USMP

Extensión y Proyección Universitaria

Facultad de Ingeniería y Arquitectura



Curso

Generative Business Intelligence GenBI

Horario: Sábados de 10:15 a 14:45 h.

Duración: 24 horas académicas

Modalidad: Virtual

Certificado: A nombre de la USMP

Inversión:

Precio general: S/. 550.00

Alumnos, egresados y docentes USMP: S/. 450.00

Descripción del curso

En un contexto donde la inteligencia artificial generativa está redefiniendo la forma en que las organizaciones consultan, explican y toman decisiones sobre sus datos, Generative Business Intelligence (GenBI) surge como una disciplina clave para conectar Business Intelligence, arquitectura de datos, gobierno del dato e IA generativa.

El curso propone una formación aplicada para diseñar, construir y evaluar sistemas analíticos conversacionales capaces de responder preguntas de negocio en lenguaje natural, generar consultas SQL controladas, recuperar contexto semántico y mantener trazabilidad sobre las respuestas. El enfoque es open-source first: primero se construyen los componentes base con tecnologías abiertas y luego se muestra cómo industrializar el patrón en AWS como laboratorio cloud final.

La propuesta no se limita a enseñar prompts o uso de copilotos. Busca formar profesionales capaces de comprender la arquitectura detrás de GenBI: capa semántica, Text-to-SQL, RAG sobre metadatos, validación, observabilidad, golden queries y despliegue gobernado en la nube.

Objetivos

- Comprender la evolución del Business Intelligence tradicional hacia GenBI, Agentic Analytics y Decision Intelligence.
- Diseñar una capa semántica de negocio que incluya métricas certificadas, dimensiones, grano, sinónimos, reglas y restricciones analíticas.
- Construir un motor Text-to-SQL empresarial con generación controlada, validación sintáctica, validación semántica y trazabilidad.
- Implementar una capa de conocimiento basada en glosarios, metadatos, documentación y búsqueda vectorial para mejorar el contexto de las respuestas.
- Orquestar flujos GenBI con herramientas modernas como LangChain y LangGraph, incorporando CrewAI a nivel introductorio para comprender patrones multiagente.
- Evaluar la fiabilidad del sistema mediante golden queries, execution match, revisión de respuestas, suites de regresión y manejo de casos donde el sistema no debe responder.
- Desplegar un componente GenBI en AWS como laboratorio final de industrialización cloud, usando servicios como S3, Glue Catalog, Athena, Bedrock, QuickSight y Lake Formation.

Dirigido a

- Estudiantes universitarios de últimos ciclos y estudiantes de postgrado vinculados a tecnología, ingeniería, administración, sistemas, ciencia de datos, analítica o transformación digital.
- Profesionales de Business Intelligence, Data Analytics, Data Engineering, Data Architecture, Data Science, IA o desarrollo de software que buscan construir soluciones analíticas impulsadas por IA generativa.
- Consultores, líderes de datos, responsables de reporting, product owners, analistas de negocio y profesionales que desean entender cómo diseñar plataformas GenBI gobernadas y no solo consumir copilotos existentes.
- Profesionales de áreas de negocio con base analítica que necesiten evaluar oportunidades, riesgos y casos de uso de IA generativa aplicada a datos empresariales.

Metodología

- Enfoque teórico-práctico: explicación conceptual, demostraciones técnicas, laboratorios y construcción incremental de un MVP GenBI.
- Aprendizaje basado en casos: el participante trabajará sobre un escenario empresarial con métricas, datos, reglas de negocio y preguntas analíticas.
- Modelo open-source first: construcción inicial con Python, PostgreSQL, Streamlit, LangChain/LangGraph, Qdrant o ChromaDB y Langfuse.
- Laboratorio cloud final en AWS para mostrar cómo industrializar el patrón usando servicios gestionados.
- Asistencia requerida: 100%, de acuerdo con los lineamientos de la institución.

TABLA DE CONTENIDOS

El curso se organiza en ocho sesiones de tres horas. Cada sesión aporta un componente del MVP final y permite avanzar desde la comprensión conceptual hasta la implementación y evaluación de una solución GenBI.

Módulo 1
Fundamentos de GenBI y evolución del BI (3 h)

- Evolución de BI tradicional, dashboards y self-service analytics hacia GenBI.
- Concepto de analítica conversacional gobernada y diferencia frente a chatbots o prompts aislados.
- Casos de mercado: copilotos analíticos, ThoughtSpot, Power BI Copilot, Tableau Pulse, QuickSight Q y Snowflake Cortex Analyst.
- Riesgos principales: métricas ambiguas, respuestas no trazables, SQL incorrecto y falta de gobierno.
- Entregable: Mapa conceptual de GenBI y análisis de un caso empresarial.

Módulo 2
Arquitectura GenBI y capa semántica (3 h)

- Arquitectura mínima: pregunta de negocio, context builder, semantic layer, LLM, SQL validator, motor de consulta, respuesta y observabilidad.
- Diseño de métricas oficiales, dimensiones, grano, jerarquías, sinónimos y reglas de negocio.
- Contratos semánticos en YAML/JSON y referencias a dbt Semantic Layer y Cube.
- Relación entre semantic layer, gobierno del dato y confiabilidad analítica.
- Entregable: Contrato semántico inicial para un dominio de negocio.

Módulo 3
LLMs aplicados a analítica y respuestas estructuradas (3 h)

- Fundamentos de LLMs: tokens, contexto, embeddings, prompts, límites y alucinaciones.
- Prompt engineering aplicado a datos y generación de respuestas estructuradas.
- Structured outputs y tool/function calling para convertir lenguaje natural en acciones controladas.
- Diseño de prompts con contexto semántico y restricciones empresariales.
- Entregable: Asistente analítico básico con Python y modelo compatible.

Módulo 4
Text-to-SQL empresarial controlado (3 h)

- Diferencia entre SQL válido y SQL correcto para el negocio.
- Allowlists de tablas, columnas, filtros y operaciones permitidas.
- Validación sintáctica, semántica y de seguridad antes de ejecutar consultas.
- Explicación del SQL generado, trazabilidad y manejo de errores.
- Entregable: Motor Text-to-SQL con PostgreSQL, Python y SQL Validator.

Módulo 5
Knowledge layer, RAG y búsqueda vectorial (3 h)

- RAG aplicado a glosarios, documentación, metadatos, reglas y ejemplos de preguntas.
- Uso de Qdrant o ChromaDB para recuperación semántica.
- Context engineering para reducir ambigüedad y mejorar el grounding.
- Trazabilidad del contexto utilizado por el modelo.
- Entregable: Capa de conocimiento con glosario y metadatos recuperables.

Módulo 6
Orquestación y agentes analíticos (3 h)

- LangChain para chains, tools y retrievers.
- LangGraph para flujos con estado, nodos y control empresarial.
- CrewAI como introducción a roles y colaboración entre agentes especializados.
- Diseño de nodos: interpretar, recuperar contexto, generar SQL, validar, ejecutar y explicar.
- Human-in-the-loop y límites de la autonomía en analítica empresarial.
- Entregable: Pipeline GenBI orquestado y trazable con agente de validación.

Módulo 7
Evaluación, fiabilidad y golden queries (3 h)

- Problema del performance cliff: sistemas que funcionan en demos pero fallan en esquemas reales por ambigüedad, joins complejos y reglas implícitas.
- Construcción de golden queries con pregunta de negocio, SQL de referencia y resultado esperado.
- Métricas de evaluación: execution match, revisión de componentes SQL y evaluación asistida de respuestas en lenguaje natural.
- Suites de regresión, detección de deriva entre versiones de prompt/modelo y mejora continua de la capa semántica.
- Diseño del fallback: cuándo el sistema debe decir “no puedo responder con seguridad”.
- Entregable: Eval harness con 15-20 golden queries para la versión intensiva del curso.

Módulo 8
Industrialización en AWS y proyecto final (3 h)

- Mapa de equivalencia entre el stack open-source y servicios AWS.
- S3 como almacenamiento, Glue Catalog como catálogo, Athena como motor SQL, Bedrock como capa GenAI, QuickSight como visualización y Lake Formation como gobierno.
- Consideraciones de seguridad, costos, permisos, trazabilidad y escalabilidad.
- Integración del MVP y presentación ejecutiva del caso de negocio.
- Entregable: MVP GenBI evaluado y presentación ejecutiva.

Docente: MSc. Ing. Jean Carlo Jesus Vallejos Pongo

Ingeniero de Computación y Sistemas por la Universidad de San Martín de Porres. Máster en Big Data por La Salle, Barcelona, España; Máster en Inteligencia Artificial por la Universidad Católica San Antonio de Murcia UCAM; y Executive MBA por ESADE Business School.
Cuenta con sólida experiencia en arquitectura de datos, Business Intelligence, Big Data, analítica avanzada e inteligencia artificial generativa, habiendo participado en iniciativas para sectores como retail, banca, farma, automoción, energía e industria.
Ha liderado iniciativas de transformación analítica, diseño de plataformas Data Lake y Lakehouse, modernización de ecosistemas BI, implementación de arquitecturas cloud y desarrollo de soluciones orientadas a la toma de decisiones basada en datos. Asimismo, ha participado en proyectos vinculados a inteligencia artificial generativa, implementando sistemas inteligentes y capas semánticas a nivel empresarial.
Actualmente se desempeña como Chief AI Architect en NTT DATA España, combinando experiencia técnica, visión estratégica y docencia universitaria en la Universidad de San Martín de Porres, la Universidad de Barcelona y la Universitat Politècnica de Catalunya.
Su enfoque académico y profesional se centra en conectar negocio, datos e inteligencia artificial para diseñar soluciones analíticas confiables, gobernadas y escalables, alineadas con los nuevos retos de Generative BI.

Inicio de clases: 21 de noviembre
Fin de clases: 12 de diciembre

Inscríbete dando clic aquí

Se entregará certificado y material del curso.

*Cronograma y plana docente sujeta a variación.

Informes:

Oficina de Extensión y Proyección Universitaria
epu_fia@usmp.pe
(511) 208 6000 anexo: 1130

Av. La Fontana 1250. Urb. Santa Patricia, La Molina