



Inteligencia artificial avanzada para la ciencia de datos II

TC3007C.501

## **Reto Metodología**

### **Integrantes**

Sebastian Rodriguez Salinas (A00827463)

Alan Mondragón Rivas (A01734565)

Elmer Osiel Ávila Vargas (A00826359)

Carlos E. Lucio Domínguez (A00828524)

Diego Solis Higuera (A00827847)

4 de octubre de 2022

## Requerimientos One Page

---

### Problema

Actualmente, el acceso a la información y datos del INEGI puede ser complicado para los usuarios: es difícil encontrar las bases de datos, algunos links no funcionan y en ocasiones, la página web se vuelve muy lenta. Se ha intentado mejorar el acceso de la información con un rediseño de la UI, pero el flujo de navegación y desempeño de la página se mantienen igual, perjudicando la experiencia de usuario y dificultando el acceso a la información.

### Solución

El proyecto consiste en desarrollar una aplicación web que aproveche y despliegue información de manera eficiente sobre la ubicación de negocios cercanos a la ubicación del usuario, por medio de comandos de voz, o bien, utilizando un chatbot. Esto permitirá al usuario encontrar comercios de distintos giros y tamaños, y potencialmente encontrar los comercios que ofrezcan los productos o servicios que desee el usuario.

RF. Requerimientos funcionales	
RF1	El sistema desplegará información en gráficas utilizando bases de datos del INEGI.
RF2	El sistema permitirá al usuario realizar búsquedas acerca de negocios cercanos a partir de su ubicación.
RF3	El sistema desplegará el tamaño de los establecimientos.
RF4	El sistema permitirá al usuario interactuar con un chatbot mediante texto para realizar peticiones de búsquedas.

RNF. Requerimientos no funcionales	
RNF1	El sistema utilizará una API para conectarse al modelo de NLP.
RNF2	El sistema se conectará a una o múltiples bases de datos para obtener información.
RNF3	El sistema utilizará una API de mapa para desplegar información.
RNF4	El despliegue del sistema será en una aplicación web.

RD. Requerimientos diferidos	
RD1	El sistema permitirá al usuario realizar las solicitudes de búsqueda en función de los productos y servicios que ofrecen los negocios.

## **Memorandum of Understanding**

---

This Memorandum of Understanding (MOU) is entered on the 5 day of October in 2022 (herein referred to as the "Effective Date") into by and between: “Equipo 4” of “Tecnológico de Monterrey” from the course “Inteligencia artificial avanzada para la ciencia de datos II (Gpo 501)” (herein referred to as "Partner A") and Dr. Abel Coronado of “INEGI” (herein referred to as "Partner B"). Parties may be referenced individually as "Party" and collectively as "Parties".

### **A. Purpose**

The purpose of this MOU is to establish a good-faith foundation between the Parties for future collaborative efforts that are mutually beneficial. The Parties agree to work together in a cooperative and coordinated manner to achieve each Party's individual desires and the collective desires of the partnership.

This MOU is designed to detail the specifics of the working relationship between the Parties to vet the success of the potential business collaboration. This MOU does not obligate the Parties to provide funds or payment. This MOU does not bind Parties to any legal obligations.

### **B. Roles and responsibilities**

To achieve Parties' mutual desires, each party agrees to the following roles and responsibilities.

Partner A Roles and Responsibilities shall include the development of a front-end interface meant to visualize the results achieved from the project created. Also the development of an algorithmic model meant to predict specific criterias based on the INEGI database. As well as the implementation of an “neutral language processor created for the accessibility purpose of the course.

Partner B Roles and Responsibilities shall include allowing the access of partner A to the INEGI database out of which the project will take part from.

Parties agree to uphold their roles and responsibilities in a committed, good-faith manner.

### **C. Resources**

To further the business relationship between the Parties, the Parties agree to provide the following resources.

Party A shall provide a final resulting product contemplating the NLP, prediction model, and interface.

Party B shall provide the database out of which the final product produced by the partner will create the prediction model.

Parties agree to provide the resources above at a minimum. The Parties may agree to provide additional resources in future agreements.

#### **D. Confidentiality**

Parties agree that they each use confidential, sensitive information to achieve their individual business goals. Due to the nature of the Parties' businesses and intent to establish a working relationship together, confidential information may be shared between the Parties.

The Parties agree to keep all confidential information and trade secrets in the strictest of confidence during the working relationship. Parties may not share any disclosed confidential information with unauthorized third parties. Parties may, at their discretion, enter into a Non-Disclosure Agreement to ensure sensitive information and trade secrets are kept confidential.

#### **E. Entire agreement**

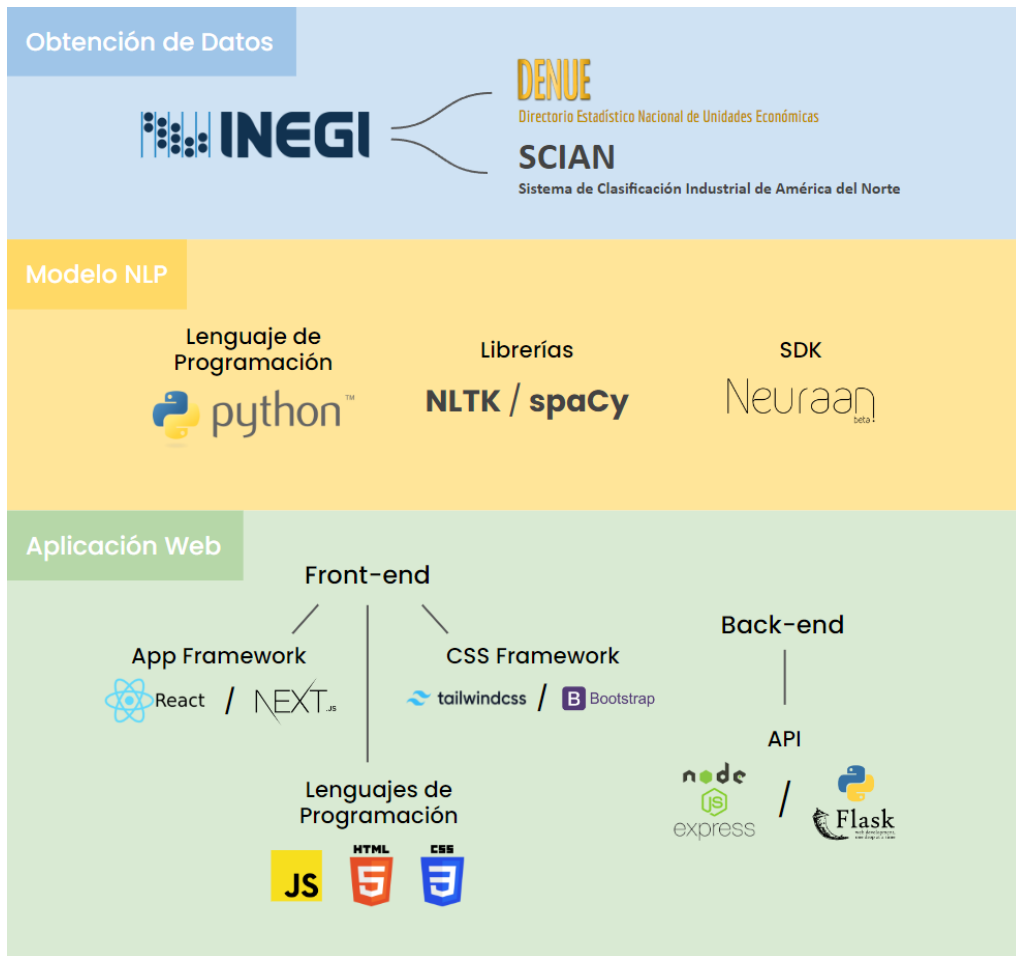
The Parties agree this MOU represents the most current agreement between the Parties and supersedes all other written or oral agreements. If Parties wish to update the terms or otherwise adjust provisions of this MOU, Parties shall do so by the drafting and signing of a new MOU or partnership contract.

## Mapeo de Recursos y Herramientas

---

### Diagrama de Recursos y Herramientas del Proyecto

En el siguiente diagrama se muestran los recursos, herramientas y tecnologías que serán utilizados para llevar a cabo nuestra solución.



Como es posible notar, los datos sobre los negocios y establecimientos serán proveídos por el INEGI mediante las bases de datos DENUE y SCIAN. Para la programación del modelo NLP se utilizará Python como lenguaje de programación, NLTK y spaCy como librerías de procesamiento de lenguaje natural y el SDK de Neuraan. Finalmente, para la solución web, en la parte de front-end tenemos como opciones de frameworks utilizar React o NextJS y para CSS nos apoyaremos de recursos como Tailwind y Bootstrap, mientras que en back-end, planeamos realizar una API utilizando NodeJS o Flask que facilite la comunicación entre la aplicación web y el modelo NLP.

## Plan Inicial Basado en la Metodología

### Metodología Seleccionada: CRISP-DM

CRISP-DM resulta ser una metodología adecuada y comúnmente utilizada para los proyectos de ciencia de datos. Se adapta de forma excelente ante las necesidades del proyecto que se busca realizar pues permite flexibilidad en el calendario y dar un enfoque ágil al desarrollo del mismo ya que no presenta una secuencia de fases estricta, se puede avanzar y retroceder entre ellas, así como realizar ciclos en caso de ser necesario. CRISP-DM consta de 6 fases mismas que se muestran en el siguiente diagrama:



Diagrama. Metodología CRISP-DM.

### CRISP-DM enfocado al reto:

- *Entendimiento del Negocio*: Comprender las necesidades del INEGI, entender a fondo cuál es el reto al que se busca atacar y a partir de ello definir una solución. Una vez realizado esto, se estructurará una definición de requerimientos para el proyecto/solución, la metodología y una planeación de desarrollo para el mismo.
- *Entendimiento de los Datos*: Entendimiento y visualización inicial de los datos provistos por el INEGI y con ello definir cuales son esenciales para el modelo así como entender las necesidades para los siguientes pasos, es decir, comprender cuál será la preparación necesaria para los datos.

Liga al cronograma: [+ Diagrama de Gantt](#)