Exploración de Problemas para el Proyecto

# Introducción

Este documento tiene como propósito identificar posibles problemas que el equipo puede abordar en el proyecto, formulando preguntas de negocio claras y asegurando la disponibilidad de los datos necesarios. Esta exploración inicial permite minimizar riesgos asociados a la falta o calidad de los datos, facilitando el desarrollo de soluciones viables.

A continuación, se presentan problemáticas potenciales junto con sus preguntas de negocio y fuentes de datos, seguido de una evaluación de su disponibilidad.

# Problemas Potenciales y Preguntas de Negocio

**Problema 1: Clasificación inicial del origen de siniestros de riesgos laborales usando NLP**

* **Pregunta de negocio:** ¿Es posible automatizar la clasificación inicial de siniestros como de origen laboral o común utilizando procesamiento de lenguaje natural (NLP) para reducir la carga operativa y mejorar los tiempos de respuesta?
* **Conjunto de datos asociado:** Formatos de reporte de siniestros (FURAT y FURAL), datos del trabajador, información de la empresa afiliada y detalles del siniestro.

3. Evaluación de Disponibilidad de Datos

**Problema 1:**

* ¿Los datos están estructurados o no estructurados?

Los formatos FURAT y FURAL contienen datos semiestructurados en formularios, con texto libre y campos específicos.

* ¿Se requiere procesamiento o limpieza antes del análisis?

Sí, es necesario preprocesar los textos, eliminar ruido, normalizar términos y estructurar la información para su uso en modelos NLP.

* ¿Existen restricciones de acceso o privacidad en los datos?

Sí, los datos contienen información sensible de trabajadores y empresas, por lo que se deben aplicar medidas de anonimización y cumplimiento de regulaciones de privacidad.

# 4. Conclusión

Este problema es viable para abordar con NLP, dado que existen datos disponibles y en volúmenes suficientes para desarrollar un modelo de clasificación. Sin embargo, se requiere un procesamiento previo de los textos y garantizar el cumplimiento de normativas de privacidad antes de su implementación.