# Evolución del procesamiento de Datos

El procesamiento de datos se presenta como el conjunto de operaciones que permiten recolectar, transformar, almacenar y analizar varias informaciones, con el fin de generar resultados útiles para la toma de decisiones. A lo largo del tiempo, esta actividad ha pasado de realizarse de forma manual a utilizar tecnologías avanzadas; que permiten tratar grandes volúmenes de datos en tiempo real. El objetivo del procesamiento de datos es optimizar la gestión de la información, facilitando su uso eficiente en diversas áreas como la administración, la ciencia, la salud, la educación y la industria. Esta evolución ha estado acompañada del desarrollo de sistemas computacionales, arquitecturas de bases de datos y herramientas de inteligencia artificial que han transformado por completo la forma en que se maneja la información [1].

# Primeros sistemas de almacenamiento y tratamiento manual

Durante las primeras etapas del procesamiento de datos, las actividades se realizaban de manera completamente manual. Los registros físicos, las fichas archivadas y los libros contables eran los medios utilizados por empresas y gobiernos para conservar la información. Este sistema resultaba propenso a errores humanos, pérdida de datos y una lentitud considerable en los procesos de búsqueda y análisis. El almacenamiento en archivos planos durante las décadas de 1950 y 1960 representó un avance, ya que se comenzaron a utilizar dispositivos magnéticos como cintas o discos para guardar información de manera secuencial. Sin embargo, esta modalidad también presentaba limitaciones evidentes, especialmente en lo relacionado a la redundancia, dependencia del software, y poca flexibilidad para el acceso simultáneo [1].

#### Surgimiento de los primeros sistemas de bases de datos

El desarrollo de sistemas de gestión de bases de datos (SGBD) surgió como una respuesta a las deficiencias de los archivos planos. Los primeros modelos utilizados fueron el modelo jerárquico y el modelo en red, que permitieron una organización más estructurada de los datos, aunque con rigidez en la navegación y dependencia del diseño físico. La verdadera revolución llegó con el modelo relacional, propuesto por Edgar F. Codd en la década de 1970, el cual estableció una separación clara entre los datos y las aplicaciones, permitiendo consultas flexibles mediante el lenguaje SQL. Este modelo facilitó la independencia de los datos, la integridad referencial y la normalización, dando paso a sistemas más robustos como Oracle, MySQL y PostgreSQL [1].

#### Procesamiento distribuido y nuevas arquitecturas

Con el crecimiento exponencial de la información en las décadas siguientes, surgió la necesidad de escalar los sistemas más allá de una sola máquina. En este contexto, el procesamiento distribuido se volvió fundamental. A través de arquitecturas como cliente-servidor y, más tarde, con la incorporación de redes de alta velocidad, se pudo dividir el

trabajo de procesamiento entre varios nodos, aumentando así el rendimiento del sistema. Este enfoque permitió la aparición de tecnologías como Hadoop y su sistema de archivos distribuido (HDFS), así como la ejecución de tareas paralelas mediante frameworks como MapReduce. De esta manera, se superaron las limitaciones físicas de los servidores tradicionales y se amplió la capacidad de procesamiento para grandes volúmenes de información [2].

# La era del Big Data y procesamiento en tiempo real

A partir del siglo XXI, el procesamiento de datos ha evolucionado hacia esquemas más complejos y dinámicos, caracterizados por la generación constante de grandes cantidades de información en formatos variados. Este fenómeno dio origen al concepto de Big Data, el cual se define por cinco características clave: volumen, velocidad, variedad, veracidad y valor. En este nuevo escenario, los sistemas deben ser capaces de capturar, almacenar y analizar datos en tiempo real, lo que ha impulsado la creación de plataformas como Apache Kafka, Apache Spark y tecnologías en la nube como Google BigQuery. Este tipo de procesamiento no solo permite analizar eventos conforme ocurren, sino también realizar predicciones, segmentaciones y decisiones automatizadas con base en patrones detectados en los datos [2].

# Inteligencia artificial y análisis predictivo

La última etapa de esta evolución está marcada por la incorporación de técnicas de inteligencia artificial en los sistemas de procesamiento de datos. Actualmente, los datos no solo se almacenan y organizan, sino que también se utilizan para entrenar modelos de aprendizaje automático que permiten realizar análisis predictivos, clasificaciones automatizadas y recomendaciones personalizadas. Estos modelos requieren bases de datos robustas, limpias y bien estructuradas, lo que refuerza la importancia de una buena gestión de datos desde el origen. La combinación entre la ciencia de datos, los sistemas distribuidos y la inteligencia artificial ha dado lugar a soluciones avanzadas que permiten a las organizaciones anticiparse a tendencias, prevenir fallos, optimizar procesos y mejorar la toma de decisiones a gran escala [2].





