

1. Datos Generales de la asignatura

Nombre de la asignatura:	Bases de Datos NOSQL
Clave de la asignatura:	CDF - 2301
SATCA¹:	3 - 2 - 5
Carrera:	Ingeniería en Tecnologías de la Información y Comunicaciones

2. Presentación

Caracterización de la asignatura

El desarrollo de nuevas tecnologías en telecomunicaciones y almacenamiento de información han ocasionado que los sistemas informáticos también evolucionen hacia nuevos requerimientos. Actualmente las aplicaciones generan o tratan con enormes volúmenes de datos; que además que cambian constantemente y cuya estructura interna, pueden ser de tipo estructurada, semiestructurada o no estructurada. Por otra parte, también se requiere que los datos sean accesibles de forma concurrente desde distintos dispositivos y que tengan tener una alta disponibilidad para ser compartidos en ambientes colaborativos.

Por lo anterior, para responder a los nuevos requerimientos de los sistemas de información, esta asignatura aporta al perfil del Ingeniero en Tecnologías de la Información y Comunicaciones, la capacidad para tratar datos complejos caracterizados por tener estructuras o esquemas irregulares, parciales, dinámicos o desconocidos, mediante el uso de bases de datos NOSQL (Not Only SQL); aportando también las herramientas necesarias para desarrollar aplicaciones modernas que requieran un manejo de bases de datos flexibles, escalables, funcionales y de alto rendimiento, que además puedan ser operadas en ambientes colaborativos.

Intención didáctica

Las bases de datos relacionales (estructuradas o SQL) no se diseñaron para poder hacer frente a la escalabilidad y agilidad que necesitan las aplicaciones actuales, ni para aprovechar de los sistemas de procesamiento y almacenamiento disponible hoy en día. Para hacer frente a esos desafíos, surgen las bases de datos NOSQL, cuya principal característica es la de almacenar los datos en un formato diferente a las tablas relacionales, y que se pueden consultar empleando una API de lenguaje natural, lenguajes de consulta estructurados declarativos o lenguajes de consulta mediante ejemplo; de ahí que sean referidas como “No

¹ Sistema de Asignación y Transferencia de Créditos Académicos

Solo SQL *-Not Only SQL-*". A diferencia de las bases de datos tradicionales SQL, cuya estructura se encuentra bien definida, los escenarios actuales de manejo de información demandan el tratamiento de datos cuya **estructura no es definida a priori**, por ejemplo, imágenes, archivos multimedia o de audio, o los datos de tipo texto.

La asignatura se organiza en 4 temas, iniciando con la introducción que incluye un repaso del modelo relacional (SQL), y de la teoría de grafos y árboles; antecedentes necesarios para identificar las diferencias SQL vs NOSQL, así como para identificar los conceptos matemáticos que se utilizarán en tratamiento y consultas de datos NOSQL.

El tema dos aborda el tratamiento de datos semiestructurados empleando documentos XML, caracterizados por tener una estructura arborescente que les permite la flexibilidad necesaria para poder representar cualquier tipo de información; y que a la fecha se han convertido en un formato universal para el intercambio de información en la Web. Este tema incluye la construcción, manipulación y consulta en documentos XML.

El tema tres está destinado a las bases de datos, donde los datos se representan como un documento de tipo JSON para facilitar el almacenamiento y la consulta de información; además de aprovechar el modelo jerárquico, flexible y semiestructurado para permitir la evolución de la bases de datos de acuerdo a sus necesidades.

Por último, en el tema cuatro se presentan las bases de datos orientas a grafos, con el propósito de potenciar el desarrollo de aplicaciones que trabajen con conjuntos de datos altamente conectados o relacionados, por ejemplo: redes sociales, aplicaciones para la detección de fraudes o grafos de conocimiento.

3. Participantes en el diseño y seguimiento curricular del programa

Lugar y fecha de elaboración o revisión	Participantes	Observaciones
Instituto Tecnológico de Toluca, septiembre de 2022	M. en T. E. Luis Antonio Estrada Manuel Dr. Federico del Razo López Dr. Roberto Alejo Eleuterio Dra. Eréndira Rendón Lara	Miembros de la Academia de Sistemas Computacionales del Instituto Tecnológico de Toluca

4. Competencias a desarrollar

Competencias específicas de la asignatura
<ul style="list-style-type: none"> • Conoce e implementa programas para manipular y procesar información estructurada (SQL). • Conoce y aplica los fundamentos teóricos de la Teoría de Grafos, aplicados al procesamiento de datos semi estructurados y de las bases de datos orientas a grafos. • Conoce los fundamentos de XML para procesar información semiestructurada con las APIs SAX y DOM; además de ser capaz de realizar consultas con XPath. • Analiza, desarrolla e implementa una base de datos orientada a documentos usando MongoDB • Analiza, desarrolla e implementa una base de datos orientada a grafos empleando Neo4j • Analiza, desarrolla e implementa bases de datos estructuradas, semiestructuradas y no estructuradas, para dar solución a problemas reales.

5. Competencias previas

Las adquiridas en las asignaturas de:

- Bases de Datos Distribuidas
- Estructuras y Organización de Datos
- Fundamentos de Base de Datos
- Taller de Base de Datos
- Sistemas Operativos II

6. Temario

No.	Temas	Subtemas
1	Introducción	1.1 Bases de Datos Relacionales 1.2 Grafos y Árboles
2	Bases de Datos Semiestructuradas (XML)	2.1 Estructura XML 2.2 Construcción de documentos XML 2.3 APIs para la manipulación de documentos XML: SAX y DOM 2.4 Consulta de Documentos XML con XPath
3	Bases de Datos NOSQL	3.1 Bases de Datos Orientadas a Documentos 3.2 JSON 3.3 Construcción de una Base de Datos con MongoDB
4	Bases de Datos Orientadas a Grafos	4.1 Relaciones SQL vs NoSQL vs Bases de Datos Orientadas Grafos 4.2 Modelado de Datos con Grafos 4.3 Construcción de una Base de Datos con Neo4j

7. Actividades de aprendizaje de los temas

Tema 1: Introducción	
Competencias	Actividades de aprendizaje
<p>Específicas:</p> <ul style="list-style-type: none"> • Conoce e implementa programas para manipular y procesar información estructurada (SQL). • Conoce y aplica los fundamentos teóricos de la Teoría de Grafos, aplicados al procesamiento de datos semi estructurados y de las bases de datos orientas a grafos. <p>Genéricas:</p> <ul style="list-style-type: none"> • Habilidades de investigación. • Capacidad de análisis y de síntesis. • Habilidad para trabajar en forma autónoma y en equipo. • Comunicación oral y escrita. • Flexibilidad. • Gestión de tiempo. • Colaboración. 	<ul style="list-style-type: none"> • Configuración del entorno de trabajo: <ul style="list-style-type: none"> ○ Sistema Operativo (Windows, Linux), ○ Virtualización (VirtualBox, Vagrant) ○ Editor/Depurador (Code Visual Studio, Sublime Text) ○ Manejador de Base de Datos SQL (SQLite, PostgreSQL) • Analizar y comprender los conceptos de la teoría de grafos • Elaboración de Mapas Mentales • Desarrollo e implementación de una Base de Datos Estructurada.
Tema 2: Bases de Datos Semiestructuradas (XML)	
Competencias	Actividades de aprendizaje
<p>Específicas:</p> <ul style="list-style-type: none"> • Conoce los fundamentos de XML para procesar información semiestructurada con las APIs SAX y DOM; además de ser capaz de realizar consultas con XPath. <p>Genéricas:</p> <ul style="list-style-type: none"> • Habilidades de investigación. • Capacidad de análisis y de síntesis. • Habilidad para trabajar en forma autónoma y en equipo. • Comunicación oral y escrita. • Flexibilidad. • Gestión de tiempo. • Colaboración. 	<ul style="list-style-type: none"> • Analizar y comprender los conceptos de XML, las APIs de manipulación y XPath para consulta. • Elaboración de Mapas Mentales • Desarrollo de programas para manipular y procesar información SQL y XML

Tema 3: Bases de Datos NOSQL	
Competencias	Actividades de aprendizaje
<p>Específicas:</p> <ul style="list-style-type: none"> • Analiza, desarrolla e implementa una base de datos orientada a documentos usando MongoDB <p>Genéricas:</p> <ul style="list-style-type: none"> • Habilidades de investigación. • Capacidad de análisis y de síntesis. • Habilidad para trabajar en forma autónoma y en equipo. • Comunicación oral y escrita. • Flexibilidad. • Gestión de tiempo. • Colaboración. 	<ul style="list-style-type: none"> • Instalación y configuración de MongoDB, para desarrollar una base de datos orientada a documentos. • Elaboración de Mapas Mentales • Desarrollo e implementación de una base de datos en MongoDB
Tema 4: Bases de Datos Orientadas a Grafos	
Competencias	Actividades de aprendizaje
<p>Específicas:</p> <ul style="list-style-type: none"> • Analiza, desarrolla e implementa una base de datos orientada a grafos empleando Neo4j <p>Genéricas:</p> <ul style="list-style-type: none"> • Habilidades de investigación. • Capacidad de análisis y de síntesis. • Habilidad para trabajar en forma autónoma y en equipo. • Comunicación oral y escrita. • Flexibilidad. • Gestión de tiempo. • Colaboración. 	<ul style="list-style-type: none"> • Instalación y configuración de Neo4j, para desarrollar una base de datos orientada a grafos. • Elaboración de Mapas Mentales • Desarrollo e implementación de una base de datos en Neo4i

8. Práctica(s)

- **Práctica 1.** Instalación y configuración del ambiente de trabajo (Sistema Operativo, Python, IDE, APIs, Base de Datos Relacional)
- **Práctica 2.** Implementación de una Bases de Datos Relacional.
- **Práctica 3.** Desarrollo de programas para manipular conjuntos de datos XML
- **Práctica 4.** Implementación de una Bases de Datos en MongoDB.
- **Práctica 5.** Implementación de una Bases de Datos en Neo4j

9. Proyecto de asignatura

El objetivo del proyecto que planteé el docente que imparta esta asignatura, es demostrar el desarrollo y alcance de la(s) competencia(s) de la asignatura, considerando las siguientes fases:

- **Fundamentación:** marco referencial (teórico, conceptual, contextual, legal) en el cual se fundamenta el proyecto de acuerdo con un diagnóstico realizado, mismo que permite a los estudiantes lograr la comprensión de la realidad o situación objeto de estudio para definir un proceso de intervención o hacer el diseño de un modelo.
- **Planeación:** con base en el diagnóstico en esta fase se realiza el diseño del proyecto por parte de los estudiantes con asesoría del docente; implica planificar un proceso: de intervención empresarial, social o comunitario, el diseño de un modelo, entre otros, según el tipo de proyecto, las actividades a realizar los recursos requeridos y el cronograma de trabajo.
- **Ejecución:** consiste en el desarrollo de la planeación del proyecto realizada por parte de los estudiantes con asesoría del docente, es decir en la intervención (social, empresarial), o construcción del modelo propuesto según el tipo de proyecto, es la fase de mayor duración que implica el desempeño de las competencias genéricas y específicas a desarrollar.
- **Evaluación:** es la fase final que aplica un juicio de valor en el contexto laboral-profesión, social e investigativo, ésta se debe realizar a través del reconocimiento de logros y aspectos a mejorar se estará promoviendo el concepto de “evaluación para la mejora continua”, la metacognición, el desarrollo del pensamiento crítico y reflexivo en los estudiantes.

10. Evaluación por competencias

La evaluación debe ser continua y formativa por lo que se debe considerar el desempeño en cada una de las actividades de aprendizaje, haciendo especial énfasis en:

- Uso de rúbricas de evaluación y tablas de cotejo
- Aplicación de exámenes teóricos y prácticos
- Entrega a tiempo de productos de aprendizaje
- Realizar una retroalimentación efectiva de los diferentes productos de aprendizaje
- Calendarización congruente de la entrega de los productos de aprendizaje para los diversos momentos de las evaluaciones diagnóstica, formativa y sumativa

11. Fuentes de información

1. Moller A., Schwartzbach M., (2005), An Introduction to XML and Web Technologies, Addison–Wesley
2. Stamile, C., Marzullo A. (2021) Graph Machine Learning, Packt Publishing.
3. Jones, C. A., Drake, F. L., Lewin, L. (2002). Python and XML: [XML processing with Python]. Beijing [u.a.]: O'Reilly & Associates. ISBN: 0596001282 9780596001285
4. Banker, K., Bakkum, P., Verch, S., Garrett, D., Hawkins, T. (2016). MongoDB in Action. Manning. ISBN: 1617291609 9781617291609
5. Kelly, A., & McCreary, D. (2013). Making Sense of NoSQL (2013). Manning Publications. Retrieved from <https://www.perlego.com/book/1469465/making-sense-of-nosql-pdf>
6. Bechberger D., Perryman P., (2021). Exploring Graph Databases; Manning Publications
7. Robinson, I., Webber, J. & Efreem, E. (2015). Graph databases O'Reilly Media