



---

**DOCUMENTO PRESENTADO:**

**BUENAS PRÁCTICAS PARA EL DESARROLLO E IMPLEMENTACIÓN DE SOLUCIONES DE INTELIGENCIA DE NEGOCIOS PARA APOYAR LOS PROCESOS DE TOMA DE DECISIONES Y MEJORAR LA EFICIENTE Y EFICACIA DE LOS PROCESOS DE LAS ORGANIZACIONES EN EL SECTOR JUSTICIA**

**PREÁMBULO:**

Teniendo presente:

Que uno de los objetivos del proyecto “Aplicación de innovaciones tecnológicas en los procesos judiciales, para apoyar el acceso a la justicia por parte de los ciudadanos en los países Iberoamericanos” es “Conocer y posteriormente difundir cómo incide el uso de herramientas innovadoras, como Big Data, la inteligencia artificial y la inteligencia de negocios. Este levantamiento de información formará parte de la evaluación de las mejores prácticas y experiencias en la implementación que serán publicadas en el Portal.”

Que el propósito es desarrollar una guía práctica que se integre de forma transparente a cualquier ambiente sin requerir una adaptación especial, ya sea de software, hardware, procesos o tareas, para los países interesados en desarrollar o fortalecer dentro de su inventario, soluciones para apoyar los procesos de toma de decisiones y mejorar la eficiente y eficacia de los procesos de las organizaciones en el sector justicia.

**En consecuencia:**

Impulsamos a los Poderes Judiciales a implementar o reforzar mediante estas buenas prácticas el desarrollo de soluciones de Inteligencia de Negocios en los procesos institucionales para apoyar los procesos de toma de decisiones y mejorar la eficiente y eficacia de los procesos de las organizaciones en el sector justicia

**En consideración:**

A continuación, se describen las mejores prácticas de los siguientes proyectos tecnológicos:

- I. Inteligencia de Negocios
- II. Inteligencia Artificial
- III. Expediente Digital

**INTELIGENCIA DE NEGOCIOS**

**CAPÍTULO PRIMERO: ASPECTOS GENERALES**

**1. Objetivos Pretendidos**

El objetivo principal de estas buenas prácticas es fortalecer los procesos para implementar soluciones de Inteligencia de Negocios, con base en la experiencia adquirida en este tipo de proyectos.

Los objetivos específicos son:

- a) Divulgar la noción sobre Inteligencia de Negocios.
- b) Proponer una estructura mínima de los instrumentos para la implementación de soluciones de Inteligencia de Negocio.

**2. Noción de Inteligencia de Negocios:**



---

La Inteligencia de Negocios es la implementación de tecnologías, aplicaciones, técnicas y prácticas para recolectar, unificar, analizar datos, con el propósito de generar conocimiento para responder preguntas de diferentes tipos, estos son:

- **Descriptivas:** Cuanto invierte el Poder Judicial en vehículos, en alquiler de edificios, en proyectos tecnológicos.
- **Asociativas:** ¿El estrés térmico de las oficinas impacta el rendimiento de las personas? ¿Decremente la calidad del servicio al usuario?
- **Comparativas:** ¿Dos juzgados con la misma cantidad de personal resuelve diferente cantidad de casos?
- **Predictivas:** ¿Cuántos casos llegaremos a resolver durante el 2022 si mantenemos las mismas condiciones?

### 3. Marco de trabajo en soluciones de BI:

Hoy en día no podemos esperar hasta que el usuario identifique y solicite una solución tecnológica, hoy debemos ir un paso adelante y llegar al usuario con una herramienta capaz de maximizar el uso de la información para potenciar el rendimiento, para esto se debe establecer un marco de trabajo claro y ágil. Este marco de trabajo se compone de tres partes fundamentales de preparación:

- **Identificar las oportunidades de negocio:** Este paso consiste básicamente en identificar procesos o tareas donde se puede explotar la Inteligencia de Negocios y el Big Data. La innovación es un factor clave para el éxito de estas iniciativas, porque si en el presente se sigue haciendo todo de la misma forma, no tendremos un futuro prometedor. Para esta exploración de oportunidades de negocio se recomienda ir a las personas que trabajan en el proceso, preguntar que hacen, como lo hacen y como se podría mejorar.
- **Definir forma de trabajo:** Cada oportunidad tiene características únicas, por esta razón es indispensable definir la forma que se utilizará para aprovechar cada oportunidad identificada, aquí se involucran las técnicas y procesos para garantizar la calidad de la información, la velocidad, la seguridad y el rendimiento requerido.
- **Crear el plan de trabajo:** En este paso se definen las necesidades de hardware, software y recurso humano para materializar la solución.

Con este marco de trabajo se mitigan los posibles riesgos que generarían la implementación del uso de soluciones de Inteligencia de Negocios, por ejemplo, la carga en la red interna por el trasiego de datos provocando denegación de servicios, el acceso inapropiado a la información genera o la creación tardía de información.

### 4. Roles del equipo de trabajo para soluciones de BI:

Para por orientar la organización hacia una cultura de datos donde se pueda actuar a partir del conocimiento construido con las herramientas de Inteligencia de la Información, se debe contar con un equipo de trabajo con diferentes habilidades y roles, algunos de estos son:

Para el negocio:

1. **Analista de negocios:** Normalmente son los intermediarios entre las partes interesadas del negocio (como los usuarios o jefes) y el departamento de tecnología (como los desarrolladores de soluciones de BI). Su función es mejorar los procesos de negocios detallando y diseñando los requerimientos para desarrollar funciones nuevas o mejoras a las existentes.
2. **Analista de datos:** Se centran en la recopilación y preparación de datos, ayudan a los usuarios en la creación de informes para el análisis de la información generada.

Para el desarrollo:



1. **Ingenieros de datos:** Son responsables de obtener, limpiar y transformar datos para que los analistas puedan accederlos y analizarlos. Son responsables de monitorear y garantizar el rendimiento, la escala, las cargas máximas y el registro de los procesos de cargas de datos.
2. **Científicos de datos:** También realizan las tareas de los ingenieros de datos, con la diferencia que pueden utilizar procesos de Machine Learning para transformar los datos, también pueden ser responsables de crear herramientas de Inteligencia de Negocios que los analistas utilizan.
3. **Especialista en visualizaciones:** Se encargan de diseñar las opciones para mostrar los convertidos en conocimientos, este rol se convierte en el más estratégico, porque el grado de aceptación de una solución de Inteligencia de Negocios está estrechamente ligada a lo atractiva y clara que esta sea. Los especialistas en visualizaciones cuentan con estrategias para diseñar informes y tableros dinámicos con los colores y elementos óptimos según las personas o procesos que utilizarán las soluciones.

Para la implementación:

1. **Administrador de la plataforma de Inteligencia de Negocios:** Es el encargado de dar soporte y mantenimiento a la plataforma, es un intermediario entre el equipo de Inteligencia de Negocios y el área de Infraestructura, algunas de sus funciones son:
  - a. Actualización del software de la plataforma.
  - b. Gestiona la solución los problemas físicos de la plataforma.
  - c. Asigna permisos para el acceso a las soluciones de Inteligencia de negocios.
  - d. Trabaja en conjunto con el equipo de Inteligencia de Negocios para solucionar problemas de rendimiento.
  - e. Instala nuevas soluciones o actualizaciones en los ambientes de producción
  - f. Respaldos de bases de datos, script y aplicativos.

Es importante aclarar que la cantidad de personas y combinación de roles por personas dependen del tamaño, madurez y dominio de la organización.

## 5. Calidad de los datos

Para las soluciones de Inteligencia de Negocios, los datos son un insumo fundamental, pero el concepto de calidad de los datos depende mucho del contexto en el que se está trabajando, por ejemplo una dirección puede componerse de tres datos país, provincia y distrito; si en esta información el dato del distrito está incompleto, los analistas de datos pueden hacer análisis a nivel de país y provincia, considerándose esta información de calidad, pero si se quiere hacer un análisis por distrito, la información no tiene una calidad aceptable. Por esta razón se define que para que la información sea de calidad debe contar con las siguientes características:

1. **Accesible:** Esta característica involucra tener acceso a los datos y también a la interpretación de estos, por ejemplo, el acceso a un archivo donde hay una serie de datos, pero no conozco exactamente qué valores se guardan en cada columna, es un archivo inaccesible.
2. **Precisa:** Los valores deben ser correctos. Por ejemplo, un dato calculado incorrectamente, una fecha mal escrita o una dirección desactualizada son datos inexactos.
3. **Consistente:** Esto significa que los datos pueden mezclarse con otros datos importantes de manera precisa, por ejemplo, combinar la información de los casos terminados de un juez por mes con el salario de ese mismo periodo.
4. **Completa:** No deben existir datos vacíos o con formatos incorrectos
5. **Clara:** Cada columna o fila debe estar bien definida sin interpretaciones ambiguas, por ejemplo, número de casos resulta ser un campo ambiguo, lo correcto sería número de casos terminados o número de casos terminados en materia penal.
6. **Relevante:** Los datos deben tener una clara relación con la solución de Inteligencia de Negocios.
7. **Oportuna:** La información debe estar disponible y procesada en el momento que se necesite.

Es importante considerar que los datos deben cumplir con estas características sin importar su origen. Hoy en días a información puede venir desde fuentes tradicionales como bases de datos o archivos estructurados, pero también pueden obtenerse datos, desde equipos especializados, sensores, redes sociales, audios, videos, imágenes y demás.



## 6. Regulaciones nacionales e internaciones para el manejo de los datos

Existen lineamientos globales y locales para el tratamiento de la información previo a su publicación en Internet o por cualquier medio con acceso general, por esta razón es indispensable conocer estos lineamientos a detalle. Algunos ejemplos globales son:

1. **Estándar ISO 27001:** Para los Sistemas de Gestión de la Seguridad de la Información, permite el aseguramiento, la confidencialidad e integridad de los datos, así como de los sistemas que la procesan.
2. **GDPR:** Es el Reglamento General de Protección de Datos vigente desde el año 2018, que tiene como objetivo reforzar el control y la protección de datos personales de todos los ciudadanos y residentes de la Unión Europea

Algunas regulaciones locales son:

- 1) **Costa Rica:** Ley de Protección de la Persona frente al Tratamiento de sus Datos Personales (N° 8968)
- 2) **Portugal:** Ley N° 67/98 del 26 de octubre - Datos Personales
- 3) **Uruguay:** Ley de Protección de Datos Personales y Acción de Habeas Data (N°18331)
- 4) **Puerto Rico:** Ley de Información al Ciudadano sobre la Seguridad de Bancos de Información” (“Ley Núm. 11-2005”)
- 5) **Panamá:** Protección de Datos Personales - Proyecto de Ley N° 665
- 6) **Nicaragua:** Ley de Protección de Datos Personales (N° 787)
- 7) **Paraguay:** LEY N° 1682 – Información de carácter privado
- 8) **Mexico:** Ley Federal de Protección de Datos Personales en Posesión de los Particulares (LFPDPPP)
- 9) **El Salvador:** Ley de Acceso a la Información Pública - Decreto N° 534
- 10) **Ecuador:** Ley del Sistema Nacional De Registro de Datos Públicos (N° 162)
- 11) **Colombia:** Ley de Protección de Datos Personales (N° 1581)
- 12) **Chile:** Ley sobre Protección de la Vida Privada (N° 19628)
- 13) **Argentina:** Ley de Protección de Datos Personales (N° 25326)
- 14) **España:** Real Decreto Ley 5/2018, adaptación del GDPR

## CAPÍTULO SEGUNDO: INVESTIGACIONES DE INTELIGENCIA DE NEGOCIOS

### 7. Pruebas de Concepto por medio de Empresas Externas

Una prueba de concepto es una implementación reducida o incompleta de una oportunidad de negocio, se ejecuta con el fin de asegurar la usabilidad y rentabilidad de la tecnología propuesta para aprovechar la oportunidad de negocio.

Estas pruebas de conceptos generalmente no devengan un costo para la organización, porque su objetivo es demostrar que la solución puede subsanar la necesidad expuesta. A continuación, se detalla un formato básico de los datos y resultados esperados en una prueba de concepto:

Datos brindados por parte de la organización:

- a. **Descripción de la oportunidad de negocio:** La organización debe describir la situación que se pretende mejorar, detallando los procesos y roles que intervienen en el mismo. Además, debe brevemente describir el resultado esperado, esto permite que la empresa interesada tenga una idea de la situación sin limitar la capacidad creativa de la misma.
- b. **Insumos necesarios para trabajar la POC:** La organización debe brindar los datos necesarios para trabajar la prueba de concepto, esto es indispensable, porque en muchas ocasiones las empresas muestran soluciones con datos óptimos que garantizan el correcto funcionamiento de la demostración, sin embargo la realidad es que la información de cada Poder Judicial es única, porque se registra y almacena de diferentes formas, tal es el caso de las pruebas de concepto para transcripciones de audio a texto, empresas que garantizan



---

más del 90% de exactitud en la transcripción y cuando se utilizan los audios guardados de la organización solamente se cuenta con un 35% de texto extraído del audio, debido a formatos de los archivos o calidad de los micrófonos utilizados. Además, según el grado de criticidad de la información se debe firmar un acuerdo de confidencialidad que garantice un correcto uso de los datos brindados.

Datos brindados por parte de la empresa:

- a. **Descripción de la arquitectura propuesta:** La empresa debe entregar un documento que detalle la arquitectura de la solución, detallando todo el hardware requerido, por ejemplo, para los servidores, deben detallar procesadores, memoria RAM, espacio para almacenamiento y todas las características relevantes que facilite la dimensión en equipo requerido para la solución propuesta.
- b. **Propuesta de ejecución del proyecto:** Se debe detallar cada etapa estimada para la implementación total de la solución, también se debe especificar el tiempo estimado para cada tarea e indicar claramente que contempla la solución, por ejemplo, activación del servicio, desarrollo de la interfaz de usuario, capacitación, etcétera.
- c. **Costo:** La empresa debe puntualizar en cada uno de los posibles costos, con el propósito de materializar completamente la propuesta económica y valorar el costo-beneficio de implementar la solución. Algunos ejemplos de los posibles costos son:
  - I. Costo de la solución propuesta: Incluye el costo de cada tarea de la propuesta de ejecución del proyecto.
  - II. Costo de almacenamiento o procesamiento: En caso de aplicar se debe detallar el costo adicional de espacio o procesamiento, explicando la forma y periodicidad de pago.
  - III. Costo de soporte.
  - IV. Costo de servicios y componentes adicionales: En la Inteligencia de Negocios las diferentes soluciones en ocasiones utilizan servicios cognitivos o similares para limpiar o procesar la información, estos servicios deben detallar el costo.
  - V. Costo de licencias
- d. **Proyectos semejantes:** Si la empresa tiene implementada y con acceso irrestricto alguna solución similar, puede facilitar el acceso con el propósito de valorar la misma.

## 8. Pruebas de Concepto a lo internos

Las pruebas de concepto también pueden realizarse a lo interno, para realizar este proceso se recomienda:

- a. **Identificar las personas adecuadas:** Las pruebas de concepto generalmente están asociadas a la investigación y adaptación de alguna tecnología emergente, por esta razón las personas que trabajen en soluciones innovadoras para Inteligencia de Negocios deben tener pericia para identificar la opción que genere más valor a la organización por un costo razonable.
- b. **Identificar la oportunidad de negocio:** Se debe tener clara la situación que se espera mejorar con la implementación de la solución de Inteligencia de Negocios.
- c. **Identificar las posibles tecnologías a investigar para la oportunidad de negocio:** En el mercado actual existen muchas opciones de tecnologías para desarrollar soluciones en Inteligencia de Negocios, para seleccionarlas se recomienda valorar los siguientes criterios:
  - I. **Documentación:** Es indispensable elegir tecnologías que cuenten con mucha documentación, con el propósito de adaptar estratégicamente esta tecnología considerando los casos especiales de casos Poder Judicial.
  - II. **Estabilidad:** El objetivo de hacer pruebas de concepto es crear valor para la institución a través de una solución de Inteligencia de Negocios en un corto periodo de tiempo, por esta razón es indispensable utilizar tecnologías emergentes estables, para evitar la implementación de soluciones que podrían fallar en ambientes de producción a causa de parches para estabilizarlas, generando retrabajo.
- d. **Plan piloto:** Se recomienda una vez investigada y desarrollada la prueba de concepto, se recomienda ejecutar un pequeño plan piloto, para valorar el rendimiento de la solución y capturar de los usuarios expertos la mayor cantidad de realimentación para garantizar el éxito de la solución



- e. **Desarrollo formal e implementación:** Una vez finalizado el proceso del plan piloto se formaliza un proyecto para adaptar la prueba de concepto a la totalidad de la oportunidad de negocio analizada.

Otra consideración para la investigación y desarrollo de pruebas de conceptos desarrolladas a lo interno es:

1. **Porcentaje de POCs implementadas:** El desarrollo de pruebas de conceptos es una inversión, porque se requiere tiempo y personal calificado para desarrollarlas, por otra parte, la innovación es cara porque significa ir a la cabeza en el uso de tecnologías nuevas, donde se va a dedicar un tiempo considerable para solucionar todos los inconvenientes que pueden toparse en el camino. Por ejemplo, en el caso de Costa Rica con dos años de desarrollar pruebas de concepto solamente se han implementado el 35% de las mismas.

## **CAPÍTULO TERCERO: DESARROLLO DE SOLUCIONES DE INTELIGENCIA DE NEGOCIOS**

### **9. Definir los requerimientos de la solución de BI**

Deben hacerse entrevistas con los usuarios expertos del negocio para determinar los requerimientos específicos de información, según las tareas cotidianas de cada rol de trabajo. Es importante concentrarse en lo que los usuarios hacen (su proceso de negocio) y ayudarlos a establecer los requerimientos específicos de datos.

### **10. Establecer una metodología y herramientas de desarrollo**

En este punto es importante tener claro que el desarrollo de una solución de Inteligencia de Negocios no debe verse como una solución de desarrollo de software en general, por lo que no todas las metodologías de desarrollo podrían adaptarse a una solución de este tipo. Por ejemplo: Aunque sería factible el desarrollo de una solución de BI usando SCRUM, otras metodologías Iterativo-Incrementales se adaptan de mejor manera a una solución de BI en la que, debe tenerse un modelo dimensional previamente modelado e implementado para para la planificación y creación de un proyecto de extracción, transformación y carga (ETL por sus siglas en ingles) y la capa de presentación al usuario. Debe además seleccionarle la herramienta de desarrollo con la cual se crearán los procesos de ETL, reportes, tableros dinámicos (dashboards), Cubos de procesamiento analítico en línea (OLAP por sus siglas en ingles) y demás entregables que se definan como parte de la solución de BI. En el mercado se ofrece una amplia gama de paquetes para el desarrollo e implementación de soluciones de inteligencia de negocios, debe hacerse un estudio de mercado que permita seleccionar la marca comercial más conveniente, acorde a los planes futuros y presupuesto de la organización.

### **11. Definir arquitectura y estándares para el ambiente de desarrollo**

Aunque todas las capas de una solución de BI pueden convivir en un mismo servidor durante la etapa de desarrollo, es recomendable en la medida de lo posible tener servidores separados para cada una de las capas (DW, procesos de ETL y capa de presentación (Front-End)), para poder hacer pruebas de rendimiento de manera más acertada. Deben también evaluarse ambientes híbridos en la nube (Cloud) o instaladas en la organización (On Premises), así como ambientes complemente On Cloud, según lo permitan los recursos de la organización.

Es recomendable tener un ambiente desarrollo separado y con todas las capas de una solución de BI (Base de datos, ETL y Front-End) y nunca implementar mejoras sobre un ambiente productivo, puesto que podrían ocurrir errores que afecten su credibilidad. En cuanto a las bases de datos transaccionales que alimentan a la solución de BI, es recomendable trabajar con datos reales (copias de bases de datos de producción), puesto que el entendimiento de muchas las reglas de negocio que quedarán plasmadas en el modelo dimensional y en los procesos de ETL, se desprenden propiamente de los datos.

### **12. Definir un cronograma de trabajo**



---

Debe crearse un cronograma de trabajo según la metodología de desarrollo seleccionada y acorde a los recursos asignados al desarrollo al proyecto. Es importante que en el cronograma se establezcan tiempos para la validación constante por parte de los usuarios de la solución.

### **13. Diseño del lógico del modelo dimensional**

Esta es una de las partes más importantes en el desarrollo de una solución de Inteligencia de Negocios. Debe seleccionarse una metodología de desarrollo estándar y que se ajuste a la tecnología seleccionada. Kimball Group (por citar un ejemplo) ha establecido una metodología y recomendaciones de diseño para modelos dimensionales (dimensiones y tablas de hechos). Un mal diseño del modelo dimensional puede hacer que se apliquen mal o se ignoren por completo las reglas de negocio. Un buen diseño dimensional hace más fácil adaptar el modelo la mayoría de las tecnologías de BI en el mercado (Cubos OLAP, por citar un ejemplo), aunque hay que dejar claro que el modelo dimensional debería crearse de manera agnóstica, orientado siempre al usuario y sin pensar específicamente en las capas del front-end que va a soportar.

Es importante tener un estándar de nombres para cada tipo de tabla del modelo dimensional, por ejemplo: que todos los nombres de las tablas de dimensión inician por el nombre “Dim” y todas las tablas de hechos inician por “Fact”. Además, algunos campos en las tablas de hechos y dimensión deberían tener nombres estándar. Por ejemplo, las llaves surrogadas de las dimensiones podrían comenzar por “Sk”.

### **14. Diseño físico del modelo dimensional:**

Para el diseño físico de la base de datos que alojará el modelo dimensional creado, existe una serie de recomendaciones generales que se detallan a continuación:

- a. Nombrar los objetos en base de datos (tablas, campos, vistas) con nombres estándar que hayan sido definidos previamente. Es recomendable que la institución desarrolle un manual para nombrar objetos de bases de datos cuando se trate de proyectos de este tipo. Ejemplo: Que todas las dimensiones inician por el acrónimo “Dim”, todas las tablas de hechos por el acrónimo “Fact” y que los nombres de las columnas sean representativos de su funcionalidad, por ejemplo, las llaves surrogadas pueden iniciar por el acrónimo “Sk”.
- b. Debe asegurarse que las columnas tengan los tipos de datos correctos. Aunque una tarea previa de “data profiling” puede darnos una idea de tipos y tamaños de las columnas, debemos asegurarnos que vayan a soportar los datos actuales y los futuros.
- c. Debe especificarse como se va a manejar el cambio en cada uno de los elementos en las dimensiones. Por ejemplo: si un “empleado” cambia de “oficina”, debo crear un nuevo registro de empleado en la nueva oficina asignada, o ¿solamente actualizo el campo oficina para este empleado?
- d. Deben identificarse las jerarquías naturales en los elementos de las dimensiones, por ejemplo: Año, Semestre, Trimestre,
- e. Identificar las llaves foráneas que de las tablas de hechos contra las dimensiones y entre dimensiones en el caso de un modelo del tipo “copo de nieve”.
- f. Debe crearse un plan inicial de índices, para las dimensiones y las tablas de hechos. Es importante la actualización de índices y valorar la inclusión de nuevos índices o eliminación de algunos previamente creados, posterior a la carga inicial del modelo.
- g. Desarrollar un plan de particionamiento de tablas de hechos, que puede basarse en el tiempo como factor de partición. Las tablas de hecho son por lo general, las que contienen más registros y una partición física de dichas tablas, ayudará al tener un mejor rendimiento en las consultas de usuario.
- h. Control de nulos: Permitir atributos nulos en las dimensiones podría ser confuso para los usuarios, máxime cuando se quiera filtrar por un atributo que podría estar en nulo o agrupar por este. Por lo tanto, debe considerarse reemplazar los nulos por un texto o número que represente el faltante de dicho dato, por ejemplo “No encontrado”, “No aplica”, 0, etc.

### **15. Desarrollo de los Procesos de ETL (Extracción Transformación y Carga)**

Una vez seleccionada la tecnología con la cual se desarrollarán los procesos de ETL (extracción, transformación y carga), debe diseñarse la solución, pensando siempre en el mayor desempeño de los procesos de carga y la menor



---

afectación los Sistemas Transaccionales que serán sustento de la solución de BI. Para lo anterior, se recomiendan algunos lineamientos generales que ayudarán a crear y documentar de la mejor manera, los procesos de ETL de la solución de BI:

- a. El uso de una base de datos transitoria (Stage), donde residirán copias de las tablas de sistema transaccional que serán truncadas y cargadas, haciendo lecturas planas desde el sistema transaccional en cada proceso de carga. Tener un “Stage Area” permite hacer transformaciones sobre los datos en un entorno de base de datos separado y tener los datos de la ventana de tiempo a cargar, en caso de reiniciar el proceso de carga por fallo.
- b. Es además muy recomendable tener un estándar de nombre para cada uno de los componentes de la solución de ETL, por ejemplo, que todos los componentes de Flujo de Datos inicien con las palabras “DF\_” y así con cada tipo de componente.
- c. Crear un Mapa de Alto Nivel de la solución de ETL. Este diagrama debe especificar los orígenes de datos (con nombre de servidores y bases de datos), determinar a grosso modo donde se utilizan las herramientas de ETL para extraer, transformar y cargar datos, así como la base de datos destino donde serán entregados los datos obtenidos desde el sistema origen.
- d. Crear un diccionario de datos con mapeo de origen y destino, especificando para cada campo en las tablas de hecho y de dimensión, de qué origen de datos, tabla y campo específico proviene cada dato y a qué tabla y campo específico del modelo dimensional corresponde. Especificar en este mapeo las transformaciones aplicadas a cada campo.

Es importante que antes de crear los procesos de ETL y en general, el diseño del modelo dimensional, se haya realizado una tarea de perfilado de datos, sobre los orígenes transaccionales.

## 16. Orígenes de datos reales

Las soluciones de Inteligencia de la Información son utilizadas para hacer todo tipo de análisis que van desde exploratorios hasta predictivos, es por esta razón que tener acceso a la información real aumenta la probabilidad de desarrollar una solución exitosa. Los datos reales permiten que los analistas de datos puedan reproducir el proceso, identificar valores atípicos y otras características que aceleran el proceso de desarrollo.

Por otra parte, desarrollar soluciones de Inteligencia de negocios con orígenes de datos alimentados por sistemas en desarrollo, generan retrabajo y llegan a agotar a los ingenieros datos, científicos de datos, analistas de datos y especialistas en visualizaciones que trabajan en el desarrollo de la solución, porque cada vez que se hace un ajuste en el origen de datos se impacta directamente toda la solución de Inteligencia de Negocios.

## 17. Bitácoras

Una bitácora es un registro cronológico, actualizado regularmente con datos específicos, en soluciones de Inteligencia de Negocios, es necesario registrar dos tipos de bitácoras:

1. **Bitácoras de cargas:** Son registros de los procesos de extracción y transformación de datos de los diferentes componentes de la solución de Inteligencia de Negocios, son importantes para poder hacer control y monitoreo de los procesos, también facilitan hallazgos importantes durante la revisión de problemas.
2. **Bitácoras de consultas:** Para cada uso de los reportes de Inteligencia de Negocios se debe registrar como mínimo el usuario que utiliza el reporte, los parámetros que utiliza, cuando inicia la consulta y cuando termina de ejecutarse, con el fin de garantizar el correcto uso y consulta de los mismos.

## CAPÍTULO CUARTO: IMPLEMENTACIÓN DE SOLUCIONES DE INTELIGENCIA DE NEGOCIOS

### 18. Mediciones





---

Se recomienda con base en las bitácoras registradas generar información útil para la toma de decisiones de las soluciones de Inteligencia de Negocios, algunos ejemplos son:

1. **Cantidad de datos:** Esto permite conocer cuanto crecen las bases de datos por día y cuanta información se mueve por la red en un horario específico.
2. **Impacto de implementación:** Conociendo los usuarios que utilizan los reportes de la solución se puede estimar cuando ahorro se genera para la organización el uso de estos.

#### **19. Documentación de procesos de implementación:**

Es importante que las tareas del Administrador de la plataforma de Inteligencia de Negocios se encuentren debidamente documentadas, se debe detallar pasado a paso como se realiza una instalación de una nueva solución, quien solicita la instalación, quien la aprueba, que se debe instalar, como se debe instalar, que permisos se deben asignar y cualquier otra información que se considere necesaria para el proceso

### **CAPÍTULO QUINTO: MANTENIMIENTO DE SOLUCIONES DE INTELIGENCIA DE NEGOCIOS**

#### **20. Documentación de procesos de mantenimiento:**

El proceso de mantenimiento y monitoreo es indispensable para garantizar el rendimiento de las soluciones de Inteligencia de la Información, algunas de las tareas que se deben documentar son:

1. Se debe documentar cada proceso o tarea de monitoreo de la solución de inteligencia de negocios, en ella detallar paso a paso como hacer el monitoreo y la periodicidad con la que se realiza.
2. Se debe detallar pasado a paso como se realiza una actualización, quien solicita la instalación, quien la aprueba, que se debe instalar, como se debe instalar, que permisos se deben asignar y cualquier otra información que se considere necesaria para el proceso.
3. Documentación de los problemas presentados y sus soluciones.
4. Proceso de respaldos y mantenimiento, por ejemplo, respaldos de bases de datos, de scripts, directorios, archivos de configuración, aplicaciones, ejecución de tareas para reorganizar índices y toda tarea que se considere indispensable para que el buen funcionamiento de una solución.



---

## INTELIGENCIA ARTIFICIAL

### CAPÍTULO PRIMERO: ASPECTOS GENERALES

#### 1. Objetivos Pretendidos

El objetivo principal es plasmar en este documento la experiencia adquirida durante el proceso de análisis, desarrollo e implementación del proyecto de Inteligencia Artificial aplicada en el proceso de tramitación judicial.

Los objetivos específicos son:

- c) Difundir a través de la base del conocimiento, la experiencia en la implementación de este proyecto.
- d) Proponer una estructura mínima de los instrumentos para la implementación de soluciones de Inteligencia de Negocio.

#### 2. Noción de Inteligencia Artificial:

La inteligencia artificial (IA) es una disciplina científica de las Ciencias de la Computación que, como cualquier otra disciplina, está compuesta por una serie de conceptos, problemas y métodos/herramientas que los resuelven.

Los conceptos característicos de la IA son:

- **Autonomía:** Habilidad para ejecutar tareas en entornos complejos sin la constante guía de una persona.
- **Adaptabilidad:** Habilidad para mejorar el rendimiento, aprendiendo desde la experiencia.

En la actualidad y bajo la filosofía de la IA es importante considerar:

- Un método o herramienta de IA maneja solo una situación o tarea específica, no existiendo un sistema que pueda manejar cualquier tarea intelectual.
- Las herramientas actuales de IA muestran comportamientos inteligentes, sin ser genuinamente inteligentes ni ser conscientes de sí mismas.

Normalmente, un sistema de IA es capaz de analizar datos en grandes cantidades, identificar patrones y tendencias y, por lo tanto, formular predicciones de forma automática, con rapidez y precisión, como, por ejemplo:

- Siri funciona como un asistente personal, ya que utiliza procesamiento de lenguaje natural.
- Facebook y Google Fotos sugieren el etiquetado y agrupamiento de fotos con base en el reconocimiento de imagen.
- Amazon ofrece recomendaciones de productos basadas en modelos de canasta de compra y preferencias del usuario.
- Waze brinda información optimizada de tráfico y navegación en tiempo real.

**Machine Learning** es un campo de la Inteligencia Artificial que crea sistemas capaces de mejorar su rendimiento en una tarea específica en la medida que cuentan con mayor experiencia o información.

*Aprender* en este contexto quiere decir identificar patrones complejos dentro de un gran volumen de datos. **La máquina que realmente aprende es un algoritmo** que revisa los datos, clasificándolos y prediciendo comportamientos.

**Deep Learning** es un subcampo de Machine Learning que utiliza una secuencia de múltiples capas de procesamiento no lineal para la extracción y transformación de características. Cada capa utiliza la salida de la capa anterior como entrada, permitiendo tener múltiples niveles de representaciones que corresponden a diferentes niveles de abstracción. Estos niveles de abstracción forman una jerarquía de conceptos.

#### 3. Marco de trabajo:



---

A años de haber implementado el expediente digital para la tramitación de causas y una vez internalizado su uso, surgen nuevas necesidades e inquietudes por parte de los usuarios, lo que nos lleva a una permanente búsqueda de soluciones a través de procesos de mejora continua.

En este marco y resultado de un trabajo conjunto con los usuarios se han podido detectar oportunidades para mejorar la experiencia de trabajo, optimizando los procesos, incorporando nuevas funcionalidades con el apoyo de nuevas herramientas tecnológicas.

Este marco de trabajo se compone de tres partes fundamentales de preparación:

- **Conocer la experiencia usuaria e Identificar las oportunidades de mejora:**
- **Buscar alternativas de solución:**
- **Desarrollo del proyecto:**

#### **4. Conformación y roles del equipo de trabajo:**

El elemento principal para considerar es el trabajo conjunto con los usuarios de los sistemas de tramitación de causa, conociendo así de primera fuente su experiencia de uso y las necesidades que requiere, una vez identificadas estas, se debe modelar el proceso a través del cual, se identifican las oportunidades de mejora, estas mejoras pueden ser o no tecnológicas. En este contexto el equipo de trabajo se conformó de la siguiente forma:

1. Jefe de Proyecto
2. Usuario líder del proceso (Comisión)
3. Analista de Proceso
4. Abogado Tramitador

Producto del trabajo de este equipo se identificó una oportunidad mejora al proceso de revisión de las presentaciones para identificar en estos las solicitudes de Ordenes de no innovar a través de inteligencia artificial. En consecuencia, se incorporó a este equipo de trabajo, los siguientes profesionales:

1. Especialista técnico en Inteligencia Artificial
2. Arquitecto
3. DBA
4. Desarrolladores
5. Técnico de control de calidad

#### **5. Identificación de variables para la búsqueda de la solución**

La alta cantidad de ingresos de presentaciones que se realizan en las Cortes de Apelaciones por parte de los abogados, lo que implica contar con un equipo humano con la capacidad suficiente para leerlas al momento que ingresan, previo a continuar con su tramitación, a fin de identificar aquellos que corresponde a solicitudes de órdenes de no innovar que requieren ser tramitados de forma inmediata, ya que contienen plazos según la ley, por lo que su flujo de tramitación es especial.

Ante este escenario, se procedió a conocer la forma en la cual se realizaba este trabajo, detectando que este se basa en la identificación de palabras claves dentro del texto de la presentación que indicaban si era o no Orden de no Innovar, posterior a esto el funcionario seleccionaba en el sistema informático una marca especial que iniciaba un flujo de trabajo diferenciado.

Esta labor era realizada por 10 funcionarios de la Corte con dedicación exclusiva a esta tarea, frente a este escenario surge la necesidad de automatizar este proceso, para lo cual se inicia la búsqueda de herramientas tecnológicas, identificando como solución la aplicación de Inteligencia Artificial.



## **CAPÍTULO SEGUNDO: INVESTIGACIÓN**

Dado que no se contaba con experiencia en esta alternativa de solución, se dio inicio a investigar el mercado, con el propósito de conocer las diferentes alternativas de solución.

### **1. Búsqueda en el mercado de soluciones tecnológicas**

IBM Watson:

Watson es la suite de IBM que contiene el conjunto de servicios, aplicaciones y herramientas de IA para desarrollo empresarial. Con Watson, es posible incorporar herramientas de inteligencia artificial en cualquier industria con el objetivo de aumentar la productividad de los equipos de trabajo, liberando a los funcionarios de tareas repetitivas y capacitándolos para que se centren en el trabajo de valor crítico para la organización.

Watson cuenta con experiencia en múltiples industrias y casos de uso, siendo de interés para el Poder Judicial la incorporación de estas herramientas de IA para agilizar la labor judicial. Destacan las herramientas:

- Watson Assistant.
- Watson Machine Learning.
- Natural Language Classifier.
- Watson Explorer.

Google Cloud:

Al ofrece servicios modernos de aprendizaje automático (Machine Learning), con modelos pre-entrenados y con la posibilidad de generar modelos personalizados. Además, incorpora herramientas para construir interfaces de conversacionales (chatbots), extracción de datos desde videos e imágenes y análisis de texto usando lenguaje natural para extraer información.

Las herramientas de IA de Google tienen un uso variado en la industria, siendo las mismas aplicaciones de Google como Gmail, Translate o Photos, las que nos demuestran su potencial y eficiencia.

El análisis de texto en lenguaje natural, modelos de machine learning y chatbots son herramientas que pueden aportar valor en la automatización de Procesos Judiciales.

ROSS

Es una herramienta de investigación legal avanzada que aprovecha el poder de la inteligencia artificial para hacer que el proceso de investigación sea más eficiente.

El sistema está capacitado, entre otras cosas, para rastrear los cambios en la ley con respecto a problemas de interés y enviar notificaciones con las actualizaciones legales pertinentes.

Esta empresa ha desarrollado diferentes servicios con especialización en materias legales, para hacer los procesos de justicia mucho más eficientes.

Se espera que su experiencia permita identificar nuevos espacios de desarrollo tecnológico y creación de servicios para el Poder Judicial.

RAVN:

Es una plataforma de Inteligencia Artificial de vanguardia que permite a una serie de aplicaciones organizar, descubrir y resumir documentos automáticamente.

Esta solución abrió nuevos caminos en la resolución de delitos al ayudar a la Oficina de Fraudes Graves del Reino Unido a exponer el soborno a gran escala y la corrupción en Rolls-Royce, y está en camino de contribuir a nuevas investigaciones.

Esta experiencia práctica muestra el valor de la inteligencia artificial en el procesamiento de altos volúmenes de información.

PROMETEA ARGENTINA:

Es un sistema de Inteligencia Artificial, desarrollado en Argentina, por un equipo multidisciplinario del Ministerio Público Fiscal de la Ciudad de Buenos Aires, en conjunto con expertos en Inteligencia Artificial.

La solución puede recibir comandos por voz o por medio de un chat, interpretando el lenguaje natural para escoger un caso y proponer una sentencia en base a clasificaciones establecidas por aprendizaje de casos anteriores.



Esta solución muestra la incorporación de herramientas de IA para proponer sentencias en base al aprendizaje que se puede obtener de la historia de casos.

#### PJ-BOT COSTA RICA

El Poder Judicial de Costa Rica construyó un Chatbot para ayudar en la atención de público online. Esta solución utiliza herramientas de IA de Google para interpretar lenguaje natural y escoger la respuesta más adecuada a la pregunta realizada.

Este ejemplo demuestra la factibilidad de utilizar herramientas de IA en el ámbito Judicial para crear asistentes virtuales de uso interno o externo a la organización y que sean capaces de comunicarse en lenguaje natural.

Con los antecedentes recabados de las opciones existente en el Mercado se opta por un desarrollo propio utilizando librerías y código abierto, generando nuestra propia plataforma de IA.

### CAPÍTULO TERCERO: DESARROLLO DE LA SOLUCIÓN

1. Entender los criterios de clasificación.
2. Contar con una muestra estadísticamente representativa de documentos que aborde todas las casuísticas posibles de clasificación. El 70% de la muestra se ocupa como entrenamiento y el 30% para validar el modelo.
3. Efectuar una limpieza del texto dejando solo aquellas palabras y frases relevantes.
4. Diseño e implementación de la Arquitectura de la solución (microservicios en java y python)
5. Definir las librerías o herramientas de IA a utilizar (python)
6. Definir cómo representar el texto a través de un modelo matemático:  
El conjunto escogido es:
  - **Modelos de aprendizaje automático tradicional:**
    - *Logistic Regression.*
    - *Multinomial Naive Bayes.*
    - *Lineal Support Vector Machine.*
    - *Random Forest.*
  - **Modelos de aprendizaje automático profundo (Deep Learning):**
    - *Convolutional Neural Network (CNN).*
    - *Bidirectional Long-short Term Memory (LSTM).*
7. Escoger un modelo matemático de entrenamiento de machine learning

Luego de validar los modelos elegir el de mayor precisión.

- **Análisis**

Del proceso de evaluación, es posible notar que el modelo con las mejores métricas es el ***Linear Support Vector Classifier*** que fue entrenado con el conjunto de entrenamiento sin aplicar remuestreo. Esto debido se debe a que, y según la comunidad de la ciencia de datos, este tipo de modelos se comporta "bien" con conjuntos de datos desequilibrados, pero para este caso el desbalance es una brecha excesivamente grande, por lo que, de todas maneras, sus métricas podrían ser mejores en caso de reducirla.

En cuanto al aprendizaje profundo, la CNN es descartada, por el momento, debido a que, luego de una cierta cantidad de iteraciones se sobreajuste a los datos de entrenamiento, y por ende, no generaliza bien, lo que indica que la arquitectura propuesta es muy compleja para el problema. Por otra parte, la LSTM es descartada totalmente, ya que luego de pruebas simples, tanto su entrenamiento como su utilización para predecir, demoran bastante tiempo, lo que no es adecuado para un escenario productivo como al que se apunta.

## CAPÍTULO CUARTO: IMPLEMENTACIÓN DE SOLUCIONES

Se desarrolló una herramienta de inteligencia artificial integrada con la Oficina Judicial Virtual y el sistema de tramitación de Cortes de Apelaciones, para realizar un análisis de la presentación de escritos identificando aquellos que contienen la solicitud de una Orden de no Innovar (ONI), ya que esta labor se realiza actualmente en forma manual, permanente y repetitiva por funcionarios de las distintas Cortes de Apelaciones.

### 1. Entrenamiento:

Para realizar la fase de entrenamiento se extrajeron del sistema informático una serie de presentaciones previas con documentos que ya habían sido clasificados como ONI, para analizar su contenido, ya que la información se podía contener de diferente manera en los textos, tales como: orden de no innovar, no innovar, ordénese no innovar, artículo 192 del código de procedimiento civil, entre otros

### 2. Integración con el sistema de tramitación de causas:

Se integro al sistema de tramitación de causas, específicamente en la bandeja de tramitación del funcionario mensajes de alerta que le permiten visualizar aquellos tramites que fueron identificados por el motor de inteligencia artificial como Ordenes de no Innovar.

The image displays two screenshots of a judicial system interface. The top screenshot, titled "Revisar Oni - CIV-174-2018", shows a list of cases with a blue "ONI" label. A callout box points to this label with the text: "La marca será visible dentro de las actividades de la bandeja unificada tanto para". The bottom screenshot, titled "Confirmar Proyecto - CIV-173-2018", shows a similar list of cases with red boxes highlighting the "ONI" labels. A callout box points to these labels with the text: "La actividad de relator de cuenta contará con la advertencia".