



Universidad Tecnológica Nacional
Facultad Regional Buenos Aires

PROGRAMA ANALÍTICO DE ASIGNATURA

DEPARTAMENTO: Ingeniería en Sistemas de Información

CARRERA: Ingeniería en Sistemas de Información

NOMBRE DE LA ACTIVIDAD CURRICULAR: Tecnologías aplicadas a soluciones de datos

Año Académico: 2023

Área: Gestión Ingenieril

Bloque: Electivas

Nivel: 5º

Tipo: Electiva

Modalidad: Cuatrimestral

Cargas horarias totales:

<i>Horas reloj</i>	<i>Horas cátedra</i>	<i>Horas cátedra semanales</i>
72	96	6

COMPOSICIÓN DEL EQUIPO DE CÁTEDRA:

Profesor Adjunto: Ing. Mariano Berruezo

FUNDAMENTACIÓN

En los últimos años, el mundo de los negocios ha experimentado un crecimiento exponencial en la cantidad de datos disponibles dentro y fuera de las organizaciones. Este incremento es impulsado por: nuevos paradigmas tecnológicos y sociales que surgen a partir del uso extensivo de equipos móviles, la conexión de miles de millones de dispositivos generadores de información a través de Internet y la adopción de las redes sociales para posicionamiento de marca y canales para la comunicación con clientes, proveedores y la comunidad.

Estos cambios imponen nuevos desafíos a las organizaciones y, sobre todo, a los profesionales de Informática que deben implementar soluciones que resuelvan nuevas problemáticas en un ambiente competitivo y dinámico, donde el tiempo de implementación se reduce, y la complejidad de las tecnologías aumenta.



Universidad Tecnológica Nacional
Facultad Regional Buenos Aires

Las nuevas soluciones de datos requieren habilidades, metodologías y buenas prácticas que no siempre están alineadas con los paradigmas tradicionales. El futuro Ingeniero en Sistemas debe desarrollar habilidades para adaptarse a la dinámica de estos cambios y así anticiparse a los nuevos desafíos de negocio que se generan día a día.

La asignatura electiva Tecnologías Aplicadas a Soluciones de Datos, brinda a los estudiantes las herramientas necesarias para desempeñarse en el ámbito profesional, de cara a los nuevos desafíos de la gestión de información. La asignatura apunta a desarrollar capacidades específicas en relación a las Soluciones de Datos, que complementan los conocimientos adquiridos previamente en la currícula de la carrera.

OBJETIVOS

- Adquirir las habilidades necesarias para aplicar nuevas tecnologías de gestión, procesamiento y almacenamiento de datos, en el marco de proyectos de sistemas de alta complejidad
- Utilizar herramientas de software de última generación para la gestión, procesamiento y almacenamiento de datos, en el marco de proyectos de sistemas de alta complejidad;
- Desarrollar habilidades para la explotación de grandes volúmenes de datos no estructurados provenientes de diversas fuentes dentro del paradigma de Big Data.
- Identificar el potencial del desarrollo de soluciones de datos en la nube, implementando software como servicio (SaaS).
- Distinguir ventajas de la aplicación en nubes públicas, privadas, híbridas o distribuidas.
- Reconocer modelos de almacenamiento en memoria y su aplicación al procesamiento y análisis de grandes volúmenes de datos.
- Aplicar las capacidades avanzadas de una base de datos relacional, tales como: arquitectura, modelado, optimización, resguardo y recuperación, replicación y seguridad.
- Reconocer la importancia de la interrelación de las diversas arquitecturas de soluciones de datos, en la convergencia tecnológica y las problemáticas de la información como recurso.
- Identificar nociones básicas de aprendizaje de máquina (Machine Learning), herramientas y metodologías para implementar un proyecto de Ciencia de Datos y analíticas avanzadas.

CONTENIDOS

Contenidos analíticos

Unidad Temática 1: Tendencias y tecnologías actuales: Introducción a modelos tecnológicos para el almacenamiento, procesamiento y distribución de datos. Nuevas tendencias y desafíos en la gestión de información. Implementaciones convencionales y nuevos casos de uso. Evolución tecnológica.



Universidad Tecnológica Nacional
Facultad Regional Buenos Aires

Unidad Temática 2: Técnicas eficientes para el modelado relacional: Conceptos de álgebra relacional. Buenas prácticas de modelado: DER, Normalización, Indexamiento (Patrones de Acceso), Optimización de tipos de datos, Mejora del Rendimiento, Modelado Dimensional, Particionamiento de datos.

Unidad Temática 3: Técnicas avanzadas para la replicación de datos: Introducción al concepto de replicación de datos. Estrategias de Replicación (Homogénea, Heterogénea, Federación). Arquitecturas. Configuración de replicación en un RDBMS. Algoritmos de balanceo de carga.

Unidad Temática 4: Seguridad, resguardo y restauración en bases de datos relacionales: Nociones de seguridad en bases de datos. Técnicas de Backup y Restore. Backup físico de la base de datos, definición y ejemplos. Backup del log de transacciones, definición y ejemplos. Restore físico y restore lógico, definición y ejemplos. Otros utilitarios de backup y restore a nivel base de datos y tabla, casos de uso y mejores prácticas.

Unidad Temática 5: Bases de datos en Memoria: Modelos de almacenamiento (Columnar y Por registro). Casos de uso del modelo columnar de almacenamiento. Arquitecturas híbridas. Soluciones tecnológicas. Ejemplos.

Unidad Temática 6: Big Data: Introducción a la gestión de grandes volúmenes de datos (Big Data) y a las problemáticas asociadas a su procesamiento y análisis. Arquitectura y Ejemplos. Introducción al framework de Hadoop y Spark. Procesamiento de datos no estructurados. Arquitectura lambda. Procesamiento y correlación de grandes volúmenes de datos provenientes de diversas fuentes.

Unidad Temática 7: Bases de datos NoSQL: Bases de datos NoSQL. Teorema CAP. MongoDB. Cassandra. Comparación de bases de datos NoSQL y bases de datos Relacionales. Integración con el modelo Relacional.

Unidad Temática 8: Cloud Databases: Concepto general de computación en la nube (Cloud Computing). Cloud públicas, privadas e híbridas. Cloud Databases. Marco Regulatorio y Seguridad. Anonimización de datos. Multitenancy. Introducción al gobierno de datos en la nube.

Unidad Temática 9: Internet Of Things: Introducción al concepto de Internet de los dispositivos (Internet Of Things o IoT). Problemáticas. Desafíos tecnológicos. Claves principales para el diseño de datos en IoT. Implementaciones actuales y ejemplos.

Unidad Temática 10: Integración de modelos: Modelo Relacional como hilo conductor. Soluciones Híbridas del Modelo Relacional y otros modelos. Convergencia tecnológica y nuevas problemáticas. Ejemplos.



*Universidad Tecnológica Nacional
Facultad Regional Buenos Aires*

Unidad Temática 11: Ciencia de Datos y Machine Learning: Introducción a los conceptos de ciencia de datos, inteligencia artificial y aprendizaje profundo (deep learning). Introducción al concepto de aprendizaje de máquina (machine learning). Reseña histórica del aprendizaje de máquina. Aprendizaje supervisado y no supervisado. Metodología y herramientas disponibles en el mercado. Ejemplos.

Unidad Temática 12: Trabajo Práctico Cuatrimestral: Durante el cuatrimestre se realizará un trabajo práctico que consistirá en la investigación de un tema a acordar con la Cátedra, relacionado con los temas tratados en la materia y/o con temas innovadores o casos de éxitos vinculados con soluciones de datos. Dicho trabajo práctico tendrá dos entregables (una presentación grupal de no más de 15 minutos y un informe sobre el tema de no más de 4 carillas de extensión). Adicionalmente, durante la cursada de la materia se realizarán hasta 7 prácticas sobre los siguientes temas:

- Práctica de Replicación (Federación de datos)
- Práctica de Backup y Restore
- Práctica de Bases de datos en memoria
- Práctica de Big Data (Spark)
- Práctica de Cloud Databases (Cloudant)
- Práctica de Internet Of Things (IoT)
- Práctica de Machine Learning (Watson Studio)

BIBLIOGRAFÍA OBLIGATORIA

- C. J. Date (2004). An Introduction to Database Systems – 8th Edition . Ed. Addison-Wesley Longman, Inc.
- Celko, J. (2015). Joe Celko's SQL for Smarties: Advanced SQL Programming 5th Edition. Ed. Morgan Kaufmann Publishers.
- Karau, H., Konwinski A., Wendell, P. Zaharia, M. (2015). Learning Spark: Lightning-Fast Big Data Analysis. Ed. O'Reilly Media.
- Stonebraker, M. (1996). Object-Relational DBMSs. Ed. Morgan Kaufmann Publishers.
- Takahashi, M. Shoko Azuma (2009). Manga Guide to Databases. Ed. No Starch Press.
- White, T. (2015). Hadoop: The Definitive Guide – 4th edition. Ed. O'Reilly Media.

PÁGINAS WEB DE INTERÉS

- IBM DB2 BLU Acceleration for Analytic Workloads:
<https://www.vconferenceonline.com/event/regeventp.aspx?id=971>
- Oracle Database In Memory:
<http://www.oracle.com/technetwork/database/in-memory/overview/twp-oracle-data-base-in-memory-2245633.html>
- Technical Overview SAP HANA: <http://scn.sap.com/docs/DOC-60354>



Universidad Tecnológica Nacional
Facultad Regional Buenos Aires

- SAP HANA - In-Memory Data Management In a Nutshell:
<https://open.sap.com/courses/hana-warmup>
- Investigating Storage Solutions for Large Data (Adam Lith, 2010):
<http://publications.lib.chalmers.se/records/fulltext/123839.pdf>
- The Google File System:
<http://static.googleusercontent.com/media/research.google.com/en//archive/gfs-sosp-2003.pdf>
- Cassandra Official Web Page: <http://cassandra.apache.org/>
- HBase Official Web Page: <http://hbase.apache.org/>
- MongoDB Official Web Page: <http://www.mongodb.org/>
- From the Internet of Computers to the Internet of Things:
<http://www.vs.inf.ethz.ch/publ/papers/Internet-of-things.pdf>
- Internet de las Cosas: los 12 adelantos tecnológicos que ya se pueden ver en la Argentina y que ponen al futuro cada vez más cerca:
<http://www.lanacion.com.ar/1827297-internet-de-las-cosas-los-12-adelantos-tecnologicos-que-ya-se-pueden-ver-en-la-argentina-y-que-ponen-al-futuro-cada-vez-mas-cerca>
- Internet of Things: Where Does the Data Go?:
<http://www.wired.com/insights/2015/03/internet-things-data-go/>

CORRELATIVAS

Para cursar y rendir

- Cursadas:
 - Redes de Datos
 - Administración de Sistemas de Información
 - Simulación
 - Ingeniería y Calidad de Software
- Aprobadas:
 - Diseño de Sistemas de Información
 - Sistemas Operativos
 - Bases de Datos