



compensar

fundación
universitaria



UCompensar

EDUCACIÓN PARA AVANZAR

SYLLABUS

Técnicas de Extracción y Almacenamiento
de Datos Masivos





FUNDACIÓN UNIVERSITARIA PANAMERICANA – UNIPANAMERICANA

SYLLABUS - TÉCNICAS DE EXTRACCIÓN Y ALMACENAMIENTO DE DATOS MASIVOS

1. INFORMACIÓN GENERAL DEL CURSO

Programa Académico:	ESPECIALIZACIÓN EN BIG DATA	Nivel de Formación.	Posgrado - Especialización
Sede:	Bogotá	Código del curso:	N/A
Nombre del curso:	Técnicas de Extracción y Almacenamiento de Datos Masivos	Modalidad	Presencial y virtual
Semestre:	I	Número de Créditos:	3
Tipo del curso:	Teórico Práctico	Componente o Área de formación:	Disciplinar
Carácter del curso:	Obligatoria	Versión:	1.0
Horas Totales:	144	Horas de Trabajo Directo con Docente	36
		Horas Trabajo Independiente:	108

2. PRESENTACIÓN DEL CURSO

El presente curso es importante para el especialista en formación, ya que en él tendrá la oportunidad de conocer a profundidad y aplicar las técnicas y procedimientos existentes en cuanto a la seguridad de datos para la extracción, transformación y aseguramiento de datos masivos, así como las políticas y la legislación vigente que garantiza la integridad y seguridad de los datos, aspectos fundamentales para el Big Data.

Igualmente permite obtener un manejo importante de las técnicas y procedimientos de extracción de datos (de acuerdo con sus características), los procedimientos de seguridad que aplican en diferentes casos, la relación de estos con las políticas que existen en la organización, así como dominar las técnicas para la transformación y aseguramiento de los datos masivos en diferentes contextos o situaciones. Igualmente, será posible que el estudiante participe en situaciones o eventos simulados acerca de la integridad de los datos utilizando las técnicas, políticas y legislación vigente de la seguridad de los datos.

Para su abordaje, se realizan actividades con acompañamiento del docente-tutor y actividades de trabajo independiente de orden particular o grupal, en medio de los cuales se realizarán los desarrollos conceptuales requeridos y la aplicación en contextos, casos o situaciones reales, utilizando modelos computacionales e interpretación de los datos masivos en situaciones o casos asignados.

3. COMPETENCIAS ESPECIFICAS, TRANSVERSALES Y GENÉRICAS

Unidad de competencia	Gestiona, visualiza y asegura grandes volúmenes de datos estructurados, semiestructurados o no estructurados, provenientes de diferentes fuentes de información utilizando procesos de extracción, transformación, análisis, interpretación y visualización con el fin de integrarlos y detectar patrones, relaciones, tendencias o posibilidades que facilite la toma de decisiones en el ámbito empresarial en el cual se aplique.
-----------------------	--



Elementos de competencia	Aplica técnicas y procedimientos de seguridad de datos para la extracción, transformación y aseguramiento de datos masivos con base en técnicas, políticas y legislación vigente para la integridad y seguridad de los datos.
4. CRITERIOS DE REALIZACIÓN	
A continuación, se presentan los criterios de realización que corresponden al desarrollo del curso y los respectivos criterios de realización, asociados a los desempeños que debe evidenciar el estudiante en su actuación. Estos son la base de las rúbricas de evaluación, de los cuales se derivarán los indicadores y sus descriptores por nivel de desempeño.	
CR 1.	Interpretación de las técnicas y procedimientos de extracción de datos a partir del origen, fuente, naturaleza y tipos de los datos masivos.
CR 2.	Presentación de los procedimientos de seguridad con base en las políticas y legislación vigente para los datos masivos
CR 3.	Uso preciso y pertinente de las técnicas de extracción, transformación y aseguramiento de los datos masivos en situaciones o casos asignados.
CR 4.	Simulación de la integridad de los datos utilizando las técnicas, políticas y legislación vigente de la seguridad de los datos en las organizaciones en situaciones o casos asignados.
CR 5.	Socialización de las técnicas de extracción y almacenamiento aplicados, a partir de modelos computacionales e interpretación de los datos masivos en situaciones o casos asignados.
5. METODOLOGÍA QUE SE DESARROLLA EN EL CURSO	
<p>Etapa 0. Reconocimiento: se valora el nivel de competencia de entrada de los estudiantes y sus saberes previos sobre: Concepto de dato, seguridad de dato y de los aspectos éticos en el almacenamiento de dato. Esto se realiza a través de una actividad tipo conversatorio o entrevista colectiva a partir de la cual el docente-tutor puede establecer en qué aspectos debe detenerse, ampliar el contenido o bien qué debe integrar en el contenido del curso. Esta actividad tiene un carácter diagnóstico, por lo tanto, no es calificable.</p> <p>Metodología Activa: Aprendizaje basado en problemas. El aprendizaje basado en problemas es un proceso mediante el cual los alumnos se involucran para desarrollar grupos de trabajo, de esta forma realizan procesos de aprendizaje cíclico compuesto por diferentes etapas que permiten formular diferentes preguntas y adquirir conocimiento adquiridas por medio de la metodología del pensamiento crítico, resolución de problemas, motivación y capacidad de transferir conocimiento en nuevas situaciones.</p> <p>Etapa 1. Contextualización: mediante algunos ejemplos tomados del contexto real, el estudiante tendrá la oportunidad comprender las generalidades del proceso de extracción y almacenamiento de Datos Masivos, y los aspectos principales del dato, como: Orígenes, fuentes, naturaleza y tipos de dato. Esto permitirá no sólo reconocer los elementos conceptuales principales sino llevarlos a una ubicación en situaciones reales de estos contenidos, lo cual es fundamental para la posterior aplicación de las técnicas de extracción y almacenamiento de datos masivos.</p> <p>Metodología Activa: Aprendizaje basado en problemas. El aprendizaje basado en problemas es un proceso mediante el cual los alumnos se involucran para desarrollar grupos de trabajo, de esta forma realizan procesos de aprendizaje cíclico compuesto por diferentes etapas que permiten formular diferentes preguntas y adquirir</p>	

conocimiento adquiridas por medio de la metodología del pensamiento crítico, resolución de problemas, motivación y capacidad de transferir conocimiento en nuevas situaciones.

Etapla 2. Profundización: Para el desarrollo de los contenidos temáticos requeridos para en el fortalecimiento de la competencia, se tiene previsto un proceso de acompañamiento y tutoría directa del docente al estudiante y unas actividades de aprendizaje en los tiempos de estudio autónomo o independiente, muy pertinentes para el nivel de formación y el resultado de aprendizaje esperado.

Para los **tiempos de aprendizaje independiente**, los estudiantes desarrollan actividades de aprendizaje autónomo y social, individual o en grupo, abordando materiales y Recursos Educativos (Digitales o físicos) disponibles en material de estudio o espacios On-Line (LMS o repositorios con Recursos Educativos Digitales). Allí, los estudiantes avanzan en la comprensión, análisis y apropiación de elementos conceptuales y teóricos propios de la competencia del curso.

En **los tiempos de trabajo directo con docente** realizan la discusión de los aspectos conceptuales previamente abordados en trabajo independiente, para posteriormente realizar su aplicación en talleres que giran en torno a casos o situaciones del entorno real, los cuales contarán con coaching y espacios de Co-working, para realizar la interpretación efectiva de las técnicas y procedimientos de extracción de datos a partir del origen, fuente, naturaleza y tipos de los datos masivos, la aplicación de procedimientos de seguridad con base en las políticas y legislación vigente para los datos masivos

Igualmente, desarrollarán actividades en torno a una situación o caso asignado, en los cuales deberán utilizar las técnicas de extracción, transformación y aseguramiento de los datos masivos, así como realizar simulaciones de la integridad de los datos, para lo cual pondrán en uso las técnicas, políticas y legislación vigente de la seguridad de los datos revisadas en las sesiones.

Metodología Activa: Aprendizaje basado en problemas.

El aprendizaje basado en problemas es un proceso mediante el cual los alumnos se involucran para desarrollar grupos de trabajo, de esta forma realizan procesos de aprendizaje cíclico compuesto por diferentes etapas que permiten formular diferentes preguntas y adquirir conocimiento adquiridas por medio de la metodología del pensamiento crítico, resolución de problemas, motivación y capacidad de transferir conocimiento en nuevas situaciones.

Etapla 3: Transferencia: en este momento del curso el estudiante contará con sesiones de trabajo directo con docente (presenciales y/o virtuales mediadas por herramientas tecnológicas, según la modalidad) y actividades independientes, en las cuales realizará la preparación de la sustentación en la cual realizará la explicación auditable de los principios estandarizados (utilizando las herramientas tecnológicas seleccionadas), evidenciando el uso de los datos masivos y su dominio conceptual de la relación Big Data e Inteligencia de los negocios. Esta es una actividad formativa y calificable.

El curso contará con una actividad “Peer to peer” que representa la interacción (virtual o presencial) con un experto en la temática, del sector productivo nacional o internacional. En dicho espacio los estudiantes podrán participar de un conversatorio, un taller o una cátedra orientado por el invitado.

En un espacio final, los estudiantes tendrán la oportunidad de socializar y sustentar el proceso realizado, explicando las técnicas de extracción y almacenamiento aplicados, a partir de modelos computacionales e interpretación de los datos masivos en situaciones o casos asignados.

Al ser un proceso educativo por competencias la evaluación se convierte en formativa, permanente, integrada, es decir, hace parte de la estrategia formativa (actividades de aprendizaje) y en integral, es decir, participa en la valoración de niveles de competencia los distintos actores que intervienen en el proceso, así:

- **Autoevaluación:** El estudiante valora sus propios desempeños, identificando el nivel alcanzado en cada momento, a partir de la revisión crítica y transparente de su propio proceso, atendiendo a unos criterios de realización/desempeño establecidos para estimar el nivel de competencia, detectando con ello fortalezas, talentos, capacidades especiales al igual que dificultades, limitantes y oportunidades de mejora.
- **Coevaluación:** Los estudiantes realizan una valoración conjunta sobre la actuación del grupo, atendiendo a unos criterios de realización/desempeño establecidos en consenso; participando así en la valoración de los niveles de desempeño evidenciados.
- **Heteroevaluación:** El profesor y los expertos invitados, valoran los desempeños de los estudiantes, identificando el nivel alcanzado en cada momento, atendiendo a unos criterios de realización/desempeño establecidos para estimar el nivel de competencia.

Metodología Activa: Aprendizaje basado en problemas.

El aprendizaje basado en problemas es un proceso mediante el cual los alumnos se involucran para desarrollar grupos de trabajo, de esta forma realizan procesos de aprendizaje cíclico compuesto por diferentes etapas que permiten formular diferentes preguntas y adquirir conocimiento adquiridas por medio de la metodología del pensamiento crítico, resolución de problemas, motivación y capacidad de transferir conocimiento en nuevas situaciones.

6. CONTENIDOS TEMATICOS

Estos contenidos se refieren a los conocimientos, actitudes, habilidades, valores y principios asociados a la competencia y que se considera necesario ser apropiados e incorporados por los estudiantes a sus estructuras (cognitivas, éticas y procedimentales) particulares, en procura de la actuación idónea esperada.

Etapas 0. Reconocimiento

Concepto de dato y seguridad de dato

Aspectos éticos en el almacenamiento de dato

Etapas 1. Contextualización

- El dato: Orígenes, fuentes, naturaleza y tipos de los datos
- Generalidades del proceso de extracción y almacenamiento de Datos Masivos

Etapas 2. Profundización

- Importancia de la seguridad de los datos en el proceso de extracción y almacenamiento de datos masivos
- Políticas y legislación en seguridad de los datos masivos



- Técnicas y procedimientos de seguridad de datos masivos
- Técnicas de extracción
- Técnicas de almacenamiento y carga de datos masivos con seguridad

Etapas 3: Transferencia

- Efectividad en la interpretación de las técnicas y procedimientos de extracción de datos (de acuerdo con el origen, fuente, naturaleza y tipos de los datos masivos)
- Cómo sustentar los procedimientos de seguridad
- Simulación de la integridad de los datos en la organización
- Socialización de las técnicas de extracción y almacenamiento aplicados: uso de modelos computacionales y realización de interpretación de los datos masivos.

Recursos bibliográficos		Recurso disponible en		
		Biblioteca de la institución		Otras bibliotecas
		Físico	Digital	
<p>Nota: El número de referentes obligatorios depende de la estrategia de formación planteada en el curso. Sin embargo, todos ellos deben estar asociados a las bases de datos digitales o al catálogo de recursos bibliográficos disponibles en la biblioteca de la institución.</p> <p>La bibliografía debe contener al menos siete referentes para consulta opcional por parte del estudiante, si desea profundizar o ampliar su conocimiento; al menos tres de ellos deben ser en otros idiomas.</p>		<p>Enlace para ir al catálogo bibliográfico disponible: http://biblioteca.unipanamericana.edu.co/ianium-bin/busqueda_rapida.pl?id=20151110162718</p> <p>Enlace para ir a las colecciones disponibles: https://unipanamericanaeduco.sharepoint.com/Portal%20MiPana/SitePages/Colecciones.aspx</p> <p>Marcar con una X según corresponda</p>	<p>Enlace para ir a las bases de datos y libros digitales: https://unipanamericana.educo/login.com:2443/login</p> <p>Marcar con una X según corresponda</p>	<p>Enlace para ir a los recursos digitales externos: https://unipanamericanaeduco.sharepoint.com/Portal%20MiPana/SitePages/Informacion%20de%20Interes.aspx</p> <p>Marcar con una X según corresponda</p>
Obligatorios	<i>Trujillo, Juan Carlos. Diseño y explotación de almacenes de datos: conceptos básicos de modelado multidimensional (2013)</i>		https://ebookcentral-proquest-com.unipanamericana.basesdedatosezproxy.com/lib/unipanamericanaasp/detail.action?docID=3214871	
	<i>Peso, Navarro, Emilio del, and González, Miguel Ángel Ramos. La seguridad de los datos de carácter personal (2a. ed.), Ediciones Díaz de Santos, 2002.</i>		https://ebookcentral-proquest-com.unipanamericana.basesdedatosezproxy.com/lib/unipanamericanaasp/detail.action?docID=4795089	
	<i>Escrivá, Gascó, Gema, et al. Seguridad informática, Macmillan Iberia, S.A., 2013.</i>		https://ebookcentral-proquest-com.unipanamericana.basesdedatosezproxy.com/lib/unipanamericanaasp/detail.action?docID=3217398	



	<i>Seguridad informática / Jesús Costas</i>	ISBN: 9789588675701		
Complementarios	<i>ETL Evolution for Real- Time Data Warehousing</i>		https://www.researchgate.net/profile/Theresa_Kraft/publication/280837435_ETL_Evolution_for_Real-Time_Data_Warehousing/links/56008f5308ae07629e52af09.pdf	
	<i>Agile ETL</i>		https://reader.elsevier.com/reader/sd/pii/S2212017313001965?token=7F809B8E9FEDA2393BB84D030FAD18E38D619B98E4757033607E272C8407C143186754C1EB30C7E3AFB79C401B122F70	

Glosario:

- **Big Data**
It is a broad term for very large or complex data sets that traditional data processing applications are not insufficient to handle.
- **Legislation**
it is the law that has been created by a legislature or other government body that allows protecting the data. The term may refer to a single law, or to the collective body of enacted laws, while "statute" is also used to refer to a single law.
- **Extraction Techniques**
Information Extraction (Information Extraction) is a type of information retrieval whose objective is to automatically extract structured or semi-structured information from documents readable by a computer.
- **Data security**
refers to digital privacy protection measures that are applied to prevent unauthorized access to data, which can be found on computers, databases, and websites.
- **Data simulation**

it is the controlled statistical sampling technique, which is used together with a model, to obtain approximate answers to questions that arise in complex problems of a probabilistic type.

7. INFORMACIÓN GENERAL DEL CURSO			
Programa Académico:	ESPECIALIZACIÓN EN BIG DATA	Nivel de Formación.	Posgrado - Especialización
Sede:	Bogotá	Código del curso:	N/A
Nombre del curso:	Técnicas de Extracción y Almacenamiento de Datos Masivos	Modalidad	Presencial y virtual
Semestre:	I	Número de Créditos:	3
Tipo del curso:	Teórico Práctico	Componente o Área de formación:	Disciplinar
Carácter del curso:	Obligatoria	Versión:	1.0
Horas Totales:	144	Horas de Trabajo Directo con Docente	36
		Horas Trabajo Independiente:	108
8. PRESENTACIÓN DEL CURSO			
<p>El presente curso es importante para el especialista en formación, ya que en él tendrá la oportunidad de conocer a profundidad y aplicar las técnicas y procedimientos existentes en cuanto a la seguridad de datos para la extracción, transformación y aseguramiento de datos masivos, así como las políticas y la legislación vigente que garantiza la integralidad y seguridad de los datos, aspectos fundamentales para el Big Data.</p> <p>Igualmente permite obtener un manejo importante de las técnicas y procedimientos de extracción de datos (de acuerdo con sus características), los procedimientos de seguridad que aplican en diferentes casos, la relación de estos con las políticas que existen en la organización, así como dominar las técnicas para la transformación y aseguramiento de los datos masivos en diferentes contextos o situaciones. Igualmente, será posible que el estudiante participe en situaciones o eventos simulados acerca de la integridad de los datos utilizando las técnicas, políticas y legislación vigente de la seguridad de los datos.</p> <p>Para su abordaje, se realizan actividades con acompañamiento del docente-tutor y actividades de trabajo independiente de orden particular o grupal, en medio de los cuales se realizarán los desarrollos conceptuales requeridos y la aplicación en contextos, casos o situaciones reales, utilizando modelos computacionales e interpretación de los datos masivos en situaciones o casos asignados.</p>			
9. COMPETENCIAS ESPECIFICAS, TRANSVERSALES Y GENÉRICAS			
Unidad de competencia	Gestiona, visualiza y asegura grandes volúmenes de datos estructurados, semiestructurados o no estructurados, provenientes de diferentes fuentes de información utilizando procesos de extracción, transformación, análisis, interpretación y visualización con el fin de integrarlos		

	y detectar patrones, relaciones, tendencias o posibilidades que facilite la toma de decisiones en el ámbito empresarial en el cual se aplique.
Elementos de competencia	Aplica técnicas y procedimientos de seguridad de datos para la extracción, transformación y aseguramiento de datos masivos con base en técnicas, políticas y legislación vigente para la integridad y seguridad de los datos.
10. CRITERIOS DE REALIZACIÓN	
A continuación, se presentan los criterios de realización que corresponden al desarrollo del curso y los respectivos criterios de realización, asociados a los desempeños que debe evidenciar el estudiante en su actuación. Estos son la base de las rúbricas de evaluación, de los cuales se derivarán los indicadores y sus descriptores por nivel de desempeño.	
CR 1.	Interpretación de las técnicas y procedimientos de extracción de datos a partir del origen, fuente, naturaleza y tipos de los datos masivos.
CR 2.	Presentación de los procedimientos de seguridad con base en las políticas y legislación vigente para los datos masivos
CR 3.	Uso preciso y pertinente de las técnicas de extracción, transformación y aseguramiento de los datos masivos en situaciones o casos asignados.
CR 4.	Simulación de la integridad de los datos utilizando las técnicas, políticas y legislación vigente de la seguridad de los datos en las organizaciones en situaciones o casos asignados.
CR 5.	Socialización de las técnicas de extracción y almacenamiento aplicados, a partir de modelos computacionales e interpretación de los datos masivos en situaciones o casos asignados.
11. METODOLOGÍA QUE SE DESARROLLA EN EL CURSO	
<p>Etapas 0. Reconocimiento: se valora el nivel de competencia de entrada de los estudiantes y sus saberes previos sobre: Concepto de dato, seguridad de dato y de los aspectos éticos en el almacenamiento de dato. Esto se realiza a través de una actividad tipo conversatorio o entrevista colectiva a partir de la cual el docente-tutor puede establecer en qué aspectos debe detenerse, ampliar el contenido o bien qué debe integrar en el contenido del curso. Esta actividad tiene un carácter diagnóstico, por lo tanto, no es calificable.</p> <p>Metodología Activa: Aprendizaje basado en problemas. El aprendizaje basado en problemas es un proceso mediante el cual los alumnos se involucran para desarrollar grupos de trabajo, de esta forma realizan procesos de aprendizaje cíclico compuesto por diferentes etapas que permiten formular diferentes preguntas y adquirir conocimiento adquiridas por medio de la metodología del pensamiento crítico, resolución de problemas, motivación y capacidad de transferir conocimiento en nuevas situaciones.</p> <p>Etapas 1. Contextualización: mediante algunos ejemplos tomados del contexto real, el estudiante tendrá la oportunidad comprender las generalidades del proceso de extracción y almacenamiento de Datos Masivos, y los aspectos principales del dato, como: Orígenes, fuentes, naturaleza y tipos de dato. Esto permitirá no sólo reconocer los elementos conceptuales principales sino llevarlos a una ubicación en situaciones reales de estos contenidos, lo cual es fundamental para la posterior aplicación de las técnicas de extracción y almacenamiento de datos masivos.</p> <p>Metodología Activa: Aprendizaje basado en problemas. El aprendizaje basado en problemas es un proceso mediante el cual los alumnos se involucran para desarrollar grupos de trabajo, de esta forma realizan procesos de aprendizaje cíclico</p>	

compuesto por diferentes etapas que permiten formular diferentes preguntas y adquirir conocimiento adquiridas por medio de la metodología del pensamiento crítico, resolución de problemas, motivación y capacidad de transferir conocimiento en nuevas situaciones.

Etapla 2. Profundización: Para el desarrollo de los contenidos temáticos requeridos para en el fortalecimiento de la competencia, se tiene previsto un proceso de acompañamiento y tutoría directa del docente al estudiante y unas actividades de aprendizaje en los tiempos de estudio autónomo o independiente, muy pertinentes para el nivel de formación y el resultado de aprendizaje esperado.

Para los **tiempos de aprendizaje independiente**, los estudiantes desarrollan actividades de aprendizaje autónomo y social, individual o en grupo, abordando materiales y Recursos Educativos (Digitales o físicos) disponibles en material de estudio o espacios On-Line (LMS o repositorios con Recursos Educativos Digitales). Allí, los estudiantes avanzan en la comprensión, análisis y apropiación de elementos conceptuales y teóricos propios de la competencia del curso.

En **los tiempos de trabajo directo con docente** realizan la discusión de los aspectos conceptuales previamente abordados en trabajo independiente, para posteriormente realizar su aplicación en talleres que giran en torno a casos o situaciones del entorno real, los cuales contarán con coaching y espacios de Co-working, para realizar la interpretación efectiva de las técnicas y procedimientos de extracción de datos a partir del origen, fuente, naturaleza y tipos de los datos masivos, la aplicación de procedimientos de seguridad con base en las políticas y legislación vigente para los datos masivos

Igualmente, desarrollarán actividades en torno a una situación o caso asignado, en los cuales deberán utilizar las técnicas de extracción, transformación y aseguramiento de los datos masivos, así como realizar simulaciones de la integridad de los datos, para lo cual pondrán en uso las técnicas, políticas y legislación vigente de la seguridad de los datos revisadas en las sesiones.

Metodología Activa: Aprendizaje basado en problemas.

El aprendizaje basado en problemas es un proceso mediante el cual los alumnos se involucran para desarrollar grupos de trabajo, de esta forma realizan procesos de aprendizaje cíclico compuesto por diferentes etapas que permiten formular diferentes preguntas y adquirir conocimiento adquiridas por medio de la metodología del pensamiento crítico, resolución de problemas, motivación y capacidad de transferir conocimiento en nuevas situaciones.

Etapla 3: Transferencia: en este momento del curso el estudiante contará con sesiones de trabajo directo con docente (presenciales y/o virtuales mediadas por herramientas tecnológicas, según la modalidad) y actividades independientes, en las cuales realizará la preparación de la sustentación en la cual realizará la explicación auditable de los principios estandarizados (utilizando las herramientas tecnológicas seleccionadas), evidenciando el uso de los datos masivos y su dominio conceptual de la relación Big Data e Inteligencia de los negocios. Esta es una actividad formativa y calificable.

El curso contará con una actividad “Peer to peer” que representa la interacción (virtual o presencial) con un experto en la temática, del sector productivo nacional o internacional. En dicho espacio los estudiantes podrán participar de un conversatorio, un taller o una cátedra orientado por el invitado.

En un espacio final, los estudiantes tendrán la oportunidad de socializar y sustentar el proceso realizado, explicando las técnicas de extracción y almacenamiento aplicados, a partir de modelos computacionales e interpretación de los datos masivos en situaciones o casos asignados.

Al ser un proceso educativo por competencias la evaluación se convierte en formativa, permanente, integrada, es decir, hace parte de la estrategia formativa (actividades de aprendizaje) y en integral, es decir, participa en la valoración de niveles de competencia los distintos actores que intervienen en el proceso, así:

- **Autoevaluación:** El estudiante valora sus propios desempeños, identificando el nivel alcanzado en cada momento, a partir de la revisión crítica y transparente de su propio proceso, atendiendo a unos criterios de realización/desempeño establecidos para estimar el nivel de competencia, detectando con ello fortalezas, talentos, capacidades especiales al igual que dificultades, limitantes y oportunidades de mejora.
- **Coevaluación:** Los estudiantes realizan una valoración conjunta sobre la actuación del grupo, atendiendo a unos criterios de realización/desempeño establecidos en consenso; participando así en la valoración de los niveles de desempeño evidenciados.
- **Heteroevaluación:** El profesor y los expertos invitados, valoran los desempeños de los estudiantes, identificando el nivel alcanzado en cada momento, atendiendo a unos criterios de realización/desempeño establecidos para estimar el nivel de competencia.

Metodología Activa: Aprendizaje basado en problemas.

El aprendizaje basado en problemas es un proceso mediante el cual los alumnos se involucran para desarrollar grupos de trabajo, de esta forma realizan procesos de aprendizaje cíclico compuesto por diferentes etapas que permiten formular diferentes preguntas y adquirir conocimiento adquiridas por medio de la metodología del pensamiento crítico, resolución de problemas, motivación y capacidad de transferir conocimiento en nuevas situaciones.

12. CONTENIDOS TEMATICOS

Estos contenidos se refieren a los conocimientos, actitudes, habilidades, valores y principios asociados a la competencia y que se considera necesario ser apropiados e incorporados por los estudiantes a sus estructuras (cognitivas, éticas y procedimentales) particulares, en procura de la actuación idónea esperada.

Etapas 0. Reconocimiento

Concepto de dato y seguridad de dato

Aspectos éticos en el almacenamiento de dato

Etapas 1. Contextualización

- El dato: Orígenes, fuentes, naturaleza y tipos de los datos
- Generalidades del proceso de extracción y almacenamiento de Datos Masivos

Etapas 2. Profundización

- Importancia de la seguridad de los datos en el proceso de extracción y almacenamiento de datos masivos
- Políticas y legislación en seguridad de los datos masivos



- Técnicas y procedimientos de seguridad de datos masivos
- Técnicas de extracción
- Técnicas de almacenamiento y carga de datos masivos con seguridad

Etapas 3: Transferencia

- Efectividad en la interpretación de las técnicas y procedimientos de extracción de datos (de acuerdo con el origen, fuente, naturaleza y tipos de los datos masivos)
- Cómo sustentar los procedimientos de seguridad
- Simulación de la integridad de los datos en la organización
- Socialización de las técnicas de extracción y almacenamiento aplicados: uso de modelos computacionales y realización de interpretación de los datos masivos.

Recursos bibliográficos		Recurso disponible en		
		Biblioteca de la institución		Otras bibliotecas
		Físico	Digital	
<p>Nota: El número de referentes obligatorios depende de la estrategia de formación planteada en el curso. Sin embargo, todos ellos deben estar asociados a las bases de datos digitales o al catálogo de recursos bibliográficos disponibles en la biblioteca de la institución.</p> <p>La bibliografía debe contener al menos siete referentes para consulta opcional por parte del estudiante, si desea profundizar o ampliar su conocimiento; al menos tres de ellos deben ser en otros idiomas.</p>		<p>Enlace para ir al catálogo bibliográfico disponible: http://biblioteca.unipanamericana.edu.co/ianium-bin/busqueda_rapida.pl?id=20151110162718</p> <p>Enlace para ir a las colecciones disponibles: https://unipanamericanaeduco.sharepoint.com/Portal%20MiPana/SitePages/Colecciones.aspx</p> <p>Marcar con una X según corresponda</p>	<p>Enlace para ir a las bases de datos y libros digitales: https://unipanamericana.educo/login.com:2443/login</p> <p>Marcar con una X según corresponda</p>	<p>Enlace para ir a los recursos digitales externos: https://unipanamericanaeduco.sharepoint.com/Portal%20MiPana/SitePages/Informacion%20de%20Interes.aspx</p> <p>Marcar con una X según corresponda</p>
Obligatorios	<i>Trujillo, Juan Carlos. Diseño y explotación de almacenes de datos: conceptos básicos de modelado multidimensional (2013)</i>		https://ebookcentral-proquest-com.unipanamericana.basesdedatosezproxy.com/lib/unipanamericanaasp/detail.action?docID=3214871	
	<i>Peso, Navarro, Emilio del, and González, Miguel Ángel Ramos. La seguridad de los datos de carácter personal (2a. ed.), Ediciones Díaz de Santos, 2002.</i>		https://ebookcentral-proquest-com.unipanamericana.basesdedatosezproxy.com/lib/unipanamericanaasp/detail.action?docID=4795089	
	<i>Escrivá, Gascó, Gema, et al. Seguridad informática, Macmillan Iberia, S.A., 2013.</i>		https://ebookcentral-proquest-com.unipanamericana.basesdedatosezproxy.com/lib/unipanamericanaasp/detail.action?docID=3217398	



	<i>Seguridad informática / Jesús Costas</i>	ISBN: 9789588675701		
Complementarios	<i>ETL Evolution for Real- Time Data Warehousing</i>		https://www.researchgate.net/profile/Theresa_Kraft/publication/280837435_ETL_Evolution_for_Real-Time_Data_Warehousing/links/56008f5308ae07629e52af09.pdf	
	<i>Agile ETL</i>		https://reader.elsevier.com/reader/sd/pii/S2212017313001965?token=7F809B8E9FEDA2393BB84D030FAD18E38D619B98E4757033607E272C8407C143186754C1EB30C7E3AFB79C401B122F70	

Glosario:

- **Big Data**
It is a broad term for very large or complex data sets that traditional data processing applications are not insufficient to handle.
- **Legislation**
it is the law that has been created by a legislature or other government body that allows protecting the data. The term may refer to a single law, or to the collective body of enacted laws, while "statute" is also used to refer to a single law.
- **Extraction Techniques**
Information Extraction (Information Extraction) is a type of information retrieval whose objective is to automatically extract structured or semi-structured information from documents readable by a computer.
- **Data security**
refers to digital privacy protection measures that are applied to prevent unauthorized access to data, which can be found on computers, databases, and websites.
- **Data simulation**

it is the controlled statistical sampling technique, which is used together with a model, to obtain approximate answers to questions that arise in complex problems of a probabilistic type.



compensar

fundación
universitaria

www.ucompensar.edu.co
Bogotá, D.C. - Colombia