**Desarrollo del Procesamiento de Lenguaje Natural en la historia**

**Sistemas Inteligentes – 2020-I**

Mateo González ***a***

Alejandro Uribe ***b***

***a***1037627601; ***b***1037627601

Junio 2020

**Tabla de Contenidos**

[1. Introducción 1](#_Toc43495480)

[2. Racionalismo 2](#_Toc43495481)

[3. Aprendizaje de Máquinas 3](#_Toc43495482)

[4. Redes Neuronales 3](#_Toc43495483)

[Referencias 3](#_Toc43495484)

**Lista de Figuras**

[Figura 1. Línea de tiempo PLN. 2](#_Toc43498956)

# Introducción

El ser humano ha buscado desarrollar herramientas que le permitan codificar el lenguaje para su transmisión, procesamiento y entendimiento con el fin de llevar a cabo tareas útiles. La invención del teléfono marcó el punto de inicio en el reconocimiento moderno del lenguaje, en tanto permitió cambiar el vehículo natural del lenguaje de mecánico a eléctrico. Homer Dudley’s llevó a cabo los primeros intentos para sintetizar eléctricamente discursos en la década 1920-1930.

No obstante, los avances necesarios para hacer que lo anterior fuera posible se dieron en la década de 1940-1950 cuando se sentaron las primeras nociones de la teoría de información. Posteriormente, se dieron avances el hardware necesario para medir y digitalizar señales sonoras y finalmente, la descripción de señales sonoras en términos de coeficientes de predicción linear (LPC), lo que proporcionó una representación conveniente para este objetivo [1].

Desde una perspectiva de alto nivel, el Procesamiento del Lenguaje Natural desarrolla aplicaciones que facilitan las interacciones entre computadores y el lenguaje humano, partiendo del modelado de mecanismos cognitivos que lo entiendan y lo produzcan.

Sus aplicaciones incluyen, entre otras, el reconocimiento de expresiones, análisis léxico, análisis gramatical, traducción, respuesta a preguntas, análisis de sentimientos y generación de lenguaje natural.

Desde una perspectiva histórica, el Procesamiento del Lenguaje Natural se ha discutido desde diferentes metodologías, las cuales se pueden resumir en tres olas [2]:

* Racionalismo.
* Empirismo.
* Aprendizaje profundo.

La Figura 1 muestra los hitos que han constituido un avance en el Procesamiento del Lenguaje Natural a lo largo de su historia.

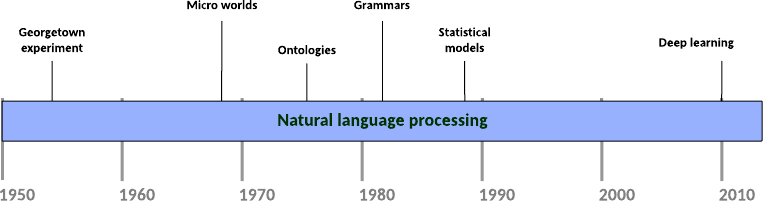


Figura 1. Línea de tiempo PLN.

Fuente: [3]

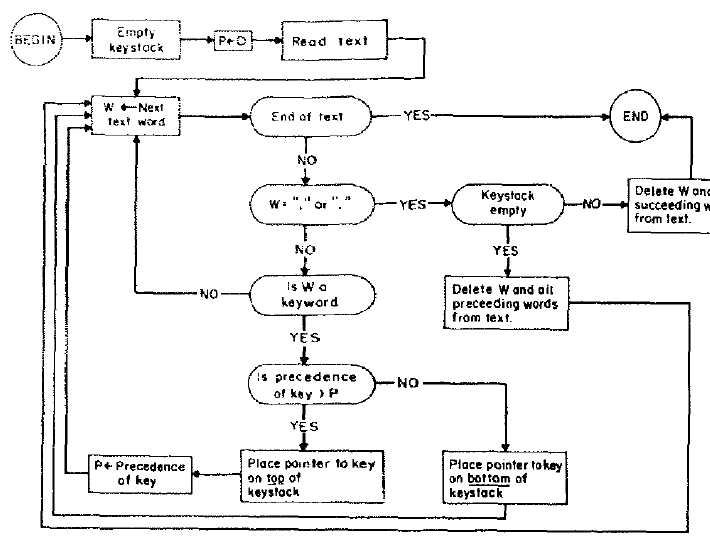
# Racionalismo

Esta ola heredó su nombre dada la extendida aceptación de los postulados de Noam Chomsky referidos a la estructura del lenguaje en los que este se encontraba instintivamente conectado al cerebro desde el nacimiento. Adicionalmente, de las propuestas racionalistas en las que se incorporaron antologías conceptuales y mecanismos de razonamiento basados en reglas.

Sin embargo, esta ola inició en la década de 1950, cuando Alan Turing propuso una prueba que permitiera evaluar la capacidad de los computadores para exhibir un comportamiento inteligente no diferenciable al de un ser humano. En dicha prueba un computador podría mantener una conversación con un interlocutor humano. En 1954, el experimento llevado a cabo en *Georgetown-IBM* logró traducir sesenta palabras del ruso al inglés, lo que constituyó el avance con más largo alcance hasta entonces.

A inicios de la década de 1980, el desarrollo de la *Inteligencia Artificial* como se conoce hoy dio sus primeros pasos. Fueron unos acercamientos iniciales que se basaban en “sistemas expertos”, los cuales se venían desarrollando desde 1970. Estos sistemas emulaban la habilidad de toma de decisiones humana haciendo uso de un conjunto de reglas lógicas complejas abstraídas de representaciones del conocimiento. Con el uso de esta herramienta se desarrollaron sistemas expertos para el Procesamiento del Lenguaje Natural como *ELIZA* y *MARGIE*.

*ELIZA* fue creado en la decada de los 60 en el MIT por Joseph Weizenbaum en un intento por superar la prueba de Turing mencionada anteriormente. A un nivel muy general, ELIZA funciona mediante la identificación de palabras claves definidas previamente en un oración, dichas palabras claves tienen asociadas reglas que descomponen la oración en partes mas pequeñas y a su vez, estas reglas tienen asociadas otras reglas para, a partir de estas partes, crear una nueva oración para enviar como respuesta. Por lo tanto el tipo de respuestas obtenidas al interactuar con ELIZA depende tanto de las palabras clave definidas, como por las reglas asociadas a cada palabra clave, a este conjunto de reglas se les denomina script. El script original creado por Weizenbaum, llamado DOCTOR, simulaba un psicologo aplicando metodologías similares a Carl Rogers. En la figura ?? se muestra el diagrama de flujo asociado con el funcionamiento de ELIZA.

 Figura ??. Diagrama de flujo de ELIZA. Tomado de [4]

Posteriormente, a finales de la década de 1980, los cuales consistían en un.

# Aprendizaje de Máquinas

Los modelos basados en reglas, como ELIZA, fueron bastante comunes durante varios años. Sin embargo, la generalización de este tipo de modelos era una tarea realmente compleja, debido a que muchas de estas reglas se asignaban basado en un contexto especifico, por lo que si se quería llevar este tipo de programas a contextos diferentes, fallaban drasticamente. A finales de los años 80 y durante la decada de los 90, dado el continuo crecimiento en el poder de procesamiento de las computadoras de la epoca, el uso de algoritmos basados en estadistica para IA adquirio gran popularidad, y el campo del procesamiento de lenguaje no fue ajeno a este. Uno de los algoritmos mas populares fue el de Caden

# Redes Neuronales

# Referencias

[1] J. Benesty, M. M. Sondhi, and Y. A. Huang, “Introduction to Speech Processing,” in *Springer Handbooks*, Springer, 2008, pp. 1–4.

[2] L. Deng and Y. Liu, “A joint introduction to natural language processing and to deep learning,” in *Deep Learning in Natural Language Processing*, Springer International Publishing, 2018, pp. 1–22.

[3] T. Jones and IBM, “Speaking Out Loud - An introduction to natural language processing,” 2017. [Online]. Available: https://w3.ibm.com/developer/learning/articles/cc-cognitive-natural-language-processing/. [Accessed: 19-Jun-2020].

[4] Joseph Weizenbaum. 1966. ELIZA—a computer program for the study of natural language communication between man and machine. Commun. ACM 9, 1 (Jan. 1966), 36–45