



**TECNOLÓGICO  
NACIONAL DE MÉXICO**



Tecnológico Nacional de México Campus Culiacán

Ingeniería en Sistemas Computacionales

Inteligencia Artificial: 18:00 – 19:00

## **“Language Models as Knowledge Bases?”**

**Resumen según Petroni et al.**

Docente: Dr. José Mario Ríos Félix

Alumno: José Alejandro Escobar Sánchez

Control: 19170573

Correo: [L19170573@culiacan.tecnm.mx](mailto:L19170573@culiacan.tecnm.mx)

## "Language Models as Knowledge Bases" Petroni, F. et al. (2019)

El concepto de Modelos de Lenguaje como Bases de Conocimiento (LMKB) propuesto por Petroni sugiere que los grandes modelos de lenguaje como GPT-3 y sus sucesores pueden ser utilizados a manera de bases de conocimiento. En este entendimiento, cada hecho del modelo puede cifrarse como una posible entrada con un dato temporal para la información que se desea obtener. Teniendo esto, se puede consultar al modelo proporcionándole una sentencia con información temporal, la cual el modelo sustituirá por la información real que se busca.

Por ejemplo, supongamos que se quiere conocer la capital de Francia. Una posible entrada para consultar el modelo sería "La capital de Francia es [MASK]". Posteriormente el modelo, que ha sido entrenado utilizando este tipo de datos, generaría la respuesta "Paris".

Esta técnica aprovecha la habilidad del modelo para generar texto con el fin de extraer información real de manera efectiva. Es un paradigma diferente a las bases de conocimiento tradicionales, las cuales normalmente dependen de relaciones explícitas y esquemas fijos. LMKB es un modelo más flexible, ya que es capaz de inferir respuestas basadas en el contexto que se le proporciona en la misma consulta.

El modelo adquiere entonces una gran proporción de las ventajas que caracterizan a los modelos de lenguaje: según Petroni (2019) "Los modelos de lenguaje tienen grandes ventajas sobre bases de conocimiento estructurados: no requieren ingeniería de schemas, permiten consultar sobre relaciones abstractas de clases, son escalables y no requieren supervisión humana en su entrenamiento".

En 2019, Petroni et al. presentaron un análisis extensivo del conocimiento relacional actualmente presente en una amplia gama de modelos de lenguaje de estado del arte pre-entrenados. Los progresos recientes en las formas de entrenar modelos de lenguaje en cuerpos de información textual han conducido a un incremento notable de mejoras en tareas relacionadas al Procesamiento de Lenguaje Natural (NLP) en modalidad "rio abajo" o "downstream". Es decir, tareas que dependen de las salidas provenientes de tareas anteriores en la secuencia de ejecución. La idea principal es que estos modelos pueden estar también almacenando información y conocimiento relacional presente en los cuerpos de datos que se utilizan para el entrenamiento de estos. Los modelos pueden entonces ser capaces de responder consultas estructuradas de forma "llenar el vacío", conocidas también como preguntas con respuestas incrustadas o sentencias Cloze.

En la práctica, a menudo es necesario extraer información relacional de textos extensos u otras fuentes para poblar bases de conocimientos. Este proceso requiere pipelines complejos de Procesamiento de Lenguaje Natural que involucren extracción de entidades, solución de co-referencias, lindeo de entidades y extracción de relaciones. Estos componentes requieren la correcta supervisión de datos y fijación de schemas. En lugar de esto, es posible solucionar estas necesidades de consulta mediante preguntas a modelos de lenguaje natural para que regresen información relacional que han aprendido.

En su análisis de distintos modelos, Petroni encontró que BERT contiene conocimiento relacional competitivo con métodos tradicionales de Procesamiento de Lenguaje Natural. También tiene un rendimiento considerable en proporcionar respuestas a preguntas de dominio abierto, demostrando un gran potencial como sistemas QA.