

Procesamiento de Lenguaje Natural (NLP)

Fundamentos y Horizontes.

1

- **¿Qué es NLP?** Disciplina de IA que permite a las máquinas comprender, interpretar y generar lenguaje humano.
- **Objetivo:** Cerrar la brecha de comunicación entre humanos y máquinas.
- **Componentes Clave del Lenguaje:** Sintaxis, Semántica, Pragmática, Discurso.
- **Importancia:** Impulsa la automatización, análisis de datos no estructurados y la interacción humano-máquina.

Procesamiento de Lenguaje Natural (NLP)

Definición y Alcance

1

“Se define como la tecnología de aprendizaje automático (Machine Learning) que permite a las computadoras interpretar, manipular y comprender el lenguaje humano.”

Comprender, interpretar y generar texto (o voz).

Capacidad de manejar diferencias en dialectos e irregularidades gramaticales.

Experiencia y accesibilidad

Lingüística computacional, Machine Learning, Deep Learning

Procesamiento de Lenguaje Natural (NLP)

Objetivos

1

Evolución: Del procesamiento hasta la comprensión y generación.

Búsqueda de un nivel de interacción mas conectivo.

Sistemas de IA mas creativos y proactivos (en lugar de solo reactivos)



Procesamiento de Lenguaje Natural (NLP)

Componentes Clave

1

- **Sintaxis:** Estructura gramatical

“El gato negro duerme tranquilamente”

- **Semántica:** Significado de las palabras

“Fui al **banco** a retirar dinero”

“Me senté en el **banco** de la plaza”

“Yo te **banco**”

- **Pragmática:** Comprensión del contexto

“Sos un genio”

“Tirar la toalla”

- **Discurso:** Narrativas coherentes.

“María fue a la tienda. Ella compró pan.”

DISCURSO

(Coherencia de múltiples oraciones, narrativas, resolución de correferencia)

PRAGMÁTICA

(Significado en contexto, intención, sarcasmo, modismos, referencias culturales)

SEMÁNTICA

(Significado de palabras, frases, oraciones, desambiguación de sentido)

SINTAXIS

(Estructura gramatical, orden de palabras, partes de la oración, frases)

Procesamiento de Lenguaje Natural (NLP)

Importancia

Importancia en aplicaciones generalizadas.

Atención al cliente automatizada

Traducción de idiomas en tiempo real.

Organizaciones:

- Procesamiento y análisis de archivo de grandes documentos.
- Análisis feedback de clientes
- Analisis de grabaciones de centros de llamadas.

NLP potencia a los **Modelos generativos**: Crear texto similar al humano.

Manipulación y comprensión de **datos no estructurados**.

Procesamiento de Lenguaje Natural (NLP)

Orígenes

2

1950s El test de **Turing** (Alan Turing)

1960s-1980s La **era simbólica** (Sistemas basados en Reglas)

1980s-2000s Revolución Estadística (**Modelos Probabilísticos y Machine Learning**)

2000s-2010s **RNNs, LSTMs y CNNs**

2017-Actualidad **LLMs y Modelos GPT**



Procesamiento de Lenguaje Natural (NLP)

Orígenes – 1950s Alan Turing

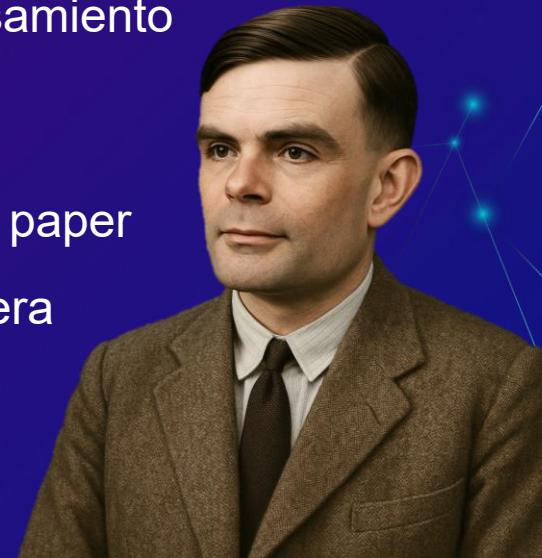
2

Padre de la ciencia de la computación

Paper científico “**Computing Machinery and Intelligence**” [1950]

Establece fundamentos teóricos y prácticos para el procesamiento del lenguaje en la IA

Experimento **Georgetown-IBM** en 1954 (impulsado por el paper presentado por Alan Turing) marcando el comienzo de la era simbólica.



Procesamiento de Lenguaje Natural (NLP)

1960s-1980s La era simbólica

2

NLP basado en Reglas.

Creación **manual** de directrices.

ELIZA (1964) Programa Psicoterapeuta - Joseph Weizenbaum [MIT]

SHRDLU (1968) [MIT] Participaba en una discusión sobre objetos

LUNAR (1971) [NASA] proporcionaba respuestas a preguntas en lenguaje natural sobre muestras de rocas lunares

Limitaciones: ambigüedad y contexto. Tiempo. Especialistas.

Escalabilidad.



Procesamiento de Lenguaje Natural (NLP)

1980s-2000s Revolución Estadística

Cambio de Paradigma: **NLP Estadístico**

- Modelos probabilísticos
- Machine Learning

Factores: Capacidad Computacional y disponibilidad de big data.

1990 **HMMs** (Modelos Ocultos de Markov): Se empleaban para predecir las etiquetas gramaticales de las palabras.

“El Perro Corre” | observaciones y etiquetas ocultas

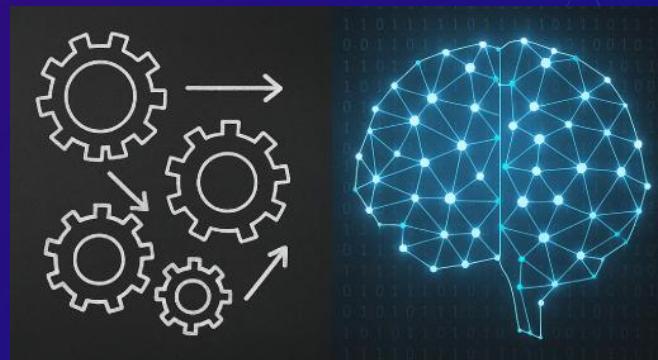
2000 **SMT** – Traducción automática estadística (Google)

Procesamiento de Lenguaje Natural (NLP)

1980s-2000s NLP Estadístico

2

“Es una lección crítica en el desarrollo de la IA: cuando un dominio es demasiado complejo para reglas diseñadas por humanos, el aprendizaje a partir de los datos, especialmente con métodos probabilísticos, ofrece una solución más robusta y escalable”



Procesamiento de Lenguaje Natural (NLP)

2000s-2010s La era de las redes neuronales

2

NLP Basados en redes neuronales

Perceptrón (1958) | Perceptrón Multicapa (1965)

CNNs Redes Neuronales Convolucionales (1980)

RNNs Redes Neuronales Recurrentes (1990)

LSTMs Long Short-Term Memory (1997)

2010 Momento Bisagra: Aprendizaje Profundo (Deep Learning)

NLP Neuronal aprenden las características directamente de los datos.

... y todo esto fue dando curso a ...

Procesamiento de Lenguaje Natural (NLP)

2017-Actualidad LLMs y Modelos GPT

2

Arquitecturas Transformer

"Attention Is All You Need" (2017) [Google]

Mecanismo de auto-atencion

GPT (Generative Pre-trained Transformer)

BERT (Bidirectional Encoder Representations from Transformers).

