



**Modelo de Marketplace de datos para la universidad Simón Bolívar,
mediante el uso y aplicación de herramientas de Microsoft 365 y Azure.**

Andrés Javier Pallares

Zaida Patricia Ojeda

Universidad EAN

Facultad de ingeniería

Programa Maestría en Inteligencia de Negocios

Bogotá, Colombia

2024

Modelo de Marketplace de datos para la universidad Simón Bolívar, mediante el uso y aplicación de herramientas de Microsoft 365 y Azure.

Andrés Javier Pallares Campos

Zaida Patricia Ojeda Guzmán

Trabajo de grado presentado como requisito para optar al título de:

Magister en Inteligencia de Negocios

Director (a):

Mgtr. Luis Armando Cobo Campo

Modalidad:

Trabajo Dirigido

Universidad EAN

Facultad de ingeniería

Programa Maestría en Inteligencia de Negocios

Bogotá, Colombia

2024

Nota de aceptación

Firma del jurado

Firma del jurado

Firma del director del trabajo de grado

Agradecimientos

Agradecemos profundamente a nuestras familias por su incondicional apoyo y aliento durante todo este proceso de investigación y redacción. Su constante respaldo nos ha dado la fuerza y la motivación necesarias para alcanzar este logro. También deseamos expresar nuestra gratitud a los profesores que, a lo largo de nuestra formación académica, nos han brindado su valioso conocimiento y orientación, inspirándonos a alcanzar nuestras metas y superar desafíos. Su dedicación y compromiso han sido fundamentales en nuestro crecimiento personal y profesional.

Resumen

Las instituciones educativas generan una gran cantidad de datos, provenientes de diferentes fuentes, como los sistemas administrativos, los sistemas de aprendizaje, los sistemas de investigación, y los datos de los estudiantes y profesores.

Tradicionalmente, la centralización de datos en instituciones educativas se ha enfocado en una única área de negocio, como la administración académica o la investigación. Esto limita el acceso y conexión entre las diferentes dependencias, lo que dificulta la generación de información valiosa y aportante para los procesos institucionales.

Para superar este desafío, se ha propuesto un diseño de Marketplace de datos en la Universidad Simón Bolívar, el cual, representa un avance significativo en la gestión de la información generada por diversas fuentes dentro de la institución. Este enfoque innovador permite consolidar datos dispersos de sistemas administrativos, de aprendizaje e investigación, promoviendo la integración y conexión entre diferentes áreas de la universidad.

El Marketplace de datos rompe con la tradicional fragmentación de datos al proporcionar una plataforma unificada para que estudiantes, profesores y administrativos accedan de manera ágil y eficiente a información relevante. Esto fomenta la colaboración interdisciplinaria y la generación de conocimiento transversal, fortaleciendo la capacidad institucional para tomar decisiones fundamentadas.

Además, esta centralización no solo optimiza los procesos internos, sino que también impulsa la innovación y mejora continua al identificar patrones y tendencias previamente desapercibidas. El diseño de la arquitectura del Marketplace abarca fases críticas, desde la ingesta y procesamiento de datos hasta el aprendizaje automático y los servicios de reportes, utilizando herramientas como Power BI para la visualización interactiva de datos, resaltando la integración con Azure y office 365.

En resumen, la adopción de un Marketplace de datos en la Universidad Simón Bolívar representa un paso fundamental hacia la transformación digital, maximizando el valor de los recursos disponibles y promoviendo la eficiencia, la innovación y el éxito en la misión educativa de la institución.

Palabras clave:

Marketplace de datos: Un Marketplace de datos es un servicio que proporciona acceso a datos de diferentes fuentes a través de una interfaz unificada. Esto facilita el uso de los datos para realizar análisis, aprendizaje automático y autogestión de reportes.

Análisis de datos: El proceso de recopilar, limpiar, analizar e interpretar datos para obtener información.

Seguridad de datos: El conjunto de medidas que se toman para proteger los datos de la pérdida, el acceso no autorizado o la alteración. Esto incluye medidas técnicas, administrativas y físicas.

Infraestructura de datos: El conjunto de componentes y sistemas que se utilizan para almacenar, procesar y acceder a los datos. Esto incluye hardware, software, redes y personas.

Abstract

Educational institutions generate a large amount of data, coming from different sources, such as administrative systems, learning systems, research systems, and student and teacher data.

Traditionally, data centralization in educational institutions has focused on a single business area, such as academic administration or research. This limits access and connection between different agencies, making it difficult to generate valuable information that contributes to institutional processes.

To overcome this challenge, different technological solutions have been proposed, such as the implementation of a Data Marketplace. Data centralization is a growing trend in educational institutions. Technology solutions, such as the Data Marketplace, can help institutions leverage data more effectively to improve their processes and make better decisions.

Keywords:

Data Marketplace: A Data Marketplace is a service that provides access to data from different sources through a unified interface. This makes it easier to use the data for analysis, machine learning, and self-reporting.

Data Analysis: The process of collecting, cleaning, analyzing, and interpreting data to obtain information.

Data security: The set of measures taken to protect data from loss, unauthorized access or alteration. This includes technical, administrative and physical measures.

Data infrastructure: The set of components and systems used to store, process, and access data. This includes hardware, software, networks and people.

Tabla de contenido

	<u>pág.</u>
LISTA DE FIGURAS.....	10
LISTA DE TABLAS.....	10
1. INTRODUCCIÓN.....	11
2. OBJETIVOS.....	16
2.1. OBJETIVO GENERAL.....	16
2.2. OBJETIVOS ESPECÍFICOS	16
3. JUSTIFICACIÓN	18
4. MARCO DE REFERENCIA.....	20
4.1. ANTECEDENTES	20
4.2. GENERALIDADES DE LOS MARKETPLACE DE DATOS	21
4.3. ENFOQUE CONCEPTUAL	26
4.4. CARACTERÍSTICAS DE LOS DATA LAKE	27
4.5. BENEFICIOS DE LOS DATA LAKE	29
4.6. HERRAMIENTAS E IMPLEMENTACIÓN Y NORMATIVIDAD	32
4.7. NORMATIVIDAD Y PRIVACIDAD DE LOS DATOS	33
4.8. ANALÍTICA DE DATOS	35
5. MARCO INSTITUCIONAL	39
5.1. PRESENTACIÓN GENERAL DE LA EMPRESA	39
5.2. REFERENTES ESTRATÉGICOS.....	39
5.3. ESTRUCTURA ORGANIZACIONAL	40
5.4. PRODUCTOS O SERVICIOS OFERTADOS	42
5.5. ANÁLISIS DEL SECTOR	42
6. DISEÑO METODOLÓGICO.....	45
6.1. TIPO DE INVESTIGACIÓN	45
6.2. ANÁLISIS EXTERNO.....	46
6.3. ANÁLISIS INTERNO	49
7. DIAGNÓSTICO ORGANIZACIONAL	50
7.1. PROCESAMIENTO ESTADÍSTICO DE DATOS	51
7.2. ANÁLISIS DE DATOS.....	55
8. PLAN DE INTERVENCIÓN	58
8.1. DISEÑO DE LA ARQUITECTURA DEL MARKETPLACE DE DATOS.....	58
.....	59
9. RECOMENDACIONES Y CONCLUSIONES	82
9.1. RECOMENDACIONES.....	82

9.2.	CONCLUSIONES	83
10.	REFERENCIAS	89
A.	ANEXO. NOMBRE DEL ANEXO	93

Lista de figuras

	<u>Pág.</u>
Ilustración 1. Capas data lake	29
Ilustración 2. Elementos clave	30
Ilustración 3. Estructura Organizacional	41
Ilustración 4.. Análisis PESTEL	46
Ilustración 5.Datos que se recolectan en la universidad	52
Ilustración 6.Desafíos que se presentan en el proceso.....	53
Ilustración 7.Tipos de Análisis usados en el proceso.....	54
Ilustración 8.Acciones para mejorar la integración de datos	55
Ilustración 9.Eschema de ingesta de datos al Marketplace	59
Ilustración 10.Eschema de ingesta, procesamiento y storage de datos al Marketplace..	60
Ilustración 11.Eschema de aprendizaje automático al MarketPlace	62
Ilustración 12.Eschema de autoservicio de reportes en el Marketplace.....	66
Ilustración 13.Marketplace, parte A: fuente de datos y Data Lake	67
Ilustración 14.Marketplace, parte B: Aprendizaje Automático	68
Ilustración 15.Marketplace, parte C: Autoservicio Reporte	68
Ilustración 16.Componentes de seguridad del Marketplaces	69
Ilustración 17.Data management.....	70
Ilustración 18.Propuesta actualizada, parte A: Seguridad, fuente de datos y Data Lake.	71
Ilustración 19.Propuesta actualizada, parte B: Aprendizaje automático y gobierno de datos	72
Ilustración 20.Propuesta actualizada, parte C: Autoservicio Reporte.....	73
Ilustración 21.Cronograma del proyecto fase de definición.....	80
Ilustración 22 Cronograma del proyecto fase de diseño	81
Ilustración 23 Cronograma del proyecto fase de operación	81

Lista de tablas

	<u>Pág.</u>
Tabla 1 Oferta Académica	42
Tabla 2 Costos de capacidad y abastecimiento en la nube de Azure.....	76
Tabla 3. Tabla comparativa de universidades y plataformas.....	84

1. Introducción

Tema de investigación: la centralización de datos en la eficiencia de los procesos administrativos en instituciones educativas. Este tema de investigación es relevante para el contexto actual, ya que las instituciones educativas están cada vez más adoptando soluciones de centralización de datos para mejorar la eficiencia de sus procesos administrativos. El estudio permitirá evaluar el impacto de la centralización de datos en la eficiencia de los procesos administrativos en instituciones educativas, lo que podrá ayudar a las instituciones a tomar decisiones informadas sobre la implementación de esta tecnología.

Problema de investigación: La multiplicidad de sistemas de información que posee la universidad Simón Bolívar genera grandes volúmenes de información, causando un mayor esfuerzo para integrar los datos, inconvenientes para la generación de reportes, carga adicional de trabajo que impacta la calidad y tiempos de respuesta de los procesos. La utilización adecuada de los datos contribuye a la utilización de los recursos de manera eficiente.

A partir de la observación y experiencia propia, los funcionarios de la universidad, usualmente, se quejan del tiempo de espera en la entrega de reportes y la falta de mecanismo para que los datos sean de fácil acceso. Especialmente, cuando los informes deben ser enviados al ministerio de educación, para ser evaluados. Si los tiempos de espera son mayores que los que el proceso espera o considera apropiados, se genera insatisfacción sin importar el tiempo que ha transcurrido.

Actualmente, se tienen iniciativas de Data Warehouse que generan información valiosa, a pesar de ello, la mayor desventaja son los reportes enviados a las entidades

externas, la demora en la actualización de los datos, la falta de mecanismos de integración, el esfuerzo para mantener la equivalencia del dato, los formatos exigidos, el proceso preanalítico y la entrega de los resultados. Adicionalmente, algunos reportes, no son de fácil acceso y es difícil mantener la calidad de los datos entre el sistema académico principal y las diferentes aplicaciones que posee la universidad.

Teniendo en cuenta lo anteriormente expuesto, los desafíos que se presentan a nivel de Data Warehouse son mencionados por diferentes autores en sus publicaciones y trabajos de grado, a continuación, se presentan algunos casos que brindan un soporte con más fundamento a lo mencionado. En el trabajo de grado denominado “diseño e implementación de un Data Warehouse para la creación de informes gerenciales”, de la fundación universitaria los Libertadores, se expone que dedicar tiempo y esfuerzo a una fase de análisis profunda es crucial para el éxito de cualquier proyecto. Esta etapa permite identificar y prevenir problemas que podrían surgir en etapas posteriores, evitando así costosas reestructuraciones de procesos, mapeos o reportes. Los inconvenientes no siempre son evidentes al inicio del proyecto, a veces, solo se detectan una vez que el reporte está terminado, como incongruencias en los datos o información no agregable que conduce a resultados incorrectos; en estos casos, se requiere retroceder y corregir los errores en las etapas anteriores, lo que implica un gasto de tiempo y recursos (Gutierrez Mendieta, 2017).

Otro trabajo de grado que se alinea al tema en mención es el presentado en la Universidad Nacional de Córdoba (UNC), la cual utiliza un sistema de Business Intelligence con tecnología costosa, lo que dificulta su mantenimiento. La licencia y el soporte de la tecnología actual son demasiado costosos para la situación económica actual de la universidad, lo que dificulta la obtención de fondos para su mantenimiento.

Para solucionar este problema, la UNC busca implementar una solución Open Source Business Intelligence; esta solución permitiría a la universidad reducir costos de licencia y mantenimiento, obtener mayor independencia de proveedores y tener una mayor flexibilidad y escalabilidad (Del Giudice & Della, 2010).

En la universidad Nacional de Cajamarca, se publicó un trabajo de grado denominado: “Mejoramiento del soporte a la toma de decisiones en los procesos de admisión, matrícula y egreso de la Universidad Nacional de Cajamarca utilizando un data Warehouse”, la problemática principal se centra en la necesidad de establecer y mantener políticas de información en las universidades que aborden las demandas de datos, así como en la colaboración activa de diversos actores para asegurar los recursos necesarios y adaptar las políticas universitarias a las necesidades sociales y económicas locales. Esto se ve reflejado en la dificultad para garantizar un soporte efectivo en la toma de decisiones en los procesos de admisión, matrícula y graduación en la Universidad Nacional de Cajamarca, debido a la falta de un sistema adecuado para gestionar y analizar la información relevante. Se propone la implementación de un almacén de datos (Data Warehouse) para proporcionar apoyo en la toma de decisiones, mediante la creación de Data Marts que satisfagan las necesidades de información específicas. Los beneficios esperados incluyen una toma de decisiones más segura, eficaz y eficiente, respaldada por datos confiables y oportunos. El Sistema de Soporte de Decisiones desarrollado se ajusta a las necesidades de análisis de los responsables de tomar decisiones, brindando una base sólida para la mejora continua (Camacho & Angello, 2014).

En otro trabajo de investigación del año 2023, se presenta una tesis denominada “Data Warehouse en Azure con Metodología Hefesto para la eficiencia operativa: caso

Universidad Nacional Hermilio Valdizán”, en la cual la problemática principal que se aborda en esta investigación es la necesidad de modernizar y estandarizar los procesos operativos en la Universidad Nacional Hermilio Valdizán. Se identifica un problema en la eficiencia de estos procesos, que actualmente son mayormente manuales y carecen de una estructura tecnológica unificada. Esto no solo afecta la productividad y eficiencia de las actividades académicas y administrativas, sino que también puede generar dificultades en la toma de decisiones y aumentar los costos asociados con la gestión de recursos humanos y la operación general de la institución. La implementación del Data Warehouse en Azure con la Metodología Hefesto se propone como una solución para abordar esta problemática y mejorar la eficiencia operativa en la universidad (Angel, 2023).

Por último, en la revista científica de informática ENCRIPSTAR, se publicó un proyecto en el año 2022, con nombre “Sistema de reportes gerenciales basados en almacenes de datos, para el proceso de matriculación del Sistema de Gestión Académico de la Universidad Laica Eloy Alfaro de Manabí”, en este proyecto la problemática principal radica en la falta de un sistema de reportes basado en BI, lo que dificulta la toma de decisiones debido a la falta de acceso oportuno a información relevante y procesable en la universidad. La implementación del almacén de datos y la plataforma de inteligencia de negocios busca abordar esta problemática y mejorar la eficiencia y efectividad en la gestión académica. El proyecto incluye actividades como la obtención de los requerimientos del negocio, la identificación de indicadores clave de rendimiento (KPI) del área, el análisis de diversas fuentes de información interna y el desarrollo de un modelado dimensional basado en el esquema estrella de Kimball. Para garantizar una integración oportuna, se realiza un proceso de extracción, transformación y carga (ETL) desde dos fuentes de datos diferentes. La creación de este DW permite que el personal

administrativo del sistema de gestión académica acceda a la información relevante a través de herramientas de procesamiento analítico en línea (OLAP). Además, se complementa con herramientas para la generación de informes y la creación de un panel de control (dashboard). La integración de estos elementos resulta en una plataforma de inteligencia de negocios que respalda los requisitos de información y análisis asociados al proceso de matriculación (Carreño, Valencia, Cedeño, 2022).

La formulación de la (s) pregunta (s) de investigación.

¿El diseño de un Marketplace mejorará el proceso de integración de datos de los diferentes sistemas de información que posee la universidad Simón Bolívar, haciendo uso y aplicación de herramientas de Microsoft 365 y Azure?

2. Objetivos

2.1. Objetivo general

Diseñar un modelo de Marketplace de datos para la universidad Simón Bolívar, mediante el uso y aplicación de herramientas de Microsoft 365 y Azure, que alimente y actualice a un único catálogo de datos y que permita realizar procesos de analítica de datos, aprendizaje automático y autogestión de reportes.

2.2. Objetivos específicos

- Definir la arquitectura del Marketplace de datos para el almacenamiento y procesamiento de la información académica con las herramientas del ecosistema Microsoft 365 y Azure.
- Realizar un análisis comparativo de Marketplaces de datos en otras universidades de la región del Atlántico.
- Proponer un diseño para acceder a los datos en un servicio estándar en modalidad de autogestión y autoservicio.
- Definir el esquema de integraciones y los ambientes para la gestión de conocimiento.

3. Alcance y limitaciones

El alcance esta enmarcado en los siguientes aspectos importantes:

- Diseño de un modelo de Marketplace de datos para integrar sistemas de información en la Universidad Simón Bolívar.
- Centralización y gestión eficiente de datos a través de herramientas de Microsoft 365 y Azure.
- Mejora en la accesibilidad y coherencia de la información para la administración y la comunidad educativa.

Limitaciones potenciales:

- Resistencia al cambio por parte de usuarios acostumbrados a sistemas existentes.
- Requerimientos de capacitación para el personal en el uso de nuevas herramientas.
- Posibles obstáculos técnicos durante la integración de sistemas.

4. Justificación

Este proyecto está dirigido al beneficio de centralizar la información para obtener resultados que aporten en la mejora de procesos académicos y administrativos que maneja la universidad; los avances a nivel de tecnología y el surgimiento de herramientas a nivel de análisis de datos son un recurso al que recurre la entidad para tomar decisiones y a la vez identificar los riesgos existentes y minimizarlos.

La universidad se verá beneficiada con el diseño del Marketplace en varios aspectos, como el hecho que la información tenga la oportunidad de ser compartida de manera más amplia a nivel interno de la organización entre dependencias, generando una sinergia que contribuya a mejorar la accesibilidad y confiabilidad de los datos; el hecho de compartir los datos de manera efectiva contribuirá a que se implementen mejores estrategias sobre el Core de negocio, para lograr así un mejor posicionamiento sobre las entidades de la misma índole.

Los impactos que se podrán evidenciar con esta propuesta se verán reflejados a nivel organizacional en la reducción de costos, gracias a la simplificación y unificación de procesos que se ejecutaran de manera automatizada y con margen de error casi nulo; se evidenciará un aumento de la productividad y la eficiencia entre las diferentes áreas ya que se reutilizarán los datos como insumos entre actividades lo cual trae una reducción de tiempos en tareas que normalmente se hacían con más utilización de tiempo, así mismo “se facilita la supervisión y el gobierno de datos gracias a la descentralización y automatización de procesos” (Ruiz,2021).

El resultado de este diseño propuesto se alinea de manera estratégica con las políticas y metas de la universidad en cuanto a la mejora continua, posicionamiento y

fidelización de clientes; identificando y priorizando los objetivos a conseguir el impacto positivo será evidenciable a corto plazo, en muchos aspectos como: recursos, resultados esperados y en si a nivel de organización.

5. Marco de referencia

5.1. Antecedentes

Teniendo en cuenta el manejo de la información en instituciones educativas se evidencia que la centralización de los datos está enfocada en una única área de negocio, lo cual de alguna manera limita el acceso y conexión que se requiere entre las diferentes dependencias para generar información que sea extraída de varias fuentes y se conviertan en un insumo valioso y aportante para los procesos que manejan este tipo de instituciones; a continuación, se presentan algunos aportes asociados al manejo de este tipo de tecnologías.

Con el crecimiento exponencial de los datos, las organizaciones se han visto obligadas a mejorar su infraestructura y tecnologías para brindar cada vez más almacenamiento a los datos y maximizar así los recursos para generar valor a la información, teniendo en cuenta los beneficios ofrecidos por las nuevas tecnologías, en la universidad internacional SEK de Ecuador se propuso la implementación de un Data Lake con los datos del LMS de Canvas de la universidad, con la finalidad de contar con un repositorio que brinde más flexibilidad y exactitud al momento de realizar análisis de datos y aprendizaje automático y en si las aplicaciones que se manejan a nivel de la institución. (Mosquera, 2022).

En RISTI, la revista ibérica de sistemas de información y comunicación se publicó un artículo en febrero de 2021, que plantea el diseño de un Data Lake para la dirección de análisis universitaria, esta propuesta radica en un sistema de almacenamiento para la gestión de datos asociados a los objetivos que tiene la institución a nivel universitario; en la descripción del artículo se evidencia la definición de la arquitectura a implementar, así

como la distribución de fases en las que se almacenaran y trataran los datos para que puedan finalmente ser analizados e influyan en la toma de decisiones. Gupta (s. f.)

Como complemento a las investigaciones anteriores, se puede mencionar un artículo publicado por la revista Espacios en el año 2019, en este se presentó un trabajo denominado “Big Data e inteligencia de negocios aplicados al estudio de mercado de estudios posgraduales de la universidad autónoma de Colombia”, este artículo describe la innovadora iniciativa de la Universidad Autónoma de Colombia en el ámbito de la inteligencia de negocios y el Big Data. El objetivo central es determinar las necesidades de los estudiantes egresados de la universidad para ofrecer programas de posgrado que respondan a las demandas del mercado actual y del futuro próximo. Para ello, se ha implementado el método científico, seleccionando cuidadosamente las herramientas de apoyo y realizando pruebas con datos recopilados de diversas fuentes: redes sociales, encuestas y bases de datos estructuradas. Si bien no puede ser tomado propiamente como el diseño de un Data Lake si está asociado en su estructura ya que hace aplicación de modelo BI para evaluar y determinar las necesidades del sector productivo y también las afinidades e intereses de los egresados al momento de seleccionar programas en estudios a nivel posgradual, lo cual se convierte en una estrategia de la Universidad Autónoma para ofertar programas acorde con la demanda y tendencias y así disminuir la incertidumbre en las diferentes convocatorias. (Castillo & moreno, 2019).

5.2. Generalidades de los Marketplace de datos

Un Marketplace de datos es una plataforma en línea que permite a los usuarios vender, comprar y compartir datos de diferentes tipos (Zhang, Chen, 2018). Es una herramienta útil para las empresas y organizaciones que desean acceder a datos de alta calidad y confiabilidad para sus investigaciones y análisis de negocios. Este mercado en

línea puede ayudar a reducir los costos y el tiempo requerido para recopilar datos de diversas fuentes, este documento proporciona un marco teórico sobre el Marketplace de datos, su función, la evolución de los datos, la importancia de los Marketplace de datos en la era de la analítica así como los desafíos y oportunidades que brindan. (Hand, 2019).

4.2.1. Función del Marketplace

El Marketplace de datos es un lugar donde los usuarios pueden vender, comprar y compartir datos. Los vendedores pueden ofrecer datos de diferentes tipos, como datos financieros, datos de redes sociales, datos de marketing, datos de transacciones y muchos otros. Los compradores pueden acceder a estos datos para sus investigaciones y análisis de negocios. Además, el Marketplace de datos también permite a los usuarios compartir sus datos y obtener una compensación por ellos. Este mercado en línea ofrece una forma eficiente de acceder a datos de alta calidad y confiabilidad para las empresas y organizaciones (Witten, Hall, 2016).

4.2.2. Evolución del Marketplace de datos

El Marketplace de datos ha evolucionado significativamente en los últimos años. Antes, la recopilación de datos era un proceso costoso y lento, lo que dificultaba el acceso a datos de alta calidad. La tecnología ha cambiado esto al permitir la recopilación, el análisis y el almacenamiento de grandes cantidades de datos. Con el surgimiento de la analítica de datos, el Marketplace se ha vuelto cada vez más importante. Este Marketplace ahora incluye tanto datos estructurados como no estructurados y se ha expandido a diferentes industrias, como la salud, la banca, las finanzas y muchas otras. (Maheshwari, 2018).

4.2.3. Importancia del Marketplace de datos en la era de la analítica

El Marketplace de datos es importante en la era de la analítica debido a que los datos son la base de los análisis de negocios y la toma de decisiones. Los datos pueden proporcionar información valiosa sobre los patrones de consumo, las tendencias del mercado, las preferencias de los clientes, entre otras cosas. Además, los datos pueden ser utilizados para mejorar los procesos empresariales y la toma de decisiones. En la era de la analítica, los datos son una de las herramientas más valiosas para las empresas y organizaciones (Chiang & Wainwright, 2019).

4.2.4. Elementos de un Marketplace

Un data Marketplace es un mercado en línea donde se pueden ofrecer conjuntos de datos a los usuarios interesados. Los elementos esenciales de un data Marketplace incluyen:

- **Datos:** Los datos se refieren a información numérica, alfabética o simbólica que se recolecta a partir de la observación, medición o registro de eventos o fenómenos (Leedy, Ormrod, 2013). Pueden ser datos estructurados o no estructurados, en bruto o procesados, de diferentes fuentes y formatos. Es el elemento central de un Marketplace de datos, que se ofrecen para compartirlos, intercambiarlos, comprarlos y/o venderlos.
- **Proveedores de datos:** Son los propietarios de los datos que se ofrecen en el Marketplace de datos. Según diversos autores, los proveedores de datos son entidades que se encargan de recolectar, procesar y distribuir información a otras organizaciones o individuos que la necesitan para tomar decisiones informadas (International Association of Privacy Professionals, 2021). Algunos ejemplos de proveedores de datos son: comScore, Registraduría Nacional del Estado Civil

(RNEC), Departamento Administrativo Nacional de Estadística (DANE), ICFES (Instituto Colombiano para la Evaluación de la Educación).

- **Compradores de datos:** Son los clientes que compran los datos ofrecidos en el data Marketplace. Estos pueden ser empresas, organizaciones gubernamentales, instituciones académicas o individuos.
- **Herramientas de análisis:** Las herramientas de análisis son una parte integral de un data Marketplace, ya que permiten a los compradores de datos procesar, visualizar y analizar los datos e información con el objetivo de facilitar su interpretación y comprensión. Esta definición se basa en la explicación de Chang (2013) en su libro "Visualización de datos y diseño de información: introducción a la visualización de datos".
- **Seguridad y privacidad:** La seguridad y la privacidad son aspectos críticos en un data Marketplace, debe ser considerada desde el inicio del diseño de un sistema, producto o servicio, y no como una medida posterior o adicional. Es por ello, que el nivel de seguridad y protección de esos datos debe ser proporcional al riesgo de su uso no autorizado o inadecuado (ISO/IEC, 2018).
- **Políticas y regulaciones:** Un data Marketplace debe cumplir con las políticas y regulaciones relevantes, como las leyes de protección de datos y las leyes antimonopolio. Un concepto clave en las políticas y regulaciones de datos es el de "privacidad de datos", que se refiere al derecho de las personas a controlar la recopilación, uso y divulgación de su información personal. (Solove, Schwartz, 2011)

- **Plataforma tecnológica del Marketplace:** un conjunto de herramientas, software y hardware que permiten la recopilación, procesamiento, almacenamiento y análisis de datos en grandes volúmenes y de diversas fuentes. La plataforma tecnológica es el marco que permite la operación del data Marketplace. Debe ser segura, escalable y fácil de usar tanto para los usuarios funcionales, técnicos, proveedores y compradores de datos. (Evans, Hagel, 2018)
- **Integración:** la plataforma debe permitir la integración de datos con otras herramientas de análisis y software para facilitar su uso. Esta integración puede incluir la eliminación de datos duplicados, la corrección de errores de formato y la estandarización de datos de diferentes sistemas. (Kimball, 2002).

En resumen, los elementos clave de un data Marketplace son los datos, los proveedores y compradores de datos, la plataforma, la seguridad, los pagos, las regulaciones, la integración y la privacidad.

4.2.5. Desafíos y oportunidades con el Marketplace de datos

El Marketplace de datos enfrenta muchos desafíos, como la privacidad de los datos, la calidad de los datos, la seguridad de los datos, la propiedad de los datos y la interoperabilidad de los datos. La privacidad de los datos es un problema importante, ya que la recopilación y el uso de datos pueden violar la privacidad de las personas. La calidad de los datos también es un problema, ya que los datos pueden contener errores o ser incompletos. La seguridad de los datos es un problema importante debido a los riesgos de los ataques cibernéticos. La propiedad de los datos también puede ser un problema, ya que los datos pueden estar sujetos a derechos de autor y otras leyes. La

interoperabilidad de los datos es otro desafío, ya que los datos pueden provenir de diferentes fuentes y formatos. (Provost & Fawcett, 2013)

5.3. Enfoque conceptual

Este enfoque presenta una descripción y fundamentos que le dan soporte a la investigación, a continuación, se describen algunos términos que brindan una conceptualización asociada al tema que se aborda.

Aprendizaje automático

El aprendizaje automático (ML) es el subapartado de la inteligencia artificial (IA) que se centra en desarrollar sistemas que aprenden, o mejoran el rendimiento, en función de los datos que consumen. Inteligencia artificial es un término amplio que se refiere a sistemas o máquinas que imitan la inteligencia humana (ORACLE, 2023).

Arquitectura de datos

La arquitectura de datos es un elemento fundamental para que los sistemas de gestión de información y de organización empresarial tengan éxito. En ella se integran los modelos, las políticas y las reglas que rigen qué datos se van a recopilar; cómo van a ser almacenados, clasificados y explotados mediante la infraestructura tecnológica disponible (Cognodata, 2022).

Big Data

El **big data** describe grandes volúmenes de datos, difíciles o imposibles de procesar a través de métodos tradicionales, que sirven para obtener ideas y tomar mejores decisiones de negocios. No importa a qué sector pertenezcan, cada vez es más habitual

que las empresas colombianas recurran al análisis de sus datos para que les ayuden a optar por movimientos estratégicos más inteligentes (UNIR, 2022).

Datos gobernados

La gobernanza es la combinación de los controles, las funciones y los procesos repetibles que generan confianza en los datos y el análisis en su organización. En los entornos de inteligencia de negocios tradicionales, la gobernanza a menudo se ve como una forma de restringir el acceso o bloquear los datos o el contenido. (TABLEAU, 2022).

Marketplace

Los Marketplace de datos tienen una relación sólida con los metadatos. Un Marketplace de datos ofrece a los consumidores de datos una experiencia de compra de datos intuitiva, segura, centralizada y estandarizada. Acerca los datos a los analistas y científicos de datos mediante el uso de los metadatos subyacentes. También realiza un seguimiento de todos los productos de datos, que a menudo se almacenan en una variedad de dominios de datos (Microsoft, 2023).

5.4. Características de los data lake

De acuerdo con Quero (2020) los data lakes han venido tomando mucha importancia con los sistemas big data, ya que por la gran cantidad de datos que se extraen de los sistemas de información y dispositivos se debe garantizar que tengan un sitio donde almacenarlos en sus formatos de origen, esto dado que los datos que producen valor a una organización pueden evolucionar según el contexto y escenario donde se usen. Dentro de las acciones que permite un data lake están:

- Almacenar datos en bruto, sin procesar.

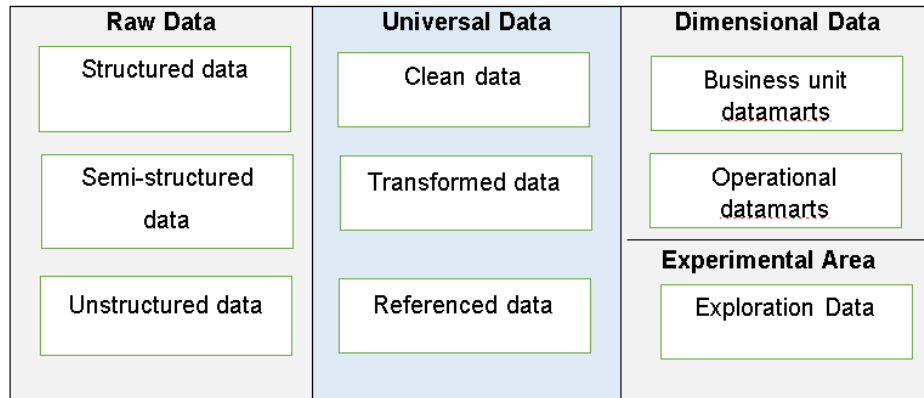
- Proveer los datos que se necesitan en los diferentes perfiles de usuario.
- Facilidad de adaptar el sistema analítico en función de las necesidades (Quero, 2020, p. 12).

Los data lake sugieren una división que puede ser adaptada de acuerdo con las necesidades que se requieran, no existe un único modelo para estructurarlo y también hay una dependencia de los requisitos, pero de manera general se podría presentar de la siguiente manera: Raw Data, Universal Data, Dimensional Data, Experimental Área y Traditional EDW.

- Data Raw: es considerado la entrada o insumo que se aproxima a los sistemas de origen, la Raw Data guarda la información como esta en los sistemas de donde proviene, si se precisa aplica una transformación sobre ellos, se le conoce como capa RDV.
- Universal Data: esta capa conocida como capa UDV evidencia el consumo de data de RDV, los datos pasan por un proceso de transformación y calidad, con la finalidad de conseguir información que este adaptada para el análisis.
- Dimensional Data: capa denominada capa DDV, aquí se crean los Data Marts, en esta capa se crean los Data Marts, aquí se permite el análisis de datos “Self - service”, en esta capa también se pueden realizar visualizaciones, reportes y se realiza combinación de datos realizada por usuarios finales.
- Experimental Area: es un área usada por usuarios avanzados es decir expertos analíticos y/o científicos de datos, con la finalidad de encontrar nuevos

conocimientos y a la vez generar nuevos modelos o algoritmos de aprendizaje avanzados (Cordova y Wenceslao, 2022, p.18).

Ilustración 1. Capas data lake



Fuente: PGA – Modelo de Arquitectura To Be

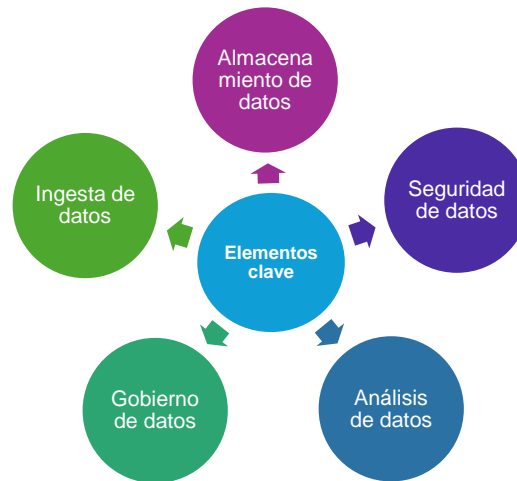
Podría indicarse que las capas están acopladas para cumplir la función que está destinada a cumplir en términos de procesamiento, gobernanza y consumo, cada experto de acuerdo con la necesidad determinara las capas o zonas requeridas e ira diferenciando las características de los datos en la medida que fluyen a través del data lake.

5.5. Beneficios de los data lake

Según Banafa (2021), un data lake es un repositorio de datos masivo que puede almacenar grandes series de datos, esto se diferencia de los data warehouse que son almacenes de datos convencionales optimizados para realizar análisis de datos en función de los atributos que tienen los datos que se guardan en ellos y usan un nivel de agregación para omitir o rechazar los datos que estén por debajo de ese nivel, por su lado los data lake están diseñados para retener todos los atributos teniendo en cuenta que no se conoce el ámbito de uso de estos

Dentro de los elementos clave para identificar los componentes de la arquitectura de los data lake están: el almacenamiento de datos, el gobierno de datos, la ingesta de datos, la seguridad de los datos y el análisis de los datos, ver ilustración 2.

Ilustración 2. Elementos clave



Nota: Elementos clave de un Data Lake. Tomado de (<https://www.dataversity.net/>, 2022)

Dentro de los beneficios o ventajas que brindan los data lake están:

- Los data lake les permiten a las organizaciones acceder de manera rápida a todos los datos.
- Los datos que están almacenados en los data lake no están limitados a datos relacionales o transaccionales.
- Evitan la necesidad de estar transfiriendo datos.
- Ofrece muchas posibilidades para los usuarios, desligando de alguna manera las limitaciones que se pueden presentar asociadas a las tecnologías de información.

- Los data lake agilizan los ciclos de desarrollo facilitando crear aplicaciones de manera rápida.
- Contribuye a la produccionización y a la analítica de datos avanzada.
- Brinda escalabilidad y flexibilidad contribuyendo con la reducción de costos operativos.
- Se genera valor a partir de datos ilimitados.
- Los data lake permiten centralizar diferentes fuentes de información.
- Facilita el acceso a los usuarios independientemente de su ubicación organizacional y/o geográfica (Banafa, 2021).

Los beneficios y ventajas que presentan los data lake son múltiples teniendo en cuenta que centralizan la información de forma óptima, dando una flexibilidad al momento de interactuar con ellos y generando valor sobre el flujo de datos para llegar a realizar análisis avanzados sobre los procesos que se manejan en una entidad, a su vez brindan varias formas de consultar los datos lo cual evita la rigidez de acceso a los mismos; el data lake puede integrar varias aplicaciones que permitirán procesar y consumir los datos con un propósito específico esto le brinda al usuario un fácil acceso a la información porque los datos siguen conservando su forma original y se almacenan de esa misma manera.

5.6. Herramientas e implementación y normatividad

Para diseñar un Marketplace de datos se debe incorporar un conjunto de soluciones tecnológicas innovadoras, disruptivas y diferenciales que ofrezcan no sólo las funcionalidades adecuadas sino también una serie de capacidades de valor añadido que permitan a la organización sacarle el máximo valor a este tipo de iniciativas. (DAMA, 2021)

Entre las tecnologías a implementar, se destacan:

- Cloud computing: es un modelo de servicio que permite acceder a recursos informáticos, como servidores, almacenamiento y aplicaciones, a través de Internet. Las soluciones de cloud computing permiten el almacenamiento y procesamiento de grandes volúmenes de datos de manera eficiente y escalable (Puttini, Mahmood, 2013).
- APIs (interfaces de programación de aplicaciones): son conjuntos de reglas, protocolos y herramientas que permiten la comunicación entre diferentes aplicaciones. Las APIs son fundamentales para la integración de sistemas y la automatización de procesos, lo que facilita el movimiento y la ingesta de datos entre diferentes plataformas y dispositivos. (Smith, 2020).
- ETL: "Extract, Transform, Load" que se refiere al proceso de extraer datos de múltiples fuentes, transformarlos para cumplir con ciertas reglas y finalmente cargarlos en un destino de almacenamiento de datos (Kimball, Ross, 2013)
- Azure: es una plataforma de nube pública que ofrece una amplia gama de servicios de computación en la nube, incluyendo el procesamiento, almacenamiento y análisis de datos, así como la gestión de aplicaciones y la creación de soluciones de inteligencia artificial y aprendizaje automático. Los

servicios de Azure están disponibles en todo el mundo y se ejecutan en una infraestructura global escalable y altamente disponible (Microsoft, 2021).

- Office365: es una suite de productividad en la nube que permite a los usuarios trabajar de forma colaborativa y eficiente desde cualquier lugar y en cualquier dispositivo con acceso a internet. Office 365 incluye aplicaciones como Word, Excel, PowerPoint, OneNote, Outlook, Teams y SharePoint, entre otras (Microsoft, 2024).

Cuadros de Mandos: herramienta de gestión que permite visualizar de manera clara y concisa los indicadores clave de desempeño de una empresa o de un proyecto, con el fin de tomar decisiones informadas y mejorar la eficiencia y eficacia de las operaciones (Loshin, 2012).

5.7. Normatividad y privacidad de los datos

La privacidad de los datos y su normatividad son conceptos críticos en la era digital actual. La privacidad de los datos se refiere al derecho de los individuos a controlar y proteger su información personal, incluyendo su recopilación, almacenamiento, uso y divulgación. La normatividad de los datos se refiere a las leyes, reglamentos y políticas que regulan el uso de los datos personales y protegen la privacidad de las personas.

La privacidad de los datos ha sido un tema importante desde que comenzó la era de la información y las comunicaciones. Con el creciente uso de tecnologías digitales y la explosión de los datos generados por ellas, la privacidad de los datos se ha vuelto aún más importante. Las personas están cada vez más preocupadas por el uso que se hace de sus datos personales y cómo se utilizan para fines comerciales, gubernamentales o de otro tipo.

En este contexto, se han desarrollado una serie de leyes y normas que buscan proteger la privacidad de los datos y establecer un marco para su uso responsable. En los Estados Unidos, por ejemplo, la Ley de Privacidad de la Información en Línea (Online Privacy Protection Act) establece requisitos para la recopilación, uso y divulgación de información personal en línea. De manera similar, la Ley de Protección de Datos de la UE (GDPR) establece reglas estrictas sobre el uso de los datos personales de los ciudadanos europeos.

Además de estas leyes, muchas empresas han establecido políticas internas de privacidad de los datos para garantizar que los datos personales de sus clientes se manejen de manera responsable. Estas políticas pueden incluir medidas de seguridad para proteger los datos de acceso no autorizado, así como procedimientos para el manejo de solicitudes de los usuarios para acceder, modificar o eliminar sus datos personales.

La normatividad de los datos también puede incluir requisitos para la divulgación de información sobre el uso de los datos. Por ejemplo, la Ley de Divulgación de Información (Freedom of Information Act) de los Estados Unidos establece requisitos para la divulgación de información por parte del gobierno federal. Del mismo modo, la Ley de Protección del Consumidor (Consumer Protection Act) de Sudáfrica establece requisitos para la divulgación de información en los contratos de los consumidores.

En conclusión, la privacidad de los datos y su normatividad son conceptos críticos en la era digital actual. La privacidad de los datos se refiere al derecho de los individuos a controlar y proteger su información personal, mientras que la normatividad de los datos se refiere a las leyes, reglamentos y políticas que regulan el uso de los datos personales y protegen la privacidad de las personas. Existen leyes y políticas a nivel nacional e

internacional que buscan proteger la privacidad de los datos y establecer un marco para su uso responsable (Contaduría General de la Nación, 2023).

5.8. Analítica de datos

5.8.1. Procesos de analítica de datos

De acuerdo con Enrique Rodal, en su podcast industrias 4.0 sobre los procesos de analítica de datos se relaciona con el término big data, ya que este asocia los términos de recolección, gestión y análisis de a gran velocidad de grandes volúmenes de datos generados por usuarios o también por máquinas, que teniendo en cuenta su tamaño se dificulta para las herramientas de software tradicionales realizar un procesamiento de la información.

La información evoluciona y se transforma, si se realiza un buen análisis ayuda en los procesos de cada organización, reduciendo tiempos y costos, también contribuye a que se tenga una efectiva trazabilidad de los servicios o productos que tienen las empresas.

El autor sugiere que antes de poner en marcha cualquier proceso de analítica de datos, lo primero es definir una estrategia que este dirigida a brindar soluciones sobre una necesidad identificada y que se pretenda resolver con ese plan estratégico, es decir que haya claridad sobre lo que se quiere monitorizar, evaluar los costos de realizar un proceso de analítica de datos, así mismo recomiendan los expertos primero poner en marcha pequeños procesos de small data que sean escalables, esto permitirá ir evaluando resultados y si son favorables se podrá seguir ampliando un proyecto a otras máquinas o procesos que se esté manejando; adicional que si trabajan proyectos pequeños los resultados serán visibles de manera más clara que un proyecto más

grande. También es importante contar con la fundamentación clara de conceptos como estadística, matemáticas, técnicas de recolección de datos, limpieza de datos, minería y programación para alinear grandes volúmenes de datos y generar información (Rodal, 2020).

En conclusión, al momento de establecer procesos de analítica de datos, se deben evaluar los factores antes mencionados y evaluar su viabilidad con respecto a los costos, teniendo la viabilidad abordar procesos pequeños que se puedan probando gradualmente.

5.8.2. Tipos de analítica de datos

Cidei (2015), plantea que la analítica de datos es una disciplina que utiliza técnicas estadísticas y de minería de datos para analizar grandes conjuntos de datos con el fin de extraer información útil y significativa. La analítica de datos se puede clasificar en diferentes tipos, según la forma en que se utilizan los datos y los objetivos que se persiguen. A continuación, se describen algunos de los tipos más comunes de analítica de datos:

- **Analítica Descriptiva:** La analítica descriptiva se utiliza para entender y describir los datos en su forma actual. Este tipo de análisis se enfoca en responder preguntas como "¿Qué ha pasado?", "¿Cómo son los datos?" y "¿Cuáles son los patrones de los datos?". El análisis descriptivo se basa en técnicas como la visualización de datos y los resúmenes estadísticos.
- **Analítica Predictiva:** La analítica predictiva se utiliza para predecir eventos futuros utilizando datos históricos. Este tipo de análisis se enfoca en responder preguntas como "¿Qué podría pasar en el futuro?", "¿Cuál es la probabilidad de

que algo suceda?" y "¿Qué factores pueden afectar el resultado?". El análisis predictivo se basa en técnicas como la regresión y el análisis de series de tiempo.

- **Analítica Prescriptiva:** La analítica prescriptiva se utiliza para recomendar acciones específicas para lograr un resultado deseado. Este tipo de análisis se enfoca en responder preguntas como "¿Qué deberíamos hacer?" y "¿Cómo podemos lograr nuestros objetivos?". El análisis prescriptivo se basa en técnicas como la optimización y el análisis de simulación.

Concluyendo, existen diferentes tipos de analítica de datos que se utilizan para entender y extraer información significativa de grandes conjuntos de datos. Cada tipo de analítica se enfoca en un objetivo específico y utiliza diferentes técnicas y herramientas. La elección del tipo de analítica de datos a utilizar dependerá de los objetivos y necesidades específicas de cada proyecto (Cidei, 2015).

5.8.3. ***Funciones Analíticas***

Las funciones más relevantes que se pueden realizar con analítica de datos están enmarcadas de la siguiente manera:

- **Regresión lineal:** La regresión lineal es una técnica estadística utilizada para modelar la relación entre dos variables continuas. Se puede utilizar para predecir el valor de una variable dependiente basándose en el valor de una o más variables independientes (Universidad de los Andes, 2023).
- **Clustering:** El clustering es una técnica de análisis de datos que implica la agrupación de objetos similares en grupos o clusters. Esta técnica se utiliza a menudo en la segmentación de clientes y en la exploración de patrones en grandes conjuntos de datos (Fernández, 2022).

- **Análisis de componentes principales (PCA):** El análisis de componentes principales es una técnica utilizada para reducir la dimensionalidad de los datos mediante la identificación de las variables que explican la mayor cantidad de varianza en los datos. Esta técnica se utiliza a menudo en la exploración de datos y en la detección de patrones (Interactivechaos, 2021).
- **Análisis de series temporales:** El análisis de series temporales es una técnica utilizada para modelar y predecir los patrones en los datos que cambian con el tiempo. Se utiliza a menudo en la predicción de la demanda y en la planificación de la producción (Consultores, 2023).
- **Minería de datos:** La minería de datos es una técnica utilizada para descubrir patrones en grandes conjuntos de datos. Se utiliza a menudo en la detección de fraudes y en la segmentación de clientes (Jesús, 2023).

6. Marco institucional

6.1. Presentación general de la empresa

La universidad Simón Bolívar, ubicada en la ciudad de Barranquilla, es una institución de Educación Superior de origen privado, sin ánimo de lucro, dedicada a la formación integral, al desarrollo de la investigación e innovación. “Su función social está inspirada en el ideario del Libertador Simón Bolívar de una Educación con equidad fundamentada en valores, promovida por una comunidad académica idónea” (Identidad Institucional, 2014).

Dentro de su estructura organizacional cuenta aproximadamente con 1800 trabajadores, entre los que se destacan: funcionarios académicos y administrativos, directivos, auxiliares, entre otros. Se caracteriza por mantener un sello de alta calidad que otorga el Ministerio de Educación, gracias a los esfuerzos y compromisos que realizan todos sus funcionarios para brindar una educación de calidad.

Se encuentra en el sector educativo posicionada como la mejor universidad de la región caribe en términos de investigación y producción científica.

6.2. Referentes estratégicos

Misión: Somos una Universidad sin ánimo de lucro, dedicada a la formación integral, al desarrollo de la investigación e innovación que articulada con el Estado, el sector productivo y la comunidad en general, responde al compromiso con el desarrollo del entorno social, político, cultural y económico. Nuestra función social está inspirada en el ideario del Libertador Simón Bolívar de una Educación con equidad fundamentada en valores, promovida por una comunidad académica idónea. La Universidad Simón Bolívar

valora y cultiva la identidad, la cultura y el respeto por el ancestro (Plan Estratégico de Desarrollo (2018-2022), 2018).

Visión En 2032, la Universidad Simón Bolívar será reconocida por la calidad de sus servicios educativos, la visibilidad de sus resultados de investigación e innovación en los ámbitos nacional e internacional y la contribución al desarrollo de las regiones Caribe y Oriental del país (Plan Estratégico de Desarrollo (2018-2022), 2018).

Principios Y Valores: Ética, Integralidad, Libertad De Enseñanza, Investigación Y Aprendizaje, Responsabilidad, Convivencia, Compromiso, Calidad, Participación, Accesibilidad. (Plan Estratégico de Desarrollo (2018-2022), 2018)

Diferenciación o nicho: En coherencia con el propósito misional, la Universidad acoge a sectores socioeconómicos vulnerables y, además, teniendo en cuenta las especificidades y exigencias de la oferta académica se aplican mecanismos universales y equitativos de ingreso, que son conocidos por los aspirantes. Además, se ofrece una opción de incentivos a los estudiantes y la oferta de actividades que les permite formarse integralmente en lo cultural y lo deportivo. (Plan Estratégico de Desarrollo (2018-2022), 2018)

6.3. Estructura organizacional

Niveles de la Estructura Organizacional. En la estructura académico-administrativa se establecen cuatro niveles de cargos: Estratégico, Táctico, Operacional, Auxiliar, tal como se puede observar en la ilustración 3.

Ilustración 3. Estructura Organizacional



Nota: Descripción de la estructura Organizacional de la Universidad. Tomado de (Universidad Simon Bolívar, 2018)

1. Nivel Estratégico. Los cargos que constituyen este nivel tienen la responsabilidad de definir la proyección institucional a partir de las declaraciones expresadas en la misión y visión. Según el nivel de responsabilidad el nivel estratégico se divide en: Alta Dirección y Dirección General: responsable de proponer reformas y/o nuevas políticas institucionales. Forman parte de la Alta Dirección los cargos de: Rector y Vicerrectores. (Estructura Organizacional, 2019)

2. Nivel Táctico. Nivel que responde por la planeación específica de un área y/o proceso y hace seguimiento a los resultados para garantizar su funcionamiento; a partir del marco de referencia elaborado por el nivel estratégico. Forman parte del nivel táctico: directores de programa, directores de Departamento, directores de Centro, jefes de departamento y jefes de Oficina. (Estructura Organizacional, 2019)

3. Nivel operacional: Nivel responsable de ejecutar las acciones establecidas en la planeación estratégica y/o en el área al que pertenece. Forman parte del nivel Operacional: Profesores, Coordinadores, Asistentes, Analistas y Profesionales de áreas administrativas. (Estructura Organizacional, 2019)

4. Nivel Auxiliar: Nivel responsable de brindar soporte y apoyo a los procesos de gestión operacional en la Institución. Forman parte del nivel auxiliar: Auxiliares, secretarías, personal de mantenimiento y logística de eventos. (Estructura Organizacional, 2019)

5.4. Productos o servicios ofertados

Servicios educativos: La universidad Simón Bolívar, ofrece una amplia oferta académica para la formación profesional en diferentes modalidades y niveles de formación. En la tabla 1, se puede observar variedad en su oferta académica según el nivel de formación, entre los programas ofertados se destacan: medicina, enfermería, ingenierías, derecho y programas de administración y negocios. La siguiente tabla describe la cantidad de programas ofertados de acuerdo con el nivel de formación.

Tabla 1 *Oferta Académica*

Nivel de Formación	Numero de Programas Ofrecidos
Profesional	29
Especialización	29
Maestría	22
Doctorado	7

Nota: Fuente (Universidad Simón Bolívar, 2018)

5.5. Análisis del sector

La comprensión del contexto en la universidad Simón Bolívar

En la Universidad Simón Bolívar se comprende la interacción con el entorno como un conjunto de relaciones sistemáticas que dan origen a las acciones de orden estratégico

que, como Institución de Educación Superior, le permiten avanzar en materia de calidad de su quehacer misional, el cual comprende el mandato consagrado en la Ley 30 de 1992, en lo referente a la docencia, la investigación y la proyección social o extensión. Para una mejor comprensión de estas relaciones, se observa cómo los procesos misionales se alimentan o parten de una comprensión de las necesidades y expectativas sociales y económicas situadas en el contexto, a partir de las cuales se fijan los objetivos y metas de la Universidad. (Plan Estratégico de Desarrollo (2018-2022), 2018)

Tendencias en Colombia y sus retos en educación superior

A nivel nacional, se ofrecen diversas posturas que señalan las dinámicas de la educación superior colombiana, sus tendencias y retos. Hacia 2012, se realizó el acostumbrado estudio de la Organización para la Cooperación y el Desarrollo Económico (OCDE) que evalúa los sistemas de educación superior de potenciales miembros. Este trabajo, elaborado recién finalizaba el período de la Revolución Educativa y del fallido intento de reforma a la educación superior, proponía a más de la sugerencia de replantear la reforma educativa, otras recomendaciones, algunas que tocan la dinámica de las instituciones como la Universidad Simón Bolívar:

- Adelantar procesos que permitan dar cabida a más estudiantes en el sistema.
- Desarrollar procesos de cualificación enfocados en tecnologías emergentes.
- Revisión externa de la oferta y la demanda de egresados de educación superior en contraste con sus índices de empleo y los niveles de ingresos relacionados con el campo de estudio.

- La rendición de cuentas a la sociedad por parte de las instituciones, en función de la consecución de los objetivos nacionales, así como en la definición de mecanismos e indicadores de desempeño.

- Composición y funcionamiento de los órganos de gobierno institucionales.

- Control financiero y uso más creativo de los sistemas nacionales de datos, de modo que la toma de decisiones en todos los niveles del sistema de educación superior se fundamente más en datos confiables. (Plan Estratégico de Desarrollo (2018-2022), 2018)

El contexto regional y local: No obstante, a los fines planteados en el recorrido antes descrito, es evidente que la región Caribe, el departamento del Atlántico y el distrito de Barranquilla, donde se sitúa la matriz de la Universidad Simón Bolívar, ofrecen un panorama muy diverso al del concierto nacional. Aunque el departamento del Atlántico ostenta una tasa de cobertura en educación superior por encima del promedio nacional (52,8%), situada en el 59,5% a 2017, según fuente de la Subdirección de Desarrollo Sectorial del Viceministerio de Educación Superior, a 2017 todos los departamentos de la región Caribe se ubican por debajo: Bolívar (37,1%), Cesar (34,5%), Córdoba (23,9%), La Guajira (21,8%), Magdalena (29,9%), San Andrés (20,5%) y Sucre (30,1%). Con una población mayoritariamente de ingresos bajos, la educación superior se constituye en un espacio de movilidad social en el que los jóvenes procedentes de familias de escasos recursos pueden encontrar la oportunidad para mejorar capacidades. El reto es pues lograr abordar los procesos formativos con pertinencia, tesis sostenida a lo largo de este apartado; para ello, el llamado es a verificar las apuestas de desarrollos locales y departamentales. (Plan Estratégico de Desarrollo (2018-2022), 2018).

7. Diseño metodológico

6.1. Tipo de investigación

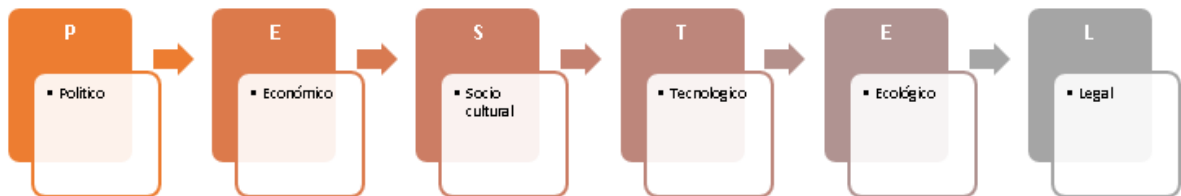
Los tipos de investigación para este trabajo se describen según su propósito como investigación aplicada ya que el desarrollo de la propuesta se está llevando bajo la modalidad de trabajo dirigido y teniendo en cuenta el objetivo general planteado se enfoca a dar solución a una problemática evidenciada en los procesos de servicios que se prestan en la universidad Simón Bolívar específicamente en las áreas financiera y bienestar universitario (Nieto, 2018). Según su grado de profundidad es un tipo de investigación descriptiva porque se reúnen o consolidan una serie de datos que pueden dar respuesta a preguntas planteadas asociadas a la problemática mencionada; según las fuentes de datos se considera una investigación mixta cuantitativa y cualitativa, cualitativa porque los datos a recopilar en esta investigación irán enfocados a mejorar los procesos con el diseño de una solución que permita obtener resultados precisos y exactos, en lo que se refiere a la cualitativa la razón es por las fuentes de datos que se encuentran en libros y material donde se presentan técnicas y herramientas actualizadas (Roberto, 2014).

Según su inferencia esta investigación es considerada inductiva que va de lo particular a lo general, porque realiza un análisis partiendo de hechos y el entorno, haciendo uso de la observación para así llegar a obtener conclusiones significativas, este método ayuda a generar conocimientos nuevos desde el equipo investigador; según la temporalidad es transversal porque se realiza un estudio sobre un caso o situación específica en un momento en concreto, los aspectos de esta metodología son útiles para describir la afectación de una variable a una muestra de población específica (PrepaUno, s.f.).

6.2. Análisis externo

La herramienta seleccionada para realizar el análisis externo será el análisis PESTEL, porque es la base y el punto de partida de cualquier proyecto a desarrollar, realizando un plan de negocio o un plan estratégico, este análisis permite valorar que influencia tiene los factores externos dentro de la entidad seleccionada que es la universidad, los factores externos son los que pueden estar afectando de manera positiva o negativa a la empresa, pero no se tiene influencia sobre ellos, este tipo de análisis tiene beneficios como la proactividad que genera, la sencillez para aplicarlo, la adaptabilidad que presenta al cambio y es un recurso complementario para evaluar el entorno. Los factores que estudia este análisis se representan como lo ilustra la imagen, los cuales se toman como base para realizar el análisis externo.

Ilustración 4.. Análisis PESTEL



Nota: Factores análisis PESTEL. Tomado de (Trenza, 2020)

La primera P hace referencia a lo político, es decir cómo afectan los factores de esta índole a la organización, seguido se encuentran los factores económicos, estos se relacionan en todos los aspectos y miden el grado de crisis que pueden llegar a impactar positiva o negativamente en la empresa, después el aspecto socio cultural que es la relación con las tendencias socio culturales y su afectación, posteriormente el aspecto tecnológico que hace referencia a cómo afecta la tecnología y su avance a la empresa, el aspecto ecológico busca determinar cómo influye esta tendencia en la entidad y por último está el aspecto legal que evalúa hasta qué punto la legislación condiciona la

actividad, perjudicando o favoreciendo a la misma, todo lo que se oriente a lo negativo son las amenazas y en positivo representa las oportunidades de crecimiento y mejora en una organización (Trenza, 2020).

Al analizar el entorno externo de la Universidad Simón Bolívar en Barranquilla, es importante considerar una serie de factores que influyen en su funcionamiento y desarrollo; a través del análisis PESTEL se puede identificar y comprender estos factores clave:

1. Factores Políticos:

- La Universidad Simón Bolívar en Barranquilla está sujeta a las políticas gubernamentales y regulaciones relacionadas con la educación superior en Colombia, los cambios en estas políticas pueden impactar en la autonomía y financiación de la universidad, además, la relación con las autoridades locales y regionales puede influir en iniciativas de colaboración y proyectos de desarrollo en la comunidad.

2. Factores Económicos:

- Las condiciones económicas locales y regionales tienen un impacto directo en la financiación y el funcionamiento de la universidad, la disponibilidad de recursos financieros y las tendencias económicas pueden afectar la viabilidad de programas y proyectos; la demanda de educación superior puede estar influenciada por factores económicos, como el acceso a empleo y oportunidades laborales en la región.

3. Factores Socioculturales:

- Las preferencias y necesidades de la comunidad estudiantil y de la sociedad en Barranquilla son importantes consideraciones para la Universidad Simón Bolívar. Adaptarse a las tendencias culturales y sociales locales es fundamental para mantener la relevancia y atractivo de la institución, así como tener en cuenta la diversidad demográfica y cultural de Barranquilla presenta oportunidades para promover la inclusión y la equidad en la educación superior.

4. Factores Tecnológicos:

- La tecnología juega un papel cada vez más importante en la enseñanza, el aprendizaje y la gestión universitaria. La Universidad Simón Bolívar debe mantenerse actualizada con las últimas tendencias tecnológicas para ofrecer una experiencia educativa de calidad, integrando tecnología en la infraestructura y los procesos universitarios puede mejorar la eficiencia y la calidad de los servicios ofrecidos.

5. Factores Ecológicos:

- La Universidad Simón Bolívar puede contribuir al desarrollo sostenible y la conservación del medio ambiente en la región, esto mediante iniciativas de gestión ambiental y prácticas sostenibles que tengan un impacto positivo en la comunidad y el entorno local, pero debe estar preparada para enfrentar desafíos ambientales, como la mitigación de riesgos naturales y la adaptación al cambio climático.

6. Factores Legales:

- Cumplir con las leyes y regulaciones locales y nacionales es fundamental para la Universidad Simón Bolívar. Aspectos legales relacionados con la contratación de personal, la protección de datos y la propiedad intelectual deben ser considerados en todas las actividades institucionales.

6.3. Análisis interno

Es importante aplicar un instrumento como la observación en la universidad porque permiten recopilar información cualitativa y en profundidad sobre las experiencias, opiniones y perspectivas de los funcionarios, profesores y personal administrativo. Estos análisis proporcionan insights valiosos para mejorar los programas educativos, identificar necesidades y áreas de mejora y promover la participación y el compromiso en la comunidad universitaria.

Para el análisis interno de esta investigación también se realizarán encuestas ya que es un procedimiento óptimo para recopilar la información, obtener muestras representativas y así mismo ayudan a evidenciar los objetivos propuestos para el proyecto, la información obtenida a través de estos instrumentos será un insumo valioso y se recopilara de dos de las áreas organizacionales de la universidad que se son bienestar universitario y el área financiera, en estas áreas están los principales stakeholders que son líderes de área que manejan los procesos actualmente y evidencian la problemática al no tener centralizada la información.

8. Diagnóstico organizacional

Las etapas definidas para realizar el diagnóstico pueden enmarcarse en 4 fases: la primera la recolección de la información en la cual mediante la observación se realizó un análisis observacional especialmente en las áreas de bienestar universitario y el área financiera, esto con el objetivo de comprender la interacción entre los stakeholders y las dificultades que enfrentan al no tener centralizada la información, a su vez se diseñaron encuestas para aplicar a los líderes de área y obtener así información cuantitativa sobre las necesidades, expectativas y opiniones con respecto a la gestión de los datos.

Con esta información se entra en la segunda etapa que es el análisis de la información, con el análisis cualitativo se identificaron patrones de comportamiento y percepciones frente a la problemática presentada y en cuanto al cuantitativo se identificaron variables, tendencias que influyen directamente en la toma de decisiones. De aquí se puede pasar a la fase 3 que es la triangulación de resultados, en la cual se identifican los puntos de convergencia y a la vez los de divergencia entre las diferentes perspectivas fortaleciendo la confiabilidad del diagnóstico y por último la fase 4 en la cual se interpretan y se generan los resultados del análisis, generando insights valiosos sobre las necesidades de los stakeholders, identificando las áreas de mejora y las oportunidades para optimizar la información en la Universidad Simón Bolívar. En cuando a la metodología se utilizó la observación, encuesta y se realizó análisis de contenido y estadístico.

8.1. Procesamiento estadístico de datos

El objetivo al realizar este procesamiento estadístico de datos es analizar el comportamiento de los usuarios involucrados para identificar las necesidades y visualizar las oportunidades de mejora.

La unidad de análisis está compuesta por los líderes de proceso y directores de departamento. Los líderes de proceso los conforman los funcionarios que hacen parte del área financiera integrada por seis personas, el área de bienestar universitario integrado por ocho personas y dos directores de procesos que tienen el control estratégico sobre sus respectivas áreas.

Durante el desarrollo de la investigación, se implementó un enfoque para el procesamiento de datos y la especificación de la población y muestra, asegurando la integridad y validez de los hallazgos, entre los cuales se especifican los siguientes:

Aplicación del instrumento y participación de personas:

- Se aplicó un instrumento de recolección de datos diseñado específicamente para este estudio, este instrumento fue utilizado para encuestar a un total de 15 participantes, todos ellos líderes de proceso y directores de departamento en la Universidad Simón Bolívar en Barranquilla.

Determinación de la muestra:

- La muestra se determinó mediante un cálculo estadístico que tuvo en cuenta un nivel de confianza del 95% y un margen de error del 5%. Esto resultó en una muestra de 15 personas, seleccionadas de manera estratégica para representar adecuadamente la diversidad de roles y responsabilidades dentro de la institución.

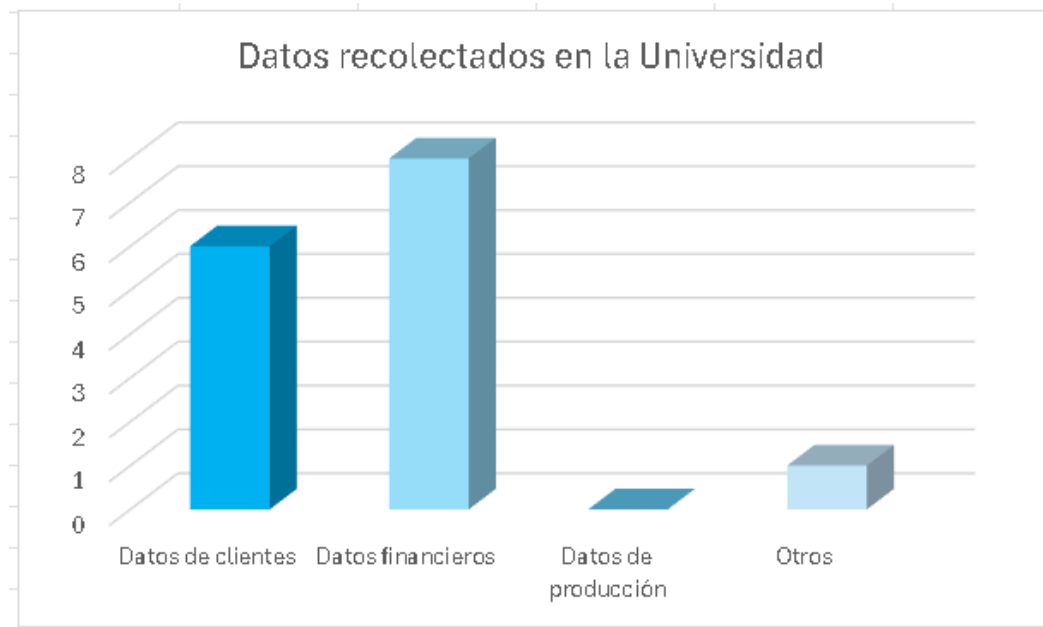
Construcción y validación del instrumento:

- El instrumento de recolección de datos se construyó teniendo en cuenta los objetivos específicos de la investigación y las variables de interés. Se llevó a cabo una revisión exhaustiva por parte de expertos en el campo para garantizar su relevancia y validez.

De acuerdo con el instrumento aplicado se representan los siguientes datos recolectados que son considerados relevantes en el proceso.

La ilustración 5 presenta una representación visual de los datos recolectados en la universidad, organizados en diferentes categorías; en primer lugar, se observan los datos financieros, seguidos por los datos de los clientes, mientras que otras categorías de datos se ubican en posiciones intermedias, finalmente, se destacan los datos de producción como la última categoría en la secuencia. Esta disposición sugiere una priorización en la recolección y posiblemente en el análisis de los datos financieros y de clientes, indicando su importancia en comparación con otras áreas de interés dentro de la universidad.

Ilustración 5. Datos que se recolectan en la universidad

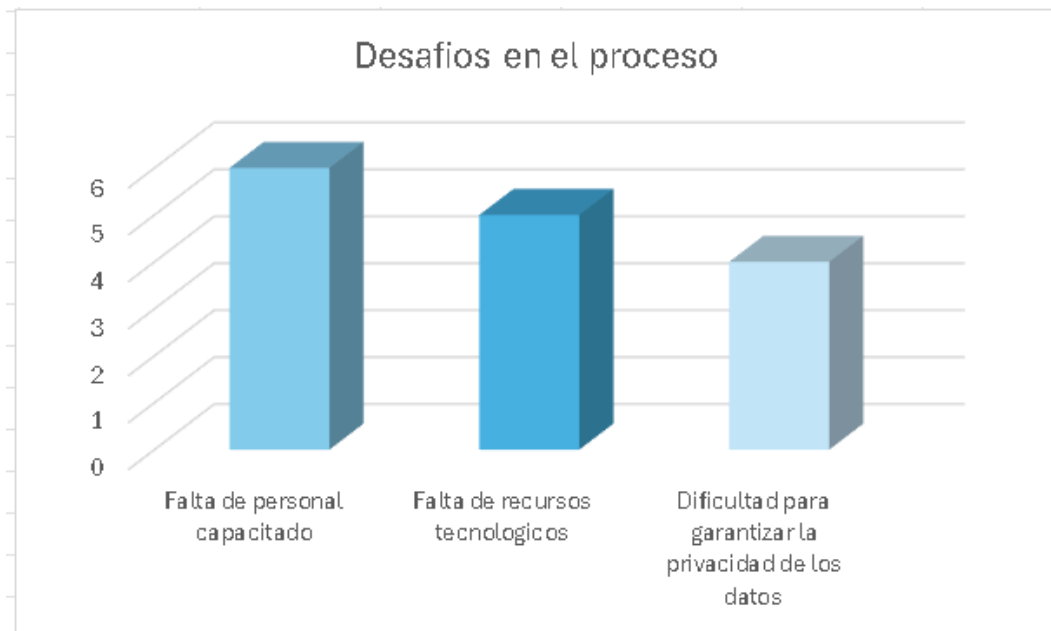


Nota: Distribución de los datos recolectados en la universidad. Elaboración propia.

El gráfico 6 ilustra los principales desafíos en el proceso, identificando tres áreas críticas. En primer lugar, se destaca la falta de personal capacitado como el desafío más significativo, seguido estrechamente por la escasez de recursos tecnológicos. Ambos desafíos representan obstáculos importantes en la eficiencia y efectividad del proceso; por último, se menciona la dificultad para garantizar la privacidad de los datos,

subrayando la preocupación por la seguridad y la protección de la información sensible, esta representación pone de relieve los aspectos clave que requieren atención y acción prioritaria para mejorar el proceso en cuestión.

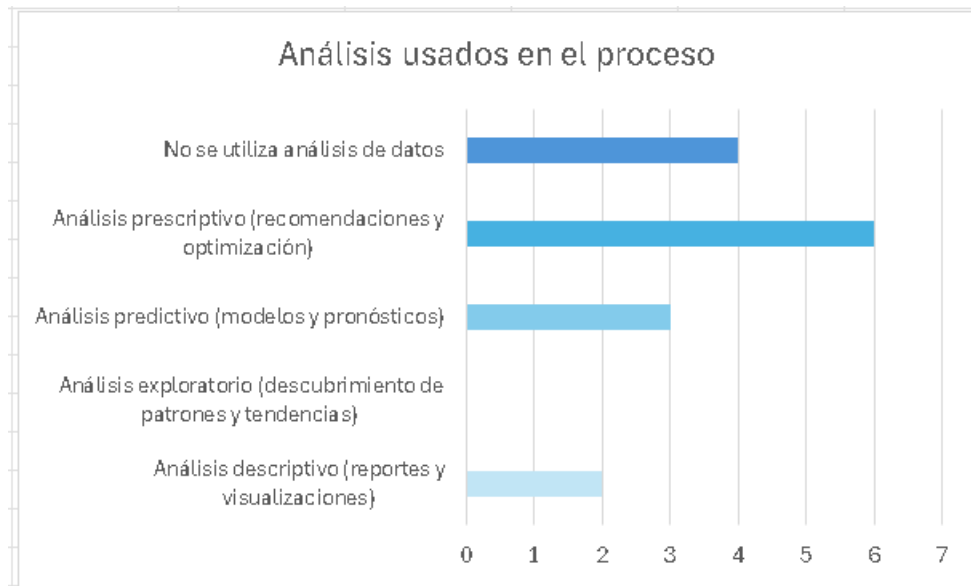
Ilustración 6. Desafíos que se presentan en el proceso



Nota: Distribución de los desafíos presentados. Elaboración propia.

El gráfico 7, muestra los diferentes tipos de análisis utilizados por la universidad, con el análisis descriptivo como el más frecuente, seguido por la ausencia de análisis de datos, el análisis predictivo y, por último, el análisis prescriptivo. Esto sugiere una tendencia predominante hacia la descripción y comprensión de los datos, seguida por un menor enfoque en la predicción y la prescripción de acciones basadas en ellos.

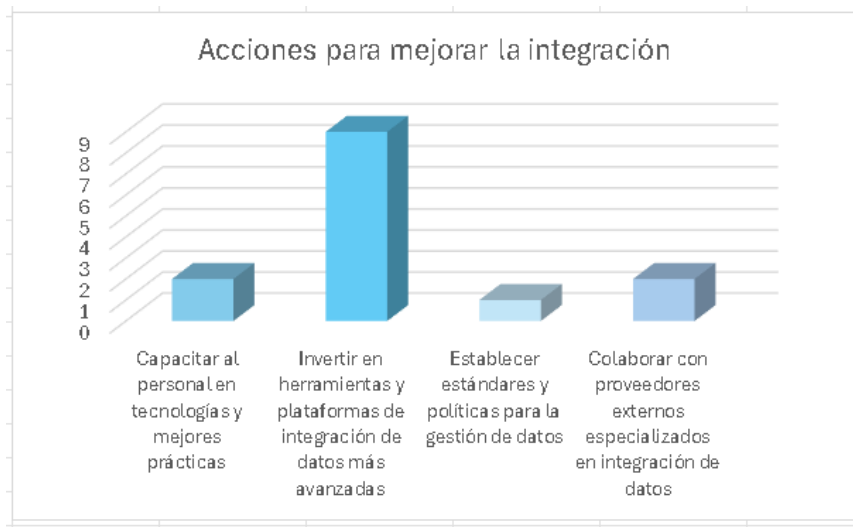
Ilustración 7. Tipos de Análisis usados en el proceso



Nota: Análisis usados. Elaboración propia.

La ilustración 8 detalla las acciones propuestas para mejorar la integración de datos dentro del entorno universitario. En primer lugar, se destaca la necesidad de invertir en herramientas y plataformas de integración de datos más avanzadas como la principal acción; esta inversión sugiere un enfoque en la adopción de tecnologías más modernas para optimizar el proceso de integración; seguido de cerca, se encuentra la capacitación del personal en tecnologías y mejores prácticas como una acción igualmente importante, indicando el reconocimiento de la importancia del recurso humano en el éxito de la integración de datos. En un nivel casi equiparable, se menciona la colaboración con proveedores externos especializados en integración de datos, resaltando la búsqueda de experticia externa para complementar los esfuerzos internos, por último, se propone establecer estándares y políticas para la gestión de datos como una medida fundamental para garantizar la coherencia y la eficacia en el proceso de integración. Esta representación visual enfatiza las múltiples facetas y enfoques necesarios para abordar de manera integral los desafíos de la integración de datos en el contexto universitario.

Ilustración 8. Acciones para mejorar la integración de datos



Nota: Acciones propuesta de mejora. Elaboración propia.

8.2. Análisis de datos

La Universidad Simón Bolívar se enfrenta a un desafío significativo en la gestión de los datos generados por sus múltiples sistemas de información. La actual fragmentación de los datos dificulta su integración, análisis y uso eficiente, lo que se traduce en:

- **Dificultades para la generación de reportes:** Los procesos para crear reportes son lentos y tediosos, lo que impacta negativamente en la toma de decisiones.
- **Carga de trabajo adicional:** El personal dedica un tiempo considerable a la gestión manual de datos, lo que reduce su productividad en otras áreas.
- **Calidad de los datos:** La falta de mecanismos de control y verificación puede afectar la calidad y confiabilidad de la información.

- **Tiempos de respuesta:** Los usuarios deben esperar largos periodos para obtener la información que necesitan, lo que genera insatisfacción e impacta en la eficiencia de los procesos.

A pesar de estas dificultades, existen fortalezas y oportunidades que pueden ser aprovechadas para mejorar la situación:

- **Iniciativas de Data Warehouse:** La universidad ha realizado inversiones en la construcción de un Data Warehouse, lo que constituye una base sólida para el desarrollo de un Marketplace de datos.
- **Conocimiento técnico:** La universidad cuenta con personal con experiencia en gestión de datos y tecnologías de la información, lo que facilita la implementación del proyecto.
- **Herramientas Microsoft 365 y Azure:** La universidad cuenta con licencias de estas herramientas, que pueden ser utilizadas para la creación e implementación del Marketplace de datos.

La implementación de un Marketplace de datos en la Universidad Simón Bolívar se presenta como una oportunidad para:

- **Centralizar la información:** Permitiendo un acceso único y seguro a los datos desde diferentes departamentos y unidades.
- **Agilizar la generación de reportes:** Automatizando la integración y análisis de datos para la creación de reportes dinámicos e intuitivos.
- **Reducir la carga de trabajo:** Liberando al personal de tareas manuales repetitivas y permitiéndoles enfocarse en actividades de mayor valor.

- **Mejorar la calidad de los datos:** Implementando mecanismos de control y limpieza para asegurar la confiabilidad de la información.
- **Disminuir los tiempos de respuesta:** Facilitando el acceso a la información de manera rápida y eficiente.

En definitiva, el desarrollo de un Marketplace de datos en la Universidad Simón Bolívar representa una inversión estratégica que permitirá optimizar la gestión de los datos, mejorar la toma de decisiones y aumentar la eficiencia de los procesos en la institución.

9. Plan de intervención

Este apartado como objetivo presentar un plan de intervención para el proyecto Modelo de Marketplace de datos mediante el uso y aplicación de herramientas de Microsoft, el plan de intervención que se propone en este trabajo busca implementar una serie de estrategias y acciones que se detallan a continuación:

9.1. Diseño de la Arquitectura del Marketplace De Datos

Evaluar alternativas tecnológicas para el diseño del marketplace de datos.

Se optó por Microsoft Azure como la plataforma para el diseño, considerando una variedad más amplia de componentes específicos según los requisitos de información del proyecto. Esto abarca las fases de ingesta de datos, procesamiento y almacenamiento, aprendizaje automático y finalmente autoservicios de reportes.

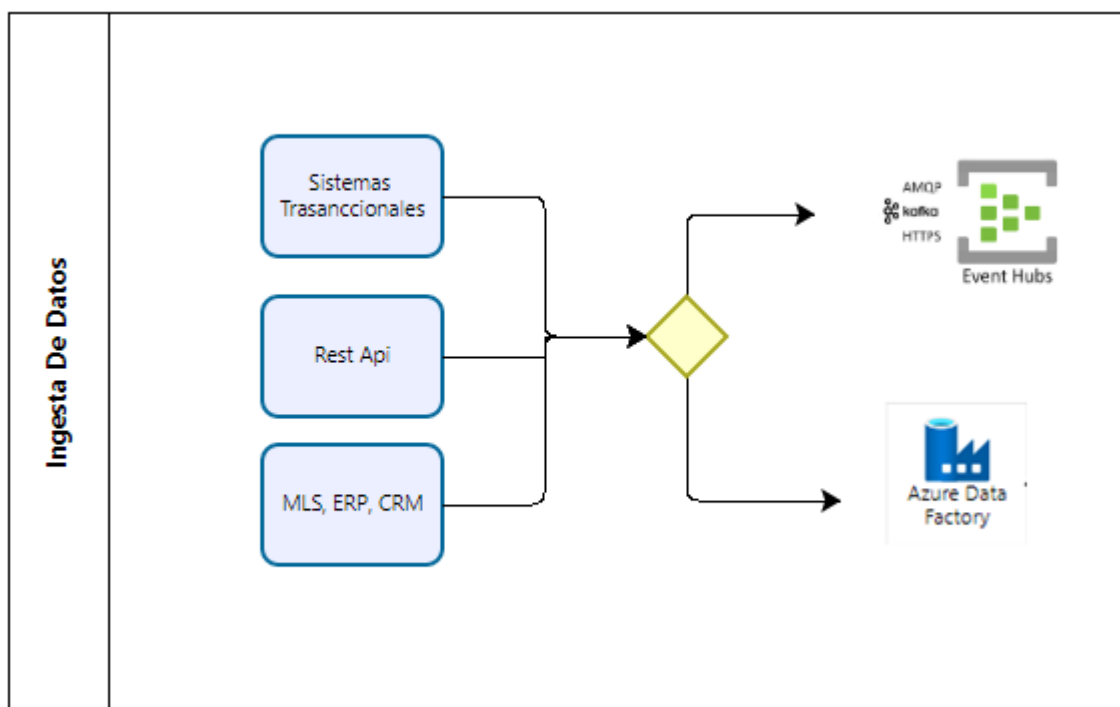
En cuanto a la visualización, se destaca el uso de Power BI, que facilita la creación y presentación de paneles para la visualización interactiva de datos. Para gestionar y exponer de manera segura las API propias, se sugiere la implementación de Azure API Management. Además, se propone la utilización de Power Apps para el desarrollo rápido de aplicaciones personalizadas sin la necesidad de habilidades de programación intensivas. Considerando las necesidades de automatización, se plantea el empleo de Azure Logic Apps, que permite automatizar y orquestar flujos de trabajo y procesos empresariales.

Componentes de ingesta de datos

En esta parte, se exponen los resultados derivados del diseño de componentes, considerando las fuentes de datos empleadas por la universidad. Se llevó a cabo un

análisis específico de los tipos de datos asociados a cada fuente, con el propósito de elegir los componentes más apropiados para su procesamiento. En la ilustración 9 que se presenta a continuación, se representa la sección correspondiente a la ingesta de datos:

Ilustración 9. Esquema de ingesta de datos al Marketplace



Nota : Esquema de ingesta de datos al MarketPlace. Fuente: Elaboración propia.

Se incluyen los elementos Event Hubs y Azure Data Factory para dirigir la información hacia el Data Lake.

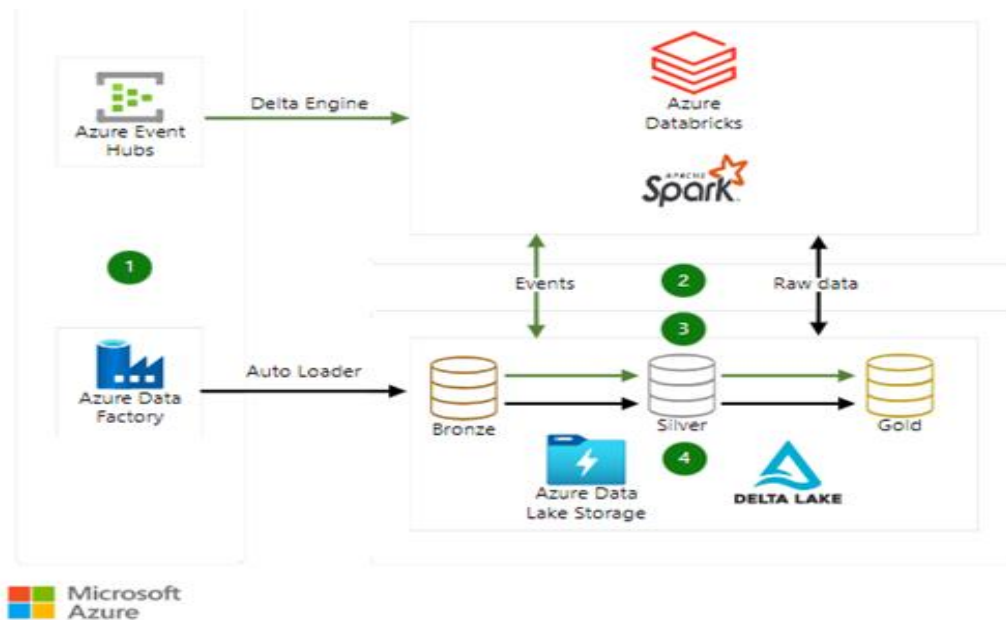
- **Event Hubs:** Este componente de Azure facilita la ingestión y el procesamiento de grandes volúmenes de eventos y datos en tiempo real. Se utiliza para recibir, almacenar y distribuir eventos de manera escalable.

- **Azure Data Factory:** Este servicio de Azure se emplea para orquestar y automatizar procesos de extracción, transformación y carga (ETL). Facilita la gestión de flujos de trabajo de datos desde diversas fuentes hasta destinos, permitiendo la integración de información de manera eficiente. Su versatilidad es clave para la construcción y ejecución de pipelines de datos en entornos de Big data.

Componentes de procesamiento y almacenamiento

En este contexto, se abordan las estrategias y herramientas seleccionadas para el procesamiento eficiente y el almacenamiento adecuado de la información recolectada, asegurando así una gestión efectiva de los datos en el entorno universitario. A continuación, en la ilustración 10, se presenta la sección del diagrama correspondiente al procesamiento y almacenamiento de los datos:

Ilustración 10. Esquema de ingesta, procesamiento y storage de datos al Marketplace



Nota: Esquema de ingesta, procesamiento y storage de datos al Marketplace. Fuente: (Microsoft, 2023).

Se incluyen los elementos Apache Spark, Azure Databricks, Data Lake Storage y Delta Lake en Data Lake Storage para almacenar y procesar toda la información.

Apache Spark: es un motor de procesamiento de datos en memoria y de código abierto. Apache Spark se utiliza comúnmente para análisis de datos, procesamiento de Big data y machine learning.

Data Lake Storage: reúne datos de streaming y por lotes, incluidos datos estructurados, no estructurados y semiestructurados, como registros, archivos y medios.

Azure Databricks: limpia y transforma los conjuntos de datos sin estructura y los combina con datos estructurados de bases de datos operativas o de almacenamientos de datos.

Delta Lake en Data Lake Storage admite transacciones ACID para ofrecer confiabilidad y está optimizado para la ingesta, el procesamiento y las consultas eficaces.

Cada servicio ingiere datos en un formato común para garantizar la coherencia. La arquitectura usa un lago de datos compartido basado en el formato abierto de Delta Lake. Los datos sin procesar se ingieren desde distintos orígenes de procesamiento por lotes y streaming para formar una plataforma de datos unificada. La plataforma se puede usar para casos de uso posteriores, como análisis, informes de BI, ciencia de datos, inteligencia artificial y aprendizaje automático (Microsoft, 2024).

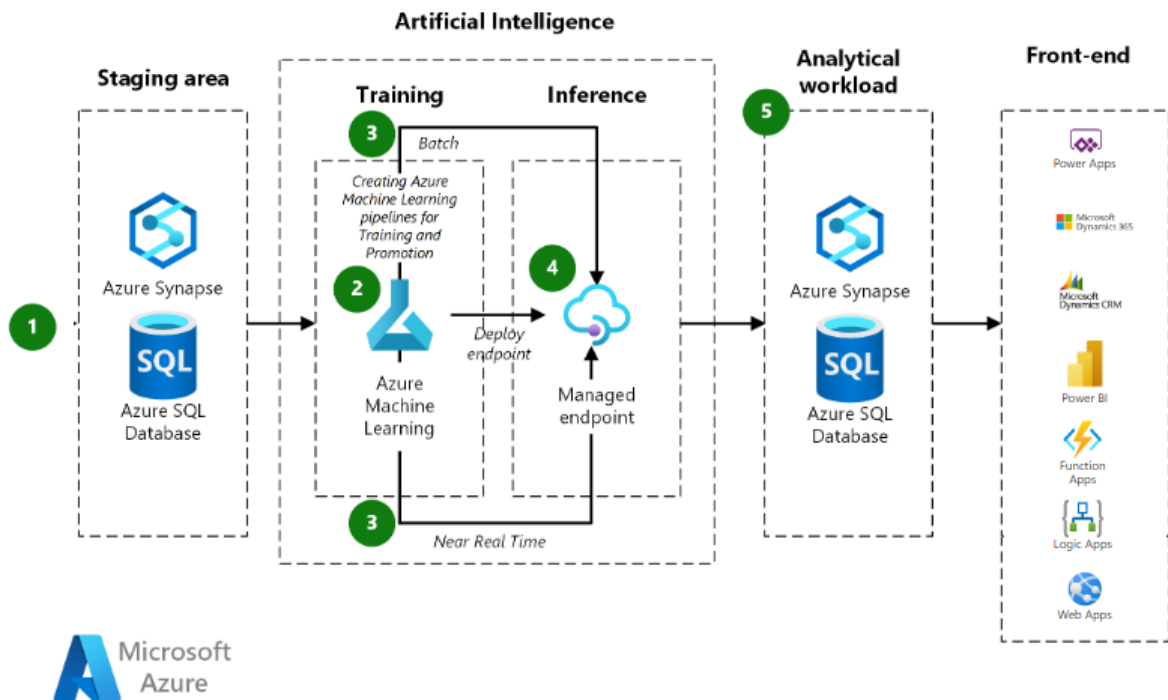
El modelo de capas Bronze, Silver y Gold optimiza el procesamiento y acceso flexible a datos. Las tablas Bronze admiten datos sin procesar, convertidos al formato Delta Lake. Silver almacena datos optimizados para BI y ciencia de datos, procesados mediante flujos y ETL. Gold contiene datos enriquecidos, listos para análisis, accesibles a través de

varias herramientas como PySpark, Koalas y Power BI. Este enfoque permite una gestión eficiente y escalable de datos en Azure Databricks.

Componentes de aprendizaje automático

En el contexto de los Componentes de aprendizaje automático, se exploran las estrategias y herramientas empleadas para desarrollar modelos predictivos y analíticos. Estos componentes permiten el procesamiento eficiente de grandes volúmenes de datos, así como el almacenamiento de modelos entrenados para su utilización en diversas aplicaciones dentro del entorno universitario. A continuación, en la ilustración 11, se presenta la sección del diagrama correspondiente al aprendizaje automático:

Ilustración 11. Esquema de aprendizaje automático al MarketPlace



Nota: Esquema de aprendizaje automático al MarketPlace. Fuente: (Microsoft, 2023).

Se incluyen los elementos Azure Machine Learning, Azure Synapse Analytics, Synapse SQL y Azure SQL Database para los aprendizajes automáticos.

- Azure Machine Learning es un servicio de ML empresarial para compilar e implementar modelos rápidamente. Proporciona a los usuarios de todos los niveles de aptitud con un diseñador de código bajo, ML automatizado (AutoML) y un entorno de Jupyter Notebook hospedado que admite diversos IDE.
- Azure Synapse Analytics es un servicio de análisis que unifica integración de datos, almacenamiento de datos empresariales y análisis de macrodatos.
- Synapse SQL es un sistema de consultas distribuidas para T-SQL que habilita escenarios de almacenamiento y virtualización de datos, además de ampliar T-SQL para responder a escenarios de streaming y ML. Ofrece modelos de recursos sin servidor y dedicados.
- Azure SQL Database es una solución de base de datos en la nube que elimina la necesidad de administrar y mantener servidores de base de datos.

En ocasiones, los desafíos del aprendizaje automático (ML) resultan demasiado complejos para ser abordados por un único modelo. La adopción de múltiples modelos, práctica común en diversos sectores, puede contribuir significativamente a mejorar los resultados, a través de Azure Machine Learning, es factible diseñar una canalización que englobe el entrenamiento, la implementación por lotes y en tiempo real de estos modelos.

En un entorno educativo universitario, este enfoque de varios modelos podría aplicarse a la predicción de resultados académicos de estudiantes. Cada modelo podría centrarse en diferentes variables y aspectos, como la asistencia, las calificaciones previas, el rendimiento en pruebas estandarizadas, etc. Al tener conjuntos de datos

específicos para cada modelo, se podría obtener una evaluación más precisa del rendimiento académico de los estudiantes y proporcionar intervenciones personalizadas según las necesidades individuales (Microsoft, 2024).

Finalmente, se expone la capa de frontend del marketplace de datos, actúa como la interfaz visual a través de la cual los usuarios exploran, buscan y acceden a conjuntos de datos. Debe ser intuitiva, atractiva y fácil de navegar. Los elementos clave incluyen:

Interfaz de Usuario (UI): Diseño atractivo y fácil de usar que permite a los usuarios explorar catálogos de datos, realizar búsquedas avanzadas y visualizar información detallada sobre conjuntos de datos.

Funcionalidad de Búsqueda: Implementación de un sistema de búsqueda efectivo que permita a los usuarios encontrar rápidamente los datos que necesitan, con filtros y categorías relevantes.

Perfiles de Usuario: Áreas personalizadas para que los usuarios gestionen sus perfiles, guarden conjuntos de datos favoritos, revisen el historial de descargas y gestionen configuraciones.

Visualización de Datos: Herramientas integradas para visualizar datos de manera efectiva, como gráficos interactivos, tablas dinámicas y vistas previas de datos, facilitando la toma de decisiones informadas.

Proceso de Compra/Descarga: Implementación de un flujo de trabajo claro y sencillo para la adquisición o descarga de datos, incluyendo opciones de pago seguro y métodos de descarga eficientes.

Seguridad y Acceso: Garantizar la seguridad de los datos mediante protocolos de autenticación seguros y proporcionar opciones de acceso controlado según los permisos del usuario.

Compatibilidad Multiplataforma: Asegurarse de que la interfaz sea compatible con diversos dispositivos y navegadores para una accesibilidad óptima.

Feedback y Evaluaciones: Incluir sistemas de retroalimentación y evaluación para que los usuarios puedan compartir sus experiencias, lo que contribuye a la mejora continua de la plataforma.

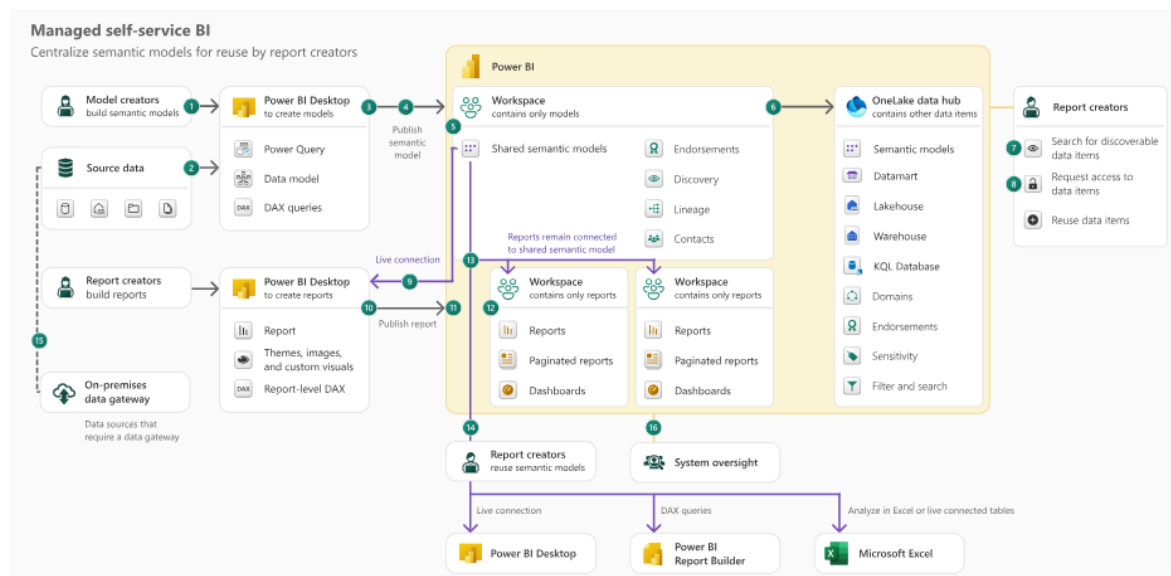
En resumen, la capa de frontend en un marketplace de datos crea la interfaz esencial para que los usuarios interactúen con la plataforma de manera efectiva y eficiente.

Componentes de autoservicio de reportes

En el contexto de los Componentes de autoservicio de reportes, se exploran las herramientas y estrategias destinadas a permitir a los usuarios generar informes de manera independiente. Estos componentes facilitan la creación ágil y personalizada de reportes, permitiendo a los usuarios explorar y visualizar datos de manera eficiente. La implementación de herramientas de autoservicio garantiza una experiencia de informes más accesible y adaptable en el entorno universitario.

A continuación, en la ilustración 12, se presenta la sección del diagrama correspondiente al autoservicio de reportes:

Ilustración 12. Esquema de autoservicio de reportes en el Marketplace



Nota: Esquema de autoservicio de reportes en el Marketplace. Fuente: (Microsoft, 2023).

En el proceso de desarrollo y utilización de modelos semánticos en Power BI, los creadores utilizan Power BI Desktop para construir modelos, conectarse a fuentes de datos y publicar en el servicio Power BI. Se fomenta la reutilización y certificación de modelos semánticos, facilitando su descubrimiento en el centro de datos OneLake. Los creadores de informes utilizan estos modelos en Power BI Desktop para desarrollar informes conectados a dichos modelos. Los informes se publican en un espacio de trabajo protegido, manteniendo la conexión con los modelos semánticos compartidos. La supervisión y monitoreo son realizados por administradores para garantizar la eficiencia y seguridad en el servicio Power BI (Microsoft, 2024).

En la ilustración 13, se presenta el modelo completo del Marketplaces de datos con los componentes de ingesta, procesamiento y análisis, almacenamiento de datos, aprendizaje automático, visualización y consumo, Se ha dividido la imagen del marketplace en tres secciones distintas para una mejor claridad y comprensión. Cada parte representa diferentes aspectos o componentes del marketplace, permitiendo así un análisis más detallado y exhaustivo de su funcionamiento y estructura,

Ilustración 13. Marketplace, parte A: fuente de datos y Data Lake

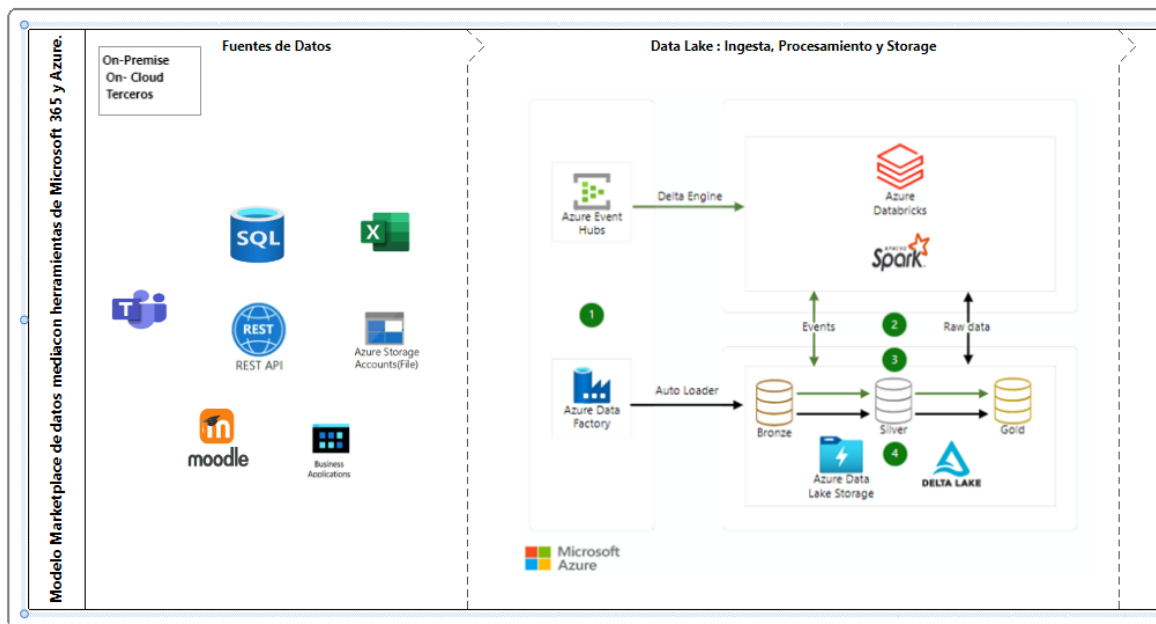


Ilustración 14. Marketplace, parte B: Aprendizaje Automático

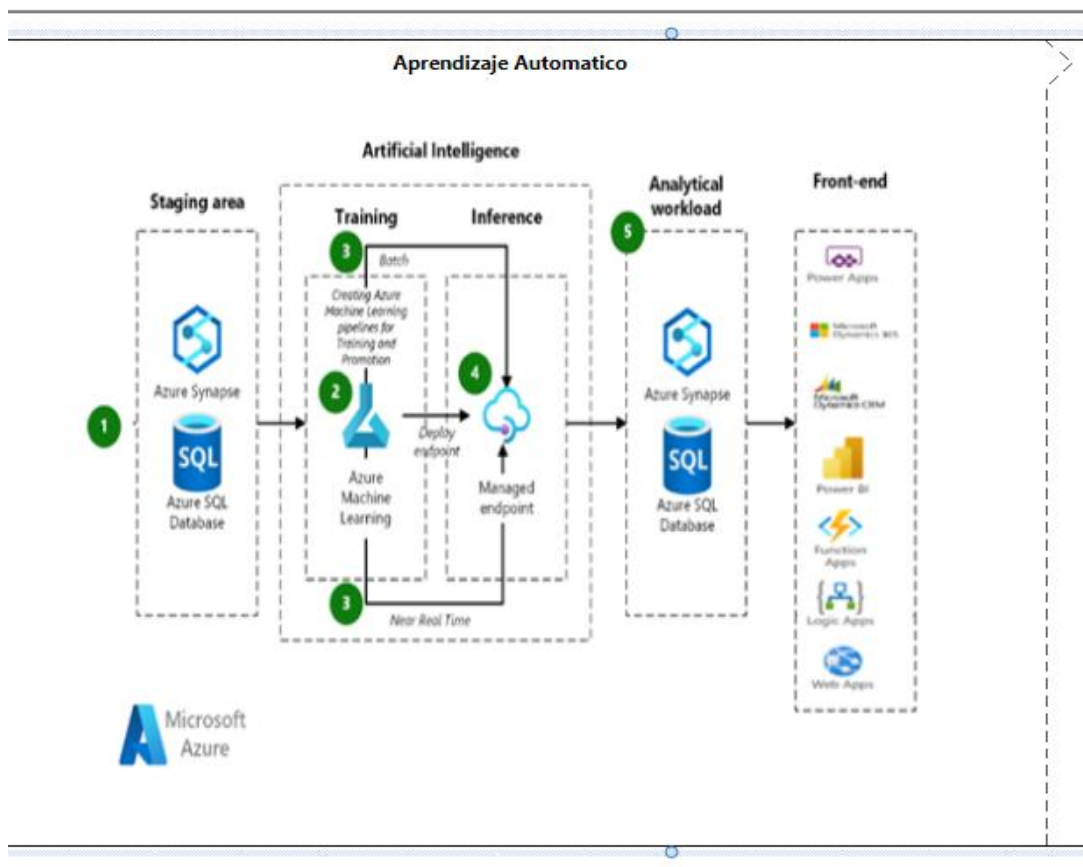
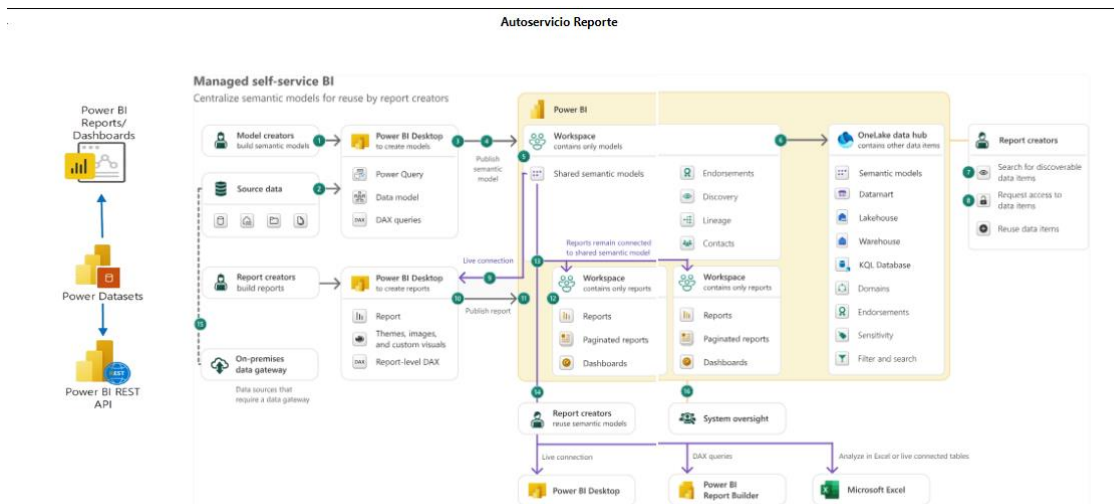


Ilustración 15. Marketplace, parte C: Autoservicio Reporte



Nota: Marketplace. Fuente: (Microsoft, 2023).

Seguridad y Gobierno de datos

Es imprescindible especificar los elementos vinculados a la seguridad de la solución y asegurar la gobernanza de datos del Marketplaces de datos.

En cuanto a seguridad, es fundamental para garantizar la integridad, confidencialidad y disponibilidad de la información intercambiada entre múltiples usuarios y proveedores.

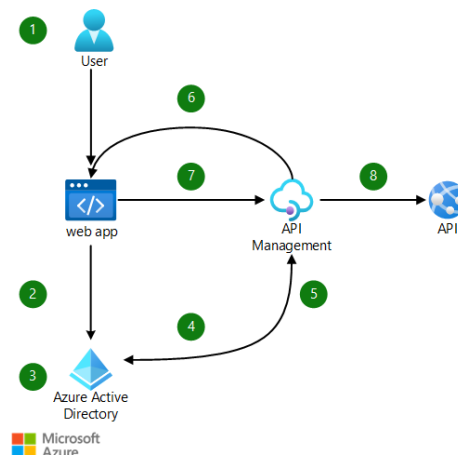
Es por ello que se incluyen los elementos:

- Api Management: permite a las organizaciones crear, publicar, gestionar y asegurar APIs de manera eficiente.

- Azure Active Directory (Azure AD): s el servicio de gestión de identidades y acceso en la nube proporcionado por Microsoft Azure.

A continuación, en la ilustración 14, se presentan los componentes de seguridad:

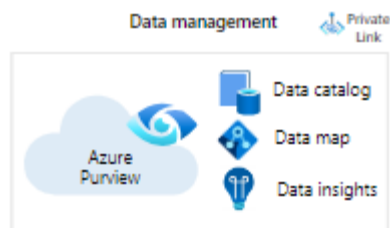
Ilustración 16. Componentes de seguridad del Marketplaces



Nota: Marketplace. Fuente: (Microsoft, 2023).

En la ilustración 15, se referencia la imagen de gobierno de datos en el Marketplace, lo cual es crucial para garantizar la calidad, seguridad y cumplimiento normativo de la información. Implica establecer políticas, procedimientos y controles para gestionar la calidad de los datos, proteger la privacidad y cumplir con regulaciones, asegurando la confianza y la eficacia en la gestión de datos. Para lo cual, se agregó, Azure Purview , proporciona una conexión de seguridad mejorada a las canalizaciones de ingesta, almacenamiento y análisis de Data Lake para catalogar automáticamente los recursos de datos (Microsoft, 2024).

Ilustración 17.Data management



Nota: Marketplace. Fuente: (Microsoft, 2023).

En consonancia con los apartados previamente desarrollados en este plan de intervención, se presenta a continuación en la ilustración 16 una propuesta completa, en la cual se enfatiza la importancia de mantener una infraestructura actualizada que garantice la seguridad y la integridad de los datos, Esto implica no solo la implementación de medidas de seguridad robustas, sino también la creación y mantenimiento de un catálogo de datos completo y preciso.

Este catálogo, detallado en secciones anteriores, sirve como una herramienta fundamental para gestionar eficazmente la información dentro del contexto del proyecto, facilitando su acceso, uso y mantenimiento en todo momento.

Ilustración 18. Propuesta actualizada, parte A: Seguridad, fuente de datos y Data Lake

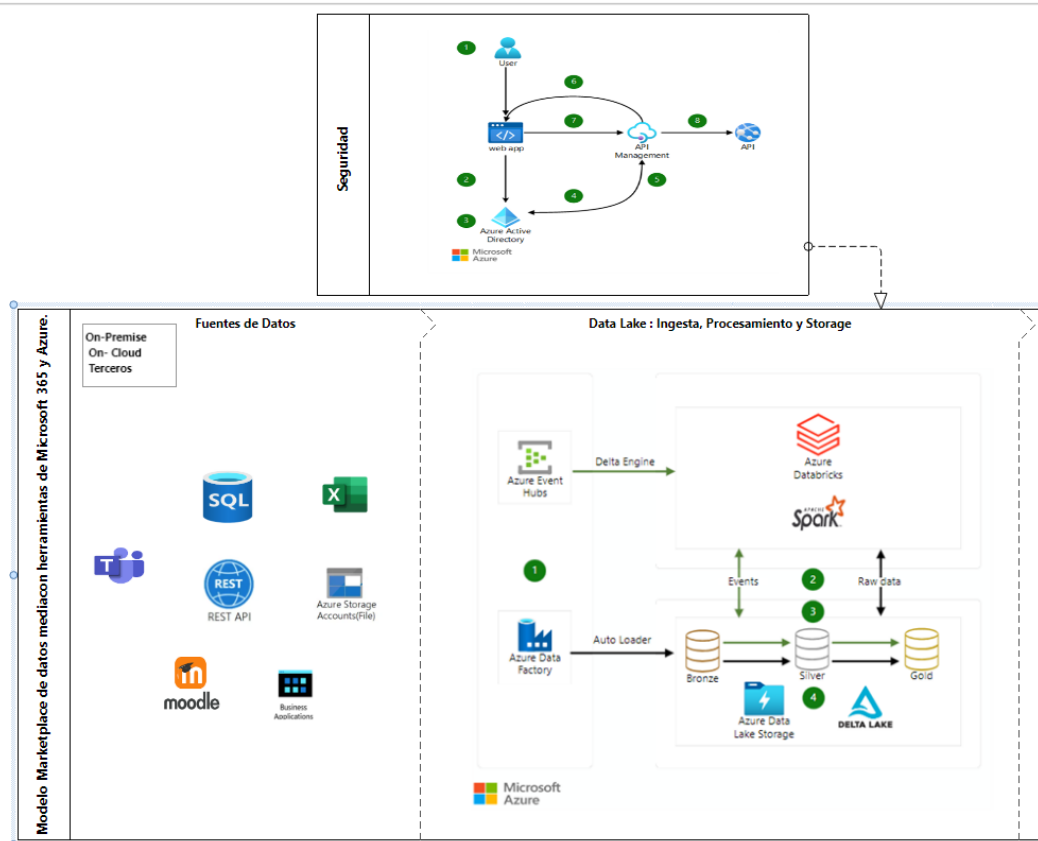


Ilustración 19. Propuesta actualizada, parte B: Aprendizaje automático y gobierno de datos

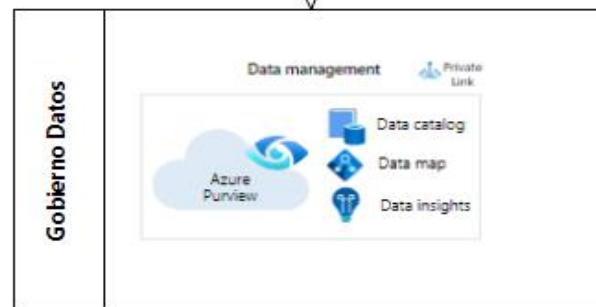
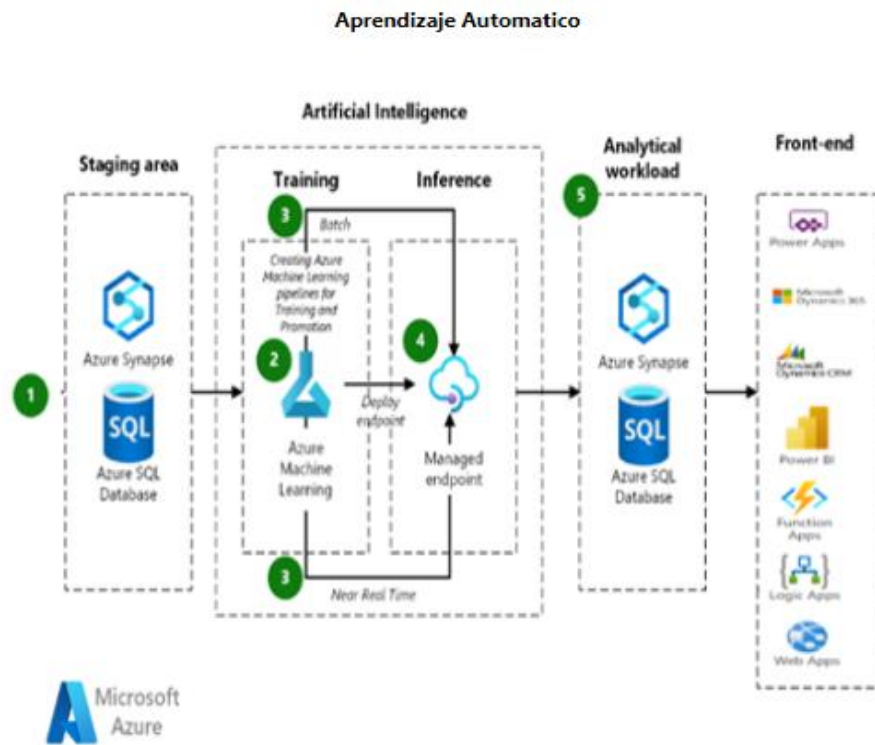
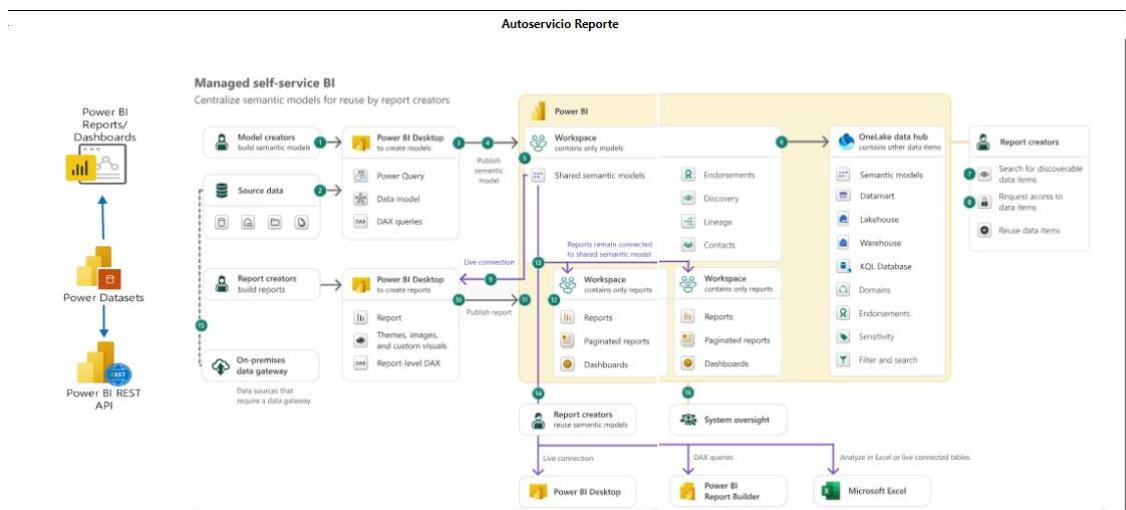


Ilustración 20. Propuesta actualizada, parte C: Autoservicio Reporte



Este modelo de Marketplace de Datos propuesto integra una serie de componentes clave que trabajan en conjunto para facilitar la ingestión, procesamiento, análisis y compartición de datos entre los diferentes usuarios dentro de la universidad. A continuación, se describen los componentes principales de este Marketplace de Datos:

- **Azure Event Hubs:** Se usaría para recopilar datos en tiempo real de diversas fuentes dentro de la universidad, como sensores de laboratorio, dispositivos IoT en el campus y sistemas de información académica.
- **Azure Databricks:** Sería utilizado para analizar grandes volúmenes de datos generados en la universidad, desde registros estudiantiles hasta datos de investigación, y para implementar modelos de machine learning que puedan mejorar la eficiencia y la toma de decisiones.
- **Azure Data Factory:** Automatizaría la preparación y transformación de datos, lo que permitiría integrar y limpiar datos de diferentes sistemas universitarios para su posterior análisis y almacenamiento.

- **Azure Data Lake Storage Gen1:** Proporcionaría un repositorio centralizado y seguro para almacenar todos los datos universitarios, tanto crudos como procesados, facilitando el acceso y la gestión de la información.
- **Azure Synapse Analytics:** Permitiría realizar análisis avanzados sobre los datos almacenados, desde la identificación de tendencias en el desempeño estudiantil hasta la optimización de recursos en la investigación académica.
- **Azure SQL Database:** Se usaría para almacenar metadatos, registros académicos y otros datos estructurados, proporcionando un almacenamiento seguro y de alto rendimiento para aplicaciones universitarias críticas.
- **Azure Machine Learning:** Sería aprovechado para desarrollar y desplegar modelos de análisis predictivo y de recomendación, ayudando a mejorar la experiencia educativa y la investigación en la universidad.
- **Azure Functions:** Se emplearían para automatizar tareas repetitivas o basadas en eventos, como la limpieza de datos, la generación de alertas sobre el rendimiento académico o la ejecución de flujos de trabajo específicos.
- **Azure Logic Apps:** Permitiría crear flujos de trabajo automatizados que conecten diferentes servicios y sistemas dentro de la universidad, mejorando la eficiencia operativa y la integración entre aplicaciones.
- **Azure App Service:** Sería utilizado para desarrollar e implementar la interfaz de usuario del Marketplace de Datos, proporcionando una experiencia intuitiva y accesible para los usuarios universitarios.
- **Azure Power BI Embedded:** Se integraría en el Marketplace de Datos para crear paneles interactivos y visualizaciones de datos, permitiendo a estudiantes,

profesores y administradores explorar y analizar la información de manera eficiente.

- **Microsoft Purview:** Se emplearía para establecer políticas de seguridad y privacidad, así como para garantizar el cumplimiento normativo en la gestión y el uso de los datos universitarios, asegurando la integridad y confidencialidad de la información.

En resumen, la integración de estos componentes dentro del Marketplace de Datos proporciona una solución completa que mejora la eficiencia y productividad de los procesos académicos, promoviendo la colaboración y la innovación en el ámbito universitario.

Plan de inversión e infraestructura:

La tecnología en la nube de Azure ofrece una solución rentable para la infraestructura tecnológica, con capacidad de escalabilidad y abastecimiento bajo demanda. Con un convenio universitario exclusivo, como el que tiene la Universidad Simón Bolívar, y licencias especializadas, los costos se reducen hasta un 30% en comparación con el modelo empresarial estándar. Esta propuesta destaca cómo Azure puede impulsar el crecimiento y la competitividad de la Universidad Simón Bolívar, maximizando el retorno de inversión mientras reduce los costos operativos.

Tabla 2 Costos de capacidad y abastecimiento en la nube de Azure

Proceso	Servicio	Región	Descripción	Costo Estimado Mensual en dólares
Ingesta, Procesamiento y Storage	Event Hubs	East US	Basic tier: 1 Throughput unit(s) x 730 Hours, 10 million Ingress events	\$10,95
	Azure Databricks	West US	All-Purpose Compute Workload, Premium Tier, 1 D3V2 (4 vCPU(s), 14 GB RAM) x 730 Hours, Pay as you go, 0.75 DBU x 730 Hours	\$504,80
	Azure Data Factory	East US	Azure Data Factory V2 Type, Data Pipeline Service Type,	\$500,00
	Azure Data Lake Storage Gen1	East US 2	Pay-as-you-go: 20 GB Storage, 5000 Read Transactions, 1000 Write Transactions	\$80,00

Aprendizaje Automatico	Azure Synapse Analytics	East US	Tier: Compute Optimized Gen2, Dedicated SQL Pools: DWU 100, 1 year reserved, 1 TB of storage with Geo-redundant disaster recovery; Este de EE. UU. Region, 100 GB of data collected per day, 7 days of Hot Cache, 30 days of total retention, 7 times estimated data compression, 730 Hours of 2 x Extra Small (2 vCores) Engine Instances, 730 Hours of 2 x 1 vCore Data Management Instances	\$1.601,86
	Azure SQL Database	East US	Single Database, vCore, General Purpose, Provisioned, Standard-series (Gen 5), Locally Redundant, 1 - 2 vCore Database(s) x 730 Hours, 32 GB Storage, RA-GRS Backup Storage Redundancy, 0 GB Point-In-Time Restore, 0 x 5 GB Long Term Retention	\$372,97
	Azure Machine Learning	East US 2	1 D3 v2 (4 Core(s), 14 GB RAM) x 730 Hours, Pay as you go	\$167,17
Autoservicio y Autogestión	Azure Functions	West US	Consumption tier, Pay as you go, 128 MB memory, 100 milliseconds execution time, 0 executions/mo	\$0,00

	Logic Apps	West US	Workloads: Standard plan, 1 WS1 (1 vCores, 3.5 GB RAM) x 730 Hours, 0 Standard Connector Calls per day x 1 day, 0 Enterprise Connector Calls per day x 1 day; Integration Service Environment: Premium tier, 0 Base Units x 730 Hours, 0 Scale Units x 730 Hours; Integration Accounts: 0 Standard Integration Accounts x 730 Hours, 0 Basic Integration Accounts x 730 Hours.	\$180,97
	App Service	West US	Basic Tier; 1 B1 (1 Core(s), 1.75 GB RAM, 10 GB Storage) x 730 Hours; Windows OS; 0 SNI SSL Connections; 0 IP SSL Connections; 0 Custom Domains; 0 Standard SLL Certificates; 0 Wildcard SSL Certificates	\$54,75
	Power BI Embedded	East US 2	1 node(s) x 730 Hours, Node type: A1, 1 Virtual Core(s), 3GB RAM, 1-300 Peak renders/hour	\$735,91
Gobierno de Datos	Microsoft Purview	East US	Elastic Data Map: 1 Capacity Unit hour, 730 hours, Automated Scanning and Classification: 1 Total scan duration in minute x 32 Total	\$453,67

			vCores across scans (For other data sources), Other features: 730 Resources Set hours, Microsoft Purview Data Catalog: C0 Service, Microsoft Purview Data Policy: 0 Policies x 730 Hours	
Seguridad e Identidad	Azure Active Directory External Identities	West US	Premium P1 tier: 50,000 monthly active user(s), 0 SMS/Phone Events	\$0,00
	API Management	East US	Developer tier, 1 unit(s), 730 Hours, 0 x 730 Hours x 5 overage workspaces	\$48,03
Support		Support Licensing Program	Microsoft Education (30% off)	\$0,00
		Total		\$3.732,83

Fuente: Elaboración propia, basado de la calculadora de Azure (AZURE, 2024)

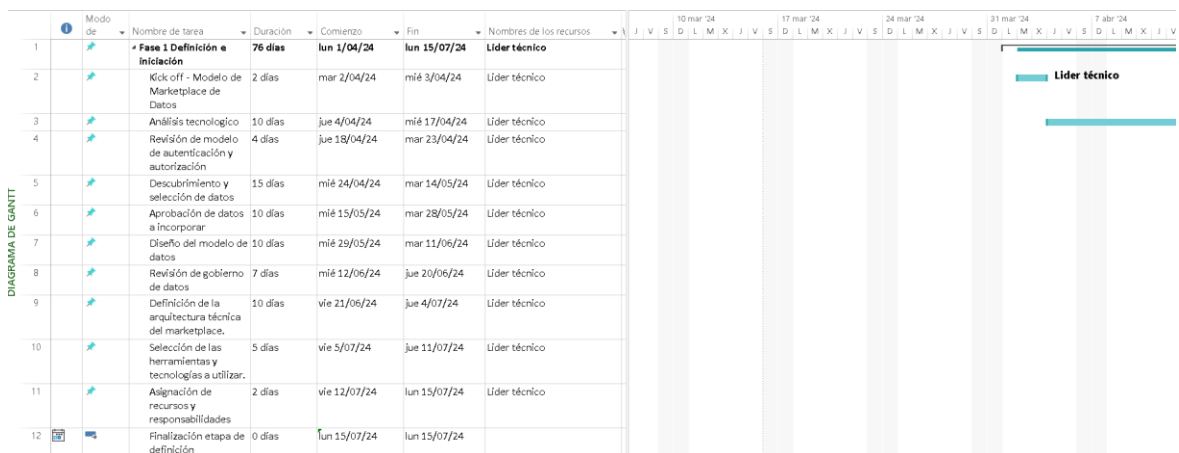
Esta tabla proporciona un desglose de los diferentes servicios de Azure y sus costos mensuales estimados para la Universidad Simón Bolívar. Los costos específicos variarán dependiendo de los requisitos y la escala del uso de los servicios.

Cronograma

Las ilustraciones 21,22 y 23 que se encuentra a continuación, presentan el cronograma del proyecto aplicado a la universidad Simón Bolívar, que detalla las actividades planificadas y su secuencia temporal. Este cronograma proporciona una visión general de la planificación y ejecución de las diversas etapas del proyecto, desde su inicio hasta su conclusión.

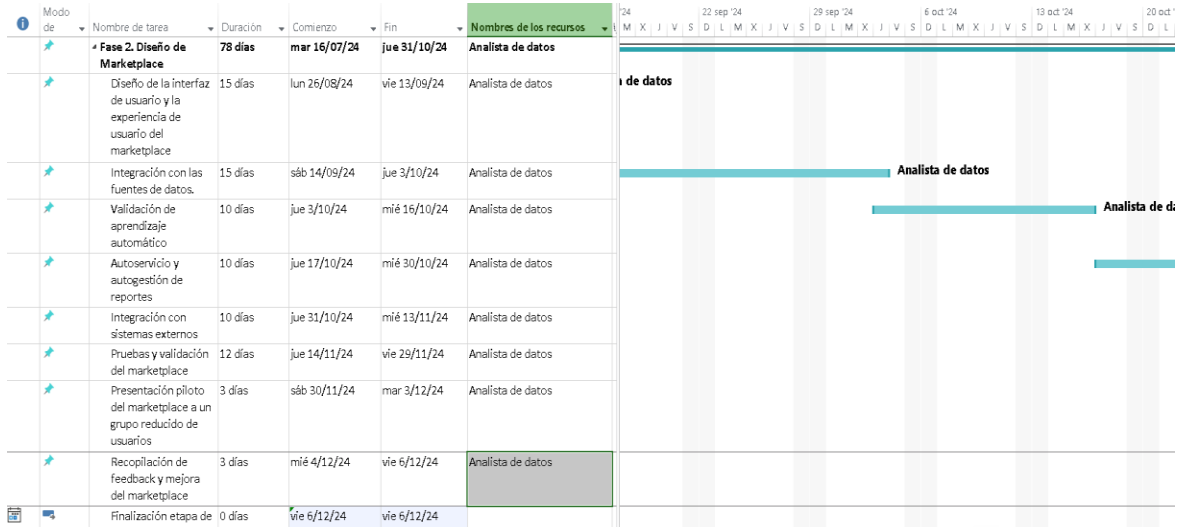
El cronograma se ha configurado para garantizar la claridad y la coherencia en la representación de las actividades y los roles a lo largo de todas las fases del proyecto. Cada fase del proyecto, desde la definición e iniciación hasta la operación, se ha detallado exhaustivamente, incluyendo todas las tareas planificadas, roles específicos y responsabilidades asociadas. Además, se han identificado y marcado los hitos clave que marcan puntos críticos de progreso; esto ha permitido proporcionar una visión técnica integral de la planificación y ejecución del proyecto, asegurando una asignación clara de tareas y plazos para el equipo.

Ilustración 21. Cronograma del proyecto fase de definición



Nota: Elaboración propia basado en las actividades a ejecutar.

Ilustración 22 Cronograma del proyecto fase de diseño



Nota: Elaboración propia basado en las actividades a ejecutar.

Ilustración 23 Cronograma del proyecto fase de operación



Nota: Elaboración propia basado en las actividades a ejecutar.

10. Recomendaciones y Conclusiones

A continuación, se presentan las recomendaciones para la implementación del plan de intervención propuesto y las conclusiones de cierre del trabajo.

10.1. Recomendaciones

Las recomendaciones que se establecen a continuación se dan teniendo en cuenta el contexto de la Universidad Simón Bolívar y el proyecto de diseño de un Modelo de Marketplace de Datos utilizando herramientas de Microsoft 365 y Azure.

- **Exploración de Frameworks de Inteligencia Artificial:** Se recomienda investigar y evaluar la viabilidad de integrar frameworks de inteligencia artificial adicionales al Modelo de Marketplace de Datos. Esto podría incluir el análisis de frameworks como TensorFlow, PyTorch o Apache MXNet para mejorar la capacidad predictiva y analítica del sistema, así como para explorar nuevas posibilidades de automatización y optimización de procesos dentro del Marketplace.
- **Análisis de Impacto Socioeconómico:** Se sugiere realizar un estudio exhaustivo sobre el impacto socioeconómico del Modelo de Marketplace de Datos en la comunidad universitaria y en la región circundante. Esto podría implicar la evaluación de cómo el acceso a datos y herramientas analíticas influye en la toma de decisiones, la innovación y el desarrollo local. Además, sería beneficioso investigar cómo el Marketplace podría contribuir al crecimiento económico y la competitividad de la universidad y su entorno.
- **Integración de Aspectos de Ciberseguridad:** Dada la importancia de la seguridad de los datos en cualquier plataforma tecnológica, se recomienda explorar la integración de medidas de ciberseguridad avanzadas en el Modelo

de Marketplace de Datos. Esto podría incluir la implementación de protocolos de encriptación robustos, la detección proactiva de amenazas y vulnerabilidades, así como la educación continua sobre buenas prácticas de seguridad para los usuarios finales.

- **Estudio de Impacto Ambiental:** Se sugiere llevar a cabo un análisis del impacto ambiental asociado con la implementación y operación del Modelo de Marketplace de Datos. Esto implica evaluar el consumo de recursos, como energía y hardware, así como los efectos indirectos en términos de huella de carbono y gestión de residuos electrónicos. Identificar oportunidades para optimizar la eficiencia energética y reducir el impacto ambiental podría contribuir a la sostenibilidad del proyecto a largo plazo.

10.2. Conclusiones

El proyecto del Modelo de Marketplace de Datos para la Universidad Simón Bolívar, mediante el uso y aplicación de herramientas de Microsoft 365 y Azure, ha alcanzado con éxito los objetivos específicos establecidos, lo que ha resultado en un avance significativo en la gestión de datos académicos y el impulso de la innovación tecnológica en la institución.

El primer objetivo, de definir una arquitectura robusta para el Marketplace de datos, se logró de manera exitosa mediante la utilización de herramientas y servicios disponibles en Microsoft 365 y Azure. Esta arquitectura proporciona una sólida base para el almacenamiento, procesamiento y análisis de datos, asegurando la escalabilidad y eficiencia del sistema.

En relación con el segundo objetivo, se realizó un análisis comparativo de plataformas que tienen Marketplaces de datos en otras universidades del Atlántico, lo que permitió identificar la oportunidad única de implementar y desarrollar sistemas innovadores que impulsen la colaboración y el intercambio de datos entre instituciones académicas, industrias y la comunidad en general.

Tabla 3. Tabla comparativa de universidades y plataformas

Criterio	Universidad Del Norte	Universidad de la Costa	Universidad del Atlántico	Universidad Libre
Plataforma utilizada	Banner	Emma	Academusoft	Siul
Marketplace de Datos	NO	NO	NO	NO
Facilidades	Interfaz intuitiva y sencilla, navegación fácil, proceso de registro rápido.	Interfaz moderna y atractiva, navegación fluida, proceso de registro simplificado.	Interfaz amigable y accesible, navegación rápida, proceso de registro automático.	Interfaz intuitiva y personalizable, navegación dinámica, proceso de registro automático
Diseño y usabilidad	Diseño limpio y minimalista, colores agradables, buena organización de la información.	Diseño moderno y profesional, colores vibrantes, experiencia de usuario fluida.	Diseño atractivo y funcional, colores corporativos, accesibilidad para usuarios con discapacidad.	Diseño elegante y vanguardista, personalización de la interfaz, experiencia de usuario altamente intuitiva.
Seguridad y confiabilidad	Protección de datos personales, sistema de pagos seguro, verificación de identidad de usuarios.	Encriptación de datos, sistema de pagos confiable.	Múltiples niveles de seguridad, sistema de pagos dinámico	Protección contra fraudes, Sistema de pagos confiable.

Soporte técnico	Soporte por correo electrónico, chat en línea y teléfono durante el horario laboral.	Soporte por correo electrónico, chat en línea.	Soporte por correo electrónico, chat en línea y redes sociales.	Soporte por correo electrónico, chat en línea, redes sociales.
-----------------	--	--	---	--

Fuente Elaboración propia, consulta paginas universidades

El tercer objetivo, referente al diseño para acceder a los datos en modalidad de autogestión y autoservicio, se materializó mediante la propuesta de un enfoque que promueve la democratización de la información dentro de la universidad. Al proporcionar a los usuarios finales la capacidad de acceder y analizar datos de manera independiente, se fomenta la transparencia y la colaboración en la toma de decisiones académicas y administrativas.

Por otro lado, la definición del esquema de integraciones y ambientes para la gestión de conocimiento cumplió con el cuarto objetivo, garantizando la interoperabilidad y coherencia en el intercambio de información dentro del Marketplace de datos. Esta estructura facilita la integración de datos provenientes de diversas fuentes y sistemas, promoviendo una visión unificada y completa de la información académica y administrativa.

En conjunto, estos logros demuestran el compromiso de la Universidad Simón Bolívar con la excelencia en la gestión de datos y la innovación tecnológica. Se espera que el Marketplace de Datos no solo mejore los procesos internos de la institución, sino que también fomente la investigación colaborativa y el desarrollo de soluciones basadas en datos en el ámbito universitario.

El diseño de un Marketplace de Datos en la Universidad Simón Bolívar representa una oportunidad significativa para optimizar la gestión de datos dentro de la institución. Este proyecto se ha enfocado en explorar y contextualizar el diseño de un Modelo de Marketplace de Datos específicamente adaptado a las necesidades de la Universidad Simón Bolívar, aprovechando las herramientas y servicios disponibles en Microsoft 365 y Azure.

Durante este proceso, se ha realizado un exhaustivo examen de modelos de componentes técnicos y buenas prácticas asociadas al Marketplace de Datos, así como de sus beneficios, limitaciones y desafíos en el contexto universitario. El resultado ha sido el desarrollo de un modelo personalizado que tiene en cuenta aspectos técnicos, la disponibilidad de fuentes de datos y las expectativas relacionadas con el potencial educativo y administrativo de la institución.

El Marketplace de Datos proporciona un entorno para la experimentación con los datos, haciendo uso de diversas técnicas y herramientas disponibles en Microsoft 365 y Azure. Esto permite descubrir el valor único de los datos y proporcionar información crucial para la toma de decisiones académicas y administrativas, así como para la mejora continua de los procesos universitarios.

Es importante destacar que el modelo propuesto es adaptable y evolutivo, ajustándose a medida que la universidad cambia y surgen nuevas necesidades. Se espera que este proceso continuo de exploración genere nuevas ideas y oportunidades para mejorar tanto los procesos académicos como administrativos de la institución.

Finalmente, cada etapa del proyecto implica la adopción de decisiones estratégicas destinadas a brindar datos diferenciados a los usuarios finales, incluyendo profesores, estudiantes y personal administrativo, con el fin de enriquecer tanto la experiencia académica como la administrativa en la institución.

La implementación del modelo propuesto, respaldada por la eficacia de los componentes de Azure, demuestra su capacidad para mejorar significativamente la integración de datos en la universidad Simón Bolívar, la eficiencia que demuestran los servicios de Azure, permite evidenciar que el Marketplace facilitaría la consolidación y gestión de datos de manera más efectiva y rentable para la universidad.

Al utilizar los servicios de Azure, la universidad puede optimizar sus recursos y reducir los costos asociados con la infraestructura y el mantenimiento de sistemas de información fragmentados. Además, la centralización de datos promovida por el Marketplace aumentaría la accesibilidad y la coherencia de la información, beneficiando tanto a la administración como a la comunidad educativa en general. En resumen, la adopción del modelo propuesto representa una solución práctica y rentable para mejorar la gestión de datos en la universidad.

11. Referencias

Banafa, A. (2021, June 9). ¿Qué es un lago de datos? Ventajas y contras. OpenMind.

<https://www.bbvaopenmind.com/tecnologia/mundo-digital/un-lago-de-datos-una-oportunidad-o-un-sueno-para-el-big-data/>

Beltrán García, V. E. (2022). Implementación de un Data Lake de los datos de uso del LMS Canvas de la Universidad Internacional SEK. Repositorio.uisek.edu.ec.

<https://repositorio.uisek.edu.ec/handle/123456789/4673>

Bermúdez Durana, J. A. (2017). Privacidad y protección de datos personales. Bogotá, Colombia: Leyer.

Brockwell, P. J., & Davis, R. A. (2013). Introduction to Time Series and Forecasting. Springer Science & Business Media.

Castillo, R., & Moreno, F. E. (2019). Big Data e inteligencia de negocios aplicados al estudio de mercado de estudios posgraduales de la Universidad Autónoma de Colombia. Revista ESPACIOS, 40(15).

<https://www.revistaespacios.com/a19v40n15/19401508.html> .

Centro de documentos. (2019). Unisimon.edu.co.

<https://www.unisimon.edu.co/servicios/centrodedocumentos>

Chang, K. (2013). Visualización de datos y diseño de información: introducción a la visualización de datos. Barcelona: Gustavo Gili.

Cognodata. (2022, January 14). Arquitectura de datos: la base de una estrategia diferenciadora - Cognodata. Cognodata. <https://www.cognodata.com/arquitectura-datos-estrategia-diferenciadora/>

DataMarketplace. (2021, 1 junio). The Essential Elements of a Data Marketplace.

Datos gobernados y confiables. (n.d.). Help.tableau.com. Retrieved March 7, 2024, from

https://help.tableau.com/current/blueprint/es-es/bp_governance_enables.htm

Esteban Nieto, N. (2018). Tipos de Investigación. Universidad Santo Domingo de Guzmán. <http://repositorio.usdg.edu.pe/handle/USDG/34>

Fernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, M. del P. (2014). Metodología de la investigación (6a. ed.). México D.F.: McGraw-Hill.

García, A. J., & Vásquez, D. A. (2019). Plataforma tecnológica de datos para la gestión de información en empresas. *Ingenium*, 13(27), 9-16.

Gutiérrez, D. K. (2018). Diseño e implementación de un data Warehouse para la creación de informes gerenciales. Recuperado de: <http://hdl.handle.net/11371/1571>

Inmon, W.H., & Kimball, R. (2002). *The Data Warehouse and ETL Toolkit: Practical Techniques for Extracting, Cleaning, Conforming, and Delivering Data*. John Wiley & Sons.

International Association of Privacy Professionals (2021). Data Providers. Recuperado de <https://iapp.org/resources/article/data-providers/>

Investopedia (s.f.). Data Provider. Recuperado de <https://www.investopedia.com/terms/d/data-provider.asp>

ISO/IEC. (2018). ISO/IEC 29100:2018 Information technology -- Security techniques -- Privacy framework. <https://www.iso.org/standard/45170.html>

Jolliffe, I. T. (2002). *Principal component analysis*. John Wiley & Sons.

Larose, D. T., & Larose, C. D. (2019). *Discovering knowledge in data: an introduction to data mining* (3rd ed.). John Wiley & Sons.

Leedy, P. D., & Ormrod, J. E. (2013). *Practical research: Planning and design* (10th ed.). Boston, MA: Pearson.

Secretaría del Senado de la República de Colombia. (1992). *Leyes desde 1992*.

(http://www.secretariassenado.gov.co/senado/basedoc/ley_0030_1992.html)

Manrique, G. A. (2019). Tipos de analítica. [Podcast]. Repositorio Institucional UNAD. <https://repository.unad.edu.co/handle/10596/40616>

Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to linear regression analysis*. John Wiley & Sons.

ORACLE. (2023). Que es el aprendizaje automático. Obtenido de Que es el aprendizaje automático: <https://www.oracle.com/co/artificial-intelligence/machine-learning/what-is-machine-learning/>

Piethinstrengholt. (2023, May 9). Marketplace de datos - Cloud Adoption Framework.

Learn.microsoft.com. <https://learn.microsoft.com/es-es/azure/cloud-adoption-framework/scenarios/cloud-scale-analytics/architectures/data-mesh-data-marketplace>

Pin-Carreño, B. J., Valencia-Riasco, S. A., & Cedeño-Cedeño, C. E. (2022). Sistema de reportes gerenciales basados en almacenes de datos, para el proceso de matriculación del Sistema de Gestión Académico de la Universidad Laica Eloy Alfaro de Manabí. *Revista Científica De Informática ENCRIPtar* - ISSN: 2737-6389., 5(9), 25–38. Recuperado a partir de

<https://publicacionescd.uleam.edu.ec/index.php/encriptar/article/view/336>

Proyecto de migración de datos hacia un Data Lake para una entidad de seguros PDF Online. (2022, July 12). Tesis Del Perú. <https://www.tesisdelperu.com/proyecto-de-migracion-de-datos-hacia-un-data-lake-para-una-entidad-de-seguros-3122766>

Quero, O. (2020). *Huge Data: la gestión masiva de los datos*. Barcelona: OBS Business School.

Rodal, E. (2020, 14 octubre). Principales etapas de los procesos de analítica de datos. Podcast Industria 4.0. <https://www.podcastindustria40.com/analitica-de-datos-industria/>

Ruiz, C., & De Anjana Data. (n.d.). Cómo construir un data Marketplace para tu organización artículo data Marketplace septiembre 2021 “cómo construir de forma exitosa un data Marketplace interno para tu organización.” Retrieved March 7, 2024, from <https://anjanadata.com/wp-content/uploads/2021/09/DATA-MARKETPLACE-ESPAN%CC%83OL.pdf>

Solove, D. J., & Schwartz, P. M. (2011). Política de privacidad y protección de datos. Editorial Atelier.

Statista (2021). Leading data providers worldwide in 2020, by revenue. Recuperado de <https://www.statista.com/statistics/865430/leading-data-providers-worldwide-by-revenue/>

UNIR. (2022). ¿Qué es el big data y cuál es su importancia? Obtenido de ¿Qué es el big data y cuál es su importancia?: <https://colombia.unir.net/actualidad-unir/que-es-big-data/>

Universidad de la Costa CUC. (s. f.). *Universidad de la Costa CUC - Inicio*. <https://www.cuc.edu.co/>

inicio. (s. f.). Uninorte. <https://www.uninorte.edu.co/>

Universidad Libre - Seccional Barranquilla. (s. f.). <https://www.unilibrebaq.edu.co/>

Universidad del Atlántico. (2024, 9 abril). *Universidad del Atlántico - ¡La mejor educación al alcance de todos!* Universidad del Atlántico - ¡La Mejor Educación Al Alcance de Todos! <https://www.uniatlantico.edu.co/>

A. Anexo. Nombre del anexo

Incluya tantos anexos como requiera. Un anexo es un documento que complementa el trabajo generando valor para el lector. Ejemplo: formatos de encuesta, informes de soporte, documentos institucionales, etc. Lístelos, con sus respectivos nombres, en páginas independientes como A, B, C, etc.

