



ASIGNATURA:	TECNOLOGÍAS APLICADAS SOLUCIONES DE DATOS	A CÓDIGO:	
DEPARTAMENTO:	INGENIERÍA EN SISTEMAS DE INFORMACIÓN	CLASE:	Cuatrimestral
ÁREA:	INGENIERÍA EN SISTEMAS DE INFORMACIÓN	HORAS SEM.:	6 hs.
BLOQUE:	ELECTIVAS	HORAS / AÑO:	Reloj 72hs./ Cátedra 96hs

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.

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; mediante el uso de
- 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.

Programa analítico:

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.

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.



*Universidad Tecnológica Nacional
Facultad Regional Buenos Aires*

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.

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)

Distribución de carga horaria entre actividades teóricas y prácticas:

Tipo de actividad	Carga horaria total en hs. reloj	Carga horaria total en hs. cátedra
Teórica	50	60
Formación Práctica	22	36
Formación Experimental	0	0
Resolución de Problemas	0	0
Proyectos de Diseño	0	0
Práctica Supervisada	6	8
Total	72	96

Articulación Horizontal y vertical con otras materias

La asignatura Tecnologías Aplicadas a Soluciones de Datos se articula en forma vertical con tres (3) asignaturas que la preceden en el plan de estudio, específicamente Diseño de Sistemas, Gestión de Datos y Sistemas Operativos.

Cada estudiante deberá tener cursada y regularizada cada una de estas asignaturas al momento de comenzar la cursada. El modelado de datos, el lenguaje SQL y las nociones de sistemas operativos modernos, son conocimientos previos que se requieren para poder abordar e incorporar eficientemente el contenido impartido en la asignatura Tecnologías Aplicadas a Soluciones de Datos.

Además, los conocimientos adquiridos en Tecnologías Aplicadas a Soluciones de Datos serán de gran utilidad para extender los alcances de otras asignaturas de niveles previos, ya que imparte conocimientos y criterios abarcativos que permiten unificar los contenidos incorporados individualmente en materias previas vinculadas a la gestión de datos.

En cuanto a la articulación horizontal, Tecnologías Aplicadas a Soluciones de Datos brinda conocimientos que son compatibles y complementarios con conceptos y contenidos de otras asignaturas electivas, tales como Implementación de Base de Datos NoSQL, Inteligencia de Negocios e Inteligencia Artificial.

Cronograma estimado de clases:



Unidad temática	Duración en horas cátedra
1	6
2	6
3	6
4	6
5	12
6	9
7	9
8	6
9	6
10	6
11	6
12	18

Bibliografía:

- C. J. Date (2004). An Introduction to Database Systems – 8th Edition .Addison-Wesley Longman, Inc.
- Dr. Michael Stonebraker (1996). Object-Relational DBMSs. Morgan Kaufmann Publishers.
- Holden Karau, Andy Konwinski, Patrick Wendell, Matei Zaharia (2015). Learning Spark: Lightning-Fast Big Data Analysis. O'Reilly Media.
- Joe Celko (2015). Joe Celko's SQL for Smarties: Advanced SQL Programming 5th Edition. Morgan Kaufmann Publishers.
- Mana Takahashi, Shoko Azuma (2009). Manga Guide to Databases. No Starch Press.
- Tom White (2015). Hadoop: The Definitive Guide – 4th edition. 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-database-in-memory-2245633.html>
- Technical Overview SAP HANA: <http://scn.sap.com/docs/DOC-60354>
- 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-sosp2003.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>



*Universidad Tecnológica Nacional
Facultad Regional Buenos Aires*

- 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:

Cursadas: Redes de Información
Administración de Recursos
Simulación
Ingeniería en Software

Aprobadas: Diseño de Sistemas
Sistemas Operativos
Gestión de Datos

PARA RENDIR:

Aprobadas: Redes de Información
Administración de Recursos
Simulación
Ingeniería en Software