



# **AWS Y EL LENGUAJE NATURAL**



## Amazon Comprehend

Amazon Comprehend es un servicio de procesamiento de lenguaje natural (NLP) que utiliza el machine learning para analizar textos. Amazon Comprehend proporciona reconocimiento de entidades personalizadas, clasificación personalizada, extracción de frases clave, análisis de opiniones, reconocimiento de entidades y más API para que pueda integrar fácilmente el procesamiento de lenguaje natural en las aplicaciones. Basta con llamar a las API de Amazon Comprehend en una aplicación y proporcionar la ubicación del documento o el texto de origen. Las API producen entidades, frases clave, opiniones e idioma en formato JSON, que puede usar en su aplicación.





## Reconocimiento de entidades personalizadas

El reconocimiento de entidades personalizadas le permite personalizar Amazon Comprehend para identificar términos específicos de su dominio. Mediante AutoML, Comprehend aprenderá a partir de un pequeño conjunto de ejemplos (por ejemplo, una lista de números de póliza, números de reclamación o NSS) y, a continuación, formará un modelo privado y personalizado para reconocer estos términos, como los números de reclamación, en cualquier otro bloque de texto dentro de archivos PDF, texto sin formato o documentos de Microsoft Word, sin necesidad de machine learning.

**Ejemplo:** en este ejemplo, a una empresa de seguros le gustaría analizar los documentos de texto correspondientes a entidades específicas de su empresa, los números de póliza.

**Texto de muestra:** Hola, me llamo Sam Ford y quisiera presentar una reclamación por accidente de coche. Mi código de póliza es 456-YQT.

Entidad	Categoría	Recuento	Confianza
456-YQT	Policy_ID	1	0.95



## Clasificación personalizada

La API de clasificación personalizada le permite crear fácilmente modelos de clasificación personalizada de textos utilizando las etiquetas específicas de su empresa sin tener que aprender ML. Por ejemplo, su organización de soporte al cliente puede usar la clasificación personalizada para categorizar automáticamente las solicitudes entrantes por tipo de problema según cómo el cliente haya descrito el inconveniente. Con el modelo personalizado, es fácil moderar los comentarios de los sitios web, evaluar la opinión de los clientes y organizar los documentos de los grupos de trabajo.

**Ejemplo:** supongamos que desea organizar los comentarios de atención al cliente de una compañía aérea. Quiere organizar cada comentario en las categorías "Preguntas sobre la cuenta", "Reembolsos de billetes" y "Reclamaciones de vuelo". Para formar el servicio, cree un archivo CSV que contenga texto de ejemplo de cada problema y etiquete cada ejemplo con una de las tres etiquetas que corresponda. El servicio formará automáticamente un modelo personalizado en su nombre. Para utilizar su modelo para analizar todas las llamadas al día siguiente, envíe cada archivo de texto al servicio y recibirá los resultados etiquetados junto con la confianza de la etiqueta correspondiente.

Texto	Etiqueta	Puntuación de confianza
Línea 0	Pregunta sobre la cuenta	0,92
Línea 1	Reembolso de boletos	1
Línea 2	Quejas sobre el vuelo	1
Línea 3	Quejas sobre el vuelo	0,91
Doc5.csv	Reembolso de boletos	1



## Reconocimiento de entidad

La API de reconocimiento de entidades devuelve las entidades nombradas (“Personas”, “Lugares”, “Ubicaciones”, etc.) que se categorizan de manera automática a partir del texto proporcionado.

**Texto de muestra:** Amazon.com, Inc. se encuentra ubicada en Seattle, WA y fue fundada el 5 de julio de 1994 por Jeff Bezos. Permite a los clientes comprar de todo, desde libros a batidoras. Seattle está al norte de Portland y al sur de Vancouver, BC. Otras compañías importantes con sede en Seattle son Starbucks y Boeing.

Entidad	Categoría	Confianza
Amazon.com, Inc.	Organización	0,96
Seattle, WA	Ubicación	0,96
5 de julio de 1994	Fecha	0,99
Jeff Bezos	Persona	0,99
Seattle	Ubicación	0,98
Portland	Ubicación	0,99
Vancouver, BC	Ubicación	0,97
Starbucks	Organización	0,91
Boeing	Organización	0,99

## Otras funciones

### Análisis de opiniones

La API de análisis de opiniones produce la opinión general de un texto (positiva, negativa, neutra o mixta).



## **Opinión focalizada**

La opinión focalizada proporciona información más granular mediante la identificación de opiniones (positivas, negativas, neutrales o mixtas) en las entidades de un texto.

## **Extracción de frases clave**

La API de extracción de frases clave produce frases clave o temas de conversación y una puntuación de confianza que respalda que se trata de una frase clave

## **Detección de toxicidad**

La detección de toxicidad de Comprehend proporciona una solución simple basada en NLP para la detección de contenido tóxico en documentos de texto. La capacidad está disponible de fábrica para moderar la conversación entre pares en plataformas en línea y las entradas y salidas de IA generativa.

## **Detección de eventos**

Comprehend Events le permite extraer la estructura de eventos de un documento, en el que sintetiza las páginas de un texto a datos procesados fácilmente para el consumo por aplicaciones de IA o herramientas de visualización de gráficos. Esta API le permite responder a las preguntas “quién”, “qué”, “cuándo” y “dónde” en



grandes conjuntos de documentos, a escala y sin experiencia previa en NLP. Use Comprehend Events para extraer detalles pormenorizados sobre eventos reales y entidades asociadas expresadas en texto sin estructurar.

## Detección de idioma

La API de detección de idioma identifica automáticamente texto escrito en más de 100 idiomas y devuelve el idioma dominante con una puntuación de confianza para respaldar que un idioma es dominante.

## Análisis sintáctico

La API Syntax de Amazon Comprehend permite a los clientes analizar texto con tokenización y categorías gramaticales, e identificar etiquetas y límites de palabras, como sustantivos y adjetivos, dentro del texto

**Ejemplo de muestra:** ¡Me encanta mi flamante y ágil Kindle Fire!

Texto	Etiqueta
Me	Pronombre
Encanta	Verbo
Mi	Pronombre
Rapidez	Adjetivo
,	Puntuación
Novedades	Adjetivo
Kindle	Nombre propio
Fire	Nombre propio
.	Puntuación



## **Compatibilidad con varios idiomas**

Amazon Comprehend puede llevar a cabo análisis de textos en alemán, inglés, español, italiano, portugués, francés, japonés, coreano, hindi, árabe, chino (simplificado) y chino (tradicional). Para crear aplicaciones en otros idiomas, los clientes pueden usar Amazon Translate para traducir el texto a un idioma compatible con Comprehend y, a continuación, usar Comprehend para analizar textos.