

**Universidad ORT Uruguay
Facultad de Ingeniería**

NOVADOC

Entregado como requisito para la obtención del título de Maestría en Big Data

Gabriela Conde – 155883

Rodrigo Morás – 144520

Tutor: Hugo Mora

2023

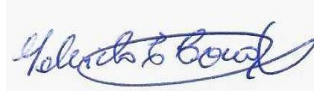
Declaración de autoría

Nosotros, Gabriela Conde y Rodrigo Morás, declaramos que el trabajo que se presenta en esa obra es de nuestra propia mano. Podemos asegurar que:

- La obra fue producida en su totalidad mientras realizábamos Maestría en Big Data;
- Cuando hemos consultado el trabajo publicado por otros, lo hemos atribuido con claridad;
- Cuando hemos citado obras de otros, hemos indicado las fuentes. Con excepción de estas citas, la obra es enteramente nuestra;
- En la obra, hemos acusado recibo de las ayudas recibidas;
- Cuando la obra se basa en trabajo realizado conjuntamente con otros, hemos explicado claramente qué fue contribuido por otros, y qué fue contribuido por nosotros;
- Ninguna parte de este trabajo ha sido publicada previamente a su entrega, excepto donde se han realizado las aclaraciones correspondientes.



Carlos Rodrigo Morás
Muracciole
29/03/2023



Gabriela E. Conde
29/03/2023

Dedicatoria

Sin el apoyo y comprensión de nuestras familias este trabajo no hubiera sido posible. Dedicamos a ellos este trabajo y les realizamos un muy especial agradecimiento.

Agradecimientos

Agradecemos a nuestro tutor Hugo Mora y al coordinador académico de la Maestría en Big Data Dr. Sergio Yovine por su guía y aportes durante el desarrollo de la tesis. Queremos agradecer también a todo el personal del Banco Central del Uruguay que participó en las diferentes etapas del proyecto y realizó numerosos aportes que contribuyeron enormemente para su desarrollo.

Abstract

El Banco Central del Uruguay (BCU) es responsable de regular y supervisar las entidades y actividades financieras, plasmando en Resoluciones su voluntad respecto del mejor funcionamiento del sistema. Estas Resoluciones se encuentran disponibles en la web institucional en formato PDF, lo que dificulta su búsqueda y extracción de información. El objetivo de este trabajo es crear una herramienta llamada NOVADOC que extraiga y almacene la información relevante en un formato estructurado, permitiendo búsquedas según patrones de lenguaje y siendo fácil de usar para los usuarios. La herramienta utiliza técnicas de Big Data como RPA, OCR, MongoDB y NLP, y se desarrolló siguiendo la metodología SCRUM. NOVADOC ha sido exitosamente implementada en el BCU, con un gran número de usuarios que la utilizan y reportando tiempos de búsqueda más cortos que otras alternativas de software. Debido a su éxito, se están evaluando posibles extensiones a la herramienta para realizar búsquedas vinculadas entre documentos.

Palabras clave

NLP, OCR, RPA, IA, SCRUM.

Índice

1.	Resumen ejecutivo.....	11
2.	Introducción.....	13
2.1.	El contexto	13
2.2.	El problema.....	15
2.3.	La solución propuesta	19
3.	Marco teórico.....	21
3.1.	Bases de datos NoSQL	22
3.1.1.	Clave-valor	24
3.1.2.	Documento	26
3.1.1.1.	MongoDB	29
3.1.3.	Columnares.....	29
3.1.4.	Grafos	31
3.2.	Automatización Robótica de Procesos.....	33
3.2.1.	Funcionamiento de RPA.....	34
3.2.2.	Principales usos y usuarios de RPA	34
3.2.3.	Principales proveedores de RPA	35
3.3.	Reconocimiento óptico de caracteres.....	36
3.3.1.	Funcionamiento de OCR	37
3.3.2.	Principales usuarios y usos de OCR.....	38
3.3.3.	Algunas herramientas disponibles de OCR.....	38
3.4.	Procesamiento de lenguaje natural	39
3.4.1.	Funcionamiento de NLP	39
3.4.2.	Principales usuarios y usos de NLP.....	40
3.4.3.	Algunas herramientas disponibles de NLP.....	41
3.5.	Metodología SCRUM.....	42

3.5.1.	Protocolos o eventos de SCRUM	44
3.6.	Gobierno de datos	46
3.6.1.	Gobierno de datos como forma de generar valor para la organización	48
3.6.2.	Creación de un programa de gobierno de datos	48
3.6.3.	Marco para un gobierno de datos	50
3.6.4.	Hoja de ruta para implementar gobierno de datos	52
4.	Solución	55
4.1.	Desarrollo.....	55
4.1.1.	Tecnologías.....	56
4.2.	Algunas dificultades encontradas durante el transcurso del proyecto	60
4.3.	Criterios de éxito.....	61
4.4.	Lecciones aprendidas	63
4.4.1.	Trabajo previo al comienzo del proyecto	63
4.4.2.	Documentación	64
4.4.3.	Tácticas para el involucramiento de los usuarios	64
4.4.4.	Investigación y capacitación en nuevas tecnologías.....	65
4.4.5.	Seguridad, registro de accesos y búsquedas	65
4.4.6.	Incorporar visiones de expertos al proyecto	65
5.	Conclusiones.....	67
6.	Referencias bibliográficas	68
	ANEXO 1: Resolución Fideicomiso Financiero	70
	ANEXO 2: Uso de NOVADOC	70

Índice de tablas

Tabla 1 Principales aplicaciones empresariales de RPA.....	34
Tabla 2 Principales aplicaciones empresariales de OCR.....	38
Tabla 3 Alcance del marco de gobierno de datos.....	50
Tabla 4 Componentes habituales de un marco de gobierno de datos.....	52
Tabla 5 Criterios de éxito del proyecto.....	61
Tabla 6 Consultas a NOVADOC.....	72

Índice de ilustraciones

Ilustración 1 Consulta web sobre fideicomisos de oferta pública vigentes	16
Ilustración 2 Ciclo interno de fideicomisos financieros.	17
Ilustración 3 Ejemplo de base clave-valor	25
Ilustración 4 Ejemplo de JSON y XML	27
Ilustración 5 Ejemplo de base columnar	30
Ilustración 6 Ejemplo Base grafo	32
Ilustración 7 El ciclo de SCRUM	43
Ilustración 8 Elementos del gobierno de datos	49
Ilustración 9 Esquema de la solución NOVADOC	55
Ilustración 10 Resultado de una búsqueda de usuario en NOVADOC	56
Ilustración 11 Construcción de robots en UiPath	58
Ilustración 12 Mensaje de correo enviado por NOVADOC	59
Ilustración 13 Mensaje a los desarrolladores de NOVADOC	60

1. Resumen ejecutivo

Dado su naturaleza de entidad pública y en particular de ente regulador de entidades y actividades financieras, el Banco Central del Uruguay (BCU) a través de la Superintendencia de Servicios Financieros (SSF) dicta actos administrativos vinculados al funcionamiento de estos y que son plasmados en Resoluciones adoptadas siguiendo los preceptos establecidos por la legislación vigente¹.

Las Resoluciones a las que refiere el párrafo precedente son generalmente disponibilizadas en la web institucional en formato de archivo pdf². Este formato de tipo no estructurado hace más compleja la búsqueda y extracción de información, viéndose por tanto afectada la eficiencia -y potencialmente la calidad- de la toma de decisiones y la construcción de memoria institucional. Generar una herramienta que justamente contribuya a ganar en eficiencia, así como a validar la homogeneidad en el tratamiento de planteos similares son las principales motivaciones de este trabajo.

Dentro de las actividades financieras reguladas, existe un grupo que representa un porcentaje significativo y es aquel de los denominados **Fideicomisos Financieros** (FF). De acuerdo a los datos del ente regulador, en 2019 se solicitó la inscripción de 17 FF (7,7%), de 10 en 2020 (5,4%) y de 14 en 2021 (6,2%) (los valores entre paréntesis corresponden a la participación de dichos trámites dentro del total de autorizaciones realizadas por la SSF). Estos instrumentos tienen un impacto no menor en la actividad económica nacional y su relevamiento oportuno es fundamental para la construcción de estadísticas, función a la que el BCU contribuye de modo significativo.

La herramienta desarrollada se denomina NOVADOC. Esencialmente NOVADOC extrae de la web institucional los archivos pdf con Resoluciones, los transforma y almacena en un formato estructurado la información relevante para la construcción de estadísticas nacionales, a la vez que permite realizar búsquedas según

¹ IMPO – Centro de Información Oficial, “Decreto 500/991”, octubre 1991, [Online], Available: <https://www.impo.com.uy/bases/decretos/500-1991>

²Las Resoluciones que se publican en la página web del BCU son aquellas clasificadas como públicas según la Ley No. 18.381. IMPO – Centro de Información Oficial, “Ley No. 18.381”, noviembre 2008, [Online], Available: <https://www.impo.com.uy/bases/leyes/18381-2008>

patrones de lenguaje y a la cual se accede usando una aplicación web de un modo totalmente amigable para los usuarios.

NOVADOC fue desarrollada usando tanto herramientas como metodologías de gestión de proyecto vinculadas al mundo de la Big Data. Es así que la extracción de archivos pdf de la web institucional se realiza mediante automatización robótica de procesos (Robotic Process Automatization – RPA). Para algunos documentos de larga data se aplicaron técnicas de tipo reconocimiento óptico de caracteres (Optical Character Recognition - OCR). Todos los archivos son almacenados usando la base de datos de tipo no estructurado MongoDB y las búsquedas de texto se realizan utilizando herramientas de procesamiento de lenguaje natural (Natural Language Processing - NLP). Finalmente, como metodología de gestión de proyectos se aplicó SCRUM.

NOVADOC fue puesta en producción en el BCU con resultados altamente satisfactorios y su uso trascendió al originalmente pensado. Esto se evidencia tanto el número de usuarios que acceden a la aplicación y la cantidad de búsquedas realizadas, así como por los menores tiempos que insumen las búsquedas de texto usando otras alternativas de software. Adicionalmente y dados estos buenos resultados, otras líneas de negocios plantearon nuevos casos de uso a ser contemplados por la herramienta así como posibles extensiones tendientes a realizar búsquedas de tipo vincular entre documentos.

2. Introducción

2.1.El contexto

De acuerdo con el artículo 3° del Texto Ordenado de la Carta Orgánica (TOCO) del BCU, entre sus finalidades se encuentra: “B) La regulación del funcionamiento y la supervisión del sistema de pagos y del sistema financiero, promoviendo su solidez, solvencia, eficiencia y desarrollo.”³

Por otra parte, el artículo 34 del TOCO demarca el perímetro regulatorio y define a qué sector del BCU le corresponderá la regulación del sistema financiero. Se transcribe a continuación en lo pertinente el artículo referido:

“Artículo 34. (Entidades supervisadas).- El Banco ejercerá la regulación y fiscalización de las entidades que integran el sistema financiero, cualquiera sea su naturaleza jurídica y dispongan o no de personería jurídica, a través de la Superintendencia de Servicios Financieros.

A estos efectos se definen como entidades integrantes del sistema financiero las siguientes:

- A) Las instituciones que integran el sistema de intermediación financiera.
- B) Entidades que presten servicios financieros de cambio, transferencias domésticas y al exterior, servicios de pago y cobranzas, servicios de cofres, créditos y otras de similar naturaleza, exceptuando a las reservadas a las instituciones de intermediación financiera.
- C) Las casas de cambio, que podrán realizar actividades de cambio y conexas, transferencias domésticas, venta de cheques de viajero y servicios de cobranza y pagos.
- D) Las Administradoras de Fondos de Ahorro Previsional y los fondos que administran.
- E) Las empresas de seguros y reaseguros y las mutuas de seguros.

³ Banco Central del Uruguay [8]

- F) Las bolsas de valores, los intermediarios de valores y las entidades de custodia o de compensación y liquidación de valores.
- G) Las administradoras de Fondos de Inversión, los fiduciarios profesionales, los Fondos de Inversión y los fideicomisos financieros de oferta pública (el resaltado pertenece a los autores).**

La Superintendencia de Servicios Financieros también reglamentará y controlará a los emisores de oferta pública de acuerdo a lo establecido en la Ley N° 16.749, de 2 de mayo de 1996.

.....”

Como surge claramente del artículo 34, los fideicomisos financieros de oferta pública integran el elenco de entidades supervisadas por la SSF y se ejercerá sobre ellos regulación y fiscalización (supervisión).

De acuerdo a la definición dada por la página web del BCU⁴, un fideicomiso “es aquel negocio jurídico por medio del cual se constituye la propiedad fiduciaria de un conjunto de derechos de propiedad u otros derechos reales o personales que son transmitidos por el fideicomitente al fiduciario para que los administre o ejerza de conformidad con las instrucciones contenidas en el fideicomiso, en beneficio de una empresa (beneficiario), que es designada en el mismo, y la restituya al cumplimiento del plazo al fideicomitente o la transmita al beneficiario.”

Adicionalmente y de acuerdo a la misma fuente, “un fideicomiso financiero es aquel negocio de fideicomiso cuyos beneficiarios son titulares de certificados de participación en el dominio fiduciario, de títulos representativos de deuda garantizados con los bienes que integran el fideicomiso, o de títulos mixtos que otorguen derechos de crédito y derechos de participación sobre el remanente. Solamente pueden ser fiduciarios en un fideicomiso financiero las entidades de intermediación financiera o las sociedades administradoras de fondos de inversión.”

⁴ Banco Central del Uruguay, “Fiduciarios Financieros y Profesionales”, marzo 2023, [Online], Available: <https://www.bcu.gub.uy/Servicios-Financieros-SSF/Paginas/Fiduciarios.aspx#:~:text=Un%20Fideicomiso%20Financiero%20es%20aquel,de%20cr%C3%A9dito%20y%20derechos%20de>

La SSF tiene publicado en la página web del BCU el proceso de autorizaciones de empresas, personas y productos que participan del sistema financiero. Este proceso está alcanzado por el sistema de gestión de calidad basado en las normas internacionales ISO y tiene como sustento las potestades expresadas en el TOCO. El documento respectivo describe las etapas del proceso las cuales incluyen: 1) Recepción de solicitud y análisis preliminar; 2) Análisis Particular; 3) Solicitudes de información; 4) Propuesta de Resolución; 5) Notificaciones y gestiones posteriores; y 6) Registro, pase a otras áreas y archivo. Adicionalmente, dentro del proceso se incluyen doce tipos de trámites de autorizaciones, quedando comprendida la inscripción de fideicomisos financieros en el tipo de trámite número tres.

Dado su naturaleza de entidad pública, la SSF dicta el acto administrativo resultante en la inscripción de un fideicomiso financiero mediante la adopción de una Resolución redactada siguiendo los preceptos establecidos por la legislación vigente⁵. Esta Resolución afecta de modo sustancial a las etapas 5) y 6) del proceso descrito en el párrafo precedente. Una vez adoptada, la misma se publica en formato pdf en el sitio web institucional.

2.2.El problema

El BCU contribuye de manera significativa a la generación de información económico-financiera y es así que uno de sus objetivos estratégicos es el de generar y difundir información y conocimiento en la materia⁶.

Como se expresara en la sección previa, es mediante una Resolución que se plasma la voluntad de la administración en relación a la inscripción de un fideicomiso financiero. Dicha Resolución contiene información muy valiosa que resulta de utilidad para diferentes usuarios. Para algunos son los aspectos documentales y formales los más relevantes mientras que para otros lo son las características económico-financieras del instrumento inscripto.

⁵ Se adjunta en el Anexo I un ejemplo de Resolución de inscripción de fideicomiso.

⁶Banco Central del Uruguay, “Estadísticas y Estudios”, marzo 2023, [Online], Available: <https://www.bcu.gub.uy/Estadisticas-e-Indicadores/Paginas/Default.aspx>

Con el diseño actual del proceso de autorizaciones, los registros vigentes de la SSF mantienen información que es relevante desde el punto de vista documental y satisfacen parcialmente las necesidades de otros usuarios. La información de carácter económico-financiero que es de utilidad para la formulación de estadísticas no resulta captada por tales registros (por ejemplo, el monto pendiente, el plazo de vigencia o la fecha de cancelación no son mostrados). En la siguiente imagen puede apreciarse una captura de pantalla de una consulta efectuada en la página web del BCU⁷ y referida a los fideicomisos de oferta pública vigentes (Ilustración 1).



Ilustración 1 Consulta web sobre fideicomisos de oferta pública vigentes

El ciclo de los fideicomisos financieros desde su surgimiento y su recorrido dentro de los distintos sectores del BCU es complejo tal como puede apreciarse en la siguiente imagen (

Ilustración 2):

⁷ Banco Central del Uruguay, “Portada/Sistema Financiero/Buscador”, marzo 2023, [Online], Available: https://www.bcu.gub.uy/Servicios-Financieros-SSF/Paginas/buscador_Registros.aspx

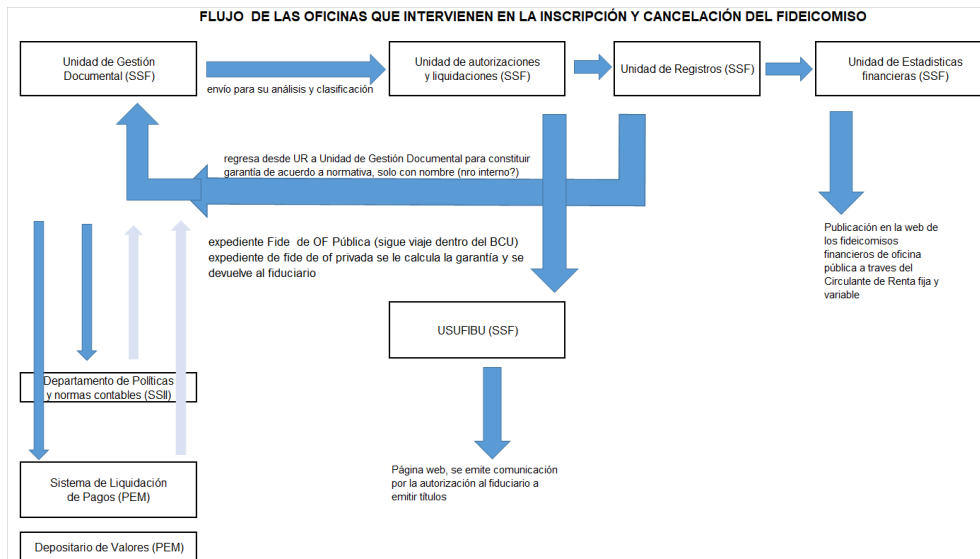


Ilustración 2 Ciclo interno de fideicomisos financieros.

Fuente: elaboración propia en base a las entrevistas realizadas.

Principales funcionalidades de las unidades presentes en el diagrama anterior

Unidad de Gestión Documental (SSF)

1. Puerta de entrada del fideicomiso al BCU. Recibe al fiduciario y hace un control administrativo de la documentación que se ingresa.
2. Activa mecanismo de constitución de la garantía mandando mails a contaduría y sistema de pagos
3. Una vez que contaduría y sistema de pagos completan sus procesos, constituye la prenda del depósito

Unidad de autorizaciones y liquidaciones (SSF)

Se analizan los documentos (el contrato entre otros). Del análisis surge validación (o no) del fideicomiso y por lo tanto se registra, y se lo clasifica en oferta pública o privada. Se pide asesoramiento a Asesoría Jurídica (AJ) u USUFIBU acerca del instrumento que se quiere emitir y sobre el monto de la garantía a constituir. Cuenta con un procedimiento detallado y flujograma.

Unidad de Registros (SSF)

1. Inscriben los fideicomisos financieros de oferta pública en el Registro de Mercado de Valores que se publica web. Para ello, ingresan la información manualmente en el SAR y luego de allí sale la información a publicarse
2. Guardan registro interno de los FF de oferta privada desde 2014.

Unidad de Estadísticas financieras (SSF)

Publica el circulante de Fideicomisos Financieros de Oferta Publica ya sean emitidos por el Sector Público o Privado

Las fuentes principales de información son:

- Nuevas emisiones: Comunicaciones emitidas por la SSF
- Información del FF: Registros, se utiliza el Prospecto informativo y Documento de emisión.
- Fuente IDI código 741 Monto en circulación, hoy en día todas las emisiones y montos en circulación salen por esa fuente.
- Otras fuentes consultadas: correo electrónico del departamento vasistema/ página web BEVSA
- Quienes usan del banco la información: U.S. Afap, etc...

Departamento de Políticas y normas contables (SSII)

Crea la cuenta especial BCU para depósito en garantía.

Sistema de Liquidación de Pagos (PEM)

Interviene pasos finales de constitución de depósito en garantía.

Depositorio de Valores (PEM)

Asigna ISIN a los instrumentos a emitir y los identifica con código BIC. Ambos códigos son independientes y se ingresan a la hora de crear tanto el instrumento como la institución, una vez que recibimos el prospecto de autorizaciones y registros.

2.3.La solución propuesta

NOVADOC surge como una solución innovadora que permite extraer información relevante de una fuente no estructurada para contemplar necesidades de usuarios distintos a los que originalmente la consumían. Fue desarrollada usando tanto herramientas como metodologías de gestión de proyecto y de gobierno de datos vinculadas al mundo de la Big Data. En primer lugar, la extracción de archivos pdf de la web institucional se realiza mediante automatización robótica de procesos (Robotic Process Automatization – RPA). En segundo lugar, para algunos documentos de larga data se aplicaron técnicas de tipo reconocimiento óptico de caracteres (Optical Character Recognition - OCR). En tercer lugar, todos los archivos son almacenados usando la base de datos de tipo no estructurado MongoDB y las búsquedas de la información relevante se realizan utilizando herramientas de procesamiento de lenguaje natural (Natural Language Processing - NLP). En cuarto lugar, como metodología de gestión de proyectos se aplicó SCRUM. Finalmente, el proyecto fue utilizado como un piloto que permitiera contribuir a incorporar el gobierno de datos en la cultura de la organización.

NOVADOC es parte de un conjunto de soluciones cuyo foco es la utilización de información por parte de múltiples usuarios desde un único lugar, contribuyendo así a mejorar el gobierno de datos de la institución. En este sentido se diseñaron e implementaron tres softwares: NOVADOC que extrae los datos desde la fuente original, un repositorio central de datos para su almacenamiento y una herramienta de visualización y descarga de la información que está en el repositorio pensada para todos los usuarios del BCU.

NOVADOC fue puesta en producción en el BCU con resultados altamente satisfactorios y su uso trascendió al originalmente pensado. Esto se evidencia tanto el número de usuarios que acceden a la aplicación y la cantidad de búsquedas realizadas así como por los menores tiempos que insumen las búsquedas de texto usando otras alternativas de software. Adicionalmente y dados estos buenos resultados, otras líneas de negocios plantearon nuevos casos de uso a ser contemplados por la herramienta así

como posibles extensiones tendientes a realizar búsquedas de tipo vincular entre documentos.

3. Marco teórico

Tanto organizaciones como individuos usando diferentes tecnologías generan datos a una gran velocidad, de gran volumen, variedad, que presentan desafíos respecto a fuentes tradicionales en cuanto a su validez pero que les adicionan un gran valor. El manejo de la Big Data de un modo consistente con los objetivos de negocio de cualquier organización plantea la necesidad de utilizar nuevas tecnologías, herramientas y metodologías de trabajo en un marco de sanas prácticas de gobierno de datos.

Entre las tecnologías y herramientas de la Big Data se encuentran aquellas que permiten extraer masivamente datos de sitios web o de bases de datos de gran tamaño (Application Programming Interfaces – APIs por su acrónimo en inglés, las bases de datos para manejo de datos no estructurados o NoSQL, RPA (Robotic Process Automation) para la realización de tareas repetitivas, las soluciones de almacenamiento y procesamiento masivo o cloud computing, el procesamiento distribuido para la validación automática de vía mecanismo de consenso que replican y sincronizan datos en diferentes ubicaciones (Distributed Ledger Technology – DLT) y aquellas que permiten realizar tareas que tradicionalmente requieren de inteligencia humana como por ejemplo el análisis de texto o procesamiento de lenguaje natural (Natural Language Processing - NLP) y el reconocimiento de imágenes (Optical Character Recognition - OCR).

Los proyectos de Big Data tienen ciclos de implementación más cortos que aquellos de los tradicionales. La implementación requiere de metodologías de gestión de proyectos ágiles que permitan detectar fallos de diseño y corregirlos rápidamente. Dentro de estas metodologías destaca SCRUM.

La gestión y la calidad de los datos adquieren una relevancia significativa en el mundo del Big Data. Es así que contar con sanas prácticas de gobierno de datos facilita la implementación de los proyectos y contribuye al desarrollo del marco de gobierno corporativo de las organizaciones.

En las secciones siguientes se describirán algunas de las herramientas, metodologías y marcos teóricos asociados a Big Data s dado su uso en el desarrollo de NOVADOC.

3.1. Bases de datos NoSQL

Entre otros, el uso masivo de la web trajo como consecuencia la generación de datos de gran volumen, a gran velocidad y de formato diverso, características esenciales del Big Data. Dichas características hacen que las tradicionales bases de datos relacionales –que se desempeñan de mejor manera cuando los datos se almacenan por fila y columna- no resulten apropiadas para contener y explotar a la Big Data. Como respuesta a esta restricción, surgen las bases de datos NoSQL, término que se atribuye a Carlo Strozzi por su trabajo de 1998. Muchas de las bases de datos NoSQL que hoy se conocen surgieron por 2009. Riak, MongoDB, HBase, Accumulo, Hypertable, Redis, Cassandra, y Neo4j fueron creadas entre 2007 y 2009.

Las bases de datos NoSQL más generalmente utilizadas comparten las características que se citan a continuación:

- **Agnósticas respecto del esquema:** una base de datos NoSQL no requiere la descripción de todos los datos posibles y sus estructuras.
- **Son no relacionales:** la información se guarda de modo agregado no siendo necesario establecer conexiones entre tablas de datos. La excepción a esto son las bases de tipo grafo.
- **Uso de hardware como commodity:** las bases de datos NoSQL no requieren de un tipo específico de hardware para su almacenamiento y procesamiento. Como consecuencia de lo anterior, pueden agregarse servidores para el almacenamiento de mayor cantidad de datos, mostrando que cuentan con potencial para la escalabilidad.
- **Son altamente distribuibles:** las bases de datos distribuidas pueden almacenar y procesar información en más de un dispositivo. Un grupo de servidores (cluster) puede ser utilizado para almacenar una única y gran base de datos NoSQL.

- **Tienen alta disponibilidad:** esto se deriva del uso de clusters que proveen tolerancia a fallos.
- **Bajos costos operativos:** la gran mayoría de ellas se construyen sobre plataformas de tipo Open Source.
- **Consistencia eventual:** los datos en múltiples nodos pueden no ser consistentes en todos ellos luego de una operación de escritura. Sin embargo, todos los nodos estarán eventualmente en un estado consistente.
- **Propiedades BASE (no ACID, o A(tomicity)C(onsistency)I(solation)D(urability)):** cumplir con BASE requiere que la mantenga una alta disponibilidad en caso de que falle un nodo de la red, no siendo requerido que mantenga un estado consistente cuando ocurra una actualización.
- El acceso a los datos es vía API.
- **Auto sharding y replicación:** para soportar el escalamiento horizontal y proveer de alta disponibilidad, las bases No SQL emplean técnicas de replicación y sharding mediante las cuales las bases se particionan de modo horizontal y luego son copiadas en múltiples nodos.
- **Cache integrado:** esto remueve la necesidad de una capa adicional de cache.
- **Soportan consultas distribuidas:** las bases mantienen la consistencia entre múltiples shards
- **Persistencia políglota:** usar bases NoSQL no excluye el uso de relacionales y los datos pueden persistirse en ambos tipos de almacenamiento.

Las características de alta escalabilidad y disponibilidad son las principales ventajas de las bases NoSQL. Las mismas, sin embargo, no están exentas de ciertas desventajas:

- **Foco reducido:** están pensadas básicamente para almacenar pero proveen muy pocas otras funcionalidades.
- **Son Open-source:** no existe aún un estándar confiable para dichas bases de datos.

- **Su administración representa un desafío:** la administración de las bases NoSQL es mucho más compleja que la de las relacionales. Se afirma que son difíciles de instalar y que su manejo diario es complejo.
- **Falta de herramientas de interfaz gráfica del usuario** (GUI por su sigla en inglés).
- **Respaldo:** en algunas bases de datos NoSQL como MongoDB no existe un enfoque para el respaldo de los datos de modo consistente.
- **Documentos de gran tamaño:** algunas bases como MongoDB y CouchDB almacenan datos en formato JSON. Ello implica que los documentos tienen un gran tamaño y por tanto agregarles claves descriptivas genera perjuicios, ya que incrementa el tamaño de la base.

Existen varios tipos de bases NoSQL los cuales pueden ser resumidos en cuatro categorías:

- Clave-valor
- Documento
- Columnares
- Grafos

Se describen a continuación dichos tipos, así como sus usos más frecuentes.

3.1.1. Clave-valor

Estas bases almacenan los datos como pares clave-valor que actúan como tablas hash. La tabla es una lista de valores donde cada valor se identifica por su clave. Los valores pueden ser de tipo variado, desde datos provenientes de sensores hasta videos.

Estas bases de datos tienen como beneficio su simplicidad. El manejo de los valores se deja en manos de la aplicación que utiliza los datos. En general no mantienen índices, lo que hace que su actualización sea muy rápida. Dada la simplicidad referida, estas bases de datos son altamente escalables.

Una base de datos de tipo clave-valor resulta apropiada en los siguientes casos:

- Los datos a almacenar son de tipo no estructurado
- Se requiere alta velocidad en la lectura/escritura
- El valor se identifica completamente vía la clave
- El valor es una entidad única que no depende de otros valores
- Los valores tienen una estructura relativamente simple o son binarios
- Los patrones de búsqueda son sencillos (insertar, seleccionar, borrar, principalmente)
- Los valores almacenados se manejan a nivel de la capa de aplicación

Existen casos en los que las bases clave-valor no resultan apropiadas:

- Las aplicaciones requieren búsquedas o filtrados que usan atributos de los valores almacenados
- Existen relaciones entre diferentes entradas de tipo clave-valor
- Se requiere la manipulación de varias claves para una sola operación
- Se requiere consistencia de esquema para diferentes valores
- Se requiere actualizar diferentes atributos individuales de la clave

Como ejemplo de bases de datos claves valor se pueden citar Riak, Redis, y Amazon Dynamo DB.

Se muestra a continuación un ejemplo de almacenamiento NoSQL de tipo clave-valor (Ilustración 3).

Clave	Valor
001	JJ, 10.10.78, Queja por servicio
785	197864485
126	<CustomerId>2980</CustomerId><Total>50.25</Total>

Ilustración 3 Ejemplo de base clave-valor

Fuente: elaboración propia en base a Erl [11]

3.1.2. Documento

Estas bases de datos también almacenan datos como pares de clave valor. Se diferencian de las bases anteriormente descritas en que el valor es aquí un documento. Los documentos suelen tener una estructura anidada y pueden ser codificados con esquemas de tipo JSON o XML.

Se muestra a continuación un ejemplo de un pedido de compra de cliente en ambos formatos (Ilustración 4):

```
{ _id: 1234, customer: 52, customer_name: "Juan Castro", items: [
  qty: 2, item: {id: 456, title: "Martillo", unit_price: 500,00}, price: 1000,00},
  qty: 1, item: {id: 111, title: "Tornillo", unit_price: 10,00}, price: 10},
], delivery_address: {street: "Al este del paraíso", city: "Oz", ...}
}
```

```
<order id= "1234">
  <customer id: "52">Juan Castro</customer>
  <items>
    <item qty="2" id="456" unit_price="500,00"
price="1000,00">Martillo</item>
    <item qty="1" id="111" unit_price="10,00"
price="10,00">Tornillo</item>
  </items>
  <delivery_address>
    <Street>Al este del paraíso</Street>
    <city>Oz</city>
  .....
  <delivery_address>
```

Ilustración 4 Ejemplo de JSON y XML

Fuente: elaboración propia en base a Fowler [12]

En una base de este tipo se puede almacenar cualquier tipo de documento sin necesidad de saber previamente su estructura. Por ejemplo, puede recibirse una gran cantidad de información en modo texto o en formato binario en los formatos que se describen a seguir:

Colecciones de archivos que contienen combinaciones de:

- Archivos estructurados (tales como información sobre gastos almacenada en formato CSV)

- Archivos semi-estructurados (XML y páginas web HTML)
- Archivos no estructurados (tales como imágenes JPEG, películas, MP3s, documentos word, PDFs, y archivos de texto plano)

Los formatos no estructurados son realmente semi-estructurados. Extraer y buscar la información de la metadata puede ser útil. Algunas bases de datos NoSQL tienen incorporado soporte para extraer la metadata y el texto plano de archivos binarios. Esta capacidad resulta muy provechosa para indexar el texto plano o para buscar o proveer un manejo más estructurado de la metadata de los archivos.

Un almacenamiento de tipo documento es adecuado cuando:

- Se necesita guardar documentos de tipo semi-estructurado que contienen una estructura de tipo plana o anidada
- Se requiere evolución del esquema ya que la estructura del documento puede ser desconocida o cambiar
- Se necesita realizar búsquedas en diferentes campos de los documentos
- Es necesario almacenar objetos de dominio, tipo clientes, de modo serial
- Los patrones de búsqueda incluyen insertar, seleccionar, actualizar y borrar.

Este tipo de almacenamiento resulta inapropiado en los siguientes casos:

- Se necesita actualizar varios documentos como parte de una única transacción
- Es necesario realizar operaciones que implican la unión entre múltiples documentos o guardar datos que están normalizados
- Se requiere forzar un tipo de esquema para alcanzar la consistencia en las búsquedas debido a que la estructura del documento puede cambiar entre sucesivas corridas de las búsquedas
- El valor guardado no se describe a sí mismo y no tiene una referencia al esquema
- Se necesita guardar datos de tipo binario

Como ejemplo de bases de datos de tipo documento se encuentran MongoDB, CouchDB, y Terrastore.

Se describen en la siguiente sección las principales características de MongoDB dada la relevancia que tiene para este trabajo.

3.1.1.1. MongoDB

MongoDB es una base de datos NoSQL multiplataforma y de código abierto que se utiliza para almacenar grandes cantidades de datos de manera distribuida y flexible. Se basa en la idea de almacenar datos en documentos similares a JSON con esquemas opcionales, lo que facilita el almacenamiento y la manipulación de datos en comparación con una base de datos relacional tradicional.

MongoDB se utiliza a menudo en aplicaciones web modernas como un medio para almacenar y manipular grandes cantidades de datos en tiempo real. Está diseñado para ser escalable y de alto rendimiento, lo que lo hace adecuado para su uso en aplicaciones de alto tráfico.

Una de las características clave de MongoDB es su soporte para la indexación, que permite a los usuarios buscar y recuperar rápidamente datos de la base de datos. También admite una amplia gama de tipos de datos y tiene un lenguaje de consulta enriquecido para recuperar y manipular datos.

MongoDB se usa comúnmente junto con otras tecnologías como Node.js, un entorno de tiempo de ejecución de JavaScript, y Express.js, un marco de aplicación web para crear API y aplicaciones web. Juntas, estas herramientas forman la pila MEAN (MongoDB, Express.js, Angular.js y Node.js), que es una opción popular para crear aplicaciones web modernas.

3.1.3. Columnares

Las bases columnares almacenan los datos de un modo muy similar a como lo hacen las bases relacionales pero agrupan columnas relacionadas en una fila. Cada columna puede ser una colección en sí misma de columnas relacionadas, referida como super-columna. Cada super-columna puede contener un número arbitrario de columnas

relacionadas que se recuperan o actualizan como una sola unidad. Cada fila consiste de múltiples familias de columnas y pueden tener un diferente set de columnas. Cada fila se identifica por una clave de fila.

A seguir se muestra el ejemplo de orden de compra utilizado para las bases tipo documento en formato columnar (Ilustración 5):

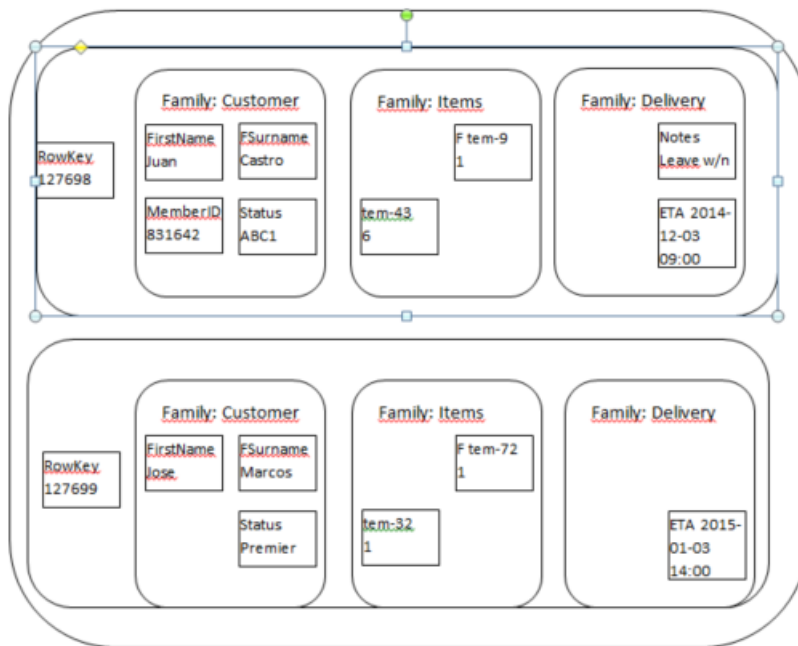


Ilustración 5 Ejemplo de base columnar

Fuente: elaboración propia en base a Fowler [12]

Una base de datos de tipo columnar resulta apropiada en los siguientes casos:

- Se necesita contar con lectura/escritura de modo aleatorio en tiempo real y los datos almacenados tienen cierta estructura definida
- Los datos representan una estructura tabular, cada fila consiste de un largo número de columnas y existen grupos anidados de datos interrelacionados
- Se requiere soporte para la evolución del esquema ya que familias de columnas pueden ser agregadas o quitadas sin momentos de tiempo ocioso del sistema

- Se accede a algunos campos de modo conjunto y se necesita que las búsquedas se realicen usando valores de los campos. Si nos datos en una columna, ese espacio no es reservado en la base.
- Los patrones de búsqueda son del tipo insertar, seleccionar, actualizar y borrar

Existen casos en los que las bases columnares no resultan apropiadas:

- Se requiere hacer uniones de las bases
- Se necesita apoyo transaccional ACID
- Se deben guardar datos de tipo binario
- Se requiere usar consultas de tipo SQL
- Los patrones de búsqueda pueden cambiar frecuentemente porque ello puede iniciar una reestructura de la forma en que las columnas se distribuyen

Como ejemplo de bases de datos de tipo columnar se encuentran Cassandra, HBase y Amazon SimpleDB.

3.1.4. Grafos

Su utilidad deriva de poder almacenar entidades interconectadas. En estas bases de datos lo relevante son las relaciones entre tales entidades. Las mismas son guardadas como nodos (o vértices) mientras que las relaciones son guardadas como aristas. Los nodos pueden tener más de un tipo de enlace entre ellos a través de múltiples aristas. Cada nodo puede tener atributos como pares clave-valor. Cada arista puede tener su propio atributo de datos como pares de clave-valor, el cual puede ser usado para filtros de consulta adicionales. Las consultas en general implican encontrar nodos interconectados basados en los atributos del nodo y/o en los atributos de las aristas (nodos transversales). La utilidad de una base de datos de tipo grafo depende del número y tipo de arista definida entre nodos. Cuanto mayor sea el número y más diversas sean las aristas, más diverso será el número de consultas que pueda manejarse. Generalmente permiten agregar nuevos tipos de nodos sin hacer cambios en la base de datos. Esto permite definir vínculos adicionales entre los nodos cuando nuevos tipos de relaciones o nodos aparecen en la base de datos.

En la siguiente imagen se puede ver un ejemplo de un almacenamiento tipo grafo (Ilustración 6):

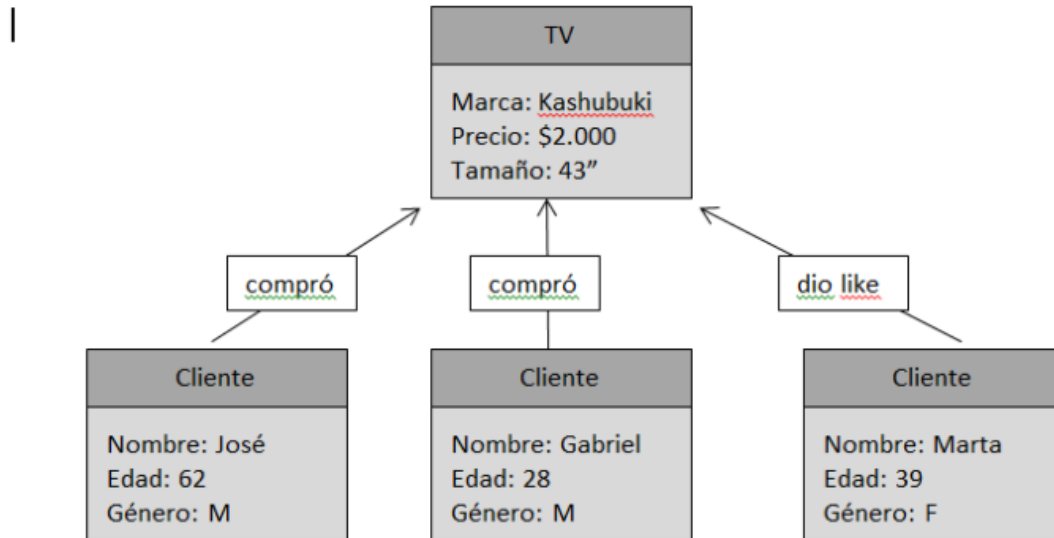


Ilustración 6 Ejemplo Base grafo

Fuente: elaboración propia en base a Erl [11]

Una base de datos de tipo grafo resulta apropiada cuando:

- Se necesitan almacenar entidades interconectadas.
- Las consultas refieren al tipo de relaciones entre cada una de las entidades más que a sus atributos.
- Se requiere encontrar grupos de entidades interconectadas.
- Se requiere encontrar distancias entre entidades en términos de distancia a un nodo transversal.
- Se requiere realizar minería de datos con el objetivo de encontrar patrones.

Existen casos en los que las bases tipo grafo no resultan apropiadas:

- Se requiere realizar actualizaciones de un gran número de atributos de tanto nodos como aristas.
- Las entidades tienen un gran número de atributos o son datos anidados.

- Se deben guardar datos de tipo binario.
- Las búsquedas basadas en la selección de los atributos del nodo o de la arista dominan a las consultas de nodo transversal.

Como ejemplo de bases de datos de tipo grafo se encuentran Neo4j, Infinite Graph y OrientDB.

3.2. Automatización Robótica de Procesos

RPA se refiere al uso de software para manejar tareas repetitivas de gran volumen que anteriormente requerían que los humanos realizaran. La tecnología RPA crea programas de software o bots para realizar tareas como ingresar datos, copiar datos entre aplicaciones, procesar transacciones, responder a las consultas de los clientes y más.

Cuando se lo combina con Inteligencia Artificial y Aprendizaje Automatizado puede capturar más contexto del contenido con el que está trabajando. Casos de uso normalmente citados son leer texto o escritura a mano con OCR; extraer entidades como nombres, términos de factura o direcciones utilizando NLP; y estimar automáticamente los daños por un siniestro en una imagen de reclamo de seguros.

RPA está diseñado para ayudar a las organizaciones a ahorrar tiempo y dinero mediante la automatización de tareas tediosas y que consumen mucho tiempo, y tiene el potencial de mejorar significativamente la eficiencia y la productividad de una empresa. Algunos de los beneficios de RPA incluyen la reducción de errores y un mejor cumplimiento, así como la capacidad de liberar a los empleados para que se centren en un trabajo de mayor valor. Como beneficio también se encuentra el que las unidades de negocio pueden implementarlo sin tener que aprender nuevas herramientas o pedir soporte a los equipos de TI y sin cambiar la infraestructura subyacente de la organización. Deberá tenerse en cuenta, sin embargo, que los bots pueden romperse cuando cambian las interfaces de la aplicación o los flujos de trabajo del proceso.

3.2.1. Funcionamiento de RPA⁸

RPA refleja la forma en que las personas están acostumbradas a interactuar y pensar en las aplicaciones de software. Los bots más simples se pueden crear registrando los clics y las pulsaciones de teclas cuando un usuario interactúa con una aplicación. Los más sofisticados utilizan visión artificial para interpretar los íconos y el diseño de la pantalla, y realizar los ajustes necesarios.

Algunas herramientas de RPA pueden usar esas grabaciones para crear bots híbridos que comienzan registrando un flujo de trabajo existente para luego generar dinámicamente una automatización del mismo en el back-end. En otras implementaciones, las herramientas de minería de procesos y de tareas se utilizan para capturar automáticamente los flujos de trabajo de los procesos de negocio que luego servirán como plantillas para las automatizaciones de RPA.

3.2.2. Principales usos y usuarios de RPA

Se utiliza en la mayoría de las actividades económicas, en particular en aquellas donde se realiza un gran número de tareas repetitivas como seguros, banca, finanzas, salud y telecomunicaciones, como puede apreciarse seguidamente (

⁸ Basado en Lawton [16] y Castellanos [10].

Tabla 1 Principales aplicaciones empresariales de RPA

Servicio al cliente	ERP	Servicios financieros	Cuidado de la salud	Recursos humanos	Gestión de la cadena de suministro
Automatizar tareas de centro de contacto como: -Verificación de firma -Carga de documentos escaneados -Verificación de documentación para autorizaciones	-Contabilidad general -Contabilidad operativa -Informes transaccionales -Presupuesto	-Pagos de divisas -Apertura y cierre de cuentas -Gestión de solicitudes de auditoría -Procesamiento de reclamos de seguros	-Manejar registros de pacientes -Atención al cliente -Administración de cuentas -Facturación -Informes Análisis	-Actualización de información de empleados -Procesamiento de envío de horas de trabajo	-Procesamiento y pagos de pedidos -Monitoreo de inventarios -Rastreo de envíos

3.2.3. Principales proveedores de RPA

- **ABBYY:** líder en desarrollo de herramientas de OCR. Últimamente ha incursionado en otros casos de uso.
- **Automation Anywhere:** proporciona una plataforma orientada a los procesos de adquisiciones, pagos, cotizaciones, recursos humanos, procesamiento de reclamos y otros procesos administrativos.
- **Blue Prism:** trabaja fundamentalmente con industrias reguladas, ofreciendo robots que se definen y administran en forma centralizada y que están alineados con el escritorio.
- **Kryon:** proporciona capacidades de automatización de ciclo completo que incluyen minería de procesos, gobernanza y módulos de inteligencia artificial que pueden ampliar las capacidades de RPA.
- **NICE:** centrado fundamentalmente en mejorar la interacción de clientes con call centers y en múltiples puntos de contacto.
- **Pegasystems:** líder en herramientas de gestión de procesos de negocio y se expandió a RPA con la adquisición de OpenSpan en 2016.
- **UiPath:** plataforma abierta para ayudar a las organizaciones a automatizar de manera eficiente los procesos comerciales. Cuenta con una serie de módulos que

permiten analizar a una organización para determinar qué procesos pueden ser automatizados. Ellos son:

- Process mining: analiza los datos del negocio para buscar fórmulas de mejora gracias a los robots y las tareas repetitivas.
- Task capture: documento los flujos de trabajo rápidamente y sin errores.
- Automation hub: administra el proceso de automatización de forma productiva.

Tanto el Gartner Magic Quadrant como el Forrester Wave nombran a UiPath como líder en el segmento RPA.

Como sus ventajas se ubican la gran capacidad de automatización robótica de procesos, su mayor agilidad y su gran escalabilidad.

3.3. Reconocimiento óptico de caracteres⁹

OCR es una tecnología que posibilita la conversión de imágenes de texto impresas o escritas en texto para que pueda ser leído por una computadora. Básicamente OCR usa algoritmos de reconocimiento de imágenes para analizar formas y patrones que en ellas se encuentran, comparándolos luego con una base de datos de caracteres conocidos y así determinar qué texto contienen. Los datos generados pueden ser así usados para realizar análisis, optimizar operaciones, automatizar procesos y mejorar la productividad.

Como se mencionaba en la sección 3.2, OCR puede combinarse con otras soluciones de inteligencia artificial. Ejemplos adicionales a los presentados precedentemente incluyen el escaneo y lectura de matrículas y señales de tránsito en vehículos autónomos o la detección de envases de productos en imágenes publicitarias. También se lo cita como una herramienta ideal para que personas con deficiencias visuales o auditivas puedan tener acceso a documentos al convertir el texto generado con OCR en Braille o en archivos de audio.

⁹ Basado en Carles [9] AWS [2] y Jandossova Troncoso [15].

Como beneficios de OCR se citan el potencial de ayudar a las organizaciones a mejorar significativamente la eficiencia operativa y la generación de texto consultable. El primero de ellos se vincula a la posibilidad de integrar los flujos de trabajo documentales con aquellos digitales dentro de la empresa. En este sentido, el software OCR posibilita, entre otros, convertir notas escritas a mano en textos y documentos editables, encontrar documentos que cumplan un cierto criterio, escanear formularios llenados a mano para revisiones y análisis automatizados. El segundo se deriva de la digitalización de los documentos nuevos y existentes y la generación de archivo con potencial para la realización de búsquedas.

Existen algunos factores que pueden limitar los beneficios del uso de OCR. Una muy mala calidad de imagen o una muy baja resolución, sumado a una tipografía de texto poco común son citados frecuentemente. Otro factor es que el texto a reconocer esté en manuscrita. Si el texto no se encuentra en escala de grises, el reconocimiento se verá dificultado. Finalmente, deberá tenerse en cuenta que existen múltiples aplicaciones que realizan OCR y que no todas cuentan con la misma efectividad.

3.3.1. Funcionamiento de OCR

El software de OCR funciona siguiendo una serie de pasos que incluyen la adquisición de la imagen, el procesamiento previo, el reconocimiento del texto y el procesamiento posterior. En el primero de los pasos, los documentos son escaneados y convertidos en datos binarios. El software analiza la imagen escaneada, clasificando las áreas claras como fondo y las oscuras como texto. En el segundo paso, el software limpia la imagen y elimina los errores para prepararla para ser leída. En el tercer paso se utilizan frecuentemente los algoritmos de Coincidencia de patrones (la imagen de un carácter o glifo se compara con uno almacenado) y de Extracción de características (se descompone al glifo según las características de líneas, círculos cerrados, dirección de línea e intersecciones de línea. Estas características se usan para encontrar la mejor coincidencia con glifos almacenados). Finalmente, los datos de texto extraídos se convierten en un archivo procesable por una computadora.

3.3.2. Principales usuarios y usos de OCR

Puede utilizarse en un gran número de actividades económicas, resultando de particular utilidad en aquellas donde se maneje un gran volumen de documentos del tipo formularios impresos, facturas, contratos y otros de tipo legal (Tabla 2).

Tabla 2 Principales aplicaciones empresariales de OCR

Servicio al cliente	Banca	Cuidado de la salud	Recursos humanos	Gestión de la cadena de suministro
-Verificación de firma -Verificación de documentación para autorizaciones	-Verificación de documentación para concesión de préstamos -Lectura de cheques -Verificación de documentos de identidad y de firma electrónica	-Procesar registros de pacientes (tratamientos, pruebas, registros hospitalarios, pagos de seguros)	-Verificación de información de empleados	-Rastreo de etiquetas, facturas, recibos y otros documentos

3.3.3. Algunas herramientas disponibles de OCR

- **ABBYY:** líder en desarrollo de herramientas de OCR. Incluye funciones para la eliminación de errores de impresión.
- **Adobe Acrobat Pro DC:** permite convertir archivos PDF a Word, Excel y otros formatos. También cuenta con una función de OCR para convertir los documentos escaneados en texto editable.
- **Nanonets:** se basa en inteligencia artificial para automatizar la extracción de datos de documentos/imágenes. Extrae datos de archivos PDF o imágenes y los exportar a archivos CSV, Excel, JSON o XML.
- **OmniPage Ultimate:** es citado como uno de los más populares por PCWorld. Incluye funciones para editar documentos PDF eliminando los errores de impresión.

- **Amazon Textract:** es un servicio en la nube de Amazon que combina OCR con aprendizaje automatizado para extraer en forma automática texto, escritura a mano y datos de documentos escaneados (PDF). Genera archivos de formatos DOCX, ODT o TXT. Carece de corrección ortográfica y solo reconoce PDFs.
- **Simple OCR:** a diferencia de las anteriores es de uso gratuito y no requiere instalación. Puede convertir imágenes y PDFs en texto editable. El usuario puede seleccionar el tipo de archivo de destino (DOCX, DOC, PDF o TXT).

3.4. Procesamiento de lenguaje natural

3.4.1. Funcionamiento de NLP

El Procesamiento del Lenguaje Natural (NLP) se refiere al conjunto de técnicas y herramientas utilizadas para hacer que las computadoras comprendan, interpretan y generen lenguaje humano. El NLP utiliza una combinación de técnicas de lingüística, aprendizaje automático y procesamiento de señales para realizar estas tareas.

El proceso de NLP generalmente comienza con la recolección de datos de lenguaje humano, como texto o audio. Luego, estos datos se preprocesan para limpiarlos y normalizarlos, lo que puede incluir eliminar caracteres no deseados, corregir errores ortográficos y convertir todo el texto en minúsculas.

A continuación, se aplican diversas técnicas de análisis de lenguaje natural para extraer información significativa de los datos. Estas técnicas pueden incluir:

- **Análisis sintáctico:** identifica la estructura gramatical de una oración, incluyendo los sustantivos, verbos, adjetivos, etc.
- **Análisis semántico:** determina el significado de las palabras y frases en contexto.
- **Análisis pragmático:** evalúa el significado en el contexto de la situación en la que se usa el lenguaje.
- **Análisis de sentimiento:** evalúa el tono emocional del texto.
- **Extracción de información:** identifica información específica, como nombres de personas o lugares, fechas, números, etc.

- Resumen automático: condensa un texto más largo en una versión más corta pero que aún retiene su significado principal.
- Generación de lenguaje natural: crea texto que suena natural y coherente, a menudo basado en algún tipo de entrada o instrucción.

El resultado final del procesamiento de lenguaje natural es un conjunto de datos que la computadora puede comprender y utilizar para realizar diversas tareas, como responder preguntas, traducir idiomas, generar respuestas automáticas, entre otras.

3.4.2. Principales usuarios y usos de NLP

El Procesamiento del Lenguaje Natural (NLP) se utiliza en una amplia variedad de campos y por diferentes tipos de usuarios. Aquí hay algunos ejemplos de los principales usuarios y usos de NLP:

- Empresas y organizaciones: las empresas utilizan NLP para analizar grandes cantidades de texto, como correos electrónicos, comentarios de clientes, transcripciones de llamadas, etc. Esto les permite comprender mejor a sus clientes, realizar análisis de sentimientos, automatizar respuestas y mejorar sus productos y servicios.
- Investigadores y académicos: los investigadores y académicos utilizan NLP para analizar grandes cantidades de texto en campos como la lingüística, la psicología, la sociología, la política, la medicina, entre otros. El análisis de texto les permite obtener información valiosa para investigaciones y estudios.
- Agencias gubernamentales: las agencias gubernamentales utilizan NLP para analizar grandes cantidades de texto, como informes de inteligencia, discursos políticos, medios de comunicación, etc. Esto les permite detectar amenazas a la seguridad nacional, realizar análisis de tendencias y tomar decisiones informadas.
- Asistentes virtuales y chatbots: los asistentes virtuales y chatbots utilizan NLP para comprender el lenguaje natural de los usuarios y responder a sus preguntas y solicitudes. Esto se utiliza en aplicaciones de atención al cliente, soporte técnico y servicios de asistencia en línea.

- Traducción automática: la traducción automática utiliza NLP para traducir texto de un idioma a otro. Esto se utiliza en aplicaciones como Google Translate, que utiliza técnicas de NLP para comprender el texto de origen y generar una traducción precisa.
- Análisis de redes sociales: las empresas y organizaciones utilizan NLP para analizar los comentarios y publicaciones en las redes sociales para comprender mejor a sus clientes, realizar análisis de sentimientos y detectar problemas tempranos.

En resumen, el NLP se utiliza en una amplia variedad de campos y por diferentes tipos de usuarios para comprender el lenguaje humano, tomar decisiones informadas y mejorar la eficiencia y la calidad de los productos y servicios.

3.4.3. Algunas herramientas disponibles de NLP

Existen muchas herramientas disponibles de Procesamiento del Lenguaje Natural (NLP) para ayudar a analizar, procesar y entender el lenguaje humano. Aquí hay algunas de las herramientas de NLP más populares:

- Natural Language Toolkit (NLTK): una biblioteca de Python gratuita y de código abierto que proporciona herramientas para el procesamiento de lenguaje natural, como análisis de sentimientos, extracción de entidades, etiquetado de partes del discurso, análisis de sintaxis y más.
- spaCy: una biblioteca de procesamiento de lenguaje natural en Python que ofrece análisis de texto en varios idiomas, etiquetado de partes del discurso, análisis sintáctico, identificación de entidades y más.
- Gensim: una biblioteca de Python para modelado de temas y análisis semántico, que incluye herramientas para la búsqueda de similitud entre documentos, identificación de temas y más.
- Google Cloud Natural Language API: una API de Google que proporciona herramientas para analizar y comprender texto, incluyendo análisis de sentimientos, extracción de entidades, análisis sintáctico y más.

- OpenNLP: OpenNLP es una biblioteca de NLP de código abierto que puede ser utilizada en proyectos .NET a través de un puente de interoperabilidad Java-.NET. OpenNLP proporciona herramientas para el etiquetado de partes del discurso, análisis sintáctico, identificación de entidades y más.
- Accord.NET: Accord.NET es un conjunto de bibliotecas y herramientas de aprendizaje automático para .NET que incluye herramientas para el procesamiento de texto y NLP. Accord.NET proporciona herramientas para la tokenización, etiquetado de partes del discurso, análisis sintáctico y más.

Estas son solo algunas de las herramientas de NLP disponibles en el mercado actualmente. Cada una ofrece diferentes capacidades y características, por lo que es importante investigar y elegir la herramienta que mejor se adapte a tus necesidades específicas de procesamiento de lenguaje natural.

3.5. Metodología SCRUM¹⁰

Scrum es un marco ágil para administrar y completar proyectos complejos. Fue desarrollado en la década de 1990 como una forma de ayudar a los equipos a trabajar juntos de manera más efectiva y entregar productos de alta calidad en un período de tiempo más corto. Al igual que un equipo de rugby (de ahí el nombre) cuando entrena para un partido, scrum anima a los equipos a aprender a través de las experiencias, a auto-organizarse mientras trata un problema y a reflexionar sobre sus victorias y derrotas en la búsqueda de la mejora continua.

Scrum está diseñado para ser simple y flexible, y se basa en la comunicación regular y la colaboración entre los miembros del equipo para tener éxito. Se usa comúnmente en el desarrollo de software, pero se puede aplicar a cualquier proyecto complejo e innovador.

Las metodologías ágiles requieren seguir ciertas prácticas, actividades, roles y entregables que pueden que no se adapten a nuestros equipos de trabajo o proyectos.

Los principios básicos de Scrum son:

¹⁰ Basado en NIMBLE [17] y ATLASSIAN [1].

- Enfatizar el trabajo en equipo y la colaboración
- Priorizar el trabajo de mayor valor
- Permitir flexibilidad y adaptabilidad frente al cambio

En Scrum, un proyecto se divide en pequeñas piezas incrementales llamadas "sprints". Cada sprint tiene una duración de tiempo de generalmente una o dos semanas, durante las cuales un equipo multifuncional trabaja para completar un conjunto específico de objetivos.

El proceso de Scrum es guiado por un "Scrum Master", que es responsable de garantizar que el equipo siga el marco de Scrum y de ayudar al equipo a eliminar cualquier obstáculo que pueda estar bloqueando su progreso. El equipo también cuenta con la ayuda de un "Product Owner", que es responsable de representar las necesidades y objetivos de las partes interesadas y de priorizar el trabajo a realizar.

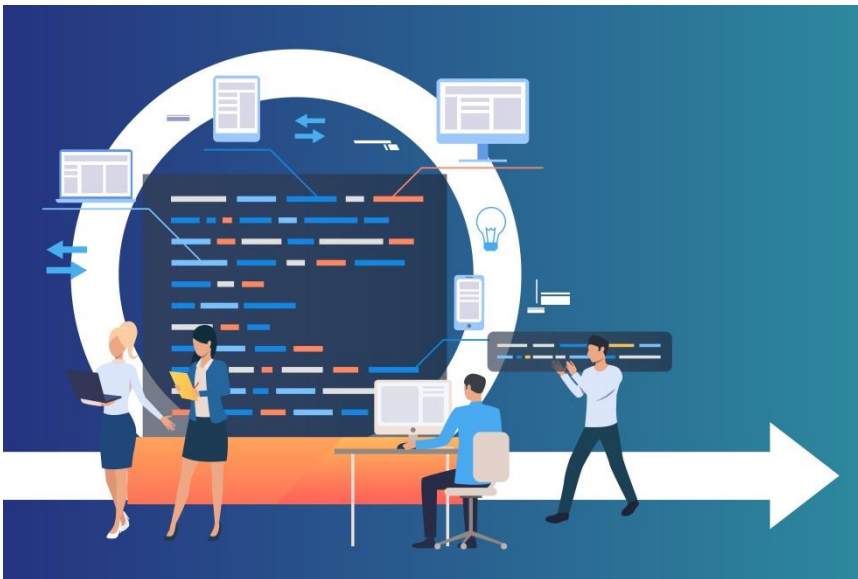


Ilustración 7 El ciclo de SCRUM

Fuente: diseñado por katemangostar / Freepik

Scrum cuenta con una serie de “artefactos” que son herramientas para solucionar problemas. Ellos son:

- **Backlog del producto:** lista principal del trabajo que debe realizar el Product Owner. Se trata de una lista dinámica de funciones, requisitos, mejoras y

correcciones que actúa como la entrada para el backlog del sprint. Es la lista de “cosas por hacer”.

- **Backlog de sprint:** se trata de la lista de elementos, historias de usuario o correcciones de errores, seleccionadas por el equipo de desarrollo para su implementación en el ciclo actual de sprint. Antes de cada sprint, el equipo elige en la reunión de planificación los elementos del backlog sobre los que trabajará.
- **Incremento:** es el producto final utilizable de un sprint

3.5.1. Protocolos o eventos de SCRUM

Algunos de los componentes más conocidos del marco de trabajo de scrum son el conjunto de eventos secuenciales, protocolos o reuniones que los equipos de scrum realizan de forma periódica.

A continuación, se muestra una lista de todos los protocolos clave en los que un equipo de scrum puede participar:

- **Organización del backlog:** este evento, que a veces se conoce como limpieza del backlog, es responsabilidad del Product Owner. Sus principales tareas son dirigir el producto hacia la visión que de él tiene y estar al tanto del mercado y los clientes. La lista se actualiza, utilizando los comentarios de los usuarios y del equipo de desarrollo para ayudar a priorizar y mantenerla limpia y así poder trabajar sobre ella en cualquier momento.
- **Sprint planning:** el trabajo a realizar (alcance) durante el sprint actual se planifica durante esta reunión por todo el equipo de desarrollo. Esta reunión es dirigida por el scrum master y es donde el equipo decide el objetivo del sprint. A continuación, se agregan historias de usuario específicas al sprint desde el backlog del producto. Estas historias siempre se alinean con el objetivo y también son acordadas por el equipo de scrum para que sean factibles de implementar durante el sprint.

Al final de la reunión de planificación, cada miembro de scrum debe tener claro qué se puede entregar en el sprint y cómo se puede entregar el incremento.

- **Sprint:** un sprint es el periodo real en que el equipo de scrum trabaja de forma conjunta para finalizar un incremento. La duración de un sprint suele ser de dos semanas, aunque algunos equipos manifiestan que les resulta más fácil una semana para el alcance o un mes para entregar un incremento valioso. Durante este periodo, el propietario del producto y el equipo de desarrollo podrán renegociar el alcance, en caso de que sea necesario.

Todos los eventos (desde la planificación hasta la retrospectiva) tienen lugar durante el sprint. Una vez que se establece un determinado intervalo de tiempo para un sprint, debe seguir siendo coherente durante todo el periodo de desarrollo. Esto ayuda al equipo a aprender de las experiencias pasadas y a aplicar ese conocimiento a futuros sprints.

- **Scrum diario o reunión rápida:** se trata de una reunión diaria de muy corta duración que tiene lugar siempre a la misma hora (normalmente, por las mañanas) y en el mismo sitio para facilitar las cosas. Esta reunión también se denomina "reunión rápida diaria" y con ello se hace hincapié en que debe ser rápida. El objetivo del scrum diario es que todos los miembros del equipo estén en sintonía, se adecuen al objetivo del sprint y dispongan de un plan para las próximas 24 horas.

La reunión rápida es el momento de expresar cualquier inquietud que se tenga con el cumplimiento del objetivo del sprint o de notificar los impedimentos existentes.

Una forma habitual de realizar la reunión rápida es que cada miembro del equipo responda tres preguntas en el contexto de alcanzar el objetivo del sprint:

- ¿Qué hice ayer?
 - ¿Qué tengo planificado para hoy?
 - ¿Hay algún obstáculo?
- **Revisión de sprint:** al final del sprint, el equipo se reúne en una sesión informal para ver una demostración o inspeccionar el incremento. El equipo de desarrollo muestra los elementos del backlog que ahora están "finalizados" a las partes

interesadas y a los compañeros de equipo para recibir comentarios. El propietario del producto puede decidir si lanza o no el incremento, aunque en la mayoría de los casos el incremento se lanza.

- **Retrospectiva de sprint:** aquí el equipo se reúne para documentar y analizar qué ha funcionado y qué no ha funcionado en un sprint, un proyecto, en las personas o relaciones, herramientas o incluso para determinados protocolos. La idea es crear un lugar donde el equipo pueda centrarse primordialmente en lo que salió bien y en lo que debe mejorarse para la próxima vez, y menos en lo que salió mal.

3.6. Gobierno de datos ¹¹

El gobierno de datos ayuda a las organizaciones a manejar de un mejor modo la disponibilidad, usabilidad, integridad, y seguridad de los datos empresariales. Su relevancia se ha puesto de manifiesto no solo con el advenimiento de la Big Data sino también por el mayor número de regulaciones que velan por la protección de los datos personales y por el creciente número de fallas de ciberseguridad.

En materia regulatoria, dentro del sector financiero global existen las regulaciones de la Securities and Exchange Commission (Comisión del Mercado de Valores), las derivadas de las Leyes Dodd-Frank (Estados Unidos) y General Data Protección Regulation (GDPR, Europa), y la Markets in Financial Instruments Directive (MiFID, Europa). La GDPR, por ejemplo, requiere que las empresas que manejan datos de ciudadanos de la Unión Europea cumplan con estrictas normas de privacidad de datos. Las multas por no cumplimiento pueden alcanzar hasta un cuatro por ciento su facturación global anual.

A nivel local –inspirada en la GDPR- también se ha emitido normativa vinculada a la privacidad de los datos. Es así que la Ley N° 18.331¹² de Protección de Datos Personales y Acción de Habeas Data define al derecho a la protección de datos

¹¹ Esta sección se base en Earl [11], Notas de estudio de la materia Fundamentos de Big Data de la Maestría en Big Data, Harmon and Gannon [14] y Taylor [19]. ..

¹² IMPO – Centro de Información Oficial, “Ley N°18331”, agosto 2008, [Online], Available: <https://www.impo.com.uy/bases/leyes/18331-2008> .

como inherente a la persona humana y por tanto comprendido en el artículo 72 de la Constitución de la República. Esa misma Ley crea la Unidad Reguladora y de Control de Datos Personales como órgano de control.

Las organizaciones y sus datos están bajo constante amenaza. Según un estudio del Departamento de Justicia de Estados Unidos, más de 4.000 ataques de *ransomware* ocurren por día. Otras pérdidas de datos se producen sin intención maliciosa y ocurren al interior de una organización por accidente o por un evento no previsto.

El gobierno de datos trasciende el manejo de la información para abarcar a varios procesos de negocio y convertirse en un facilitador del gobierno empresarial de la información (Enterprise Information Management – EIM), término que refiere a la disciplina que trata como las organizaciones actúen en base al reconocimiento del valor de los datos. Incluye el uso de la tecnología y procesos para elevar a los datos al status de activo empresarial.

El gobierno de datos acompaña un amplio set de estrategias y funciones vinculadas a los datos. Entre ellas se encuentran Acceso y disponibilidad de los datos, Integridad de los datos, Trazabilidad de los datos, Prevención de pérdida de datos, Seguridad de los datos, Sincronización de datos, y Manejo de los datos maestros.

Están comprendidos dentro de EIM varios términos que suelen confundirse. Así por ejemplo, el usual confundir manejo de datos con gobierno de datos. Mientras que el primero es la implementación técnica del gobierno de datos, el segundo tiene que ver con roles y responsabilidades, políticas, definiciones, métricas, y el ciclo de vida de los datos.

Un segundo término que requiere ser aclarado es aquel de gobierno de la información. El gobierno de datos se ocupa de los datos independientemente de su significado mientras que el gobierno de la información tiene que ver fundamentalmente con el sentido que los mismos tienen para la organización.

3.6.1. Gobierno de datos como forma de generar valor para la organización

Una estrategia de gobierno de datos inteligente acompañada de una buena solución tecnológica pueden ayudar a la organización a manejar correctamente sus datos, cumplir con los requisitos regulatorios y apoyarla en el camino a la transformación digital.

Son varias las ventajas de alcanzar un buen gobierno de datos. Algunas de ellas refieren a:

- Mejorar la calidad de los datos
- Incrementar el valor de los datos
- Aumentar el cumplimiento regulatorio
- Mejorar la toma de decisiones basadas en datos
- Mejorar el desempeño organizacional
- Favorecer el uso compartido de datos dentro y fuera de la organización
- Incrementar el acceso y disponibilidad a los datos
- Mejorar las consultas
- Reducir datos vinculados a la pérdida de datos
- Establecer reglas para el manejo de los datos

El alcanzar estas ventajas no ocurre de manera espontánea. Las organizaciones deben invertir recursos y comprometerse para ello.

3.6.2. Creación de un programa de gobierno de datos

Los pasos requeridos para crear un programa de gobierno de datos se citan a continuación:

- Definir la visión, metas, y beneficios
- Analizar el estado actual del gobierno y la administración de datos

- Desarrollar una propuesta/hoja de ruta basada en los dos primeros pasos, incluyendo un plan inicial
- Diseñar y desarrollar el programa
- Implementar el programa
- Monitorear y medir el desempeño
- Mantener el programa

No hay una forma única de crear un programa de gobierno de datos. El diseño y la implementación serán tan complejos como lo requiera la sofisticación y la naturaleza de la organización. Algunas se centrarán en el cumplimiento regulatorio mientras que otras pondrán el énfasis en generar mayores beneficios.

El gobierno de datos ha evolucionado recientemente hasta llegar a definir su alcance, el cual comprende a los siguientes elementos (Ilustración 8) según uno de los marcos de referencia (DAMA):



Ilustración 8 Elementos del gobierno de datos

Fuente: marco DAMA internacional

Se describe en la **Tabla 3** el significado de cada uno de los elementos del alcance:

Tabla 3 Alcance del marco de gobierno de datos

Componente	Definición
Gestión de Arquitectura de datos	Definir el marco de referencia para gestionar los activos de datos
Desarrollo de datos	Análisis, diseño, implementación, testeo, puesta en producción y mantenimiento
Operación de base de datos	Proveer soporte desde la adquisición de los datos hasta su depuración
Gestión de la seguridad de datos	Asegurar la privacidad, confidencialidad y acceso apropiado a los datos
Gestión de calidad de datos	Definir y monitorear la calidad de datos
Gestión de datos maestros	Gestionar las versiones y réplicas de datos
Gestión de DW y BI	Habilitar reportes y análisis de datos
Gestión de documentos	Gestionar los datos que están por fuera de las bases de datos
Gestión de metadata	Integrar, controlar y proveer los datos de los datos

3.6.3. Marco para un gobierno de datos

Un programa de gobierno de datos debe implementarse tomando en cuenta los intereses, así como las necesidades, presupuesto, y capacidades de la organización. No necesariamente todos sus elementos tendrán el mismo nivel de relevancia en un momento dado del tiempo. Para guiar a las organizaciones en la implementación de los elementos existen los denominados marcos de gobierno de datos y todos ellos tienen en común que consideran a las personas, los procesos, y la tecnología.

La

Tabla 4 muestra los componentes más regularmente presentes en un marco de gobierno de datos, así como sus cualidades más destacadas.

Tabla 4 Componentes habituales de un marco de gobierno de datos

Liderazgo y Estrategia
Obtener el apoyo de la alta gerencia y alinear al gobierno de datos a la visión, misión, y estrategia de la organización
Roles y Responsabilidades
Identificar y empoderar a los miembros del equipo con responsabilidades sobre el programa de gobierno de datos y a otros que deban rendir cuentas por los datos.
Políticas, Procesos y Estándares
Desarrollar y hacer cumplir las políticas, procesos y estándares para el gobierno y el manejo de datos
Métricas
Medir y reportar el desempeño del programa de gobierno de datos y de los resultados de mayor alcance de las actividades de gestión de datos.
Herramientas
Desarrollar y utilizar un rango de herramientas de software para apoyar al gobierno de datos y su gestión.
Comunicación y colaboración
Adoptar métodos específicos para la comunicación de los resultados y para la colaboración en tanto la implantación del gobierno de datos como en su gestión.

3.6.4. Hoja de ruta para implementar gobierno de datos

Una hoja de ruta para implementar gobierno de datos consta de un plan detallado con una lista de tareas a realizar para mejorar el grado de avance (madurez) de un programa de gobierno de datos. Normalmente incluye las mejores prácticas en la materia y está basado en una evaluación de madurez, así como en los objetivos del negocio. Debe ser considerado como una herramienta de comunicación que proporciona un plan a la gerencia, el equipo de trabajo y a los inversores sobre cómo avanzar para alcanzar las metas de gobierno de datos.

A la hora de presentar una hoja de ruta es importante que todos los miembros de la organización puedan percibir cómo los elementos motivadores (*drivers*) de la implementación de gobierno de datos van a ser considerados. Entre los *drivers* normalmente se encuentran:

- Alcanzar definiciones empresariales sobre los datos

- Generar confianza en la calidad de los datos
- Conocer con qué activos de Business Intelligence (BI) se cuenta y retirar aquellos que ya no están siendo usados
- Comprender las métricas que aparecen en los activos de BI
- Lograr transparencia y consistencia sobre las métricas de los activos de BI
- Realizar actualizaciones de software

Normalmente en las organizaciones aparecen barreras y riesgos para la implementación exitosa de un gobierno de datos. Entre ellos se cuentan:

- Roles y responsabilidades poco claros
- Intenciones vagas y falta de capacidad para generar una estrategia
- Falta de capacidad para poner los servicios y procesos de gobierno de datos en producción
- Problemas de comunicación
- No involucramiento de la organización
- Falta de personal con las capacidades adecuadas
- Falta de apropiación de la iniciativa por parte de la alta gerencia y estilos de liderazgo
- Métricas, drivers y metas poco definidas
- Falta de preparación para un cambio cultural
- Falta de preparación del equipo de BI
- Indefinición en el enfoque de gobierno de datos
- Falta de participación de los equipos de Cumplimiento, Tecnología de la Información o del Negocio.

Una vez que se han comprendido y priorizado los *drivers* y que se ha delineado cómo evitar o mitigar riesgos es necesario establecer el grado actual de madurez de la organización y cómo va a llevarse a cabo el plan para mejorar esa evaluación.

Las organizaciones que buscan desarrollar una hoja de ruta a gobierno de datos generalmente pueden ubicarse dentro de cuatro niveles. Ellos incluyen nivel inicial, uni-dimensional o pasivo, progresivo, y completamente funcional.

En el nivel inicial pocas personas en la organización reconocen la necesidad de contar con gobierno de datos y promueven su adopción. El presupuesto para herramientas es reducido y el personal lo usa de acuerdo a su saber y entender.

En el nivel uni-dimensional, una herramienta es desplegada por un equipo, generalmente el de BI. Otros miembros de la organización como representantes del negocio, de cumplimiento, desarrolladores de BI, de ciberseguridad o infraestructura no participan de la iniciativa.

En el nivel progresivo el gobierno de datos ha sido adoptado multidimensionalmente, incluyendo a todas las áreas de la organización. A menudo se comienza con mejorar el conocimiento de la importancia de un gobierno de datos para luego pasar a calidad de datos, seguridad y políticas adecuadas.

En el nivel completamente funcional, la organización ha madurado y se encuentra trabajando en incorporar los requerimientos de gobierno de datos y su curaduría en la práctica diaria.

4. Solución

En la presente sección se describirá el funcionamiento de NOVADO así como las tecnologías empleadas, las dificultades que el proyecto enfrentó, los criterios de éxito así como lecciones aprendidas que pueden ser de utilidad para futuros proyectos innovadores. Como se comentó previamente, NOVADOC extrae de la web institucional los archivos pdf con Resoluciones, los transforma y almacena en un formato estructurado la información relevante para la construcción de estadísticas nacionales, a la vez que permite realizar búsquedas según patrones de lenguaje y a la cual se accede usando una aplicación web de un modo totalmente amigable para los usuarios. La Ilustración 9 muestra de modo esquemático cómo funciona NOVADOC.

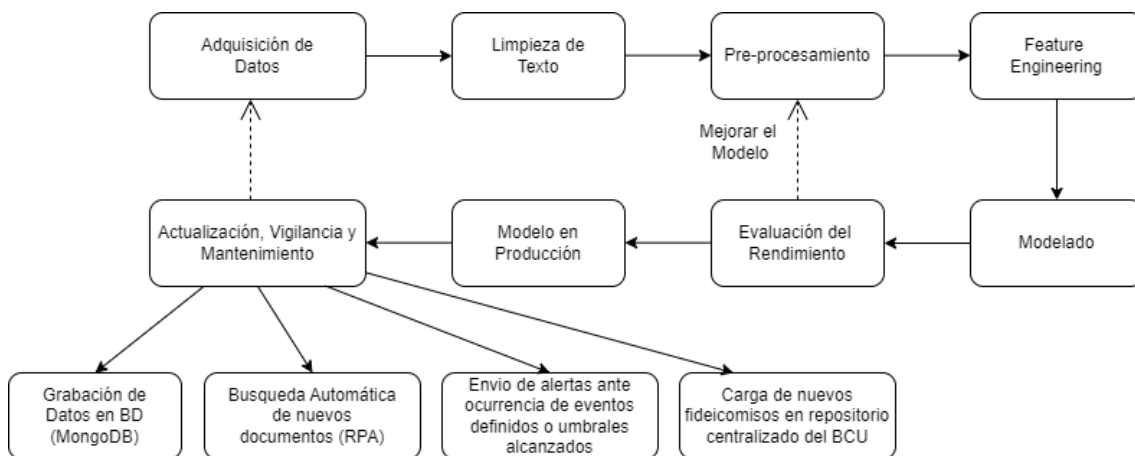


Ilustración 9 Esquema de la solución NOVADOC

Fuente: elaboración propia

4.1. Desarrollo

Esta sección describirá cómo se desarrolló la aplicación, incluyendo aquellas decisiones críticas, los desafíos enfrentados en el proyecto y su alineación a los objetivos buscados.

Se mencionarán las tecnologías utilizadas y aquellas otras que fueron evaluadas, justificando las elecciones realizadas. Es de destacar que los desafíos que el proyecto

enfrentaba por su carácter innovador y las características de la organización para la cual se desarrolló –margen reducido para extenderlo más allá del año fiscal- fueron determinantes para que las herramientas se seleccionaran al inicio del mismo.

4.1.1. Tecnologías

En cuanto a la elección del lenguaje de programación para desarrollar la aplicación web y el backend del sistema, se consideraron tres opciones: Java, C# y Python. Después de evaluar la experiencia del equipo, se optó por ASP.NET con C# debido a la disponibilidad de librerías gratuitas, documentación extensa y facilidad de uso. Para el desarrollo en C#, se seleccionó Microsoft Visual Studio Code, que cuenta con una extensión específica para este lenguaje. La Ilustración 10 muestra como el resultado de una consulta de usuario realizada en NOVADOC, la tiene se presenta en un formato totalmente amigable.

Año	Título	Path	Fecha RR-SSF	Nro. Exp.	Fecha del Contrato	Tipo Oferta	Fideicomitente	Fiduciario/a	Nombre Fideicomiso
2023	RR-SSF-2023-112.pdf	http://www.bcu.uy/ServiciosFinancieros/SSF/Resoluciones/SSF/RR-SSF-2023-112.pdf	2023-50-1-00256		07/03/2023	PRIVADA	Volmet S.A.	TMF Uruguay Administradora de Fondos de Inversión y Fideicomisos S.A.	Fideicomiso Financiero Crediton
2023	RR-SSF-2023-93.pdf	http://www.bcu.uy/ServiciosFinancieros/SSF/Resoluciones/SSF/RR-SSF-2023-93.pdf	2023-50-1-00255		17/02/2023	PRIVADA	Intendencia de Cerro Largo	República Administradora de Fondos de Inversión S.A.	Fideicomiso Financiero Cerro Largo I
2022	RR-SSF-2022-870.pdf	http://www.bcu.uy/ServiciosFinancieros/SSF/Resoluciones/SSF/RR-SSF-2022-870.pdf	2022-50-1-02014		27/12/2021	PUBLICA	Suscriptores de valores	Winterbotham Fiduciaria S.A. Administradora de Fondos de Inversión	Fideicomiso Financiero Eucalyptus del Sur
2022	RR-SSF-2022-742.pdf	http://www.bcu.uy/ServiciosFinancieros/SSF/Resoluciones/SSF/RR-SSF-2022-742.pdf	2022-50-1-02134		23/11/2022	PRIVADA	Federico Ravera	TMF Uruguay Administradora de Fondos de Inversión y Fideicomisos S.A.	Fideicomiso Financiero The Edge
2022	RR-SSF-2022-882.pdf	http://www.bcu.uy/ServiciosFinancieros/SSF/Resoluciones/SSF/RR-SSF-2022-882.pdf	2022-50-1-02050		27/10/2022	PRIVADA	Intendencia de Montevideo	República Administradora de Fondos de Inversión S.A.	Fideicomiso Financiero Montevideo Se Adelanta
2022	RR-SSF-2022-641.pdf	http://www.bcu.uy/ServiciosFinancieros/SSF/Resoluciones/SSF/RR-SSF-2022-641.pdf	2022-50-1-01350		12/10/2022	PRIVADA	Wenance Uruguay S.A.	VALONES Administradora de Fondos de Inversión y Fideicomisos S.A.	Fideicomiso Financiero Wenance I
2022	RR-SSF-2022-834.pdf	http://www.bcu.uy/ServiciosFinancieros/SSF/Resoluciones/SSF/RR-SSF-2022-834.pdf	2022-50-1-01294		11/10/2022	PUBLICA	Suscriptores de valores	República Administradora de Fondos de Inversión S.A.	Fideicomiso Financiero Fondo de Deuda para Infraestructura en Uruguay IV CAF-AM
2022	RR-SSF-2022-547.pdf	http://www.bcu.uy/ServiciosFinancieros/SSF/Resoluciones/SSF/RR-SSF-2022-547.pdf	2022-50-1-01610		24/08/2022	PRIVADA	Intendencia de Canelones	República Administradora de Fondos de Inversión S.A.	Fideicomiso Financiero Canelones 2022
2022	RR-SSF-2022-841.pdf	http://www.bcu.uy/ServiciosFinancieros/SSF/Resoluciones/SSF/RR-SSF-2022-841.pdf	2022-50-1-01868		19/08/2022	PRIVADA	Organización de las Naciones Unidas para el Desarrollo Industrial	República Administradora de Fondos de Inversión S.A.	Fideicomiso Financiero REIF
2022	RR-SSF-2022-381.pdf	http://www.bcu.uy/ServiciosFinancieros/SSF/Resoluciones/SSF/RR-SSF-2022-381.pdf	2021-50-1-01746		14/06/2022	PRIVADA	Uruguay Aracar S.A.S.	TMF Uruguay Administradora de Fondos de Inversión y Fideicomisos S.A.	Fideicomiso Financiero Aracar I
2022	RR-SSF-2022-380.pdf	http://www.bcu.uy/ServiciosFinancieros/SSF/Resoluciones/SSF/RR-SSF-2022-380.pdf	2022-50-1-01000		14/06/2021	PRIVADA	Intendencia de Maldonado	República Administradora de Fondos de Inversión S.A.	Fideicomiso Financiero Reocalización del Asentamiento Kennedy
2022	RR-SSF-2022-016.pdf	http://www.bcu.uy/ServiciosFinancieros/SSF/Resoluciones/SSF/RR-SSF-2022-016.pdf	2022-50-1-00658		09/05/2022	PRIVADA	MEVIR	Compañía Nacional Financiera Administradora de Fondos de Inversión S.A.	Fideicomiso Financiero de oferta privada Fondo Para Construcción de Viviendas
2022	RR-SSF-2022-288.pdf	http://www.bcu.uy/ServiciosFinancieros/SSF/Resoluciones/SSF/RR-SSF-2022-288.pdf	2022-50-1-00718		05/05/2022	PRIVADA	Intendencia de San José	República Administradora de Fondos de Inversión S.A.	Fideicomiso Financiero para el Financiamiento San José III
2021	RR-SSF-2021-248.pdf	http://www.bcu.uy/ServiciosFinancieros/SSF/Resoluciones/SSF/RR-SSF-2021-248.pdf	2021-50-1-00735		03/05/2021	PRIVADA	Intendencia de Maldonado	República Administradora de Fondos de Inversión S.A.	Fideicomiso Financiero Maldonado II
2021	RR-SSF-2021-179.pdf	http://www.bcu.uy/ServiciosFinancieros/SSF/Resoluciones/SSF/RR-SSF-2021-179.pdf	2021-50-1-00562		26/03/2021	PRIVADA	Noir S.A. y Múltiple S.A.	República Administradora de Fondos de Inversión S.A.	Fideicomiso Financiero Citricola

Ilustración 10 Resultado de una búsqueda de usuario en NOVADOC

Fuente: elaboración propia.

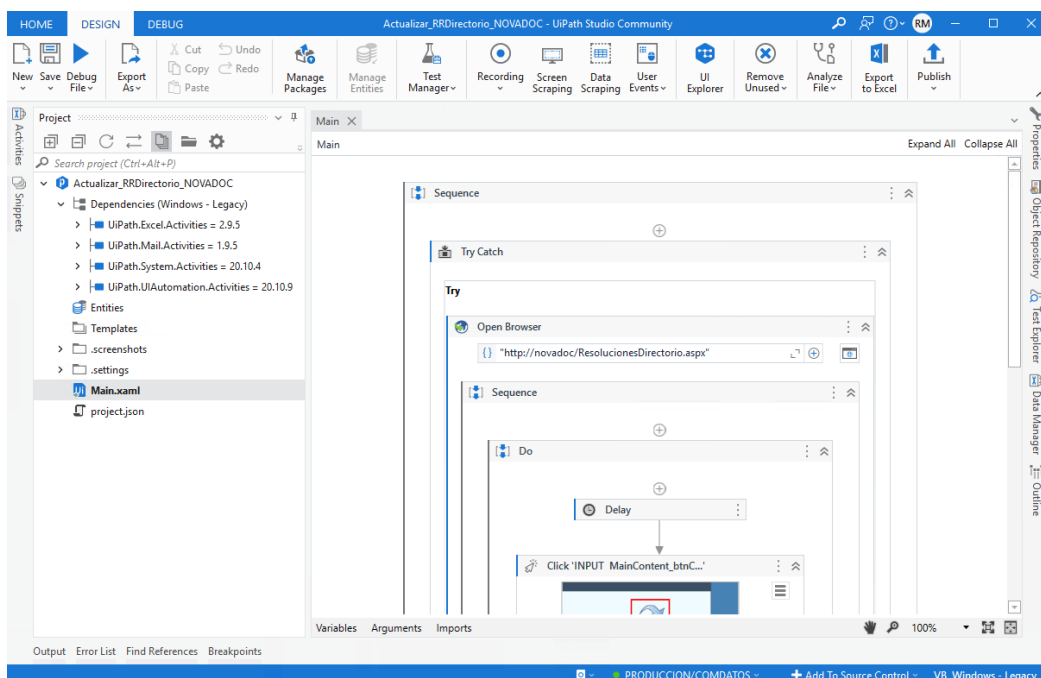
Es importante destacar que la combinación del lenguaje, el framework y las librerías seleccionadas permitieron una rápida creación del backend.

Para la lectura de documentos PDF, se decidió utilizar la tecnología OCR, probándose diversas librerías con diferentes motores de OCR y configuraciones hasta encontrar la mejor opción para pasar el contenido de los documentos PDF a texto.

Para almacenar el contenido de estos documentos, se seleccionó la base de datos documental MongoDB y su motor de consultas SQL e indexación de datos.

Para procesar los datos almacenados en la base de datos, se utilizó NLP o junto con algunas funciones de .NET. La forma de trabajo consistió en leer conjuntos de documentos con un formato similar (por ejemplo, por períodos de tiempo) y alimentar un motor de reglas para reconocer palabras clave dentro del texto de estos documentos.

Por último, para automatizar la carga de los documentos en NOVADOC de forma periódica y totalmente automática, se construyeron robots con tecnología RPA utilizando el entorno de desarrollo UiPath Studio con su orquestador en la nube para agendar los disparadores de los procesos. Este robot busca nuevas resoluciones en la página del BCU y las carga en NOVADOC si aún no se encuentran allí. La siguiente ilustración muestra parte de la construcción de los robots (Ilustración 11).



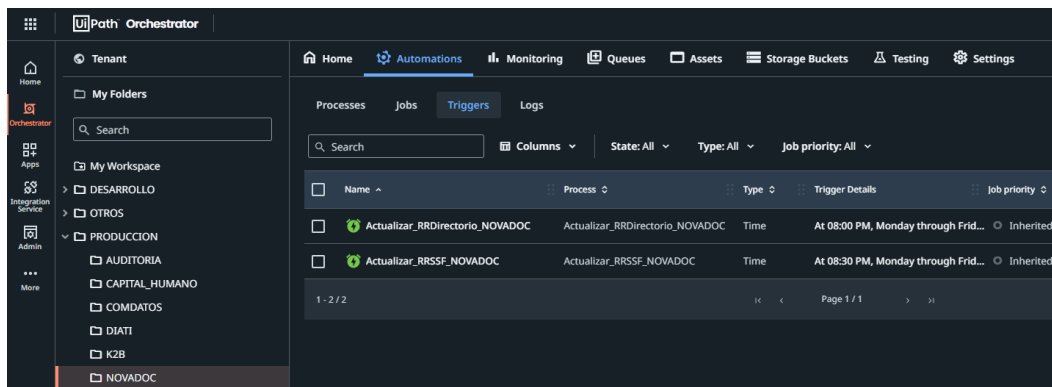
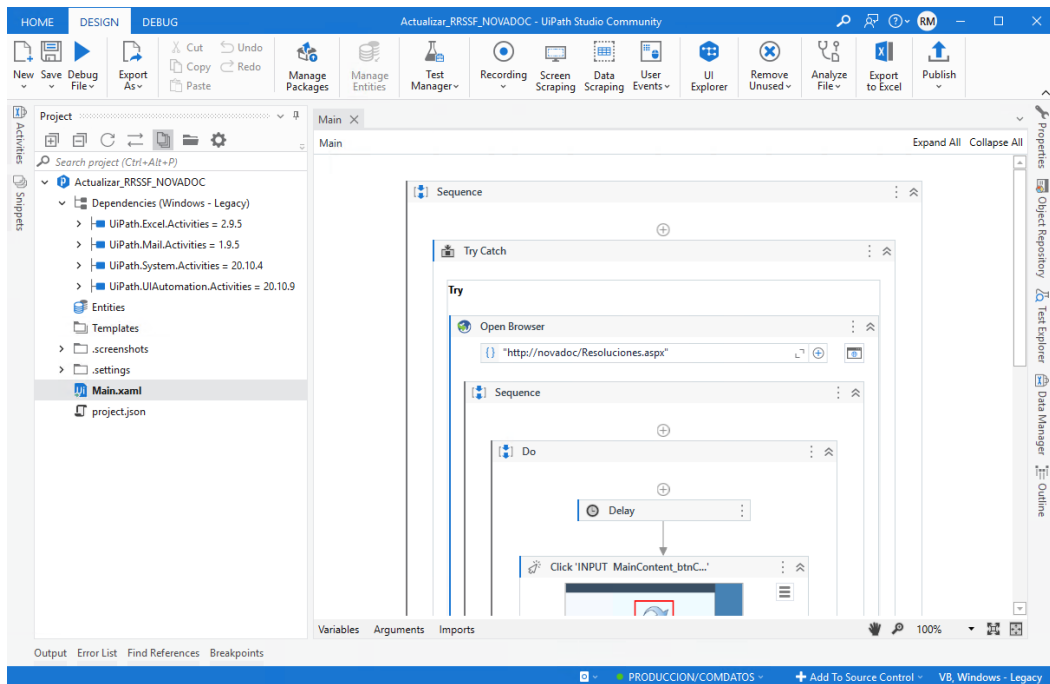


Ilustración 11 Construcción de robots en UiPath

Fuente: elaboración propia.

Durante las últimas etapas del desarrollo de NOVADOC, se adicionó la funcionalidad de enviar alertas o notificaciones por correo electrónico en caso de que se cumplan ciertas condiciones y se reconozca que se ha cargado un nuevo fideicomiso en la aplicación. Se muestra a continuación un ejemplo del mensaje emitido (Ilustración 12):

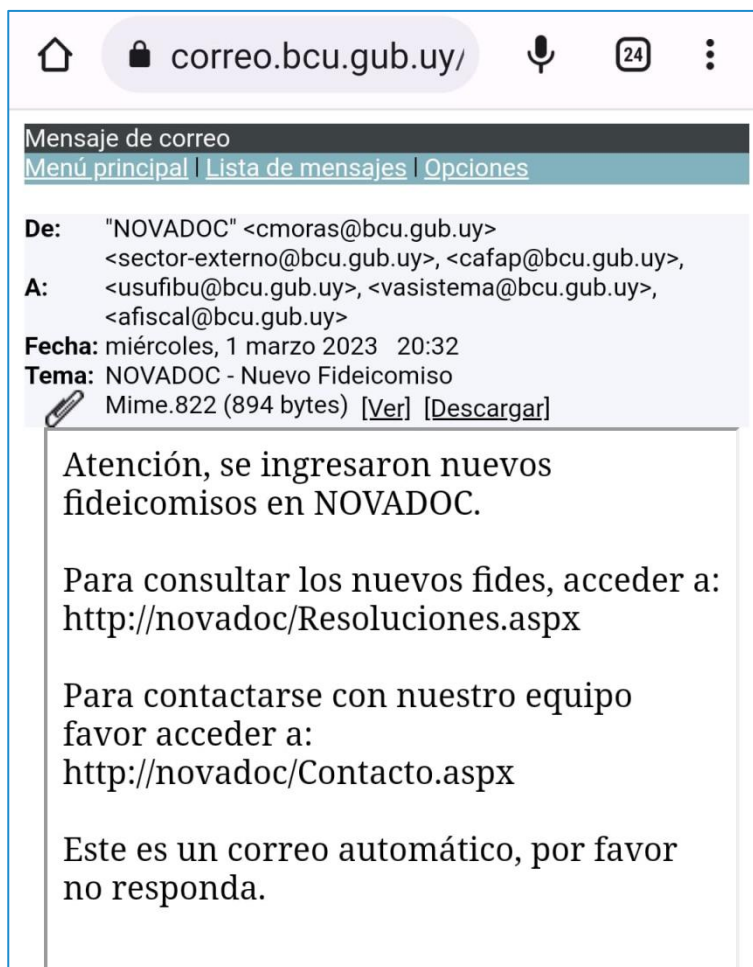


Ilustración 12 Mensaje de correo enviado por NOVADOC

Fuente: elaboración propia

También se agregó la posibilidad de poder contactarse con los desarrolladores de la aplicación en caso de querer realizar alguna sugerencia, consulta o cualquier comentario acerca de NOVADOC (Ilustración 13).



Ilustración 13 Mensaje a los desarrolladores de NOVADOC

Fuente: elaboración propia

4.2. Algunas dificultades encontradas durante el transcurso del proyecto

La falta de recursos tuvo un impacto significativo en el desarrollo del proyecto, tanto en lo relativo a recursos humanos de alta dedicación como respecto a la posibilidad de adquirir herramientas de software óptimas.

El gobierno de datos se encuentra en un estado incipiente en la organización por lo que los controles de calidad de algunos de ellos son precarios.

Se verificaron dificultades para completar los roles previstos en la metodología de gestión de proyectos. Como ejemplo, los roles de responsable de datos (data owner)

y de responsable de producto (product owner) recayeron en la misma persona, la cual tampoco tuvo posibilidades de dedicación exclusiva.

Asimismo la composición de los equipos de trabajo fue inestable durante el proyecto. La falta de estabilidad y la merma en la cantidad de integrantes en los equipos de trabajo, llevó a la concentración y sobrecarga de funciones de liderazgo y responsabilidades en pocas personas.

Limitantes económicas dificultaron la prueba de herramientas diversas de software. Restricciones presupuestarias determinaron que las herramientas de software utilizadas fueran aquellas para las que existían licencias adquiridas en la institución o que no tienen costo.

4.3. Criterios de éxito

A fin de analizar los resultados obtenidos, el grupo examinó los criterios de éxito determinados al inicio del proyecto. Estos resultados se consideran una medida objetiva del cumplimiento de las metas propuestas (**Tabla 5**).

Tabla 5 Criterios de éxito del proyecto

Objetivos académicos	
Objetivo propuesto	Desarrollar una solución para un problema de negocio real que integrara los conocimientos adquiridos en la maestría.
Criterio de éxito establecido	Contar con una solución implementada y testeada por usuarios que integrara herramientas adquiridas en la maestría y que exhibiera un desempeño superior al del estado inicial según las diferentes métricas seleccionadas.
Resultado obtenido	La solución desarrollada integra diversas herramientas del mundo de la Big Data (bases de datos no estructuradas, gobiernos de datos). Las métricas validan un desempeño superior de la solución desarrollada respecto del estado inicial.
Objetivo propuesto	Profundizar el conocimiento de tecnologías de Big Data.
Criterio de éxito	Al finalizar el proyecto, los integrantes del equipo deberán exhibir un

establecido	conocimiento superior al del punto de partida de las tecnologías relevantes.
Resultado obtenido	Los integrantes del equipo muestran mayor entendimiento de las tecnologías de Big Data, siendo capaces de explicar sus ventajas y desventajas.
Objetivo propuesto	Gestionar el proyecto siguiendo metodologías apropiadas para las tecnologías a aplicar
Criterio de éxito establecido	<ul style="list-style-type: none"> • Contar con un equipo de trabajo que cubriera los diferentes roles requeridos por la metodología. • Tener un compromiso mínimo por integrante de 15 horas semanales. • Generar retroalimentación y lecciones aprendidas en cada una de las etapas del proyecto.
Resultado obtenido	Se aplicó la metodología SCRUM de proyectos que permitió no solo detectar en forma temprana obstáculos para la culminación del proyecto sino también facilitar el compromiso de los integrantes del equipo. Si bien se identificaron los diferentes roles, algunos de ellos fueron cumplidos simultáneamente por algún miembro del equipo debido a restricciones de recursos.
Objetivos de proyecto	
Objetivo planteado	Integrar transversalmente conocimientos de diferentes áreas de negocios y del área tecnológica.
Criterio de éxito establecido	Contar con una solución tecnológica que satisficiera a los diferentes sectores del negocio.
Resultado obtenido	La solución desarrollada cuenta con un fuerte componente tecnológico e integra los requerimientos de los diferentes actores organizacionales.
Objetivo planteado	Generar una solución innovadora que mostrara a la organización los beneficios de aplicar tecnologías emergentes.
Criterio de éxito establecido	La solución incorpora tecnologías novedosas para la organización y se cuenta con mediciones de su desempeño.
Resultado obtenido	La solución es usada de modo muy significativo por los usuarios, se requirieron implementaciones adicionales por parte de otros sectores de la organización y se registró una asistencia significativa en las instancias de

difusión organizadas.	
Objetivos de producto	
Objetivo planteado	Realizar un producto que cubra las necesidades de los usuarios.
Criterio de éxito establecido	<ul style="list-style-type: none"> • Cumplir con el 100% de los requerimientos establecidos. • Lograr un nivel de satisfacción de los usuarios al finalizar el proyecto superior al 95%.
Resultado obtenido	El principal objetivo logrado, fue el desarrollar la totalidad del producto planteado, cumpliendo con la calidad requerida. Según las encuestas realizadas, se obtuvo una satisfacción de los usuarios superior al 95% logrando generar un proyecto exitoso.

El Anexo 1 muestra mediante estadísticas confeccionadas en base a los logs de los usuarios la amplia aceptación por parte de los usuarios de NOVADOC.

4.4. Lecciones aprendidas

El proyecto generó un número muy significativo de aprendizajes que constituyen en una fuente de consulta para futuras implantaciones. Se destacan en este sentido la planificación inicial, documentación por etapas, desarrollo de tácticas originales para el involucramiento de los usuarios, investigación y generación de conocimiento compartido en el equipo, incorporación de elementos de seguridad de la información, manejo de los diferentes *stakeholders*.

4.4.1. Trabajo previo al comienzo del proyecto

Como fue comentado previamente, el equipo comenzó a reunirse cerca de un año antes del inicio formal del proyecto, en donde se fueron planteando ideas, discutiendo las mismas, buscando proyectos y aplicaciones ya existentes, evaluando y descartando algunas. Luego de definida la idea central, el equipo comenzó a investigar sobre la mejor forma de implementarla, llegando a definir una lista de requerimientos iniciales.

Esto determinó que en cierta forma se haya adelantado parte del trabajo, considerando que el proyecto podía iniciarse solamente con la idea. Permitió al equipo conocer sobre el negocio, las tecnologías y ese conocimiento fue, sin dudas, muy importante para las primeras etapas del proyecto.

Las reuniones previas facilitaron el conocimiento de los integrantes del equipo así como de los roles que cada integrante podría cumplir en el proyecto.

4.4.2. Documentación

En la etapa de requerimientos se generaron una serie de documentos que fueron de mucha utilidad para la documentación académica final. Asimismo, otros documentos conteniendo los detalles de la replanificación y de los riesgos fueron claves para facilitar la generación de este documento.

La extensa documentación desarrollada no solo resultó de gran importancia para esta tesis sino también para la generación de memoria institucional en materia de gestión de proyectos innovadores. De no haberse registrado cuidadosamente cada una de las etapas, la duración del proyecto –un año calendario- habría conspirado para recordar todas las decisiones de diseño y los análisis de riesgo y costo-beneficio adoptadas. Adicionalmente, se generaron insumos muy valiosos para futuros proyectos que apliquen tecnologías innovadoras en la organización.

4.4.3. Tácticas para el involucramiento de los usuarios

Desde el inicio del desarrollo del proyecto y siguiendo la metodología SCRUM, se realizaron demos para que los usuarios pudieran comenzar a familiarizarse visualmente con el producto y de esta forma ir validando la solución desarrollada a medida que la misma tomando forma. Cada sprint tuvo como resultado una demo.

Esto motivó a los usuarios, quienes incluso comenzaron a proponer nuevas ideas sobre funcionalidades y a cuestionar sobre algunos detalles de la solución. Esta táctica novedosa facilitó mucho la implementación y la adopción por parte de los usuarios.

Instancias de demostración masiva fueron utilizadas para mostrar el producto a nivel organización. Al finalizar con las funcionalidades principales del producto, se realizó una demo dirigida a todos los funcionarios de la organización, contando con una participación aproximada de 100 personas.

4.4.4. Investigación y capacitación en nuevas tecnologías

Al inicio del proyecto, el conocimiento de las tecnologías a aplicar era dispar entre los integrantes del equipo, factor que generaba incertidumbre. Esto fue identificado como un alto riesgo para el proyecto que debía mitigarse rápidamente. Para ello, se llevó a cabo una fuerte investigación sobre las mismas que incluso incluyó un pequeño piloto en cada una de las tecnologías. Esta actividad fue clave para lo que fue la etapa de implementación, en la cual incluso se reutilizaron en gran parte los pilotos desarrollados.

4.4.5. Seguridad, registro de accesos y búsquedas

Al inicio del proyecto, se definió la forma en la cual se iba a medir el uso y los defectos de la aplicación. Se decidió llevar un archivo de logs tanto para registrar los accesos como así también las búsquedas realizadas por los diferentes usuarios. También se registraron los errores que ocurren en la aplicación y se envió alerta a los desarrolladores de NOVADOC para que puedan tomar acciones al respecto.

En cuanto a la seguridad, se optó por permitir el acceso a la aplicación mediante *login* a Windows Active Directory (WAD). Esto permitió restringir el acceso a las diferentes funcionalidades usando perfiles de usuarios.

4.4.6. Incorporar visiones de expertos al proyecto

Se realizaron múltiples entrevistas con expertos del negocio así como con representantes del sector tecnológico que no formaban parte del equipo del proyecto. Esto contribuyó enormemente a homogeneizar los conocimientos del equipo, realizar

validaciones adicionales de la solución construida y facilitar su integración al mapa de aplicaciones institucional.

5. Conclusiones

NOVADOC constituye una muy buena síntesis de las herramientas de Big Data abordadas durante el transcurso de la maestría. Integra gestión de proyectos utilizando metodologías ágiles (SCRUM), tecnologías como bases de datos no estructuradas (MongoDB), procesamiento de lenguaje natural (NLP), reconocimiento óptico de caracteres (OCR) y automatización robótica de procesos (RPA), y gobierno de datos. Es asimismo una solución aceptada y valorada por los usuarios.

El desarrollo y puesta en producción de NOVADOC produjo múltiples efectos a nivel organizacional. Entre ellos se encuentra el generar un caso de éxito en la aplicación de tecnologías y metodologías novedosas para la institución y una serie de lecciones que podrán ser incorporadas como mejores prácticas en futuros desarrollos de tipo similar. Adicionalmente, mostró la relevancia de romper estructuras de silos e incorporar a un mayor número de *stakeholders* en la gestión de proyectos de software.

Dada la necesidad cada vez mayor desde el negocio de contar con información de variado tipo para la toma de decisiones, una posible extensión del trabajo es incorporar otro tipo de documentos generados en la institución a una herramienta como NOVADOC.

Una extensión adicional consistiría en realizar búsquedas maestras que cubran todos los núcleos captados por la herramienta arrojando un resultado unificado, construido en base a vínculos. Esta extensión requeriría considerar otra base de datos, entendiéndose que las orientadas a grafos resultarían las más apropiadas.

6. Referencias bibliográficas

- [1] ATLISSIAN, “¿Qué es SCRUM?”, marzo 2023, [Online] ,Available: <https://www.atlassian.com/es/agile/scrum>
- [2] AWS, “¿Qué es el reconocimiento óptico de caracteres”, marzo 2023, [Online], Available <https://aws.amazon.com/es/what-is/ocr/>
- [3] Banco Central del Uruguay, “Memoria 2019 y Plan de Actividades 2020”, 2020, [Online], Available: https://www.bcu.gub.uy/Servicios-Financieros-SSF/Paginas/Memoria_PlanDeActividades_SSF.aspx
- [4] Banco Central del Uruguay, “Memoria 2020 y Plan de Actividades 2021”, 2021, [Online], Available: https://www.bcu.gub.uy/Servicios-Financieros-SSF/Paginas/Memoria_PlanDeActividades_SSF.aspx
- [5] Banco Central del Uruguay, “Memoria 2021 y Plan de Actividades 2022”, 2022, [Online], Available: https://www.bcu.gub.uy/Servicios-Financieros-SSF/Paginas/Memoria_PlanDeActividades_SSF.aspx
- [6] Banco Central del Uruguay, “Proceso de autorizaciones de la SSF”, [Online], Available: <https://www.bcu.gub.uy/Servicios-Financieros-SSF/Paginas/Procedimientos-y-documentaci%C3%B3n-del-proceso.aspx>
- [7] Banco Central del Uruguay, “Recopilación de Normas del Mercado de Valores”, marzo 2023, [Online], Available: https://www.bcu.gub.uy/Acerca-de-BCU/Normativa/Paginas/Reordenamiento_Merval.aspx
- [8] Banco Central del Uruguay, “Texto Ordenado de la Carta Orgánica”, noviembre 2021, [Online], Available: <https://www.bcu.gub.uy/Acerca-de-BCU/Paginas/Informacion-Institucional.aspx>
- [9] J. Carles, “Fundamentos usos y limitaciones del ocr”, junio 2014, [Online], Available: <https://geekland.eu/fundamentos-usos-limitaciones-ocr/>
- [10] A. Castellanos, “¿Qué es UiPath?”, diciembre 2020, [Online], Available: <https://www.incentro.com/es-ES/blog/que-es-uipath>
- [11] Th. Earl, W. Khatak and P. Buhler, *Big Data Fundamentals, Concepts, Drivers & Techniques*, Boston, MA: Pearson Service Tech Press, 2016.
- [12] M. Fowler, *NoSQL for Dummies*, New Jersey, NJ: Wiley, 2015.
- [13] G. Harrison, *Next Generation Databases: NoSQL and Big Data*, New York, NY: Springer Science + Business Media, 2015.
- [14] P. Harmon and T. A. Gannon, *Intelligent Data Governance for dummies*, Hoboken, NJ: John Wiley & Sons, Inc., 2022.

- [15] D. Jandosova Troncoso, “Los 20 Mejores Programas Gratuitos Y De Pago De OCR (Reconocimiento Óptico De Caracteres)”, octubre 2022, [Online], Available: <https://marketsplash.com/es/mejor-software-ocr/#AmazonTexttractH3>
- [16] G. Lawton, “Automatización Robótica de procesos (RPA)”, marzo 2023, [Online], Available: <https://www.computerweekly.com/es/definicion/Automatizacion-de-procesos-roboticos-RPA>
- [17] NIMBLE, “¿Qué es la metodología SCRUM? Y Gestión de proyectos SCRUM”, marzo 2023, [Online], Available: <https://www.digite.com/es/agile/que-es-scrum/>
- [18] A. Singh y S. Ahmad, *Data Modeling with NoSQL Database*, Boston, MA: Pearson Education, Inc., 2021.
- [19] A. G. Taylor, *Data Governance for Dummies*, Hoboken, NJ: John Wiley & Sons, Inc., 2018.

ANEXO 1: Resolución Fideicomiso Financiero

RR-SSF-2021-239 Fecha: 28/04/2021 15:51:21

Superintendencia del Uruguay

SUPERINTENDENCIA DE SERVICIOS FINANCIEROS – RESOLUCIÓN

SUPERINTENDENCIA DE SERVICIOS FINANCIEROS

VISTO: La nota presentada por TMF URUGUAY Administradora de Fondos de Inversión y Fideicomisos S.A., en su calidad de Fiduciario, solicitando la autorización de la modificación del contrato de Fideicomiso Financiero de fecha 14 de abril de 2021, modificativo del contrato de Fideicomiso Financiero denominado "Fideicomiso Financiero TIERRAS IRRIGADAS II".

RESULTANDO:

- I) Que con fecha 31/10/2018 TMF Uruguay Administradora de Fondos de Inversión y Fideicomisos S.A. en calidad de Fiduciaria, y Bearing Agro S.A. en calidad de fideicomitente constituyeron un fideicomiso financiero de oferta pública denominado "Fideicomiso Financiero TIERRAS IRRIGADAS II".
- II) Que por Resolución SSF 2018-735 de fecha 29/11/2018 la Superintendencia de Servicios Financieros resolvió Inscribir en el Registro del Mercado de Valores, los Certificados de Participación escriturales a emitirse bajo el Fideicomiso Financiero del Resultando I) por un valor de hasta U\$S 150.000.000 por parte de TMF Uruguay Administradora de Fondos de Inversión y Fideicomisos S.A. y estampar la constancia correspondiente para la inscripción del contrato de Fideicomiso en el Registro de Actos Personales.
- III) Que el 07/07/2020 la Asamblea de Titulares de Certificados de Participación emitidos en el marco del Fideicomiso del Resultando I aprobó una modificación al referido contrato de fideicomiso financiero, la cual fuera otorgada el 14/07/2020 y oportunamente inscripta en el Registro del Mercado de Valores.
- IV) Que el 06/04/2021 la Asamblea de Titulares de Certificados de Participación emitidos en el marco del Fideicomiso del Resultando I aprobó una nueva modificación al referido contrato de fideicomiso financiero, habiéndose otorgado la misma el día 14/04/2021.

CONSIDERANDO:

- I) Que la Asamblea de Titulares debidamente convocada, aprobó la modificación referida por unanimidad de presentes que representaban el 100% de los Títulos emitidos.
- II) Que no surgen objeciones para inscribir la modificación otorgada el 14/04/2021 y estampar la constancia de oferta pública en la Modificación del contrato de Fideicomiso Financiero del Visto.

ATENCIÓN: A lo dispuesto en la Ley N° 17.703 de 27 de octubre de 2003; en el Decreto N° 516/003 de 11 de diciembre de 2003; en la Recopilación de Normas

ANEXO 2:

Uso de NOVADOC

Métricas

Durante la planificación del proyecto se establecieron criterios importantes acerca de la calidad del proceso y del producto, para luego definir las métricas considerando cada uno de estos criterios. De esta manera se pudo evaluar el avance y desempeño del proyecto en distintas áreas, así como también detectar actividades que no se estuvieran

realizando correctamente. En el caso de encontrar algún desvío de las metas establecidas se podría realizar un análisis causal y, de ser necesario, llevar a cabo acciones correctivas.

Métricas de satisfacción

De forma de poder medir que se cumplió con los objetivos de calidad del producto, se realizaron encuestas de satisfacción a los usuarios.

Encuesta de satisfacción del usuario

Al final de cada sprint se realizó una demo a los usuarios mostrando las funcionalidades implementadas en el sprint. Luego de la demostración se le dio un espacio para realizar comentarios sobre el producto y si el mismo estaba cumpliendo con sus expectativas y si tenían sugerencias de mejoras para agregar en futuras versiones.

El resultado de cada sprint fue más que satisfactorio. Los usuarios remarcaron en cada una de las demostraciones realizadas que los resultados excedían sus expectativas, resaltando los beneficios de la solución e insistiendo en las posibilidades ciertas de una ampliación para otros núcleos de documentos dentro de la organización.

Detección de errores

Del análisis realizado se aprecia una satisfacción general con el producto desarrollado, detectando como mayor déficit, sobre todo para los desarrolladores, la necesidad de mensajes de error más claros o explicativos. Por este motivo fue que se agregó un archivo de registro de logs para que NOVADOC registrara más información en caso de la ocurrencia de algún error.

Registro de uso

Utilizando el registro de logs mencionado en el punto anterior, se obtuvieron métricas de utilización de NOVADOC y de las preferencias y búsquedas más comunes realizadas por los usuarios.

Tabla 6 Consultas a NOVADOC

Usuario	Cantidad de visitas
WAD\ACAVIGLIA	679
WAD\ADMINISTRATOR	543
WAD\RPABDCOM	454
WAD\ABRAGA	230
WAD\ITUVI	198
WAD\MLARRE	159
WAD\ARUDELI	158
WAD\NURIARTE	63
WAD\COLVEIRA	55
WAD\ANDELEON	51
WAD\FLDELGADO	46
WAD\IMARONAS	42
WAD\CSALVARREY	38
WAD\EROLAND	36
WAD\GUELAZQUEZ	31
WAD\ANDOMINGUEZ	26
WAD\ASILVA	26
WAD\LPINTO	26
WAD\PDELMONTE	25

WAD\MAKFERNANDEZ	24
WAD\ALEAL	20
WAD\JSANDER	18
WAD\AVIERA	15
WAD\LVIGNALI	15
WAD\CARIVERO	14
WAD\VFREIRE	14
WAD\MTABOADA	12
WAD\CHERRERA	11
WAD\MNOBELASCO	11
WAD\ROLOPEZ	11
WAD\SFREGOSI	11
WAD\AMICHALSKI	10
WAD\RSTEVENAZZI	10

Nota: se reportan en la Tabla usuarios con número de consultas mayor o igual a 10.

Por otro lado, contar con información detallada e histórica según los logs permite conocer en qué medida creció la utilización de la herramienta dentro de la organización. Se pudo observar que luego de la presentación realizada a todos los funcionarios de la institución, el uso de NOVADOC creció en un alto porcentaje -superior al 50%- y también surgieron planteos para desarrollar nuevas aplicaciones con las tecnologías utilizadas durante este proyecto.

