



---

MNA Maestría en Inteligencia Artificial Aplicada

Materia: TC5035.10 Proyecto Integrador  
(Gpo 10)

Prof. Titulares:

Dra. Grettel Barceló Alonso

Dr. Luis Eduardo Falcón Morales

## **Semana 2. Propuesta de proyecto y firma de convenios**

### **EQUIPO 2**

Matthias Sibrian Illescas - A01794249

José Ramiro Adán Charles - A00174646

Genaro Ramos Higuera - A00351269

## 1. Antecedentes

El Tecnológico de Monterrey es una institución que vela por el cumplimiento de una normativa que rige la vida académica con el objetivo de mantener el orden y de proteger los derechos de comunidad, para lo cual cuenta con área especializada que es la de Legislación Académica, la cual presenta diversas necesidades en el manejo de su información.

## 2. Entendimiento del negocio

- Formulación del problema: ¿Qué es lo que se intenta resolver?
  - Reducir la complejidad y el tiempo que toma determinar los lineamientos vigentes (o más recientes) de cualquier política.
- Contexto: ¿Por qué es importante resolver este problema?
  - Para asegurar la aplicación de las normativas correctas y vigentes y asegurar su eficacia en el mantenimiento del orden y la protección de derechos la comunidad académica.
- Objetivos: ¿Cuál es la meta prevista?
  - Reducir el tiempo de la búsqueda de lineamientos de N minutos promedio.
  - Reducir la cantidad de búsquedas (consultas) de N a 1 o 2 (pregunta al chatbot)
- Preguntas clave: ¿Cuáles son las preguntas clave que deben responderse?
  - ¿Cuál es la legislación vigente sobre esta política?
  - ¿Cuándo cambió este elemento relacionado a esta política?
  - ¿En qué período de tiempo se reflejaron estos cambios en políticas?
  - ¿Quién introdujo estos cambios y cuándo sobre esta política en particular?
  - ¿Cuáles políticas se cambiaron entre 2005 y 2010 por esta persona?
- Involucrados: ¿Quiénes participan: personal / equipos / departamentos? ¿En qué etapa y con qué nivel/tipo?
  - Área de Legislación Académica
  - Gerencia de Inteligencia Artificial
  - Dirección de Hub IA
  - Equipo 2 MNA

## 3. Entendimiento de los datos

- Descripción de los datos: ¿Cuál es el contexto de estos datos?
  - Sharepoint donde se almacenan los documentos (PDFs o Word) donde se almacenan las políticas, es decir, en forma de datos no estructurados y en continua actualización.
  - Cuerpos de texto de correos electrónicos mediante los cuales se notifican cambios y/o modificaciones a las políticas.
- ¿Qué contienen?
  - Los lineamientos de las políticas de cada tema específico.
  - Fechas relevantes como vigencia, creación, expiración, etc.
  - Correos con textos con tags correspondientes a los temas y políticas.
- Técnica de ML: Supervisado (Regresión, Clasificación), No-supervisado, Profundo
  - Este tipo de tarea involucra representar la información contenida en los documentos de una forma en que convenientemente se evidencie las relaciones entre versiones y las

relaciones entre una política y otra en el tiempo. Además, necesita una arquitectura que sea capaz de responder a consultas en lenguaje natural generadas basándose en las representaciones codificadas de las cuáles aprendió. Este tipo de arquitectura necesita aprendizaje profundo: específicamente, una red neuronal de grafos (GNN), una red neuronal recurrente (RNN) o modelos Transformadores (como BERT). La tarea de aprendizaje automático se podría clasificar como Sequence-to-Sequence Learning e Information Retrieval usando Procesamiento de Lenguaje Natural.

- Identificación de las variables: ¿Cuáles son las entradas y la salida?
  - Entradas (información NO CONFIDENCIAL)
    - Corpus de los textos de los documentos (políticas)
    - Corpus de los textos de los correos de notificaciones de políticas nuevas y/o actualizadas.
  - Salidas
    - Texto generado con la información solicitada (tema, lineamieneto política, versión, fecha, referencias, etc y según lo que el área funcional defina).

#### **4. Convenios**

Se incluye en Parte 6: Anexos el convenio firmado por estudiantes y el Dr Juan Arturo Nolzco Flores, Director de Data Science Hub y CoreLab de Data Science del Tec de Monterrey.

#### **5. Bibliografía**

Instituto Tecnológico y de Estudios Superiores de Monterrey. (s.f.). Hub de Inteligencia Artificial. Recuperado de <https://tec.mx/es/ia>.

Jurafsky, D., & Martin, J. H. (2019). Speech and Language Processing (3era Edición). Pearson.

Marcheggiani, D., & Titov, I. (2017). Encoding sentences with graph convolutional networks for semantic role labeling.

#### **6. Anexos (continúa en la siguiente página)**



Tecnológico  
de Monterrey

# COMPROMISOS DE COLABORACIÓN PARA EL PROYECTO INTEGRADOR

**Escuela de Ingeniería y Ciencias**

**Maestría en Inteligencia Artificial Aplicada**

# MNA

## GENERALES

El *Proyecto Integrador*, corresponde al proyecto final que se realiza del programa de la Maestría en Inteligencia Artificial Aplicada. Éste se desarrollará y evaluará en la materia con clave TC5035, del mismo nombre.

El objetivo principal del proyecto es resolver una problemática real, vigente, sea del sector industrial o de instituciones gubernamentales, ONG u algún otro tipo de institución. Lo anterior, previa autorización del comité.

**Para efectos del presente documento el programa de Maestría en Inteligencia Artificial Aplicada (en adelante MNA) y la Dirección de Ingeniería de Software, Vicepresidencia de Transformación Digital, del Tecnológico de Monterrey, (en adelante el cliente) reconocen los siguientes compromisos:**

## COMPROMISOS

Para realizar el proyecto integrador la MNA se compromete a lo siguiente:

1. Asignar un asesor de proyecto con experiencia y vinculación en la problemática presentada por el cliente.
2. Ofrecer las instalaciones y equipo existente requerido que contribuya a resolver la problemática identificada.
3. Proporcionar retroalimentación constante y sistemática durante el registro de la propuesta y desarrollo del proyecto en la materia TC5035.
4. Conformar el comité de evaluación una vez que se hayan alcanzado los objetivos presentados en la propuesta del Proyecto Integrador.

El cliente se compromete a:

1. Asignar a un patrocinador con reconocidas fortalezas en el área de dominio a atender por el proyecto.
2. Retroalimentar a(los) integrante(s) del proyecto con relación a los resultados parciales obtenidos.
3. Asegurar los recursos de información requeridos para la realización del proyecto.
4. Reconocer la participación del estudiante o los estudiantes, así como al Tecnológico de Monterrey, en las publicaciones o reportes que se generen del proyecto.


Ambas instituciones se comprometen a:


1. Guardar la confidencialidad con relación a los resultados obtenidos que sean sensibles/estratégicos para cualquiera de las instituciones.
2. Interacción bidireccional en reuniones de trabajo para optimizar el proceso de implementación del Proyecto Integrador.

La colaboración se realizará, (sin limitarse a ello) para la implementación del Proyecto de INFORMATION RETRIEVAL USING GENERATIVE MODELS a partir del 15 de Abril del 2024 y concluyendo tentativamente el 21 de Junio del 2024:

Alumno(s) de Proyecto Integrador

  
Matthias Sibrian Illescas  
Matricula A01794249

  
José Ramiro Adán Charles  
Matricula A00174646

  
Genaro Ramos Higuera  
Matricula A00351269

Nombre, matrícula y firma

Por la MNA

Dr. Luis Eduardo Falcón. Director Nacional de la Maestría en Inteligencia Artificial Aplicada

Nombre, puesto y Firma

Patrocinador asignado por el cliente

Ing. Manuel Terán Melgarejo  
Gerente de Plataformas de Inteligencia Artificial

  
Dr. Juan Arturo Nolasco Flores  
Data Science Hub & CoreLab Data Science

Nombre, puesto y Firma