



# INSTITUTO TECNOLÓGICO SUPERIOR DE JEREZ INGENIERÍA EN SISTEMAS COMPUTACIONALES

9° SEMESTRE  
INTELIGENCIA ARTIFICIAL  
ISC. JESÚS ARANDA GAMBOA



TEMA 4: APLICACIONES CON TÉCNICAS DE IA  
PROCESAMIENTO DE LENGUAJES NATURALES

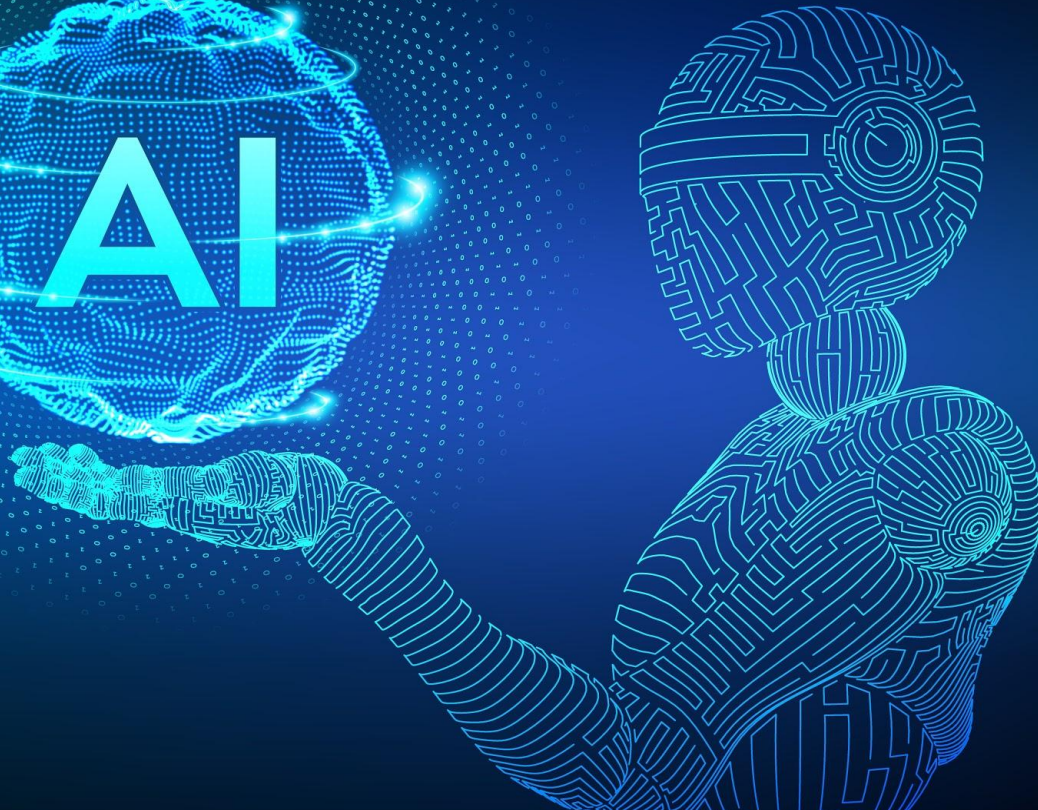
LIZA AREMY SANTANA CONTRERAS  
NO.CONTROL: 16070005



12 DE DICIEMBRE 2020  
JEREZ ZAC.



Artificial  
Intelligence



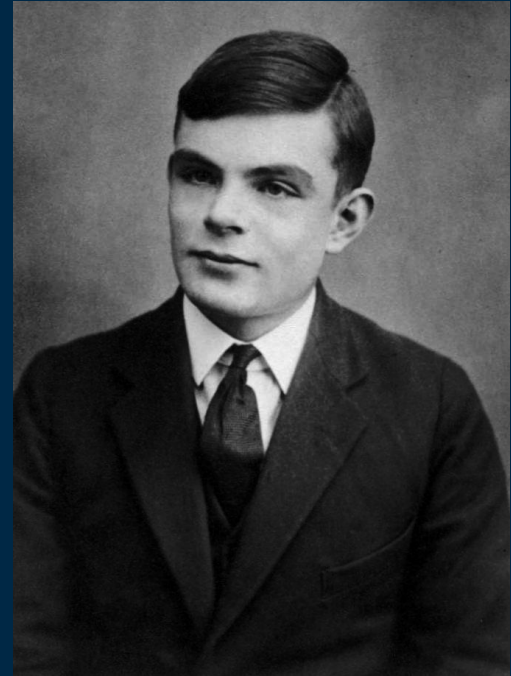
# Procesamiento de Lenguaje Natural (PLN).

# Orígenes del PNL

El PNL tiene sus orígenes en los años 50's, Alan Turing publicó "Computing machinery and intelligence".

- Test de turing: es una prueba de la habilidad de una máquina que presenta un comportamiento inteligente.

En 1954 Georgetown logró realizar traducción automática de más de sesenta oraciones del Ruso al Inglés.



# Introducción PNL

El procesamiento del lenguaje natural (PNL) es un campo de:

- Ciencias de la computación
- Inteligencia artificial
- Lingüística

Tiene como objetivo analizar las interacciones entre las computadoras y el lenguaje humano.

- Lenguaje oral
- Lenguaje escrito
- Lenguaje de señas

Se encarga de buscar mecanismos eficaces para la comunicación entre personas y máquinas utilizando el lenguaje natural.

# Lenguaje Natural

El Lenguaje Natural se utiliza a diario como medio de comunicación entre humanos.

- Castellano
- Inglés
- Francés
- Mandarin

Se caracteriza por:

- Se caracteriza por poseer una sintaxis y una gramática específica.
- Contienen muchas ambigüedades.

Los tipos de lenguaje natural se pueden clasificar:

- Lenguaje oral
- Lenguaje escrito
- Lenguaje de señas

# Modelos para PNL

El procesamiento de un lenguaje natural desde un punto de vista computacional implica un “proceso de modelización matemática”. Esto se debe a que:

- Los ordenadores sólo entienden código binario.
- Los desarrolladores crean los programas utilizando lenguajes como Java y Python.

## Lingüistas computacionales

Se encargan de preparar el modelo lingüístico para que los programadores lo implementen en un código eficiente y funcional.



# Modelos para PNL

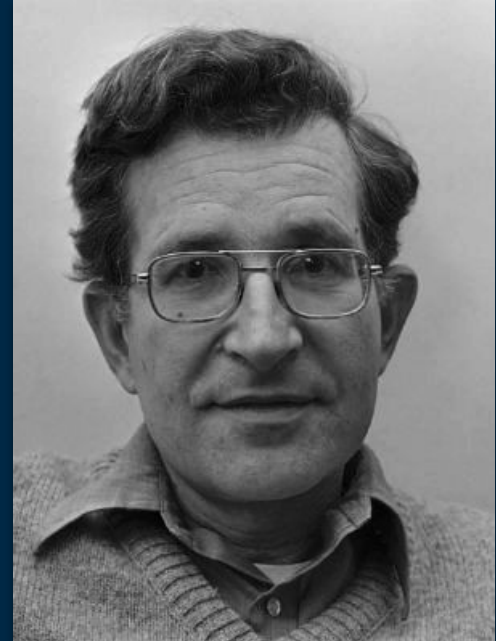
## Modelos Lógicos (gramáticas):

Se basan en:

- Reglas de reconocimiento de patrones estructurales.
- Diccionarios computacionales.

Para resolver las tareas como lo son buscar información, traducir.

Buscan reflejar la estructura lógica del lenguaje y surgen a partir de las teorías de N. Chomsky en los años 50.





# Modelos para PNL

## Modelos probabilísticos del lenguaje natural (basados en datos):

Se basan en:

- Colecciones de ejemplos y datos.

A partir de ellos se calculan las frecuencias de diferentes unidades lingüísticas:

- Letras.
- Palabras.
- Oraciones.

Para determinar su probabilidad de aparecer en un contexto determinado para predecir sin necesidad de recurrir a reglas gramaticales. Es el paradigma de aprendizaje automático.

# Fases del PNL

1. Análisis morfológico:
  - a. Extraer unidades léxicas.
2. Análisis sintáctico:
  - a. Análisis gramatical de las frases.
3. Análisis semántico:
  - a. La extracción del significado de la frase.
4. Análisis pragmático:
  - a. Determinar las referencias de los pronombres.
5. Planificación de la frase:
  - a. Estructurar cada frase del texto.
6. Generación de la frase:
  - a. La generación de la cadena de palabras.



# Ventajas del PNL

## Ambigüedad:

- A nivel léxico:
  - Una misma palabra puede tener varios significados.
- A nivel estructural:
  - Se requiere de la semántica para desambiguar.
- A nivel pragmático:
  - Cuando una oración no significa lo que realmente se está diciendo.

## Separación entre las palabras:

- Lengua hablada:
  - No se suelen hacer pausas entre palabra y palabra.
- En la lengua escrita:
  - En el chino mandarín no tienen separaciones entre las palabras.

## Errores en el análisis de datos:

- Acentos extranjeros.
- Errores léxicos.
- Errores semánticos.



# Importancia del PNL en la actualidad

Grandes volúmenes de datos:

- Textuales
- Audio
- Video



# Aplicaciones del PNL

- Traducción Automática de textos.
- Análisis del sentimiento de los textos.
- Resúmenes de textos automáticos.
- Clasificación de currículos.
- Búsqueda en internet por voz.
- Extracción de la información.
- Asistentes virtuales.
- Question-Answering.



Hi, I'm Cortana.



# Librerías para PNL

- Apache OpenNLP.
- Stanford NLP suite.
- Gate NLP library.
- Natural language toolkit (NLTK):
  - Es la biblioteca más popular para el PNL escrita en Python.
  - Es la biblioteca de procesamiento de lenguaje natural más fácil de usar.



# Conclusión

El PLN es una área de la inteligencia artificial que es sumamente importante, ya que son muchos los beneficios que traen sus diferentes de sus aplicaciones, por ejemplo, los motores de búsqueda como Google lo utilizan para determinar el tipo de usuario eres, es decir, si Google determina que eres un desarrollador muestra resultados relacionados con la programación.

Otro de los grandes beneficios del PNL son los asistentes virtuales, como Siri, Alexa ó Cortana, los cuales ayuda a ser más productivos en algunas tareas.



# Referencias

- [1] Liddy, E. D. (2001). Natural language processing.
- [2] Verdejo Maillo, María Felisa, "Comprensión del lenguaje natural: Avances, aplicaciones y tendencias", Arbor CLI, 595, (1995), 39-83.
- [3] Válles, Mari, "La web semántica y las tecnologías del lenguaje humano" en Web semántica y sistemas de información documental, Gijón, Trea, (2018).