natural Language AI, https://cloud.google.com/natural-language
[8] Cloud Natural Language pricing, https://cloud.google.com/natural-language/pricing
[9] AutoML for Nautural Language, https://cloud.google.com/natural-language/automl/docs