



Comprender qué son los datos y cuál es su ciclo de vida

El proceso ETL

Para dedicarse profesionalmente a desarrollar soluciones de IA es importante entender que el **DATO** es la base de todo. Sobre los datos, desarrollaréis soluciones en este ámbito. Por poner algún ejemplo:

- 1.- Los coches *Tesla* tienen infinidad de datos de sus sensores y a partir de estos toman decisiones en tiempo real para la conducción autónoma.
- 2.- El corrector de palabras de *whatsapp* tiene tu histórico de escritura y a partir de ahí te recomienda palabras parecidas. Mmmm, las palabras son datos.
- 3.- No te ha pasado alguna vez que ¿antes de escribir nada en el buscador de Google este ya te recomienda aquello que pretendías buscar? Tu histórico de datos (tu experiencia pasada) ayuda a esta BigTech a predecir tus necesidades. Datos, datos, datoooooss, ...

Aquí podríamos hablar largo y tendido sobre la **ética y la IA** pero lo dejamos para otro día.

Está claro que una vez aplicados los algoritmos matemáticos adecuados a partir de la materia prima podemos construir la parte "mágica" como hacen Tesla o Google pero para eso antes hace falta generar dato, cocinarlo/organizarlo/tratarlo, etcétera.

Para Google "su" dato somos nosotros, para Tesla su dato son los sensores del coche y para vosotros el dato será aquello que aporte valor a vuestra solución.

Para que las predicciones y la IA funcionen un aspecto fundamental es que los datos una vez generados, se organicen y se limpien para cargarlos en las bases de datos. Este proceso se conoce como **ETL** y está muy bien explicado en el siguiente [link](#)

#HASHTAGS *(etiquetas de ayuda para búsqueda de información relevante)*

#ETL #bases-de-datos #data-preprocessing #data-cleaning #missing-values #data-merge #fuentes-de-datos #sensores

LINKS DE INTERÉS

<https://www.youtube.com/watch?v=u3Le4lFePnQ&t=8s>

<https://platzi.com/blog/que-es-un-etl/>

<https://blog.powerdata.es/el-valor-de-la-gestion-de-datos/bid/288859/procesos-etl-extracci-n-transformaci-n-carga>

DICCIONARIO

PUNTUACIÓN

Programación: 1
Redes: 1
Seguridad: 1
Algoritmia: 4